Biological and Medical Physics, Biomedical Engineering

Heinz Fabian
Dieter Naumann *Editors*

Protein Folding and Misfolding

Shining Light by Infrared Spectroscopy

Springer

BIOLOGICAL AND MEDICAL PHYSICS, BIOMEDICAL ENGINEERING

For further volumes:
http://www.springer.com/series/3740

BIOLOGICAL AND MEDICAL PHYSICS, BIOMEDICAL ENGINEERING

The fields of biological and medical physics and biomedical engineering are broad, multidisciplinary and dynamic. They lie at the crossroads of frontier research in physics, biology, chemistry, and medicine. The Biological and Medical Physics, Biomedical Engineering Series is intended to be comprehensive, covering a broad range of topics important to the study of the physical, chemical and biological sciences. Its goal is to provide scientists and engineers with textbooks, monographs, and reference works to address the growing need for information.

Books in the series emphasize established and emergent areas of science including molecular, membrane, and mathematical biophysics; photosynthetic energy harvesting and conversion; information processing; physical principles of genetics; sensory communications; automata networks, neural networks, and cellular automata. Equally important will be coverage of applied aspects of biological and medical physics and biomedical engineering such as molecular electronic components and devices, biosensors, medicine, imaging, physical principles of renewable energy production, advanced prostheses, and environmental control and engineering.

Heinz Fabian
Dieter Naumann
Editors

Protein Folding and Misfolding

Shining Light by Infrared Spectroscopy

With 108 Figures

Springer

Editors:
Dr. Heinz Fabian
Professor Dr. Dieter Naumann
Robert Koch Institut
Nordufer 20, 13353 Berlin, Germany
E-mail: fabianH@rki.de, naumannD@rki.de

Biological and Medical Physics, Biomedical Engineering ISSN 1618-7210
ISBN 978-3-642-22229-0 e-ISBN 978-3-642-22230-6
DOI 10.1007/978-3-642-22230-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011937438

Cover design: eStudio Calamar Steinen

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The progress in understanding protein folding and misfolding is primarily due to the development of biophysical methods, which permit to probe conformational changes with high kinetic and structural resolution. A whole battery of techniques is being used to address the fundamental problems of protein folding and misfolding. The most common approaches rely on rapid-mixing methods to initiate the folding event via a sudden change in solvent conditions. Traditionally, techniques such as fluorescence, circular dichroism or visible absorption spectroscopy are applied to study the processes. In contrast to these techniques, infrared spectroscopy came into play only very recently. The significant progress made in this field to date permits to follow folding events over the timescale from picoseconds to minutes with high structural resolution. The aim of this unique book is to provide an overview of the latest developments and applications as seen by pioneers in this burgeoning field. The various chapters present representative examples on the sort of information which infrared techniques can provide and how this information is extracted from the experimental data. The discussion of the state-of-art technology, data evaluation strategies and representative applications on protein folding and misfolding should help the readers to estimate whether their particular systems are appropriate to be studied by infrared spectroscopy, and to assess the specific advantages the various infrared techniques have.

This book contains nine chapters. The introductory chapter by Gareth Morgan and Sheena Radford focuses on the array of experimental methods that are presently applied to the key questions of how folding, misfolding and aggregation of proteins are linked, both in vitro and in the environment of the cell. The second chapter by Joseph Brauner and Richard Mendelsohn presents a semi-empirical method of simulating the experimental amide I contour of a protein or peptide molecule whose atomic coordinates are available, and discusses the correlations between the amide I contour and the secondary structure of a protein. The adaptation of conventional mixing and temperature-jump technologies to the specific requirements of time-resolved FTIR spectroscopy, which enable to explore protein folding and misfolding events on the millisecond-to-minute timescale, together with

representative results on different proteins, is then described by Heinz Fabian and Dieter Naumann. The fourth chapter by Satoshi Takahashi and Tetsunari Kimura gives an overview on time-resolved FTIR spectroscopy based on continuous-flow rapid-mixing set-ups, which allow to follow protein folding events in the sub-millisecond-to-second time range. The authors describe practical issues in applying their devices to explore mechanism of secondary structure formation and protein main chain dehydration. Chapter 5 by Roland Winter and co-workers reports on pressure changes as an alternative trigger to unfold or refold proteins and to induce disaggregation of misfolded species. After describing the experimental techniques, examples of pressure-induced un- and refolding reactions of proteins as well as studies on enzyme reactions are presented. The use of laser-induced temperature-jump IR spectroscopy as a method to study α-helix and β-sheet formation in the nanosecond-to-microsecond time range is presented by Karin Hauser, with emphasis on strategies to obtain insights into folding mechanism on the level of single amino acid residues. The seventh chapter by Wolfgang Zinth and Josef Wachtveitl demonstrates that photo-switches incorporated into suitably designed amino acid sequences open up numerous new applications by applying light as trigger to initiate peptide folding. Their pioneering investigations on selected light-triggered peptides demonstrate ultrafast folding reactions, and show that these processes may span the range between picoseconds and tens of microseconds. Another way to trigger unfolding and misfolding events by light is the use of caged compounds, which is described next by Andreas Barth and co-workers with special focus on the use of caged protons for time-resolved infrared spectroscopic experiments. The light-induced release of protons generates a pH jump much more rapidly than any conventional mixing technique, thus paving the way for investigating early events of aggregation processes of peptides or proteins. The final chapter by Martin Zanni and co-workers illustrates that two-dimensional infrared spectroscopy combined with isotope labelling is an elegant and powerful tool to obtain residue-specific structural information on folding and aggregation processes of peptides and proteins. A mathematical formalism to guide the interpretation of one- and two-dimensional IR spectra of amyloid fibrils is presented, which enables the design of the best isotope labelling scheme of the peptide. The chapter ends with explanatory experiments, demonstrating the specific power of the two-dimensional IR approach.

This book is the result of the work of many colleagues who generously agreed to contribute to this book by taking time away from their other responsibilities. We wish to thank all the authors for their extremely valuable contributions. We hope that their ideas and experiences will be of interest not only for those readers already familiar with infrared spectroscopic techniques, but also inspire other colleagues in the protein community to take advantage of the possibilities described herein for

their particular research in the future. Special thanks go to our co-worker Angelika Brauer for her great help and continuous encouragement during the technical preparation of the book chapters.

Berlin, Germany *Heinz Fabian*
July 2011 *Dieter Naumann*

Contents

1 Linked Landscapes and Conformational Conversions: How Proteins Fold and Misfold 1
Gareth J. Morgan and Sheena E. Radford
1.1 Introduction 1
1.2 The Unfolded Ensemble Under Native Conditions 4
1.3 Folding and Misfolding Intermediates 6
1.4 Protofibrils, Oligomers and Toxicity 8
1.5 Amyloid Structure 9
1.6 From the Test Tube to the Cell 11
1.7 Conclusions 13
References 13

2 A Quantitative Reconstruction of the Amide I Contour in the IR Spectra of Peptides and Proteins: From Structure to Spectrum 17
Joseph W. Brauner and Richard Mendelsohn
2.1 The Approach to Simulation of the Amide I Contour 17
2.1.1 Introduction 17
2.1.2 Historical Background 18
2.1.3 Normal Coordinate Calculations 19
2.1.4 Ab Initio Force Field Calculations 20
2.1.5 The Modified GF Matrix Method 21
2.1.6 Constructing the G and F Matrices in the Coupled Oscillators of One Kind Method 22
2.1.7 Simulating the Amide I Contour 31
2.2 Applications 32
2.2.1 Isotopic Labeling 33
2.2.2 Modeling the Early Stages of Thermal Denaturation 39
2.2.3 Amide I Structure-Frequency Correlations in Globular Proteins 41
2.2.4 IRRAS Simulations 45

2.3 Conclusions and Future Prospects 50
References 51

3 Millisecond-to-Minute Protein Folding/Misfolding Events Monitored by FTIR Spectroscopy 53
Heinz Fabian and Dieter Naumann
3.1 General Considerations 53
3.2 FTIR Spectroscopy, Experimental Aspects 55
3.2.1 Proteins in Aqueous Solutions 55
3.2.2 Measurements in D_2O 56
3.2.3 FTIR Spectra of Chemical Denaturants 60
3.3 Kinetic FTIR Experiments Applying Rapid Mixing and Temperature-Jump Approaches 61
3.3.1 Rapid-Scan FTIR Spectroscopy: Advantages and Limitations 61
3.3.2 Design and Operation of a Stopped-Flow Apparatus for Measurements in Heavy Water 62
3.3.3 A Stopped-Flow Apparatus for Measurements of H_2O-Protein Solutions 63
3.3.4 T-Jump Experiments in Heavy Water 64
3.4 Examples of Applying T-Jumps onto a Protein Solution 66
3.4.1 Refolding of Wild-Type Ribonuclease T1 and Some of Its Mutants 66
3.4.2 Unfolding of the λ-Cro Repressor 72
3.5 Examples Making Use of Rapid-Mixing Methods 74
3.5.1 Refolding of α-Lactalbumin Studied by Stopped-Flow Infrared Spectroscopy After a pH-Jump 74
3.5.2 Misfolding of β_2-Microglobulin 78
3.5.3 The α-to-β Conversion Process of the Prion Protein 85
References 87

4 Watching Dynamical Events in Protein Folding in the Time Domain from Submilliseconds to Seconds: Continuous-Flow Rapid-Mixing Infrared Spectroscopy 91
Satoshi Takahashi and Tetsunari Kimura
4.1 Introduction 91
4.2 The Collapse and Search Mechanism of Protein Folding 92
4.2.1 The Protein Folding Mechanism Depends on the Chain Length 92
4.2.2 Kinetic Investigation of Protein Folding for Intermediate Proteins 93
4.3 Development of Continuous-Flow Time-Resolved Infrared Spectrometer 94
4.3.1 Comparison of Different Methods for Triggering Protein Folding Events 94

4.3.2 Development of a Continuous-Flow Cell with a T-Shaped Flow Channel 95
4.3.3 Construction of the Time-Resolved Spectrometer Based on Infrared Microscopy 98
4.4 Practical Issues for Kinetic Infrared Investigations of Protein Folding 101
4.4.1 Selection of the Initial Unfolded State 101
4.4.2 Suppression of the Aggregate Formation 101
4.4.3 Method of Spectral Analysis 102
4.5 Application to Protein Folding 103
4.5.1 Pioneering Investigations of Rapid-Mixing Infrared Spectroscopy 103
4.5.2 Apomyoglobin 104
4.5.3 Single-Chain Monellin 107
4.6 Summary and Perspective 110
References 113

5 High-Pressure Vibrational Spectroscopy Studies of the Folding, Misfolding and Amyloidogenesis of Proteins 117
Roland Winter, Matthias Pühse, and Jonas Markgraf
5.1 Introduction to High-Pressure Bioscience 117
5.2 Fundamental Concepts: Stability Diagram of Proteins 119
5.3 Experimental Methods 120
5.3.1 High-Pressure FTIR Spectroscopy 120
5.3.2 Diamond Anvil Cell Technology 122
5.3.3 Pressure Calibrants for Infrared Spectroscopy 123
5.4 Examples of Pressure Studies on Proteins and Polymers 124
5.4.1 Pressure-Induced Protein un- and Refolding Reactions 124
5.4.2 Protein Folding Kinetics 127
5.4.3 Pressure-Assisted Cold Denaturation of Proteins 130
5.4.4 Pressure Effects on Oligomeric Proteins and Chaperones 132
5.4.5 Cosolvent Effects 133
5.4.6 Aggregation/Fibrillation Reactions of Proteins 135
5.4.7 Enzymatic Reactions 138
5.4.8 Synthetic Polymers as Protein Mimetics 139
5.5 Conclusions and Outlook 143
References 144

6 Dynamics of α-Helix and β-Sheet Formation Studied by Laser-Induced Temperature-Jump IR Spectroscopy 147
Karin Hauser
6.1 Peptide Folding Dynamics 147
6.1.1 Secondary-Structure Formation 147
6.1.2 The Amide I Band as Structural Probe 148
6.1.3 Equilibrium vs. Kinetic Data 149
6.1.4 Rate Constants 150

6.2 Laser-Induced T-Jump Technique 151
6.2.1 Generation of the Heating Pulse 152
6.2.2 Photo-Acoustic Effects, Cavitation and Thermal Lensing 156
6.2.3 Experimental Setup 157
6.3 T-Jump Relaxation Kinetics 158
6.3.1 Two-State and Multistate Folders 158
6.3.2 Helix Dynamics 161
6.3.3 Hairpin Formation 161
6.4 Site-Specific Dynamics with Isotopic Editing 162
6.4.1 Site-Specific Frequency Shifts 162
6.4.2 Insights into Folding Mechanisms on the Residue Level 163
6.4.3 Single and Multiple Isotope Labels 166
References 168

7 Light-Triggered Peptide Dynamics 171
Wolfgang Zinth and Josef Wachtveitl
7.1 Introduction 171
7.2 Light-Triggered Peptides 172
7.2.1 The Photochromic Switching Unit 172
7.2.2 The Linking Group 175
7.2.3 The Peptide Moiety 177
7.3 Characterization of Light-Triggered Peptides by Stationary Spectroscopy 177
7.4 Methods for the Study of Ultrafast Structural Dynamics 180
7.5 Applications 183
7.5.1 Ultrafast Spectroscopy on Cyclic Azobenzene Peptides 183
7.5.2 Unfolding and Folding of a Light Switchable Hairpin Model Compound 184
7.5.3 Toward Light Switchable Tertiary Structures: (I) Azo-maquettes 185
7.5.4 Toward Light Switchable Tertiary Structures: (II) Azo-collagens 187
7.6 Conclusion 190
References 191

8 Time-Resolved FTIR Spectroscopy of pH-Induced Aggregation of Peptides 193
John E.T. Corrie, Alex Perálvarez-Marín, and Andreas Barth
8.1 Introduction to Infrared Difference Spectroscopy 193
8.1.1 Principles 193
8.1.2 Triggering Protein Reactions 194
8.1.3 Interpreting Difference Spectra 196
8.2 Caged Compounds 198
8.2.1 Introduction to Caged Compounds 198
8.2.2 Caged Protons 202
8.2.3 Difference Spectrum of Caged Sulfate Photolysis 204

8.3 Acidification-Induced Unfolding of Myoglobin 206
8.4 Acidification-Induced Aggregation of the Alzheimer's Peptide 207
8.4.1 Introduction to the Alzheimer's Peptide 207
8.4.2 Time-Resolved Infrared Difference Spectroscopy of the Aggregation of the Alzheimer's Peptide 208
8.5 Outlook 212
References 213

9 Examining Amyloid Structure and Kinetics with 1D and 2D Infrared Spectroscopy and Isotope Labeling 217
Lauren E. Buchanan, Emily B. Dunkelberger, and Martin T. Zanni
9.1 Introduction 217
9.2 Vibrational Modes of Amyloids 220
9.3 Isotope Labeling Schemes 225
9.4 Vibrational Dynamics of Amyloids 228
9.5 Experimental Methods 229
9.6 Experimental Data 230
9.7 Summary 236
References 236

Index 239

Contributors

A. Barth Stockholm University, Department of Biochemistry and Biophysics, Arrhenius Laboratories, 10691 Stockholm, Sweden, barth@dbb.su.se

Joseph W. Brauner Rutgers University, Newark College, 73 Warren Street, Newark, NJ 07102, USA, matjoe63@verizon.net

Lauren E. Buchanan Department of Chemistry, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, USA, lbuchanan@chem.wisc.edu

John E.T. Corrie MRC National Institute for Medical Research, The Ridgeway Mill Hill, London NW7 1AA, UK, john.corrie@btinternet.com

Emily B. Dunkelberger Department of Chemistry, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, USA, eblanco@chem.wisc.edu

Heinz Fabian Robert Koch-Institute, Biomedical Spectroscopy, Nordufer 20, 13353 Berlin, Germany, FabianH@rki.de

Karin Hauser Department of Chemistry, University of Konstanz, Universitätsstr. 10, 78464 Konstanz, Germany, Karin.Hauser@uni-konstanz.de

Tetsunari Kimura Institute for Molecular Science, 38 Nishigo-Naka, Myodaiji, Okazaki 444-8585, Japan, tkimura@ims.ac.jp

Jonas Markgraf Faculty of Chemistry, Physical Chemistry I – Biophysical Chemistry, TU Dortmund University, Otto Hahn Str. 6, 44227 Dortmund, Germany, jonasmarkgraf@gmx.de

Richard Mendelsohn Rutgers University, Newark College, 73 Warren Street, Newark, NJ 07102, USA, mendelso@andromeda.rutgers.edu

Gareth J. Morgan University of Leeds, Mount Preston Street, Leeds LS2 9JT, UK, garethjmorgan@gmail.com

Dieter Naumann Robert Koch-Institute, Biomedical Spectroscopy, Nordufer 20, 13353 Berlin, Germany, NaumannD@rki.de

Alex Perálvarez-Marín Harvard University, 52 Oxford Street, NorthWest Building, Cambridge, MA 02138, USA, peralvarezmarin@gmail.com

Matthias Pühse Faculty of Chemistry, Physical Chemistry I – Biophysical Chemistry, TU Dortmund University, Otto Hahn Str. 6, 44227 Dortmund, dr.matthias.puehse@googlemail.com

Sheena E. Radford University of Leeds, Mount Preston Street, Leeds LS2 9JT, UK, s.e.radford@leeds.ac.uk

Satoshi Takahashi Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan, st@tagen.tohoku.ac.jp

Josef Wachtveitl Institute of Physical and Theoretical Chemistry, Goethe-University Frankfurt/Main, Max von Laue-Straße 7, 60438 Frankfurt am Main, Germany, wveitl@theochem.uni-frankfurt.de

Roland Winter Faculty of Chemistry, Physical Chemistry I – Biophysical Chemistry, TU Dortmund University, Otto Hahn Str. 6, 44227 Dortmund, Germany, roland.winter@tu-dortmund.de

Martin T. Zanni Department of Chemistry, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, USA, zanni@chem.wisc.edu

Wolfgang Zinth Department of Biomolecular Optics, Faculty of Physics, Ludwig-Maximilians-University Munich, Oettingenstr. 67, 80538 Munich, Germany, wolfgang.zinth@physik.uni-muenchen.de

Chapter 1
Linked Landscapes and Conformational Conversions: How Proteins Fold and Misfold

Gareth J. Morgan and Sheena E. Radford

Abstract Polypeptide chains must make a vast, complex array of interactions in order to fold to their correct structures. The mechanisms by which they do so, and the consequences when this process fails, are the subject of intense study. A key conceptual development has been the idea that proteins fold and misfold on linked, funnelled energy landscapes, where rarely populated states (such as folding intermediates) can provide access to misfolded and sometimes aggregation-prone ensembles. Advances in experimental methodologies and computer simulations are driving an increased understanding of the forces involved in folding and aggregation. This knowledge is beginning to be used to stabilise proteins for new biological functions as well as to develop treatments for diseases of misfolding. In this introductory chapter, we focus on the array of experimental methods that are being applied to the key questions of how the folding, misfolding and aggregation of proteins are linked, both in vitro and in the more complex environment of the cell.

1.1 Introduction

A protein emerging from the ribosome must make many specific contacts and interactions if it is to fold to its correct structure. That a protein can fold to a pre-defined, low-energy state, despite the myriad of competing interactions it faces, is a spectacular triumph of evolution, and it is perhaps unsurprising that this process sometimes goes awry. Misfolding and aggregation are the basis of many diseases, and these processes have been studied in great detail with a view to developing therapeutics that target specific aspects of the disease process. In this chapter, we examine key concepts and questions in the field of protein folding

S.E. Radford (✉)
University of Leeds, Mount Preston Street, Leeds LS2 9JT, UK
e-mail: s.e.radford@leeds.ac.uk

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*,
Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_1,

and misfolding, and review the experimental methods used to investigate these processes.

How does a polypeptide chain fold into a specific structure? Different proteins fold by different structural mechanisms [1], but the underlying theme that has emerged from theoretical and experimental studies is that the folding energy landscape is funnel-shaped, so that initial favourable interactions, some of which may be present even in the initial unfolded state [2], make further folding more favourable (Fig. 1.1) [3–5]. Protein structures are mainly governed by promiscuous non-covalent interactions, which can form between any suitable pair of residues, so there are many potential ways for non-native contacts to form. The ability of proteins to form alternative, non-native interactions, termed "frustration" [6], can be visualised as roughness on the folding energy landscape, leading to the population of transient intermediates and kinetic traps. Non-native interactions in the early stages of folding can retard or disrupt the protein's search for its native state, but may also direct the beginning of the folding process [2, 7–9]. The native states of proteins are thought to be "minimally frustrated", which provides their stability [10]. However, proteins are only marginally stable: indeed, many undergo conformational changes, and some even major unfolding events, as part of their function [11, 12]. So even in their native states, proteins show conformational diversity relating to binding, catalysis or regulation [13, 14]. An emerging theme is that protein structure and stability have been modulated by the requirement for function over the course of evolution [15, 16]. This compromised stability may allow conformational changes, but is also at the root of the numerous diseases of protein misfolding [17]. Misfolded forms of proteins may be inactive, toxic, or may form aggregates such as amyloid fibrils, which are ordered aggregates found in many disorders including Alzheimer's disease and type 2 diabetes [17]. Although amyloid fibrils or other aggregates may be toxic to cells, recent work has focused on the toxicity of smaller, soluble oligomeric species, termed protofibrils, that are thought to be the underlying cause of several amyloid disorders [18].

The continuum of conformational states that proteins can populate at equilibrium, from unfolded species to partially folded, folded, misfolded and aggregated forms, defines an energy landscape on which single protein molecules can fold and misfold, and ensembles of molecules can interact to form oligomers and aggregates [19]. These two processes may be linked by unfolded species or rarely populated intermediate states that can progress into the folding or misfolding regions of the energy landscape. Figure 1.1 shows an illustration of such a linked landscape.

Five key areas of research in the areas of protein folding and misfolding are described in this chapter:

- The nature of unfolded states under non-denaturing conditions.
- The intermediates that link the folding and aggregation landscapes.
- The structural and dynamic properties of toxic oligomers.
- The structure of mature amyloid fibrils.
- How studies of folding and misfolding in vitro relate to the processes that occur in the cell.

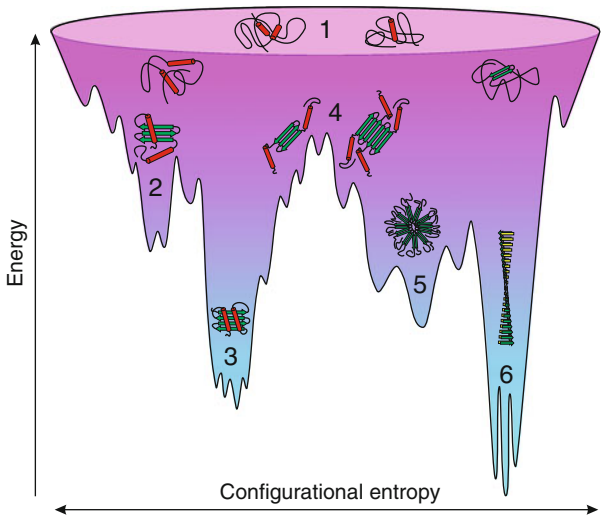

Fig. 1.1 A protein's folding and misfolding energy landscape. The energy of the protein chain is represented by the depth of the wells, and the configurational entropy of the states is represented by their width. The ruggedness of the landscape corresponds to intermediates and other metastable states. (1) Unfolded states showing varied conformations and residual structures; (2) folding intermediates; (3) native state, showing functional conformational changes; (4) misfolded intermediates and small oligomers; (5) toxic oligomers and protofibrils; (6) amyloid fibrils, which may have several distinct morphologies

We provide a brief introduction to each of these areas of research, highlighting the methodological developments that have provided recent insights. A recurring theme is the characterisation of sparsely populated, heterogeneous or dynamic conformations that are intractable to conventional high-resolution structural analysis by nuclear magnetic resonance (NMR) or X-ray crystallography. Information on these states has been gleaned from developments in methods to analyse rare conformations at equilibrium or in transition, in tandem with increasingly powerful simulations that provide detailed interpretations of these results [20,21]. A complete description of an energy landscape requires both structural and dynamic information on all the states populated. Different techniques complement each other by providing information at different resolutions, from atomic structures to overall shape and size, and on different timescales, from bond vibrations and side chain rotations to domain movements and oligomerisation. Methods used to study the folding and misfolding of proteins are shown in Fig. 1.2. These techniques can provide structural and dynamic information at various levels of resolution, and can be combined with rapid perturbation (such as continuous- or stopped-flow mixing, temperature jump and pressure jump methods) or quenching (pH jump or flash freezing) in order to study kinetic processes. As structural ensembles become more diverse, it is more difficult to extract high-resolution data; methods that report on the dynamic behaviour of proteins in solution, and thus define the size and

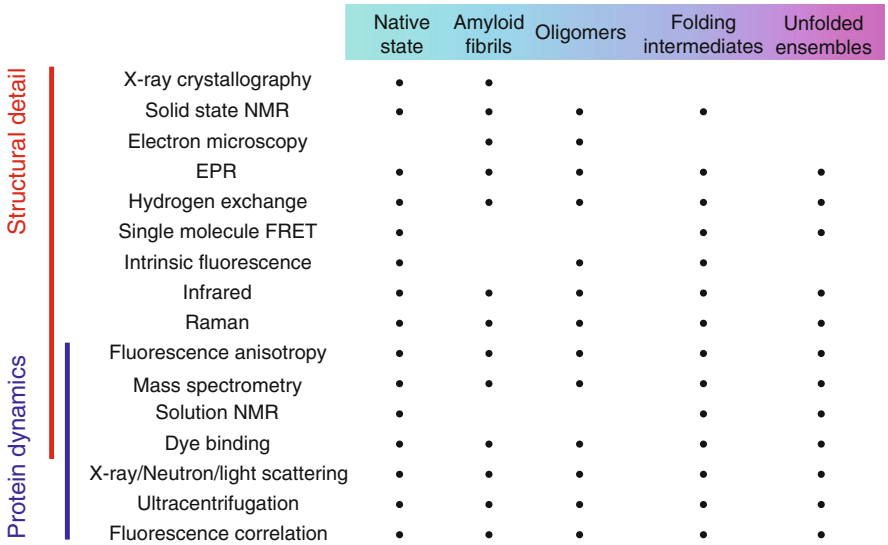

Fig. 1.2 Methods used to investigate different states in protein folding and misfolding. Colours correspond to the position of states on the landscape in Fig. 1.1. Information about both local structures and the global dynamics of individual protein molecules and multimeric assemblies can be gained via different techniques. *NMR* nuclear magnetic resonance; *EPR* electron paramagnetic resonance; *FRET* Förster resonance energy transfer

shape of oligomeric species, therefore complement techniques able to reveal higher resolution information. An important development is the use of single molecule techniques to analyse rare events against the background of the population of molecules [22]. We consider only soluble proteins (although significant progress has been made in the field of membrane protein folding in recent years, the experimental challenges involved remain considerable [23]). More detailed reviews of folding and misfolding, and the techniques used to study them, can be found elsewhere [24, 25].

1.2 The Unfolded Ensemble Under Native Conditions

Proteins sample many states at equilibrium, including partially folded, disordered and unfolded species. High-energy species are populated to a low extent, however, and unfolded states of globular proteins have mostly been accessed by altering the solution conditions so as to preferentially stabilise unfolded states by temperature, denaturant or pH [26, 27]. How these denatured states relate to the unfolded states that the protein visits under conditions that favour folding (at physiologically relevant pH, temperature and solution conditions) is also beginning to be understood [27]. A complementary approach is to mutate or chemically

modify proteins so that the native state is significantly destabilised under ambient conditions [28, 29]. An emerging theme is that unfolded proteins are not simply random coils, but are often compact, containing residual or transient structures that may represent the starting points for folding [21]. The importance of unfolded states has been highlighted by the increasing realisation that many proteins are unstructured in their active states. These "natively unfolded" or "intrinsically disordered" proteins have many functions and are especially common in binding and signalling systems [30]. Many natively unfolded proteins can fold into a specific structure upon binding to their ligands, and the mechanisms by which this occurs may be informative for understanding folding in general. In addition, many amyloid-forming or aggregation-prone proteins are disordered in their monomeric solution states [17]. These include the short peptides such as amyloid beta (Aβ) – involved in Alzheimer's disease – and the type 2 diabetes-associated islet amyloid polypeptide (IAPP). Larger, intracellular proteins such as α-synuclein and the N-terminal region of huntingtin, which are involved in Parkinson's and Huntington's diseases, respectively, are also unstructured in vitro. For these proteins, the absence of a defined globular structure is presumably important for their function, but it also allows them to access more easily aggregation-prone conformations that cause the pathologies associated with these proteins.

Solution NMR spectroscopy can provide detailed information at atomic resolution on unstructured proteins [21, 31, 32]. Although classical structural parameters such as long-range nuclear Overhauser effect (NOE) distance restraints cannot be observed for most unfolded proteins, chemical shifts, nuclear relaxation properties and residual dipolar couplings can provide significant structural information. These methods have been used to observe residual or transient structure in unfolded proteins that fluctuates on a picosecond to millisecond timescale, at the level of individual residues [33]. The transient nature of the species identified at equilibrium provides information on the diversity of states within the unfolded ensemble. Even in denaturant many proteins retain residual structure, observed through relaxation behaviour, proximity to paramagnetic moieties or residual dipolar couplings [26, 34, 35]. In the absence of denaturant, two strategies have been used to study unfolded states: either sensitive detection of rare states from within the native ensemble, or disruption of the native state by mutation. Single molecule fluorescence techniques have been particularly successful at observing rarely populated unfolded molecules [22]. This method can be used to monitor the amplitudes and timescales of distance fluctuations in unfolded states via Förster resonance energy transfer (FRET), promising very detailed descriptions of conformational fluctuations, although only a limited amount of information can be gained from a single experiment [22, 36, 37]. Studying severely destabilised variants of proteins, from which key core contacts have been removed by mutation, is a complementary method [28, 29]. Since these variants populate unfolded ensembles to a much higher extent, more mainstream structural techniques, in particular the NMR experiments described above, can be applied to gain a rich array of information [28, 29]. However, as with all mutational studies, the results need to be interpreted with caution, in case the mutations disrupt the folding landscape to too great an extent.

At high denaturant concentration, proteins behave like random coils or simple polymers, whose hydrodynamic radius is dependent on their length and degree of solvation. However, as denaturant is removed, polypeptides populate an array of more compact structures, the nature of which is more sequence dependent. These structural features may contain native-like or non-native contacts and are often only formed transiently. Interpretation of data on unfolded states is therefore best achieved by considering ensembles of structures that represent the data, rather than attempting to imagine a single dominant conformation. These residual structures have been observed in several proteins [32]. The presence of residual structure in unfolded ensembles determines the starting point for folding, and may direct the protein to fold towards its native state; alternatively, they may impede the folding process [2]. Although conformational conversions between unfolded conformations are assumed to be rapid by funnel-shaped landscape models [3], a recent intriguing study based on exhaustive simulation data suggests that unfolded states are kinetically isolated and may actually interconvert most readily via the native state [7].

1.3 Folding and Misfolding Intermediates

Transiently populated, partially folded species are ubiquitous on folding energy landscapes, and describing their properties is crucial in order to understand the folding and misfolding of proteins. Transition states and intermediates on folding pathways serve as gateways and junctions for different routes through the energy landscape, defining the likely conformational states that can be visited en route to the native state. Stable or metastable intermediates have been observed in an increasing number of small proteins, including some previously thought to fold via a two-state process, suggesting that landscape ruggedness is a ubiquitous feature of the folding of even small, simple proteins [38].

Detailed structural analysis of intermediates is difficult because of their transient nature and conformational dynamics, and this problem is exacerbated in the case of transition states, which by definition have a vanishingly short life span. Intermediates and transition states can be detected by kinetic methods. Much of the structural knowledge of these states comes from the Φ-value analyses of small proteins, whereby the effects of mutations on folding kinetics are interpreted in terms of changes in stability, and hence, structure of transient states [39]. Recently, more adventurous Φ-value analyses have been used to investigate the structural determinants of mechanical stability [40], to examine the folding of repeat proteins [41] and to monitor solvation and core packing [42]. Although Φ-values are relatively crude measures of structure, they can be used to restrain or test molecular dynamics simulations, which can reveal models of transient states in atomistic detail [20]. Φ-values obtained from continuous-flow and stopped-flow kinetic measurements of the folding of the bacterial immunity protein Im7 allowed calculations of the structures of Im7's on-pathway intermediate and two

transition state ensembles [8]. Meanwhile, ab initio calculations have been used to simulate the entire folding pathway of the 40-residue protein NTL9 over a millisecond timescale, and the folding trajectories compared with experimental data [43]. These computational methods allow previously invisible or ambiguous interactions to be identified, which in turn suggest further experimental work, and can help identify patterns between different proteins. Simulations of many folding trajectories uniquely allow important structural intermediates and metastable states to be identified that are invisible to experimental approaches. These include very early collapsed species, which direct the initial stages of folding, and high-energy intermediates that are kinetically invisible [7].

Solution NMR techniques can reveal atomic resolution information about intermediates that are visited at equilibrium. Hydrogen exchange kinetics report on the hydrogen bonding and solvent accessibility of exchange-labile hydrogen atoms within a protein, and are particularly sensitive to the presence of secondary structure elements [44]. Hierarchical unfolding of subdomains (known as "foldons") of cytochrome *c* has been described in great detail [45]. In some cases, it is possible to correlate dynamics measured by hydrogen exchange with specific folding intermediates, as described for Im7 [46]. Alternatively, folding and hydrogen exchange kinetics may report on different parts of the energy landscape, as is the case for the ribosomal protein S6 [47]. Another method that has allowed partially folded species to be observed is relaxation dispersion NMR, which allows the presence of "invisible" species populated to $<1\%$ at equilibrium to be detected via their effect on the relaxation of NMR signals [48]. Chemical exchange on the timescale of an NMR experiment (typically milliseconds) causes line broadening in the spectra, and precise quantitation of this effect allows information about the structure and dynamics of the non-native species to be extracted. This effect was first used to discover a hidden intermediate in the folding of an SH3 domain and the chemical shift changes between the native and intermediate states used to restrain a molecular dynamics simulation to generate an all-atom model of the intermediate ensemble [49]. Finally, a recent approach using flash-freezing and solid-state NMR (ssNMR) offers the opportunity to gain high-resolution information on freeze-trapped intermediate states of even fast-folding proteins [50].

An alternative approach to observing intermediates is to use single molecule techniques to detect species that are obscured by ground states at equilibrium. Single molecule FRET has been used to measure distance distributions of many individual proteins to build up a description of the protein's folding energy landscape [22]. At least in theory, these techniques allow transitions of individual molecules to be characterised. However, the extremely rapid timescales of these events present a formidable technical challenge. Recent progress by Eaton and co-workers has determined an upper time limit for these transitions of 200 μs for the protein GB1 immobilised on a surface, which is 10,000 times faster than the average folding time [51].

The soluble, globular form of a protein may be dramatically different from its conformation within an amyloid fibril. Understanding the structural transitions between these two states, therefore, requires that the intermediates and transition

states on the pathway to the misfolded state are characterised. Folding intermediates may be a crucial part of the process, as they may be aggregation-prone themselves, or have the conformational flexibility needed to access the direct precursors for amyloid formation [19, 25]. Understanding a protein's folding mechanism can therefore provide a gateway to studying the transient states that link the folding and aggregation regions of the protein's conformational energy landscape (Fig. 1.1). Such intermediates were first observed in mutants of lysozyme [52] and have now been characterised for several proteins [25, 53, 54]. A well-studied example of such a state is a folding intermediate of β_2-microglobulin (β_2m) containing a non-native *trans* proline bond at position 32. Population of this species correlates with the rate of amyloid elongation, suggesting that this conformation is a gateway to further misfolding or oligomerisation [55]. Amyloid formation by β_2m under physiological conditions can be initiated by Cu (II) binding to a species that has the same *trans* proline residue [56]. This conformation may be a useful target for drugs designed to prevent dialysis-related amyloidosis, which is caused by deposition of β_2m. This knowledge-based drug development is exemplified by attempts to prevent amyloidosis of the human protein transthyretin (TTR), which is associated with familial amyloid polyneuropathy [53]. The folding and misfolding landscape of TTR is well characterised, enabling drugs to be designed that stabilise the native tetramer of TTR, thus preventing amyloidosis. These compounds are among the first drugs for amyloid disease in clinical trials [53].

1.4 Protofibrils, Oligomers and Toxicity

For a protein to aggregate, molecules in an aggregation-competent conformation must encounter one another and form oligomeric species that eventually grow into large assemblies such as amyloid. Although amyloid fibrils have long been associated with disease, it is not clear how the mature fibrils contribute to the pathogenesis of amyloid disorders [17]. A turning point was the observation that progression of Alzheimer's disease correlates with the concentration of soluble, oligomeric forms of the 42-residue isoform of the amyloid beta peptide ($A\beta_{42}$), rather than levels of amyloid plaques [18]. Soluble oligomers, also known as protofibrils, have since been implicated as toxic species in several amyloid diseases, although the precise roles of fibrils and oligomers in the progression of disease remain unclear. A wide array of oligomeric species is formed by different proteins, which populate a heterogeneous array of oligomeric species depending on the experimental conditions. Structural information about these oligomers is limited by their heterogeneity and dynamic behaviour, and characterising these species is an important future challenge for the field. An important observation is that toxic, soluble oligomers formed from several different proteins are recognised by a conformation-specific antibody, called A11, suggesting a common underlying structure and, possibly, mechanism of toxicity [57]. Electron paramagnetic resonance (EPR) measures the mobility of nitroxide spin labels and their proximity to one another. The behaviour of a spin label

attached to $A\beta_{40}$ has been shown to vary between fibrils and two different oligomeric states, supporting the theory that A11-reactive oligomers have a distinctly different structure to both other oligomeric species and mature fibrils [58].

New experimental methods are allowing transiently populated oligomeric species en route to amyloid formation to be characterised. An important development is the use of ion mobility mass spectrometry to resolve and measure the mass and conformational properties of pre-amyloid oligomers. This technique relies on separation of ions in the gas phase, based on their mass, cross-sectional area and charge, and allows mass spectra to be deconvoluted. Diverse populations of oligomers have been observed during the fibrillogenesis of Aβ [59] and β_2m [60], and distinct species of monomeric IAPP have been identified by similar methods [61]. Importantly, the time resolution of these experiments allows the dynamics of these species to be studied: mixing samples of ^{14}N- and ^{15}N-labelled protein at different stages of oligomer formation and monitoring the rate of their interconversion [60]. Atomic resolution structural data on oligomeric intermediates is difficult to obtain, but an ssNMR investigation into the structure of a neurotoxic, spherical intermediate formed by $A\beta_{40}$ showed that the peptide had a conformation very similar to that found within mature fibrils [62].

Alongside experimental observations of the process of amyloid formation, simulations and modelling are increasing our understanding of these processes. A range of computational techniques have been applied to the study of protein aggregation, from mathematical approaches that model the entire population of species [63, 64], through coarse-grained simulations of many peptides interacting, to atomistic simulations of the precise nature of the aggregation process [65]. As in the case of folding intermediates, the combination of experimental and computational data is greatly increasing the depth of understanding of these systems. Computational methods can report on the entire population of misfolding molecules, without the ambiguity inherent in the low-resolution ensemble methods by which these processes are followed. The mechanisms involved in the initiation of amyloid fibril formation and their subsequent elongation remain poorly characterised at a molecular level. Future investigations using the techniques described above, as well as further method development, should continue to increase our knowledge of this important research area.

1.5 Amyloid Structure

Through an electron microscope, amyloid fibrils formed from a wide variety of structurally diverse proteins look remarkably similar: they are long, unbranched, straight structures with a periodic twist and are usually around 10-nm wide [17]. Combined with their common cross-β architecture, revealed by X-ray fibre diffraction [66], and the observation that almost any protein can form amyloid under suitable conditions [67], the structural similarity in amyloid fibril architecture suggests that amyloid may be a universal alternative conformation for all

polypeptides [17]. However, closer examination of the structure of fibrils formed by different proteins by ssNMR and cryo-electron microscopy (cryo-EM) has revealed a diverse array of architectures and organisation of the β-strands, even among fibrils formed from the same polypeptide sequence [68–70], leaving open the question of how similar amyloid fibrils really are at the atomic level. Since they are large, non-crystalline aggregates, amyloid fibrils are intractable to either conventional X-ray crystallography or high-resolution studies by solution NMR. However, recent developments in these techniques, as well as in ssNMR and cryo-EM, have now provided detailed structural models of several amyloid fibrils [71–75]. Further insights can be obtained from other spectroscopic methods, particularly EPR [76] and infrared (IR) spectroscopies. IR is particularly useful in the study of amyloid fibrils and their precursors, since it is sensitive to changes in β-sheet structure [77]. Many amyloid fibrils have a distinctive strong peak in the amide I band around 1,620 cm^{-1}, which can be used to follow the formation of ordered β-sheet structure [78]. Isotopic labelling with ^{2}H, ^{13}C or ^{18}O alters the resonance frequencies of bonds and can be used to increase resolution within samples. Developments in 2D IR [77] and in the specific isotopic labelling of proteins [79] hold significant promise in this field, as described elsewhere in this volume.

Cryo-EM methods allow the direct visualisation of individual amyloid fibrils. Reconstructions of amyloid fibrils from different proteins have revealed substantial variation from the simple ladder cross-β models once envisioned [68]. Although EM reconstructions of amyloid fibrils do not yet approach atomic resolution, mainly due to the inherent heterogeneity of the samples, the details of the fibril architectures revealed are beautiful, fascinating and perplexing in equal measure. For example, the diversity of amyloid architectures observed in fibrils of β_2m [68] and Aβ [69, 70], even within a single sample, suggests that fibril conformations are sensitively balanced and affected by many factors. Hydrogen exchange kinetics can be used to monitor solvent exposure and hydrogen bonding in amyloid fibrils [80]. Although the fibrils themselves are not amenable to direct analysis, the exchange reaction can be quenched and the fibrils dissociated using aprotic organic solvents. This maintains the exchange patterns of individual residues, which can then be analysed by mass spectrometry or solution NMR. These two detection methods are complementary: MS providing information about populations of molecules, while NMR can supply residue-specific information. The structured cores of several amyloid fibrils, as well as pre-fibrillar oligomers [81], have been mapped in this way, and the exchange kinetics have been interpreted as representing dynamic recycling of molecules from fibril ends [82]. EPR measurements can also provide information about the local structure and dynamics of spin labels introduced into amyloid fibrils. The structured regions of amyloid fibrils formed from β_2m were mapped by measuring spin label mobility and thiol accessibility, and the presence of spin–spin interactions suggested a parallel, in-register structure for these fibrils [83], in common with other fibrils created from unrelated protein sequences [73, 74, 76].

An important development in recent years has been the determination of the first atomic resolution structures of amyloid fibrils. X-ray crystal structures of

several short peptides in an amyloid-like conformation have been solved [71, 72]. Microcrystals with needle-like shapes were selected from solutions, which also encouraged the growth of amyloid fibrils. Their structures revealed precise conformations of different peptides that are consistent with subtle variations on the cross-β architecture, including the arrangement of β-strands within and between the β-sheets of the fibrils. Six out of eight possible topologies and multiple inter-sheet interfaces have been observed in crystals and these data have been used to develop models of other amyloid fibrils [72]. Although crystals of only small peptides have been successfully grown using this technique, the information has provided an important validation of previous models of amyloid fibrils based on the cross-β architecture.

Solid-state NMR can, in theory, resolve molecular structures of amyloid fibrils in great detail. However, spectral crowding means that the interpretation of spectra is difficult and laborious, generally requiring many samples, isotopically labelled at specific sites. Detailed models of $A\beta_{40}$ [74] and IAPP [73] have been published, which show stacked, in-register β-strands in a strand-turn-strand motif with tightly packed cores. The fungal prion HET-s, the largest protein for which the structure of an amyloid fibril has been determined, has a more complex solenoid structure, where each molecule forms two loops, contributing two β-strands to each of four β-sheets that run parallel to the fibril axis [75]. Although these structures represent a major step in our understanding of amyloid structure, it is not yet clear how the conformation of polypeptides within an amyloid fibril determines their overall gross morphology, mechanical properties or any interactions the fibrils may make in vivo. Correlating the wide array of structural information with biological effects remains a major challenge for the future.

1.6 From the Test Tube to the Cell

Much of our knowledge of protein folding and misfolding has come from studying pure, isolated proteins in vitro. However, the environment of a cell in which a newly synthesised protein must fold is far more complex, and translating knowledge of a protein's folding mechanism in vitro into the context of cell remains a challenge [84]. Experimental approaches and analytical tools are being developed to monitor protein folding and misfolding in vivo. The cellular environment influences all parts of a protein's folding and misfolding landscape, from interactions of the nascent chain with the ribosome on which it is synthesised, through binding to chaperones and functional partners, to the eventual destruction of a protein by regulated proteolysis. Figure 1.3 describes some of the states that a protein chain can form within cells and displays the dynamic interplay of different factors in maintaining cellular homeostasis.

The first interaction that a protein makes in the cell is with the ribosome on which it is synthesised. Solution state NMR of ribosome-bound nascent chains has been used to demonstrate formation of tertiary structure in domains still tethered to the ribosome [85] and has shown that the nascent chain of an SH3 domain does

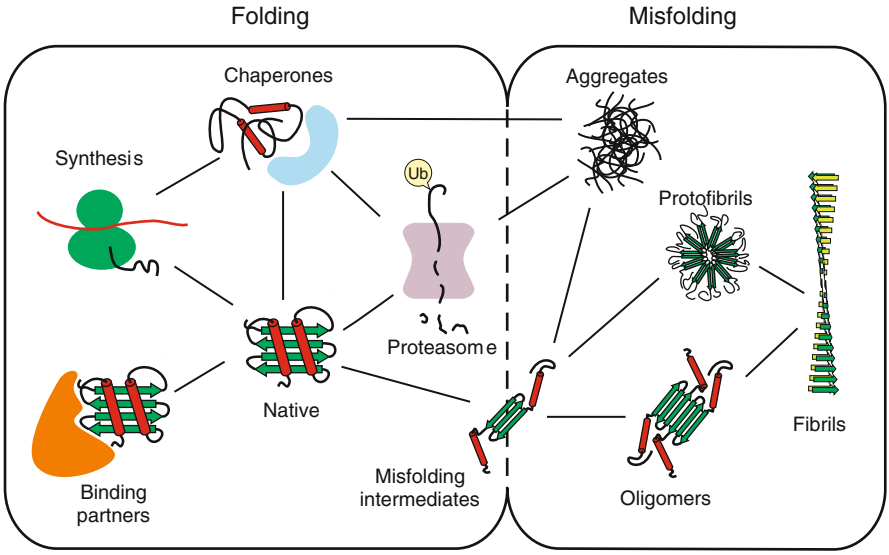

Fig. 1.3 Proteins interact with many partners in the cell, which may influence their folding and misfolding. After synthesis on ribosomes, proteins may interact with chaperones or binding partners and may be targeted for degradation by the proteasome by ubiquitination (indicated by "Ub"). Chaperones and disaggregases may prevent the formation of misfolded species, or perhaps sequester toxic oligomers into inert amyloid fibrils

not fold until it has fully emerged from the peptide exit tunnel [86]. Many proteins cannot fold efficiently alone and require the assistance of chaperones to reach their native state [87]. The chemical modification of proteins by the covalent addition of, among others, phosphate, sugars or even polyproteins such as ubiquitin plays critical roles in their functions, and the consequences of these post-translational modifications on folding have recently begun to be investigated in molecular detail. Glutathionation of superoxide dismutase, for example, promotes dimer dissociation and thereby increases amyloid propensity [88]. Direct observations of folding in vivo are complicated by the difficulty of isolating signals from the protein of interest from those arising from other cellular components. Specific fluorescent labelling using dyes has allowed direct observation of folding and unfolding events within cells [89], while proteins' movements and interactions within calls can be readily tracked using genetic fusions with fluorescent proteins [90]. Another challenge for these studies is the difficulty of perturbing the protein's environment without damaging the cell. A novel method using rapid laser temperature jump in vivo is one solution to this problem [91]. Solution NMR of labelled proteins within living cells holds great promise in the study of proteins in their native environment, including their dynamics and interactions with chaperones [92]. Initial results from these techniques show that the broad conclusions from in vitro studies remain valid within cells, but subtle differences may modulate the folding landscape in important ways.

Cells have a complex network of systems whose role is to maintain the integrity of their proteins. The holistic function, termed "protein homeostasis" or "proteostasis", of these networks is beginning to be investigated [93]. Misfolding diseases are increasingly being seen as caused by a failure of the cell's proteostasis machinery. However, the mechanisms of toxicity, or even the identity of toxic species, involved with misfolding and amyloid formation are not clear. As protein conformational diversity is revealed in more detail, and as new chaperones continue to be identified, the challenge of integrating these data into a working model of how proteins fold and misfold within cells becomes more complex [84]. A systems biology approach to the folding and export of aggregation-prone proteins from the endoplasmic reticulum has been used to rationalise the effects of destabilising mutations on different model proteins [94], and these integrative approaches offer a way to make sense of the ever-expanding array of information gained from an increasingly diverse range of experiments.

1.7 Conclusions

Our knowledge of how proteins fold continues to expand. The folding energy landscapes of many soluble proteins have been investigated in great detail through a combination of the techniques described above, and structural and kinetic information is now available for states which had previously resisted atomic characterisation. Although great progress has been made, many areas remain to be fully investigated. Open questions include the nature of unfolded states in different proteins, the structural transitions that link folding and misfolding and the mechanisms of toxicity in diseases of misfolding. The remainder of this volume describes the contribution to this important and burgeoning field that is being made using infrared methods that extend and complement the biophysical approaches described here.

References

1. A. Zarrine-Afsar, S.M. Larson, A.R. Davidson, Curr. Opin. Struct. Biol. **15**, 42 (2005)
2. M.A.C. Reed, C. Jelinska, K. Syson, M.J. Cliff, A. Splevins, T. Alizadeh, A.M. Hounslow, R.A. Staniforth, A.R. Clarke, C. Jeremy Craven, J.P. Waltho, J. Mol. Biol. **357**, 365 (2006)
3. J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Proteins **21**, 167 (1995)
4. K.A. Dill, S.B. Ozkan, M.S. Shell, T.R. Weikl, Annu. Rev. Biophys. **37**, 289 (2008)
5. M. Karplus, Fold. Des. **2**, S69 (1997)
6. D. Bryngelson, P.G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **84**, 7524 (1987)
7. G.R. Bowman, V.S. Pande, Proc. Natl. Acad. Sci. U.S.A. **107**, 10890 (2010)
8. C.T. Friel, D.A. Smith, M. Vendruscolo, J. Gsponer, S.E. Radford, Nat. Struct. Mol. Biol. **16**, 318 (2009)
9. A. Zarrine-Afsar, S. Wallin, A.M. Neculai, P. Neudecker, P.L. Howell, A.R. Davidson, H.S. Chan, Proc. Natl. Acad. Sci. U.S.A. **105**, 9999 (2008)

10. D.U. Ferreiro, J.A. Hegler, E.A. Komives, P.G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **104**, 19819 (2007)
11. S. Hammes-Schiffer, S.J. Benkovic, Annu. Rev. Biochem. **75**, 519 (2006)
12. E.M. Puchner, A. Alexandrovich, A.L. Kho, U. Hensen, L.V. Schafer, B. Brandmeier, F. Grater, H. Grubmuller, H.E. Gaub, M. Gautel, Proc. Natl. Acad. Sci. U.S.A. **105**, 13385 (2008)
13. O.F. Lange, N.A. Lakomek, C. Fares, G.F. Schroder, K.F. Walter, S. Becker, J. Meiler, H. Grubmuller, C. Griesinger, B.L. de Groot, Science **320**, 71 (2008)
14. J.S. Fraser, M.W. Clarkson, S.C. Degnan, R. Erion, D. Kern, T. Alber, Nature **462**, 669 (2009)
15. M.A. DePristo, D.M. Weinreich, D.L. Hartl, Nat. Rev. Genet. **6**, 678 (2005)
16. L. Foit, G.J. Morgan, M.J. Kern, L.R. Steimer, A.A. von Hacht, J. Titchmarsh, S.L. Warriner, S.E. Radford, J.C. Bardwell, Mol. Cell. **36**, 861 (2009)
17. F. Chiti, C.M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006)
18. J. Hardy, D.J. Selkoe, Science **297**, 353 (2002)
19. T.R. Jahn, S.E. Radford, FEBS J. **272**, 5962 (2005)
20. M. Vendruscolo, Curr. Opin. Struct. Biol. **17**, 15 (2007)
21. H.J. Dyson, P.E. Wright, Chem. Rev. **104**, 3607 (2004)
22. B. Schuler, W.A. Eaton, Curr. Opin. Struct. Biol. **18**, 16 (2008)
23. P.J. Booth, J. Clarke, Proc. Natl. Acad. Sci. U.S.A. **107**, 3947 (2010)
24. A.I. Bartlett, S.E. Radford, Nat. Struct. Mol. Biol. **16**, 582 (2009)
25. F. Chiti, C.M. Dobson, Nat. Chem. Biol. **5**, 15 (2009)
26. C.S. Le Duff, S.B. Whittaker, S.E. Radford, G.R. Moore, J. Mol. Biol. **364**, 824 (2006)
27. Y. Li, F. Picart, D.P. Raleigh, J. Mol. Biol. **349**, 839 (2005)
28. B. Anil, Y. Li, J.H. Cho, D.P. Raleigh, Biochemistry **45**, 10110 (2006)
29. T.L. Religa, J.S. Markson, U. Mayor, S.M. Freund, A.R. Fersht, Nature **437**, 1053 (2005)
30. H.J. Dyson, P.E. Wright, Nat. Rev. Mol. Cell. Biol. **6**, 197 (2005)
31. D. Eliezer, Curr. Opin. Struct. Biol. **19**, 23 (2009)
32. T. Mittag, J.D. Forman-Kay, Curr. Opin. Struct. Biol. **17**, 3 (2007)
33. A.K. Mittermaier, L.E. Kay, Trends Biochem. Sci. **34**, 601 (2009)
34. M.J. Cliff, C.J. Craven, J.P. Marston, A.M. Hounslow, A.R. Clarke, J.P. Waltho, J. Mol. Biol. **385**, 266 (2009)
35. R. Mohana-Borges, N.K. Goto, G.J. Kroon, H.J. Dyson, P.E. Wright, J. Mol. Biol. **340**, 1131 (2004)
36. S.D. Pugh, C. Gell, D.A. Smith, S.E. Radford, D.J. Brockwell, J. Mol. Biol. **398**, 132 (2010)
37. D. Nettels, A. Hoffmann, B. Schuler, J. Phys. Chem. B **12**, 6137 (2008)
38. D.J. Brockwell, S.E. Radford, Curr. Opin. Struct. Biol. **17**, 30 (2007)
39. A. Matouschek, J.T. Kellis Jr., L. Serrano, A.R. Fersht, Nature **340**, 122 (1989)
40. D.P. Sadler, E. Petrik, Y. Taniguchi, J.R. Pullen, M. Kawakami, S.E. Radford, D.J. Brockwell, J. Mol. Biol. **393**, 237 (2009)
41. N. Courtemanche, D. Barrick, Structure **16**, 705 (2008)
42. A.I. Bartlett, S.E. Radford, J. Mol. Biol. **396**, 1329 (2010)
43. V.A. Voelz, G.R. Bowman, K. Beauchamp, V.S. Pande, J. Am. Chem. Soc. **132**, 1526 (2010)
44. S.W. Englander, Annu. Rev. Biophys. **29**, 213 (2000)
45. M.M. Krishna, H. Maity, J.N. Rumbley, Y. Lin, S.W. Englander, J. Mol. Biol. **359**, 1410 (2006)
46. S.A. Gorski, C.S. Le Duff, A.P. Capaldi, A.P. Kalverda, G.S. Beddard, G.R. Moore, S.E. Radford, J.Mol. Biol. **337**, 183 (2004)
47. E. Haglund, J. Lind, T. Oman, A. Ohman, L. Maler, M. Oliveberg, Proc. Natl. Acad. Sci. U.S.A. **106**, 21619 (2009)
48. P. Neudecker, P. Lundstrom, L.E. Kay, Biophys. J. **96**, 2045 (2009)
49. D.M. Korzhnev, X. Salvatella, M. Vendruscolo, A.A. Di Nardo, A.R. Davidson, C.M. Dobson, L.E. Kay, Nature **430**, 586 (2004)
50. K.N. Hu, W.M. Yau, R. Tycko, J. Am. Chem. Soc. **132**, 24 (2010)
51. H.S. Chung, J.M. Louis, W.A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **106**, 11837 (2009)
52. D.R. Booth, M. Sunde, V. Bellotti, C.V. Robinson, W.L. Hutchinson, P.E. Fraser, P.N. Hawkins, C.M. Dobson, S.E. Radford, C.C. Blake, M.B. Pepys, Nature **385**, 787 (1997)

53. S. Connelly, S. Choi, S.M. Johnson, J.W. Kelly, I.A. Wilson, Curr. Opin. Struct. Biol. **20**, 54 (2010)
54. R.A. Staniforth, S. Giannini, L.D. Higgins, M.J. Conroy, A.M. Hounslow, R. Jerala, C.J. Craven, J.P. Waltho, EMBO J. **20**, 4774 (2001)
55. T.R. Jahn, M.J. Parker, S.W. Homans, S.E. Radford, Nat. Struct. Mol. Biol. **13**, 195 (2006)
56. C.M. Eakin, A.J. Berman, A.D. Miranker, Nat. Struct. Mol. Biol. **13**, 202 (2006)
57. R. Kayed, E. Head, J.L. Thompson, T.M. McIntire, S.C. Milton, C.W. Cotman, C.G. Glabe, Science **300**, 486 (2003)
58. J.W. Wu, L. Breydo, J.M. Isas, J. Lee, Y.G. Kuznetsov, R. Langen, C. Glabe, J. Biol. Chem. **285**, 6071 (2010)
59. S.L. Bernstein, N.F. Dupuis, N.D. Lazo, T. Wyttenbach, M.M. Condron, G. Bitan, D.B. Teplow, J.-E. Shea, B.T. Ruotolo, C.V. Robinson, M.T. Bowers, Nat. Chem. **1**, 326 (2009)
60. D.P. Smith, S.E. Radford, A.E. Ashcroft, Proc. Natl. Acad. Sci. U.S.A. **107**, 6794 (2010)
61. N.F. Dupuis, C. Wu, J.E. Shea, M.T. Bowers, J. Am. Chem. Soc. **131**, 18283 (2009)
62. S. Chimon, M.A. Shaibat, C.R. Jones, D.C. Calero, B. Aizezi, Y. Ishii, Nat. Struct. Mol. Biol. **14**, 1157 (2007)
63. W.F. Xue, S.W. Homans, S.E. Radford, Proc. Natl. Acad. Sci. U.S.A. **105**, 8926 (2008)
64. T.P.J. Knowles, C.A. Waudby, G.L. Devlin, S.I.A. Cohen, A. Aguzzi, M. Vendruscolo, E.M. Terentjev, M.E. Welland, C.M. Dobson, Science **326**, 1533 (2009)
65. B. Ma, R. Nussinov, Curr. Opin. Chem. Biol. **10**, 445 (2006)
66. M. Sunde, L.C. Serpell, M. Bartlam, P.E. Fraser, M.B. Pepys, C.C. Blake, J. Mol. Biol. **273**, 729 (1997)
67. F. Chiti, P. Webster, N. Taddei, A. Clark, M. Stefani, G. Ramponi, C.M. Dobson, Proc. Natl. Acad. Sci. U.S.A. **96**, 3590 (1999)
68. H.E. White, J.L. Hodgkinson, T.R. Jahn, S. Cohen-Krausz, W.S. Gosal, S. Muller, E.V. Orlova, S.E. Radford, H.R. Saibil, J. Mol. Biol. **389**, 8 (2009)
69. J. Meinhardt, C. Sachse, P. Hortschansky, N. Grigorieff, M. Fändrich, J. Mol. Biol. **386**, 869 (2009)
70. A.K. Paravastu, R.D. Leapman, W.M. Yau, R. Tycko, Proc. Natl. Acad. Sci. U.S.A. **105**, 18349 (2008)
71. R. Nelson, M.R. Sawaya, M. Balbirnie, A.O. Madsen, C. Riekel, R. Grothe, D. Eisenberg, Nature **435**, 773 (2005)
72. M.R. Sawaya, S. Sambashivan, R. Nelson, M.I. Ivanova, S.A. Sievers, M.I. Apostol, M.J. Thompson, M. Balbirnie, J.J. Wiltzius, H.T. McFarlane, A.O. Madsen, C. Riekel, D. Eisenberg, Nature **447**, 453 (2007)
73. S. Luca, W.M. Yau, R. Leapman, R. Tycko, Biochemistry **46**, 13505 (2007)
74. A.T. Petkova, Y. Ishii, J.J. Balbach, O.N. Antzutkin, R.D. Leapman, F. Delaglio, R. Tycko, Proc. Natl. Acad. Sci. U.S.A. **99**, 16742 (2002)
75. C. Wasmer, A. Lange, H. van Melckebeke, A.B. Siemer, R. Riek, B.H. Meier, Science **319**, 1523 (2008)
76. M. Margittai, R. Langen, Q. Rev. Biophys. **41**, 265 (2008)
77. Z. Ganim, H.S. Chung, A.W. Smith, L.P. Deflores, K.C. Jones, A. Tokmakoff, Acc. Chem. Res. **41**, 432 (2008)
78. G. Zandomeneghi, M.R. Krebs, M.G. McCammon, M. Fandrich, Protein Sci. **13**, 3314 (2004)
79. S.-H. Shim, R. Gupta, Y.L. Ling, D.B. Strasfeld, D.P. Raleigh, M.T. Zanni, Proc. Natl. Acad. Sci. U.S.A. **106**, 6614 (2009)
80. M. Hoshino, H. Katou, K. Yamaguchi, Y. Goto, Biochim. Biophys. Acta **1768**, 1886 (2007)
81. N. Carulla, M. Zhou, M. Arimon, M. Gairi, E. Giralt, C.V. Robinson, C.M. Dobson, Proc. Natl. Acad. Sci. U.S.A. **106**, 7828 (2009)
82. N. Carulla, G.L. Caddy, D.R. Hall, J. Zurdo, M. Gairi, M. Feliz, E. Giralt, C.V. Robinson, C.M. Dobson, Nature **436**, 554 (2005)
83. C.L. Ladner, M. Chen, D.P. Smith, G.W. Platt, S.E. Radford, R. Langen, J. Biol. Chem. **285**, 17137 (2010)
84. L.M. Gierasch, A. Gershenson, Nat. Chem. Biol. **5**, 774 (2009)

85. S.T. Hsu, P. Fucini, L.D. Cabrita, H. Launay, C.M. Dobson, J. Christodoulou, Proc. Natl. Acad. Sci. U.S.A. **104**, 16516 (2007)
86. C. Eichmann, S. Preissler, R. Riek, E. Deuerling, Proc. Natl. Acad. Sci. U.S.A. **107**, 9111 (2010)
87. F.U. Hartl, M. Hayer-Hartl, Nat. Struct. Mol. Biol. **16**, 574 (2009)
88. K.C. Wilcox, L. Zhou, J.K. Jordon, Y. Huang, Y. Yu, R.L. Redler, X. Chen, M. Caplow, N.V. Dokholyan, J. Biol. Chem. **284**, 13940 (2009)
89. Z. Ignatova, L.M. Gierasch, Proc. Natl. Acad. Sci. U.S.A. **101**, 523 (2004)
90. J. Lippincott-Schwartz, E. Snapp, A. Kenworthy, Nat. Rev. Mol. Cell. Biol. **2**, 444 (2001)
91. S. Ebbinghaus, A. Dhar, J.D. McDonald, M. Gruebele, Nat. Methods **7**, 319 (2010)
92. K. Inomata, A. Ohno, H. Tochio, S. Isogai, T. Tenno, I. Nakase, T. Takeuchi, S. Futaki, Y. Ito, H. Hiroaki, M. Shirakawa, Nature **458**, 106 (2009)
93. W.E. Balch, R.I. Morimoto, A. Dillin, J.W. Kelly, Science **319**, 916 (2008)
94. R.L. Wiseman, E.T. Powers, J.N. Buxbaum, J.W. Kelly, W.E. Balch, Cell **131**, 809 (2007)

Chapter 2
A Quantitative Reconstruction of the Amide I Contour in the IR Spectra of Peptides and Proteins: From Structure to Spectrum

Joseph W. Brauner and Richard Mendelsohn

Abstract The broad amide I band in the infrared spectrum of peptides and proteins is accurately simulated with the coupled oscillator model of Miyazawa. Formulas for modeling the geometry-dependent interactions between peptide groups needed to reproduce the amide I contour for a peptide or protein of known structure are given. The accuracy and utility of this approach are suggested through several examples for which protein structure cannot be acquired by standard high-resolution technologies.

2.1 The Approach to Simulation of the Amide I Contour

2.1.1 Introduction

Infrared spectroscopy (IR) has been used for over half a century to study protein structure and ligand–protein interactions. The technique cannot provide the detailed structural information available from high-resolution biophysical methods such as X-ray diffraction and NMR spectroscopy. However, IR spectra do contain structural information and data can be acquired from samples in physical states not amenable to the high-resolution methods, for example, noncrystalline solids such as the amyloid plaques of Alzheimer's disease, thin films including monolayers at the air–water interface, fibrous proteins such as collagen, tissues, and cells, and so on. In addition, modern infrared techniques such as those reviewed in this volume permit the elucidation of structural elements with high temporal resolution.

The most intense band in the mid-IR spectrum of proteins which can provide useful structural information is due to the amide I normal mode of vibration which

J.W. Brauner (✉)
Rutgers University, Newark College, 73 Warren Street, Newark, NJ 07102, USA
e-mail: matjoe63@verizon.net

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*,
Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_2, © Springer-Verlag Berlin Heidelberg 2012

is mostly a peptide group carbonyl stretch with a minor C–N stretching contribution [1–4]. The band appears in the spectral region between 1,600 and 1,700 cm^{-1}. The initial application of this band was primarily as an indicator of secondary structure in homopolypeptides and specific spectral assignments were proposed [5–7].

The purpose of this chapter is to present a semiempirical method of simulating the experimental amide I contour of a protein or peptide molecule whose atomic coordinates are available. The idea for this method or its early implementation did not originate with the authors of this chapter. However, we have managed to refine the approach to the point where good agreement between the experimental and simulated amide I contours has been achieved for a fair number of molecules possessing a variety of structures.

Having to proceed from structure to spectrum may not appear to be, initially, a great advance. Nevertheless, accurate simulation, first of all, requires that all the spatially sensitive interactions between the peptide groups which produce the broad amide I contour are accounted for and are correctly modeled. The details of this modeling are presented in this chapter.

The mathematical results of an accurate simulation permit one to know the contribution of every peptide group in the molecule to the amide I contour. Since the structural motif to which the groups belong can be deduced from the structure, this sheds light on the validity of the structure–frequency correlations, mentioned above, which have had a wide variety of applications. For example, isotope editing of peptide structures has been viewed as an experimental way to discover the spectral contribution of specific residues in the molecule. Then, by way of the structure–frequency correlations, the structural motif to which the residue belongs could be revealed. The simulation method permits us to assess the correctness of this procedure by being able to predict and interpret the spectral changes wrought by these substitutions. Isotope editing will be discussed in this chapter. In this context, we must consider how isotopic substitution affects the vibrational interactions between substituted and unsubstituted peptide groups.

Finally, if a given starting structure has its amide I contour accurately calculated, and the structure is changed by, say, a molecular mechanics simulation, then the spectral changes could be predicted and compared to the actual spectra. We illustrate this procedure in this chapter for the early stages of thermal denaturation of a collagen model peptide. We now recount the path that led to the simulation method presented in this chapter.

2.1.2 Historical Background

In 1950, Elliott and Ambrose [5] observed that the frequency of the amide I band for polypeptides in the alpha-helix conformation was about 20 cm^{-1} higher than that of polypeptides in the beta-sheet conformation. This observation led to the notion that the broad amide I contour of globular proteins was due to a small number of highly overlapped bands each arising from one of several structural motifs such as coil or sheet and each at a unique frequency. The thinking was that the different

interactions between peptide groups in different structures shifted the absorption from the amide I frequency of a hypothetical isolated peptide group. Actually, group theory indicated the infinite helix and the infinite pleated sheet to have multiple amide I modes expected to vibrate at different frequencies. The magnitude of this frequency splitting provides a sensitive test for theories of interaction.

In 1960, Miyazawa [8], in his pioneering approach, attempted to quantify the frequency shifts with his "weakly coupled oscillator model" providing a basis for a perturbation treatment on localized vibrations. In this description, the amide I motion in each peptide group is represented by an oscillator having a common unperturbed frequency. The coupling between the oscillators was ascribed to kinetic and potential energy interactions through the valence bonds and the hydrogen bonds connecting the peptide groups locked in a regular arrangement. Miyazawa derived a frequency shift formula in terms of interaction constants and phase angles between the oscillators for infinite, regular structures. No formulas for calculating the interaction constants were provided.

The values of the constants were first obtained empirically by fitting observed frequency shifts to the derived formula. The antiparallel pleated sheet structure provided a useful test case. The amide I frequencies of an infinite version of this structure were analyzed in terms of the normal modes of vibration of four peptide groups in a unit cell consisting of two groups in one chain and two groups in an adjacent chain. This model predicts one infrared inactive mode and three infrared active modes. The IR and Raman spectra of a number of molecules exhibiting this structure were obtained and the frequencies of the four bands were assigned. However, it was impossible to arrive at a consistent set of interaction constants for a number of different molecules. Normal coordinate calculations were then undertaken in an effort to resolve this quandary [9].

2.1.3 Normal Coordinate Calculations

The Wilson **GF** matrix method of normal coordinate analysis [10] is a method for solving Newton's equations of motion for a system of coupled harmonic oscillators. The oscillators represent atomic motions such as bond stretches, bond angle changes, and dihedral bond angle changes, in which the center of mass of the atoms does not move. The oscillators interact because they have atoms in common; that is, the interaction is through valence bonds. The bond stretches, bond angle changes, etc., form internal coordinates. The internal coordinates act like connected springs.

The apparently chaotic vibrational motion of a molecule, if it could be directly observed, is actually the sum of a number of orderly vibrations called normal modes of vibration. Each normal mode is represented by a normal coordinate. A change in a normal coordinate is a weighted sum of changes in a specific collection of internal coordinates. A normal mode actually involves the motion of every atom in the molecule moving in phase, but usually only a small subgroup of atoms has any appreciable amplitude.

To successfully undertake a normal coordinate calculation, the structure of the molecule must be known. A nonlinear molecule with N atoms has 3N – 6 normal modes of vibration. This same number of internal coordinates must be used. The **G** matrix is basically an inverse mass matrix where the geometry and the masses of the atoms involved in a given internal coordinate are properly taken into account. Standard formulas are available for calculating the **G** matrix elements for various kinds of internal coordinates. The **F** matrix contains the force constants of the internal coordinates and the force constants for their interactions. The force constants in the **F** matrix must be estimated initially and then adjusted by trial and error until agreement is reached between the calculated and observed molecular frequencies. The set of force constants in the symmetric **F** matrix is called the force field. The frequencies are calculated by diagonalizing the **GF** matrix product. The frequencies are related to the eigenvalues and the contributions of the internal coordinates to the normal coordinates are calculated from the eigenvectors of the **GF** matrix product. For a globular protein with a thousand or more atoms, this is almost an impossible task and it has not been reported, to our knowledge. There are simply too many force constants to be estimated and refined.

A more tractable case is a highly regular structure such as an α-helix or a β-sheet. The force field of large regular structures is built up by piecing together the force field obtained by analyzing fragments of the large structure. Even so the classical normal coordinate calculation is a highly inefficient way of obtaining the range of frequencies of the amide I mode in a polypeptide. If the molecule has ten peptide groups and one hundred atoms then there will be ten amide I frequencies out of a total of 294 normal mode frequencies. Furthermore, the classical application of the method where interactions between the internal coordinates are through valence bonds does not produce the observed range of frequencies of the amide I mode. Through hydrogen bond interactions must be modeled as a CO stretch–HO stretch interaction. But this too is not enough. Finally, through-space transition dipole–dipole interaction must be modeled as a CO stretch–CO stretch interaction. The transition dipole–dipole interaction is usually calculated in the point dipole approximation. Transition dipole strengths and relative geometries must be estimated [9].

2.1.4 Ab Initio Force Field Calculations

In this approach, quantum mechanics in the form of density functional theory is used to calculate the force field needed in the **F** matrix of the **GF** matrix method instead of trying to estimate the force constants and refine them by trial and error [11, 12]. The three-dimensional potential energy field of the molecule is calculated for an energy-optimized molecular structure and the internal coordinate force constants and their interaction force constants are calculated as second derivatives of the potential energy field with respect to an internal coordinate or a pair of internal coordinates.

Even for relatively small structures, this is a computationally intensive exercise requiring a supercomputer and days or possibly weeks of calculation time. Larger regular structures are handled by piecing together the results for smaller structures. The calculation does not include through-space transition-dipole coupling between the peptide groups, which must be added on in a somewhat ad hoc manner.

2.1.5 The Modified GF Matrix Method

In the *modified* **GF** matrix method, the oscillators being coupled represent the vibrations of one kind of normal coordinate as found from the classical application of the **GF** method to a model molecule. For molecules of a polymeric nature such as polypeptides or proteins, this normal vibration occurs at multiple locations in the molecules of interest. Coupling vibrational modes of one kind leads to a much more efficient frequency distribution calculation for that mode because the calculation is narrowly focused. It is a higher level calculation because it subsumes a number of structural details. However, it can only be applied to a mode that does not interact with other modes.

Classical normal coordinate calculations carried out by Crawford and Fletcher [13] for *N*-methylacetamide showed that the displacements of the methyl groups are negligible compared to the motions of the CONH atoms of the peptide group in the amide I mode. This suggested to Miyazawa that the amide I mode is highly localized in a peptide group and it does not interact with modes outside of the group. The six normal modes of the peptide group itself are orthogonal to each other and do not interact. The independence of the amide I mode was restated by Torii and Tasumi [14] who argued for an approximate amide I subspace within the complete vibrational mode space of a protein molecule.

However, Miyazawa [8] realized that the amide I vibrations should interact among themselves through a resonance phenomenon when they are connected by valence bonds and/or hydrogen bonds. He estimated that the interaction should be strong enough to cause a first-order perturbation to the natural frequency of a hypothetical isolated peptide group in a polypeptide chain. Thus, he envisioned a system of coupled oscillators. This approach is not unique in vibrational spectroscopy. For example, Snyder and Schachtschneider [15] effectively explained the progression of methylene wagging mode bands observed for ordered crystalline alkanes in terms of coupled oscillators of one kind. We have adopted Miyazawa's idea and have applied it to the simulation of the amide I contour of peptides and globular proteins by providing formulas to calculate the geometry-sensitive interaction force constants. In the method, the oscillators being coupled represent the amide I motions of the individual peptide groups in a polypeptide or protein. One simple harmonic oscillator represents each peptide group. Each oscillator has a mass and a force constant. The motion of an oscillator is represented by a change in its internal coordinate, which is a simple one-dimensional displacement coordinate.

The displacement coordinates of all the peptide group amide I oscillators in the molecule are taken to form an orthogonal set.

To use our approach, the structure of the molecule must be known. The realization of the model then comes from completely listing and accurately calculating the geometry-dependent interaction force constants between all the peptide groups in the molecule. There are other interactions such as through-space transition–dipole interaction, which must be taken into account.

The force constants of the oscillators and the force constants for their interactions are placed in an **F** matrix. The reciprocal masses of the oscillators are placed in a **G** matrix. The eigenvalues of the **GF** matrix product give the normal mode frequencies of the system of coupled oscillators. The eigenvectors of the **GF** matrix product give the contributions of the individual oscillators to each collective normal mode. Thanks to modern, powerful software and high-speed personal computers, the diagonalization of even large, unsymmetric matrices is trivial. Any needed ancillary calculations are easily programmed and the approximations that marred the quality of earlier simulations of the amide I contour can be avoided.

2.1.6 Constructing the G and F Matrices in the Coupled Oscillators of One Kind Method

The details of matrix construction presented here are taken from the published work [16–18] on the simulation of the amide I contour by the authors of this chapter. The ideas of other researchers in this effort are commented on in context.

2.1.6.1 The G Matrix Elements

The inverse **G** matrix, $\mathbf{G}^{-1}$, in Wilson's **GF** matrix method, relates the kinetic energy of a vibrating system to the set of the time derivatives of the internal coordinates of the system written as a column vector, $\mathbf{S}^{\cdot}$, by

$$2T = \mathbf{S}^{\cdot\mathbf{t}}\mathbf{G}^{-\mathbf{1}}\mathbf{S}^{\cdot}.$$

In terms of Cartesian mass-weighted displacement coordinates [10], $q_1 = (m_1)^{1/2} x_1, q_2 = (m_1)^{1/2} y_1, q_3 = (m_1)^{1/2} z_1, q_4 = (m_2)^{1/2} x_2$, etc., the kinetic energy of this system is given by:

$$2T = \mathbf{q}^{\cdot\mathbf{t}}\mathbf{q}^{\cdot}.$$

Since the internal coordinate of an oscillator in the model being used here is just the one-dimensional, unweighted displacement coordinate for that oscillator, then the inverse **G** matrix, at this point, would be diagonal where the elements are the masses of the respective oscillators. In our model, we assume no kinetic energy

interactions between the oscillators. All of the interaction between the oscillators will be confined to the potential energy regime. The inverse **G** matrix remains diagonal. Then the **G** matrix is also diagonal where the elements are the inverse masses of the respective oscillators.

This mass is taken to be the same for the oscillators representing all peptide groups except for those groups where the nitrogen atom comes from a proline residue or where the carbonyl carbon atom has been isotopically substituted by ^{13}C. We name the groups based on the name of the residue supplying the N atom. For Pro groups, replacement of the hydrogen atom of the peptide group in the pyrrolidine ring is modeled as increasing the mass of the oscillator. Likewise, a ^{13}C substitution in the carbonyl carbon atom increases the mass of the representative oscillator.

As explained below, in the subsequent calculations, the **G** matrix is multiplied by the common mass of the oscillators. This makes the common diagonal elements unity and the diagonal elements for Pro and ^{13}C-substituted groups a number less than one. The actual values used are given when specific cases are discussed.

2.1.6.2 The Diagonal Elements of the F Matrix

Before interactions between oscillators are taken into account, the **F** matrix diagonal elements are the force constant, k, for the harmonic oscillator, which represents each peptide group in a polypeptide or protein. This constant is taken to be the same for all the oscillators in the molecule. The underlying assumption is that the electronic structure of the peptide groups is the same before interaction. As explained below, in the subsequent calculations, the **F** matrix is divided by the common force constant of the oscillators. This makes the value of the diagonal elements unity prior to interaction.

Interaction changes the independent vibrational frequency of an oscillator even without changing its intrinsic force constant. By independent vibration frequency, we mean the vibration frequency of an oscillator in a system of interacting oscillators when all the other oscillators are held still. The independent vibration frequency of a spring changes if it is connected to other springs as is made clear in Symon's [19] exposition of coupled oscillators.

The way to take this effect of interaction into account is to add the off-diagonal elements in a given row into the diagonal element. This was not done in the work of other researchers possibly because the off-diagonal elements are only a fraction of percent of the diagonal element.

There is a second physical reason for modifying diagonal force constants based on the effect of interaction on the intrinsic force constant of the oscillator representing a group. It appears that a hydrogen bond increases the representative force constant for the peptide group supplying the CO moiety. The group supplying the NH moiety seems to be unaffected in this way. This is consistent with the minor role of the NH in-plane bend and the major role of the CO stretch in the potential energy distribution of the amide I normal mode of vibration. A similar observation was made by Torii and Tasumi [14] who found it necessary to raise by up to 1.6%

the diagonal force constant of peptide groups that have their CO bonds "pointing toward the inside of β-sheets and for those belonging to some α-helices". We found it necessary to add −10 times the through hydrogen bond interaction force constant to the diagonal force constant of the peptide group supplying the CO or acceptor moiety. The details of the calculation of the through hydrogen bond interaction force constants are presented below.

A further diagonal force constant modification is required for multistranded sheet structures. In our first simulation paper [16] on the anomalous amide I intensity distribution of ^{13}C-substituted peptides known to form antiparallel sheet structures, we found it necessary to take the through hydrogen bond interaction force constant to be about twice as large as that for similar molecules in the coil form. This assumption was necessary to reproduce the $\sim$70 cm^{-1} splitting between the intense low-frequency band and the much lower intensity high-frequency band seen for these model peptides. The assumption was made even though the O. . .H distances and the N–H, O–C angles are similar between sheet and coil structures. In our second publication [17], we came to appreciate that there is another source of interaction across hydrogen bonds, namely, a π−π interaction between π electrons in the peptide groups. The hydrogen bonds serve to align the peptide groups laterally over an extended range. It appears that the π electrons in the peptide groups can become delocalized across the entire sheet, thereby coupling together the amide I vibrations of the aligned groups. The strength of the interaction increases with the width of the sheet. This interaction is not important in helices because the peptide groups do not align as well as they do in sheet structures.

The effect is weak for a few strands, but if the molecules aggregate to form many-stranded sheets then the effect becomes substantial. This result is consistent with the findings of an ab initio study by Kubelka and Keiderling [11], which showed that the separation between the low- and high-frequency bands in the amide I contour of antiparallel sheet structures increases with the number of strands in the sheet.

The π−π interaction increases the force constant of the peptide groups supplying the CO moiety in the hydrogen bonds connecting the strands in the sheet. The groups supplying the NH moiety again seem to be unaffected in this way. The details of the π−π interaction force constant are presented in the next section. We found it necessary to add −10 times the π−π interaction force constant to the diagonal force constant of the peptide group supplying the CO or acceptor moiety.

To summarize, the **F** matrix, in the subsequent calculations, is divided by the intrinsic, common force constant for all the oscillators. This makes the diagonal elements unity prior to interaction. Then the diagonal elements are modified by two mechanisms as detailed above. First, interaction itself requires the off-diagonal elements in a given row be added into the diagonal element. Second, the diagonal element corresponding to an oscillator that supplies the CO moiety to a hydrogen bond must have −10 times the corresponding through hydrogen bond interaction force constant, which is also dimensionless, added to it. Also, if an oscillator is part of a sheet structure involving more than two strands and it supplies the CO moiety to an interstrand hydrogen bond then it must have −10 times its π−π interaction force constant added to its diagonal element.

2.1.6.3 The Off-Diagonal Elements of the F Matrix

The calculation of the off-diagonal elements of the **F** matrix as caused by various geometry-sensitive mechanisms will now be considered. The formula for transition dipole interaction is based on classical electrostatic theory. The formulas for through hydrogen bond, through valence bond, and through hydrogen bond $\pi-\pi$ interaction have been arrived at empirically based on elementary physical considerations. No detailed knowledge is claimed.

The empirical formulas were refined to obtain the best fit with the experimental amide I band for a number of different peptides and proteins. Quantum mechanical calculations may provide more exact formulas for an interaction based on a given intergroup geometry, but the geometries between peptide groups in a globular protein are quite variable. Furthermore, a number of assumptions and other sources of error go into the simulation calculation. There is inherently a certain amount of uncertainty in the atomic coordinates of an X-ray or an NMR protein structure, especially in the peptide backbone loops connecting secondary structures. Hydrogen atoms, an important member of the peptide group, are not generally determined in X-ray structures and they must be located by assuming standard peptide group geometry. Then for X-ray structures the crystal packing forces are released on placing the molecule in solution. Solvation forces may become important and may somewhat distort the X-ray determined structure. In addition, the solvent for IR measurements, usually buffered D_2O, interacts electrically with the surface peptide groups of the molecule and this may have spectral effects. Fortunately, these potential problems have not seriously materialized for the molecules we have selected for amide I simulation and good agreement between experimental and simulated amide I contours has been achieved. No doubt improvements can be made in the empirical formulas given here, but because of the large number of and the variability of the structural details in protein molecules progress may be slow.

2.1.6.4 Through-Space Transition–Dipole Interaction

The strong amide I absorption in the IR spectra of proteins is due to the large transition-dipole moment of a peptide group in this mode of vibration. The transition-dipole moment, a quantum mechanical integral and a vector quantity, is proportional to $(\partial \boldsymbol{\mu}/\partial Q)_0$, the instantaneous rate of change in the electric dipole moment of the group with respect to the normal coordinate for the amide I mode, evaluated at the equilibrium value of the normal coordinate.

The large value of this derivative is rationalized by a charge-flux model [20] based on resonance structures for the NCO subgroup. Normal coordinate analysis shows that in the amide I mode the NC bond shortens, while the CO bond lengthens. The bond length changes are correlated with electron pair shifts. The CO double bond becomes a single bond as a pair of electrons goes onto the O atom, while a pair of electrons moves off the N atom to make the NC single bond into a double bond. Thus, the N atom acquires a positive charge as the O atom acquires a negative

charge. In effect, electrons have flowed from the N atom to the O atom in one half of the vibration cycle. The charge flow reverses in the second half of the cycle. Altogether there is a substantial change in the dipole moment of the NCO group in the amide I mode. The intermediate structure where both the CO and NC bonds have partial double bond character is used to rationalize the planar geometry of the peptide group.

The dipole derivative, $\partial\boldsymbol{\mu}/\partial Q$, represents an oscillating electric dipole in a peptide group due to its amide I vibration. Physically, the oscillating electric dipole acts like an antenna that can broadcast and receive infrared radiation. The communication is sensitive to the relative orientation of the two dipoles involved. The through-space exchange of radiation between a pair of oscillating electric dipoles causes their vibrational motions to interact. The interaction is represented by a pair of off-diagonal force constants in the **F** matrix of the coupled oscillator system. Every peptide group in the molecule participates in this kind of interaction regardless of the structure in which it is located.

The oscillating dipoles can be regarded as quasistatic if the wavelength of their radiation is much greater than both their physical size and the range of their effects [21]. We take the physical size of the dipoles to be 2.2 Å, the distance between the N and O atoms in a peptide group. The electric field of a dipole falls off inversely proportional to the third power of the distance from its center. At a distance of 50 Å, its field is negligible. The wavelength of 1,650 cm^{-1} radiation is 60,606 Å.

To arrive at an expression for a through-space interaction force constant, we first calculate the total interaction potential energy of the system of dipole oscillators and then differentiate the potential energy with respect to the pair of internal coordinates for the oscillators involved. The internal coordinate of an oscillator in this model corresponds to the normal coordinate of an amide I mode of a peptide group in the classical theory. These internal coordinates are taken to form an orthogonal set in the amide I subspace.

The potential energy of a collection of static dipoles, V, is the sum of all of the pairwise interaction energies, W_{ij}, where

$$W_{ij} = \mathbf{p}_j \cdot \nabla\Phi_i .$$

$\mathbf{p}_j$ is the electric dipole moment of the jth oscillator and $\nabla\Phi_i$ is the gradient of the electric potential of dipole i at the location of dipole j. The oscillating electric dipole in the amide I mode of a peptide group is due to a flow of electrons between the N and O atoms, which are 2.2 Å apart. Therefore, dipole i of moment, $\mathbf{p}$, is located along a local z axis between $+a/2$ and $-a/2$, where $a = 2.2$ Å, and its potential expanded in the axial multipole expansion [21] as

$$\Phi(r,\theta) = pP_1(\cos\theta)/r^2 + pa^2P_3(\cos\theta)/(4r^4) + \cdots$$

P_i is the ith Legendre polynomial. r and θ are the spherical polar coordinates of the point at which the potential is evaluated in the local coordinate system. The commonly used point dipole approximation for the interaction energy results from

retaining only the first term in the expansion. This is a useful approximation when $a/r \ll 1$, but for nearest or next nearest peptide groups in the chain or nearby groups in other parts of the protein backbone this condition is not met. The approximate interaction energy can be expressed as

$$W_{ij} = \mathbf{p}_i \cdot \mathbf{p}_j / {r_{ij}}^3 - 3\left(\mathbf{p}_i \cdot \mathbf{r}_{ij}\right)\left(\mathbf{p}_j \cdot \mathbf{r}_{ij}\right)/{\mathrm{r}_{ij}}^5.$$

The first two terms in the expansion are kept as a refinement of the approximation. For axial symmetry in spherical polar coordinates, the gradient operator is as follows:

$$\nabla = \mathbf{i}_r \partial/\partial \mathrm{r} + \mathbf{i}_{\Theta}\,(1/r)\,\partial/\partial\theta.$$

Then

$$W_{ij} = \mathbf{p}_j \cdot \left[\mathbf{i}_r\left(-2\mathbf{p}_i \cos\theta/{r_{ij}}^3 - \mathbf{p}_i a^2(5\cos^3\theta - 3\cos\theta)/2{r_{ij}}^5\right)\right.$$
$$\left. + \mathbf{i}_\theta\left(-\mathbf{p}_i \sin\theta/{r_{ij}}^3 + \mathbf{p}_i a^2(-15\cos^2\theta\sin\theta + 3\sin\theta)/8{r_{ij}}^5\right)\right].$$

The force constant for this interaction $F_{ij} = (\partial^2 V/\partial q_i \partial q_j)_0 = (\partial^2 W_{ij}/\partial q_i \partial q_j)_0$, where q_i is the internal coordinate of the ith oscillator and $p_i = f(q_i)$. Then

$$F_{ij} = (\partial p_i/\partial q_i)_0 \left(\partial p_j/\partial q_j\right)_0 \mathbf{u}_j \cdot \left[\mathbf{i}_r\left(-2\cos\theta/{r_{ij}}^3 - a^2(5\cos^3\theta - 3\cos\theta)/2{r_{ij}}^5\right)\right.$$
$$\left. + \mathbf{i}_\theta\left(-\sin\theta/{r_{ij}}^3 + a^2(-15\cos^2\theta\sin\theta + 3\sin\theta)/8r_{ij}^5\right)\right],$$

where $\mathbf{u}_j$ is a unit vector along $(\partial \mathbf{p}_j/\partial q_j)_0$.

The units we use for the dipole derivative $(\partial \mathbf{p}_j/\partial q_j)_0$ are $\mathrm{cm}^{3/2}/\mathrm{s}$, which are equivalent to D Angstroms^{-1} amu$^{-1/2}$ as usually used by other researchers. These units result from taking q to be a mass-weighted displacement coordinate instead of a simple displacement coordinate; that is, q has units of $\mathrm{g}^{1/2}\,\mathrm{cm}$. This choice for the units of the dipole derivative results in the units of the force constant above being $1/\mathrm{s}^2$ instead of the proper units for a force constant of g/s^2. Since in the calculations the $\mathbf{F}$ matrix is divided by k, the force constant of an unperturbed oscillator, which makes the diagonal elements dimensionless, the off-diagonal elements must also be dimensionless. Therefore, we divide the above expression for F_{ij} by $\lambda_0 = 4\pi^2 c^2 \nu_0^2 = k/m$ where ν_0 is the frequency of a hypothetical unperturbed amide I oscillator. Generally, a value of $1{,}650\,\mathrm{cm}^{-1}$ has been used for ν_0. λ_0 has the units of 1/s and this makes the through-space transition–dipole interaction force constants dimensionless. The formula for F_{ij} is also divided by r^p, taken to be a dimensionless number, where p is some fractional power. This approach is utilized to model the dielectric effect of the matter between the oscillating electric dipoles. A value of $p = 0.9$ has been used for globular proteins and for sheet structures.

To facilitate the description of the calculation for F_{ij}, we employ the concept of a sending dipole, that is, the one whose electric field, $\nabla\Phi_i$, is being computed at some point in space and a receiving dipole, that is, the one whose dipole moment, $\mathbf{p}_j$, is interacting with the electric field of the sending dipole.

The starting point for these calculations is usually the PDB crystal coordinates of the CON atoms in the peptide groups. Transformations back and forth between local Cartesian coordinates, *x, y, z*, in the crystal and the crystal Cartesian coordinates themselves, *X, Y, Z*, are made via a transformation matrix in terms of the Euler angles, θ, Φ, and χ [10, Appendix I]. Vector rotations are effected by varying the Euler angle χ, which describes rotation around a local z axis. The orientation of the local z axis in the crystal coordinate system defines the Euler angles θ and Φ.

The sending dipole is taken to be centered on the C atom of one peptide group in the molecule. A 2.2 Å vector, which represents the oscillating electric dipole, is extended out from the C atom along the CO bond of that peptide group and then rotated 21° off the CO bond away from the CN bond in the plane of the peptide group. The normal to the plane which gives the orientation of the rotation axis is found by crossing the CN bond vector into the CO bond vector. From the crystal coordinates of the end of the transition dipole relative to the C atom, we calculate the Euler angles θ and Φ since the dipole defines a local z axis with origin on the C atom. The third Euler angle, χ, is taken to be zero in constructing the transformation matrix to be used subsequently.

The crystal coordinates of the end of the receiving dipole are found by a procedure similar to the one for the sending dipole. The receiving dipole is then backed up along itself so as to be centered on the C atom of the receiving peptide group. The receiving dipole is broken into ten contiguous pieces of 0.22 Å each. The Cartesian components of the length of a piece in the crystal coordinate system are converted into the Cartesian components of its length in the local coordinates by the transformation matrix based on the sending dipole. A vector from the center of the sending dipole to the center of a piece of the receiving dipole has its Cartesian components in the crystal system converted into the local system by the same transformation matrix. From these components, the unit vectors $\mathbf{i}_r$ and $\mathbf{i}_\theta$ are computed in the local system. These unit vectors which represent the longitudinal and transverse components of the electric field of the sending dipole are used to form a scalar product with a unit vector of the piece of the receiving dipole in the local system, that is, $\mathbf{u}_j \cdot \mathbf{i}_r$ and $\mathbf{u}_j \cdot \mathbf{i}_\theta$.

The length of the vector from the center of the sending dipole to the center of a piece of the receiving dipole is r in the equation for F_{ij}. Cosθ can be calculated by the ratio of the projection of the vector on the local z axis to itself. Sin θ is calculated from $\cos\theta$ by $\sqrt{(1-\cos^2\theta)}$ since θ goes from 0 to 180°. The interactions of the ten pieces of receiving dipole with the sending dipole are added together to get the total interaction force constant.

The rotation angle of the transition dipole off the CO bond of 21° is for nonproline-N-based groups. For *trans*-proline-N groups, the angle is taken to be −7°. For *cis*-proline-N groups, the angle is taken to be 2°. The numerical values

of the dipole derivative used in the various simulations to be described later will be given then.

The reason for breaking the receiving dipole into ten contiguous pieces is the interaction formula $W_{ij} = \mathbf{p}_j \cdot \nabla \mathbf{\Phi}_i$ pertains to the scalar product of two vectors at a point in space. This implies that $\mathbf{p}_j$ is a point dipole. However, we are using finite dipoles and the electric field of the sending dipole varies in space. By breaking the receiving dipole into smaller pieces, the variation of the electric field acting on the receiving dipole is reduced.

The **F** matrix is symmetric; therefore, the interaction between dipoles i and j is calculated only once. This avoids instabilities in the diagonalization of the **GF** matrix product caused by discrepancies between the F_{ij} and the F_{ji} elements.

2.1.6.5 Through Hydrogen Bond Interaction

The normal coordinate calculations of Krimm and Abe [9] revealed that through valence bond and through hydrogen bond interactions could not produce the observed range of frequencies for amide I. Only transition dipole interaction was able to give the observed range. Torii and Tasumi [14], perhaps influenced by the observations of Krimm and Abe, neglected through hydrogen bond and through valence bond interactions in their model calculations of the amide I band of globular proteins. They limited the interaction considered to transition dipole coupling.

We found it necessary to include both through hydrogen bond and through valence bond interaction in our simulations of the amide I contour. The resulting interaction force constants are only a fraction of a percent of the diagonal force constants, but for accurate modeling they cannot be neglected. More dramatic is the effect that hydrogen bonding has on the intrinsic force constant of the group supplying the CO moiety to the bond as discussed above.

The through hydrogen bond interaction like transition dipole interaction is geometry sensitive. It depends on the C–O, H–N angle and the CO$\cdots$HN distance. Hydrogen atoms are not located in X-ray protein structures, but they are found in NMR structures. When the locations are not supplied we have calculated them by assuming a standard published peptide group geometry, namely, a N–H bond length of 1 Å rotated 123° off the N–C bond to form a *trans*-peptide group.

The hydrogen bonds between peptide groups are "moderately" strong and the interaction is basically electrostatic, that is, a dipole–dipole interaction [22]. There is no buildup of electron density between the O and H atoms. In our calculations, the hydrogen-bonded pairs are found by a function that limits the CO$\cdots$HN distance to 2.6 Å or less and the C–O, H–N angle to 120° or more. The through hydrogen bond interaction force constant between peptide groups i and j is modeled by:

$$F_{ij} = \{[3\cos^2(\kappa - 116) - 1]/d^3\}s,$$

where κ is the C–O, H–N angle in degrees and d is the CO$\cdots$HN distance in Å. The strength factor, s, is set to -0.011. The result is taken to be a dimensionless

number. This formula was suggested by the equation for the axial component of the electric field of a point dipole. The offset of 116° was chosen empirically. It is close to the 120° angle between the unshared electron pairs on the carbonyl oxygen and the CO bond in a simple Lewis model. The location of the center of electron density on the carbonyl oxygen determines the orientation of the interacting dipole on the acceptor moiety.

2.1.6.6 Through Valence Bond Interaction

Even though the methyl groups move relatively little compared to the CONH atoms in the amide I vibration of *N*-methylacetamide, and these methyl groups correspond to the α carbon atoms in a polypeptide chain, the α carbon atoms do move, and they provide a strong connection between adjacent peptide groups in the chain. We found it necessary to include a through valence bond interaction force constant in the **F** matrix of our simulation model.

Each peptide group is separated from its neighboring group by an intervening C^{α} atom, which is covalently bonded to the N atom of the first group and the carbonyl C of the second group. The peptide groups themselves are generally planar. The Ψ, Φ angles at a given C^{α} determine the relative orientation of the adjacent peptide planes. Accuracy in the simulations requires two components to the bond-electron-mediated interaction. These are a through σ-bond term that is constant and a $\pi-\pi$ component that depends on the relative orientation of the peptide planes. We have used the following expression for the through valence bond interaction force constant between peptide groups i and j,

$$F_{ij} = (0.00244 + 0.0151 \cos\xi)\, s,$$

where $\cos\xi$ is the cosine of the angle between the normal of the first plane and the negative of the normal to the second plane. The normals are found by crossing the CN bond vector into the CO bond vector of a group. The strength factor, s, was set to 0.6. The result is taken to be a dimensionless number.

The interaction force constants generated by the through valence bond mechanism are like the through hydrogen bond interaction force constant, only a fraction of a percent of the diagonal force constant for a group, but they too cannot be neglected.

2.1.6.7 Through Hydrogen Bond $\pi-\pi$ Interaction

As mentioned in the discussion of diagonal force constant, many-stranded, antiparallel β-sheet structures exhibit an $\sim 70\,\mathrm{cm}^{-1}$ splitting between the low- and high-frequency peaks in their amide I contour. This large splitting does not appear in the amide I band of most globular proteins because they do not have sheet structures with more than a few strands.

This large splitting suggests there is a special interaction in sheet structures that is relatively weak, but becomes noticeable when the number of laterally bonded chains becomes large. The lateral bonding in sheets is through hydrogen bonds. Through hydrogen bonding has already been taken into account. What else could be happening? It appears that the relatively coplanar and aligned peptide groups in the sheet have a $\pi-\pi$ interaction between the π electrons of the peptide groups. This type of interaction is not important in other structural motifs such as coils and loops because the peptide planes do not align as well in these structures.

We have used the following expression for the through hydrogen bond $\pi-\pi$ interaction force constant between peptide groups i and j,

$$F_{ij} = (\cos\eta)^2 \ (N-2)^{1.2}\, s,$$

where $\cos\eta$ is the cosine of the angle between the normal of the first plane and the negative of the normal to the second plane. The normals are found by crossing the CN bond vector into the CO bond vector of a group. N is the number of chains laterally bonded. The strength factor, s, was set to -0.0002. The result is taken to be a dimensionless number.

2.1.7 Simulating the Amide I Contour

In constructing the **G** and **F** matrices, it is not necessary to individually specify the mass and force constant of the common oscillator. Rather the diagonal **G** matrix is multiplied by the common mass, m, and the diagonal elements of the **F** matrix are divided by the common force constant, k. The new diagonal elements in the **G** and **F** matrices are dimensionless. Thus, the **GF** matrix product has been multiplied by m/k. The off-diagonal elements in the **F** matrix are fractions of the reduced diagonal elements. Since $m/k = 1/\lambda_0$, where $\lambda_0 = 4\pi^2c^2\nu_0^2$, the eigenvalues of the **GF** matrix product equal λ_i/λ_0, where $\lambda_i = 4\pi^2c^2\nu_i^2$ and ν_i is a normal mode frequency. A value of ν_0, the frequency of the hypothetical unperturbed oscillator, is chosen in the calculation of the transition–dipole interaction force constants. The value of ν_0 affects the shape of the simulated band because it affects the value of the transition–dipole interaction force constants.

Once the **G** and **F** matrices are constructed, the eigenvalues and the eigenvector matrix, **L**, of the **GF** matrix product are found. The square root of the eigenvalues multiplied by ν_0 gives the frequencies of the normal modes of the collection of oscillators. It may be necessary to change the value of ν_0 somewhat at this point to make the simulated band coincide with the experimental band. This does not alter the simulated band shape; it merely shifts it along the frequency axis.

The intensity of each normal mode must now be calculated. First, weighted sums of the Cartesian components in the crystal space of all the transition dipoles in the molecule are computed. Recall that each transition dipole is represented by a vector 2.2 Å long rotated by a certain angle off the CO bond in each peptide group.

For a given normal mode, the weighting factors are the elements of the corresponding row of the inverse **L** matrix. The ith row pertains to the ith normal mode and the jth column pertains to the jth peptide group. This is made clear in appendix VIII of *Molecular Vibrations* [10] where one finds $\mathbf{S} = \mathbf{LQ}$. **S** is a column vector of internal coordinates, **Q** is a column vector of normal coordinates, and **L** is the eigenvector matrix. On pre-multiplying both sides of this equation by $\mathbf{L}^{-1}$, we get $\mathbf{L}^{-1}\mathbf{S} = \mathbf{Q}$. Following the rules of matrix multiplication, we see that a row of the inverse eigenvector matrix, $\mathbf{L}^{-1}$, gives the contribution of each internal coordinate to the normal coordinate, which corresponds to that row. In our model, this corresponds to the contribution of the transition dipole on each peptide group to the resultant transition moment for a given normal mode of the collective system.

The intensity of a normal mode is taken to be proportional to the sum of the squares of these summed Cartesian components. If the transition-dipole strength of a given group is different from that of the common group its Cartesian components must be scaled by the ratio of its strength to the dipolar strength of the common group. According to quantum mechanics, the intensity of a spectral line is proportional to the square of the corresponding transition-dipole moment.

The spectral lines are then broadened by centering a lineshape function on each line. Typically, we have used 70% Lorentzian and 30% Gaussian lineshape functions with a common width parameter for both functions and for all the lines of a given molecule. The width parameter is chosen to give the best fit between the experimental and calculated amide I contours. The intensities calculated above are amplitude factors in the lineshape functions. The simulated contour is the sum of all the subbands so produced.

While the correct weighting factor for calculating the resultant transition-dipole moment is $(\mathrm{L}^{-1})_{ij}$, the fractional participation of peptide group j in normal mode i is given by $[(\mathrm{L}^{-1})_{ij}]^2$. The sum of the $[(\mathrm{L}^{-1})_{ij}]^2$ for a given row is 1.

2.2 Applications

In this section, we illustrate the usefulness of amide I simulation by considering the insights it can supply when IR spectroscopy is used to study the structure or behavior of peptides and proteins. First, we consider isotopic labeling which has been anticipated to be a means of finding the amide I contribution of specific residues in a polypeptide molecule, and of identifying the secondary structure in which they are located. Isotopic labeling, however, in antiparallel sheet structures of short peptide chains produces an effect called "anomalous intensity." The origin of anomalous intensity is then explored. The simulation of substituted sheet structures is also used to show that the amide I contour can distinguish whether the chains in amyloid peptides are packed in a parallel or antiparallel fashion. In addition, the interaction and coupling of substituted and unsubstituted oscillators are considered.

Second, we consider predicting the spectral changes expected after structure modification as a way of elucidating the mechanism of the thermal denaturation of a

collagen model peptide. Third, we use data mining based on accurate simulations of the amide I contours of ribonuclease A and horse heart myoglobin to comment on the validity of widely used structure–frequency correlations and on the meaning of the new peaks bands generated by Fourier self-deconvolution of an amide I contour. Finally, we consider how knowledge of the transition-dipole components for each amide I mode of a protein can be used to determine its orientation at the air–water interface. When this is done in the presence and absence of a substrate, the orientations can indicate the location of the binding sites on the surface of the protein. The transition-dipole components can be found from the mathematical results of the simulation. These applications have all been described to some extent in our published studies.

2.2.1 Isotopic Labeling

The spectral contribution of specific peptide groups in a protein to its amide I contour or the secondary structure to which they belong cannot be gleaned from the contour itself. Yet it is desirable to have this information in studies such as protein folding, ligand, or substrate binding, and the formation of fibrous protein aggregates. The substitution of ^{13}C in the carbonyl of a peptide group is expected to lower the independent amide I vibration frequency of that group. On treating the C=O stretching mode as an isolated harmonic oscillator, a 37 cm^{-1} downward shift is predicted.

In addition, substitution is expected to "decouple" the amide I vibrations of the ^{13}C and ^{12}C groups. That is, the ^{13}C group should vibrate in a mode with very little ^{12}C participation. This is often the case but not because, as is revealed by our model calculations, the interaction force constants between the two kinds of groups have been changed. This is elaborated on below. Rather, the "isolation" of the ^{13}C oscillator results from its not responding to the unchanged interaction forces because of its reduced independent vibration frequency. However, there is one case in which there is substantial participation of ^{12}C oscillators in a mostly ^{13}C mode, namely, in short-chain, substituted, antiparallel sheet structures with a large number of laterally bound strands. The ^{12}C participation gives rise to "anomalous" intensity in the ^{13}C mode, which is considered in Sect. 2.2.1.1.

It has been anticipated that isotope editing will reveal the original spectral contribution of the substituted group and that the newly formed isolated oscillator can act as a probe of the local secondary structure in which it finds itself [23]. One attempts to extract this information by subtracting the amide I bands obtained before and after the substitution. However, as we showed in our second simulation publication [17] the difference band does not unfailingly show the original and new spectral contributions of the substituted group. Isotopic substitution does more than shift the frequency of the affected oscillator. It disrupts the coupling patterns of other oscillators that existed before the substitution and one could see the bizarre result that intensity disappears from a lower frequency and appears at a higher frequency in

the difference spectrum. In this case, the peak positions in the difference spectrum obviously do not indicate the original and new contributions of the group to the contour. This was the simulation result [17] obtained by substituting peptide group 52 located in helix 3 of ribonuclease A.

One seems to fare better when substitutions are made in sheet regions of a protein molecule possibly because in the relatively open sheet structure the substituted oscillator is less coupled to other groups. Thus, simulation of isotopic substitution in peptide group 80 located in a sheet region of ribonuclease A produced a difference band that closely matched the original and new spectral contributions of the group.

Substitution in the extended beta-sheet conformation has been observed to produce new, readily observed bands from 1,615 to 1,584 cm^{-1}. The actual frequency shifts and intensity changes depend on a number of factors such as whether the substitutions are single or multiple, the location of the substitutions in the chain, the spacing between multiple substitutions in one chain, the length of the chains, and whether the strands join in the parallel or antiparallel conformation. The new low-frequency, observable bands are good evidence that the substitution was made in a sheet region. Unexpectedly high intensity in these bands tells us that the sheet structure consists of many strands laterally bound together in an antiparallel fashion as is explained in the next section.

Substitution in group 52 in helix 3 of ribonuclease A (Fig. 2.5 in [17]) also produces a low-frequency band at 1,610 cm^{-1}, but this new band was not manifest in the difference spectrum. The reason for this is simply that the substitution of one group in a protein produces a very small change in the amide I contour. The spectral changes caused by disruptions of the coupling patterns in the unsubstituted molecule can easily mask the spectral changes due to the substituted oscillator itself.

2.2.1.1 Anomalous Intensity Distribution in the β-Sheet Conformation

When peptides known to exist in the β-sheet conformation are isotopically substituted with ^{13}C in their peptide carbonyls, some intensity of amide I in the strong low-frequency peak at around 1,621 cm^{-1} disappears and a new band at 1,610 cm^{-1} or as low as 1,584 cm^{-1} appears. Often, the new lower frequency band has a greater area than expected on the basis of the number of substitutions made. The larger than expected intensity is referred to as "anomalous." We investigated this phenomenon in our first publication [16] on the simulation of the amide I contour. The experimental contours of three isotopically substituted versions of a simple hydrophobic peptide, $K_2(LA)_6$, known to form an intermolecular antiparallel β-sheet in methanol solution were obtained. These substituted peptides showed substantial anomalous intensity.

In simulating the contour for these short-chain peptides, we took the through valence bond and through hydrogen bond interaction force constants to be constants. This was possible because a regular, antiparallel, two-stranded sheet generated from the coordinates supplied in Fraser and MacRae [24] was used to model the peptides in solution. However, we made the through hydrogen bond interaction force constant

for the two-stranded sheet model twice that used for the helical $K_2(LA)_{10}$ peptide, even though the CO $\cdots$ HN distances and the CO, HN bond angles are similar in the two conformations. The importance of $\pi-\pi$ interactions between laterally bonded peptide groups was not appreciated at that time.

Recently, we obtained an improved fit for the $K_2LA^*LA^*(LA)_4$ disubstituted peptide by using the simulation method as described in this chapter, which includes $\pi-\pi$ interactions. Figure 2.1a depicts our original comparison of experimental and simulated amide I spectra for this molecule. The vertical lines represent intensities

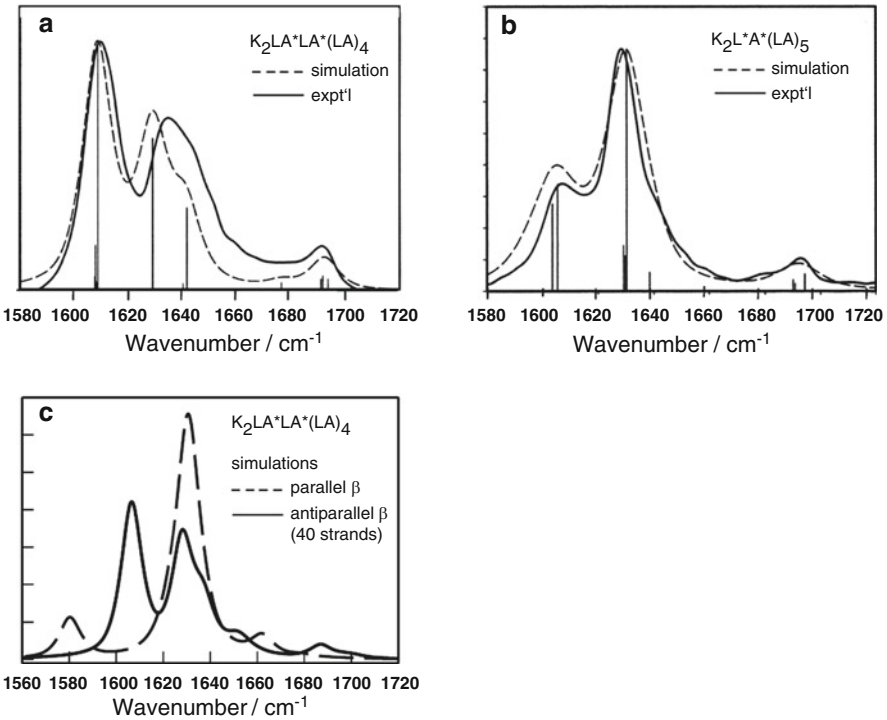

Fig. 2.1 (**a**) Experimental (*solid line*) and simulated (*dashed line*) spectra of the amide I (1,580–1,720 cm^{-1}) region for $K_2LA^*LA^*(LA)_4$. Intensities of the components generated from the model are represented by vertical *solid lines* at the calculated frequencies. The *dashed* contour is generated by broadening each component as a 30% Gaussian–70% Lorentzian function with 6 or 7 cm^{-1} full width at half height. (**b**) Experimental (*solid line*) and simulated (*dashed line*) spectra of the amide I (1,580–1,720 cm^{-1}) region for $K_2L^*A^*(LA)_5$. Intensities of the components generated from the model are represented by vertical *solid lines* at the calculated frequencies. The *dashed* contour is generated by broadening each component as a 30% Gaussian–70% Lorentzian function with 8 or 10 cm^{-1} full width at half height. (**c**) Ability to distinguish antiparallel from parallel β-sheets in a labeled peptide. Simulations of an antiparallel β-sheet (40 antiparallel strands of 14 residues each) in the amide I (1,580–1,720 cm^{-1}) region for $K_2LA^*LA^*(LA)_4$ (*solid line*) are compared to the parallel β-sheet 40 strands of 14 residues each. The origin of the intense 1,600 cm^{-1} component in the the antiparallel structure arises from $\pi-\pi$ interaction as discussed in the text

of amide I components calculated from the model at the corresponding frequencies. The lines are broadened as discussed above. The interesting feature of this spectrum (the molecule is known to be in the β form) is the anomalously large intensity at 1,611 cm^{-1}. As shown by comparison of Fig. 2.1a with Fig. 2.1b, shifting the position of the second label by one residue doubles the intensity of this low-frequency feature. However, as noted above, to achieve this fit required a doubling of the through H-bond interaction force constant for the two-stranded sheet model to twice that used for the helical $K_2(LA)_{10}$ peptide. This was difficult for us to justify on physical grounds.

To achieve a similar quality of fit with H-bond interaction constants calculated by the formula given in Sect. 2.1.6.7, we used a molecular model consisting of a 40-strand sheet with 13 peptide groups per chain. The same standard antiparallel geometry as in our first effort was used [24]. A 40-strand sheet was necessary to reproduce the observed $>70\,cm^{-1}$ split between the high- and low-frequency peaks in the amide I contour. The $\pi-\pi$ interaction is inherently weak, but becomes strong when many β strands are laterally hydrogen bonded together. The other interactions as calculated by the given formulas will not by themselves produce the broad peak separation seen in the contour of many β peptides.

The inverse **L** matrix produced by this simulation shows that the anomalous intensity in the predominantly ^{13}C modes is due to the substantial participation of ^{12}C oscillators. This is the same conclusion we reached in our first effort. The reason for the unexpected ^{12}C participation in ^{13}C modes is discussed in Sect. 2.2.1.2.

2.2.1.2 Amyloid-Peptide Packing Models

Whether the chains in β-amyloid fibrils pack in a parallel or antiparallel fashion can help resolve proposed mechanisms for their formation from properly folded proteins. However, conflicting conclusions have been reported for different model peptides. Halverson et al. [25] studied the nine amino acid peptide β34–42 (H_2N–Leu–Met–Val–Gly–Gly–Val–Val–Ile–Ala–CO_2H), which represents the carboxyl terminus of the amyloid-forming protein of Alzheimer's disease. The peptide resembles the native protein in its insolubility and cross β-fibrillar structure. FTIR analysis of the β34–42 fibrils was interpreted as indicating a predominantly antiparallel β-sheet structure. Seven ^{13}C carbonyl singly labeled analogues of this peptide were also prepared and their films analyzed by FTIR. All but the N-terminus group labeled peptide gave an anomalously intense ^{13}C band located between 1,605 and 1,615 cm^{-1}.

Paul and Axelson [26] obtained amide I contours of the amyloid protein Aβ40 in a fibrillized state by evaporating a solution of the incubated polypeptide onto an internal reflection crystal. The amide I bands of two singly ^{13}C-substituted versions of this molecule were also obtained. On the basis of a blue-shifted ^{12}C band and what they regarded as a disproportionately strong red-shifted ^{13}C band, they concluded that the chains packed in a parallel fashion. The disproportionate strength of the

^{13}C band was attributed to the proximity of the ^{13}C groups in the in the fibrils formed from the substituted molecules. However, the spectrum of the unlabeled molecule has a strong amide I′ peak at 1,624 cm^{-1} and a high-frequency shoulder at 1,685 cm^{-1} which Torii [27] points out is considered to be a marker for antiparallel packing. Our simulations [17] confirm that only the antiparallel chain packing gives rise to a high-frequency peak.

Tycko and coworkers [28] synthesized a 40-residue amyloid peptide. They used multiple-quantum, solid-state NMR to show that this peptide packs in a parallel chain structure.

To see how chain packing affects the amide I contour of substituted, short-chain peptides, we calculated (Fig. 2.1c) the band for the disubstituted $K_2LA^*LA^*(LA)_4$ peptide in the parallel configuration. The good match (compare Fig. 2.1a, c) between the experimental and simulated contours when antiparallel packing was assumed, as noted previously, gives us confidence that the molecule does indeed pack in an antiparallel form in methanol solution and that the simulation parameters are correct. The model parameters were kept the same for the two structures, only the geometries differed.

As is clear from Fig. 2.1c, there is a dramatic difference between the calculated amide I contours of the two forms. The antiparallel sheet gives rise to an extremely intense ^{13}C band at 1,606 cm^{-1} matching that in the experimental contour, while the ^{13}C band at 1,580 cm^{-1} for the parallel sheet model has no anomalous intensity whatsoever. The ^{12}C band for the antiparallel sheet at 1,628 cm^{-1} is much broader and more structured than the ^{12}C band at 1,631 cm^{-1} for the parallel sheet. The calculated contours also show a high-frequency peak at 1,665 cm^{-1} for the parallel form and at 1,685 cm^{-1} for the antiparallel form.

If we refer to a line of *trans*-chain hydrogen-bonded peptide groups in the sheet as a row then there are two rows in the parallel form, which contain only substituted groups. There are four rows in the antiparallel form that contain substituted groups alternating with unsubstituted groups. The inverse **L** matrix shows that the low-frequency modes at 1,580 cm^{-1} for the parallel form come from ^{13}C groups almost exclusively. For the antiparallel sheet structure, the inverse **L** matrix shows that the ^{13}C modes at 1,606 cm^{-1} have about a 28% contribution from ^{12}C oscillators. One mode at 1606 cm^{-1} which has 33% of the total contour intensity has 56 oscillators with a contribution of ½% or greater to the mode. Of these 56 groups, 14 are ^{12}C groups. The ^{12}C groups active in the mode are situated between two ^{13}C groups in a row. The ^{13}C group that supplies the CO moiety to an active ^{12}C oscillator has a larger contribution to the mode than many of the other ^{13}C oscillators. These combinations are located in rows near the middle of the sheet.

Even though there is a mismatch in the independent vibration frequencies between the two kinds of oscillators only because of the mass difference, the geometry of the antiparallel sheet and the strong interaction force constants elicit the participation of the ^{12}C groups in the ^{13}C low-frequency modes. The participation of ^{12}C groups in the low-frequency ^{13}C modes does not occur in the parallel chain conformation. Moreover, the participation of ^{12}C groups in the low-frequency ^{13}C modes of the antiparallel form causes their red shift from the low-frequency ^{12}C

peak to be considerably less than the corresponding red shift of the pure ^{13}C modes of the parallel form.

When the chains are 40-residues long, disubstitution does not result in dramatic anomalous intensity for the antiparallel structure. There is reduced participation by ^{12}C groups in the low-frequency ^{13}C modes. The parallel structure for these longer chains gives no anomalous intensity at all. Again, the separation between the low-frequency ^{12}C and ^{13}C modes is similar to that noted for the 14-residue chains. However, when six substitutions are made in a 40-residue chain as, for example, the six glycines in Tycko's 40-residue amyloid peptide, then once again there results considerable anomalous intensity for the antiparallel packing, but none for the parallel sheet. The separations between the low-frequency ^{12}C and ^{13}C modes continue to be much larger for the parallel packing. The coupling of ^{12}C oscillators into largely ^{13}C modes causes their red shift to be considerably less than that seen for practically pure ^{13}C modes.

2.2.1.3 Interaction and Coupling

The words interaction and coupling are often used interchangeably as we have done occasionally in this chapter. But there is a subtle difference in their meanings. Interaction leads to coupling. A vibrational interaction force constant between two oscillators means that the vibration of one oscillator exerts an oscillating force on the vibration of the other. The force can be transmitted electromagnetically, electrically, or mechanically. Two oscillators are coupled if they participate jointly and significantly in one or more normal modes of the collective system of oscillators. The degree of coupling between two oscillators depends on the total strength of the interaction between them and on the correspondence of their independent vibration frequencies. The formulas we have employed for the calculation of the four types of interaction force constants used in our simulation method and their evaluation for specific cases lead to the following points:

The through-space transition–dipole interaction force constant is not frequency dependent. The transition dipoles are taken to be quasistatic and there is no factor in the formula for this interaction force constant allowing for the frequencies of the oscillators being coupled.

The transition–dipole interaction force constants are of about the same size in helices and sheets. It is not because these constants are larger in sheets than in helices that we get the anomalous intensity seen for peptides that form β-sheets. Rather, it is the additional $\pi-\pi$ interaction in laterally broad sheets and the alternation of ^{12}C and ^{13}C oscillators in a hydrogen-bonded line across the sheet in antiparallel packing that causes the two kinds to participate jointly in the new low-frequency bands seen on isotopic substitution as was explained above.

No interaction force constants are changed in our model because of isotopic substitution. If a ^{13}C-substituted oscillator vibrates in a relatively pure mode at a lower frequency it is because its increased mass lowers its independent vibration frequency and this makes it less likely to resonate with ^{12}C oscillators. The ^{13}C

oscillator is effectively decoupled from the unsubstituted groups. This is the case in globular proteins even in sheet domains, which are never broader than five or so strands and the $\pi-\pi$ interaction is weak.

2.2.2 Modeling the Early Stages of Thermal Denaturation

In our third publication [18], our simulation method was used to produce the amide I contour corresponding to the early stages of the thermal denaturation of a collagen model peptide in aqueous solution. These were compared to the experimental bands obtained for this process. The peptide was $[(\text{Pro–Pro–Gly})_{10}]_3$ or $(\text{PPG})_{10}$. The X-ray structure of $(\text{PPG})_{10}$ is available from the Protein Data Bank as file 1KF6. In the crystalline state this molecule consists of a triple helix, which forms a tightly wound, rigid rod. This structure was modified by a 1 ns molecular mechanics simulation in explicit water at 16 °C to obtain a more realistic starting structure for a molecule in aqueous solution.

To model the effects of increasing the temperature of the system, Φ, Ψ angle changes were made at the α-carbon atoms in the three peptide chains of the triple helix starting at each outer end and proceeding inward one carbon atom at a time. The angle changes were suggested from the dihedral angles found in a molecular mechanics simulation for the molecule at 27 °C. These angle changes produced a gradual unraveling of the molecule from each end. A partially unwound structure of the $(\text{PPG})_{10}$ molecule is shown in Fig. 2.2a.

Most of the modified structures produced no unphysical van der Waals' contacts and gave rise to reasonable amide I IR bandshapes. A sequence of contours representing ten steps of simulated unwinding closely matched the experimental series of contours taken at small temperature intervals. Changes in the simulated amide I contour as function of this gradual unwinding of the triple-helical structure are shown in Fig. 2.2b, and compared (Fig. 2.2c) with corresponding experimental IR spectra over the range 12–34 °C. Good agreement between the experimental and simulated amide I bands indicates that the early stages of thermal denaturation have been correctly modeled.

This success suggests that we can use a similar procedure to model proposed conformational changes in which a known protein starting structure is altered by making appropriate dihedral angle changes at the α-carbon atoms in the protein backbone. The coordinate system rotation matrix using the Euler angles as described in Wilson, Decius, and Cross [10, Appendix I] is convenient for this purpose. The calculated amide I contours can be compared with the experimental bands.

But more was learned from this study than just the mechanism of denaturation of a collagen model peptide. In the first place, a more accurate simulation of the amide I contour of $(\text{PPG})_{10}$ was required than was obtained in our second publication [17]. This means a closer examination of the model parameters for the peptide groups. $(\text{PPG})_{10}$ turned out to be an ideal choice for a more refined assignment of group parameters. It has only two kinds of residues, and a rigid

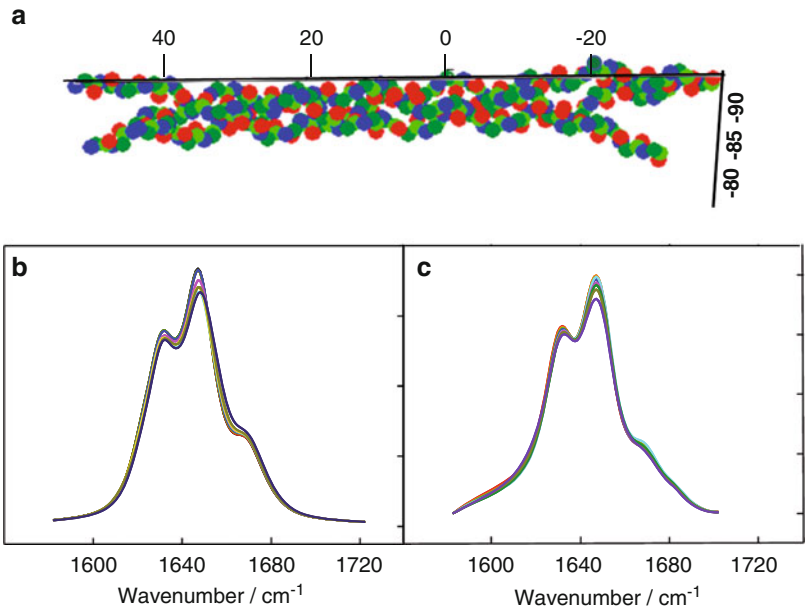

Fig. 2.2 (**a**) Structure of a partially unwound $(PPG)_{10}$ molecule after 13 steps of simulated unwinding. The scale numbers are in Å. (**b**) Simulated changes in the amide I contour for $(PPG)_{10}$ as a function of progressive gradual unraveling (the intensity if the 1,640 cm^{-1} peak decreases as the unraveling proceeds) of the triple helix, according to molecular mechanics simulations. (**c**) Experimental FTIR spectra corresponding to a temperature range of 12–34°. The intensity of the 1,640 cm^{-1} peak decreases as the temperature increases

geometry. Its amide I contour is much more idiosyncratic and detailed than that of typical globular proteins.

In [17], identical parameters were assumed for the two Pro groups, that is, the two peptide groups whose nitrogen atoms come from each of the two proline residues in the $(PPG)_{10}$ sequence. In our third simulation paper [18], distinctions between the two Pro groups and changes for the Gly group were required to achieve an excellent fit to the experimental band. The inverse mass factor for the Pro1 groups was set to 0.986, while that for the Pro2 groups was set to 0.963. The Pro2 groups are between α-carbons, both of which are part of a pyrrolidine ring while only one α-carbon atom bracketing the Pro1 group is part of such a ring. This appears to increase the effective mass of the Pro2 oscillator.

The transition-dipole strength of the Pro1 groups was kept at 400 $cm^{3/2}/s$, while that for the Pro2 groups was lowered to 300 $cm^{3/2}/s$. This reduction in strength reflects the effect of the hydrogen bond between the CO of a Pro2 group and HN of a Gly group on another chain in the triple helix. The hydrogen bond apparently decreases the electron-density flux during the amide I vibration of the Pro2 group. The transition-dipole strength of the Gly groups, that is, the peptide groups whose nitrogen atom comes from a glycine residue, was reduced to 220 $cm^{3/2}/s$ from the

value of 300 $cm^{3/2}/s$ used earlier. Apparently, the hydrogen atom on the α-carbon atom next to the nitrogen atom of a Gly group does not supply as much electron density to the Gly group as a pyrrolidine ring does for the Pro groups.

To test the validity of these parameter changes, the amide I contour of the collagen model peptide $(Pro–Hyp–Gly)_4$–Pro-Hyp–Ala–$(Pro–Hyp–Gly)_5$ was also simulated. The parameters for the two molecules should be similar. Two small changes were necessary to get an excellent fit to the experimental band. The inverse mass factor of the Pro1 groups was set to 0.976 and the transition-dipole strength of the Gly groups was set to 250 $cm^{3/2}/s$. The only justification for these new values is the excellent agreement between the simulated and experimental amide I contours. In our experience it is very unlikely that wrong parameters would give a good fit along the entire contour. It is impossible to tweak one isolated part of the curve with a parameter value change. Any change affects the whole band shape. The experimental and simulated amide I contours for $(PPG)_{10}$ and for $(Pro–Hyp–Gly)_4$–Pro–Hyp–Ala–$(Pro–Hyp–Gly)_5$ are compared in Fig. 2.3a, b. Included in the figures are simulated contributions of particular peptides.

The current approach allowed us to clarify the assignment of the three peaks that occur in the experimental amide I contour of $(PPG)_8$, which was studied by Lazarev et al. [29]. The molecule is a repeating triplet of peptide groups so it was assumed that each group in the triplet gives rise to its own band. The notation we use is elaborated in the caption to Fig. 2.3a. They assigned the peak at 1,628 cm^{-1} to the Pro1 groups and the peak at 1,642 cm^{-1} to the Pro2 groups because the intensity of the 1,642 cm^{-1} band decreases the fastest on raising the temperature of the solution of these molecules. The carbonyls of the Pro2 groups are hydrogen bonded to hydrogen atoms of the Gly groups in another chain of the triple helix. They reasoned that the breaking of these hydrogen bonds was the mechanism for the thermal denaturation of the triple helix, which would have the greatest impact on the Pro2 groups. Our simulations show that the 1,642 cm^{-1} band is primarily due to the Pro1 groups, while the Pro2 groups contribute most strongly to the 1,628 cm^{-1} peak.

2.2.3 Amide I Structure-Frequency Correlations in Globular Proteins

In our second publication on the quantitative reconstruction of the amide I contour of proteins and peptides [17], we applied our simulation method to five globular proteins. Excellent fits between the experimental and simulated bands were obtained for ribonuclease A and for horse heart myoglobin.

Bovine pancreatic RNase A possesses 124 residues and possesses an α, β structural motif with a binding pocket for RNA. The N-terminal half of the protein contains three alpha-helices (residues 3–13, 24–34, and 50–60), while the C-terminal consists of three β-hairpins (residues 61–74, 79–104, and 105–124), two of which (residues 61–74 and 105–124) exist in a four-stranded, antiparallel β-sheet

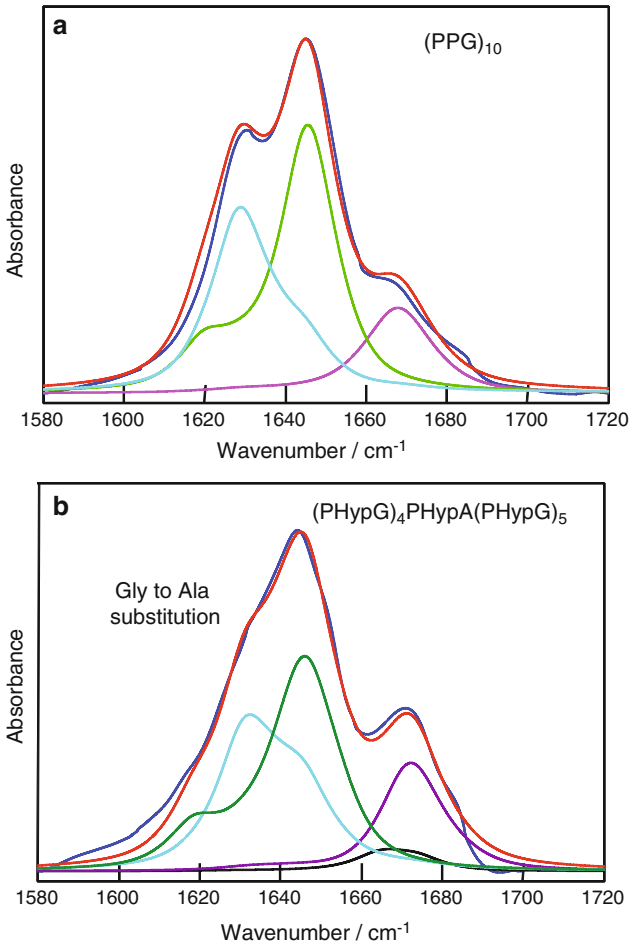

Fig. 2.3 (**a**) The experimental (*dark blue line*) and simulated (*red line*) amide I contour for triple helical $(PPG)_{10}$ in solution. The total simulation is broken into contributions of the peptide bond preceding residue Gly ($C_\gamma O_\gamma$ – *red line*), Pro1 ($C_G O_G$ – *green line*), and Pro2 ($C_X O_X$ – *blue line*) subbands. The group name comes from the residue providing the N in the peptide bond. (**b**) The experimental (*dark blue line*) and simulated (*red line*) amide I contour for triple helcal $(PPG)_{10}$ in solution. The total simulation is broken into contributions of the peptide bond preceding residue Gly ($C_\gamma O_\gamma$ – *red line*), Ala ($C_{\gamma\text{-Ala}} O_{\gamma\text{-Ala}}$ – *black line*), Pro1 ($C_G O_G$ – *green line*), and Pro2 ($C_X O_X$ – *blue line*) subbands. The group name comes from the residue providing the N in the peptide bond

structure that lies on helix 3 (residues 50–60). The β-hairpin 79–104 interacts with a β-strand (residues 42–45) to form a three-stranded, antiparallel β-sheet that lies on helix 2 (residues 24–34).

Horse heart myoglobin is a single-chain globular protein of 153 amino acids, containing a heme prosthetic group in the center around which the remaining

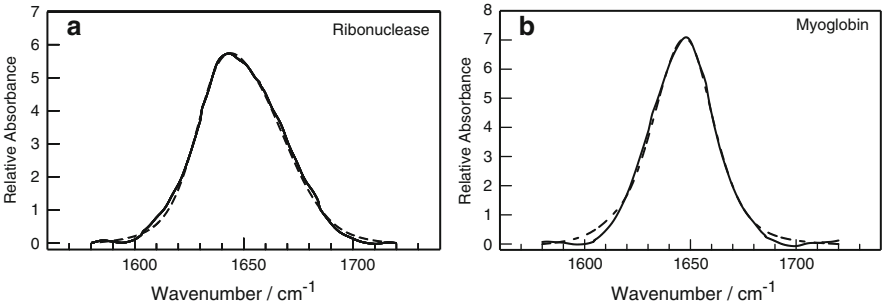

Fig. 2.4 (**a**) Experimental (*solid line*) and simulated (*dashed line*) amide I contours for ribonuclease A (Protein data bank entry, 1AFU) and (**b**) Experimental (*solid line*) and simulated (*dashed line*) amide I contours for horse heart myoglobin (Protein data bank entry, 1YMB)

apoprotein folds. Its secondary structure contains a very high proportion (75%) of α-helical secondary structure. The polypeptide consists of eight separate right-handed α-helices, connected by short nonhelical regions. In the tertiary structure, amino acid R-groups pack into the interior of the molecule and are mostly hydrophobic in character, while those exposed on the surface of the molecule are mainly hydrophilic.

Comparison of the simulated with the experimental amide I contours is shown in Fig. 2.4a, b for RNase A and myoglobin, respectively. The good agreement in each case gave us confidence to "data mine" the results. One such endeavor was to plot (Fig. 2.5a) the RNase subbands due to oscillators in each of the four basic structural motifs, namely, helix, sheet, loop, or turn for this protein. We are able to deduce the structural motif to which each peptide group belongs by examination of the hydrogen-bonding patterns in the molecule as revealed by geometrical calculations made on its structure. The results show that the subbands for the four kinds of secondary structure are much broader and overlapped than has been assumed in Fourier self-deconvolution and band fitting procedures aimed at determining the spectral contributions to the amide I contour of different secondary structures. Deconvolution does produce new peaks in the amide I contour, but these peaks do not represent bands each arising from one kind of secondary structure. Over most of the frequency range of the contour one has concurrent contributions from all four motifs. Even the individual normal modes often involve proximate peptide groups in different structures.

For ribonuclease A, we note the more tightly wound helix 3 gives a peak more than ten wavenumbers higher in frequency than the less tightly wound helices 1 and 2. Helices 1 and 2 are classical α-helices in that the carbonyl oxygen of one group is hydrogen bonded to the amine hydrogen of a peptide group three units down the chain. In helix 3, some of the hydrogen bonds are between groups separated by two as well as three units. One can actually see the tighter coiling of helix 3 by plotting the groups involved in Mathcad's 3D plotting facility and looking down the helix axis. This visualization also serves as a check as to exactly which groups belong to a given substructure.

Data mining was also carried out for myoglobin. The subbands for each of the eight α-helices were calculated and the results are plotted in Fig. 2.5b. The subbands all span a range from 1,620 to 1,680 cm^{-1}. The shorter helices give less symmetrical contours, and helix 3, which is quite short, gives a triple humped contour. The wide

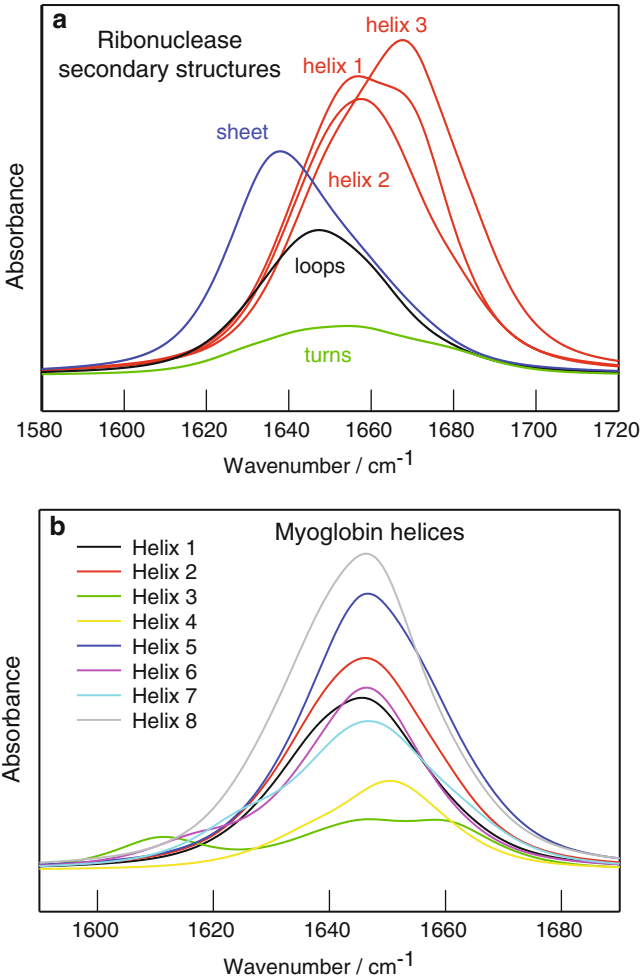

Fig. 2.5 (**a**) "Data mining" for ribonuclease A. The calculated spectral contributions for oscillators in the indicated secondary structures are noted. Helical intensities have been multiplied by 10. Helices 1 and 2 have the classical H-bonding pattern, while helix 3 is more tightly wound. The sum for several different sheet regions (total 53 peptide groups) is shown including three turns within sheet structures. Various types of turns differing in the number of groups separating the H-bonded residues and loop regions composed of 40 peptide groups yield broad bands. (**b**) "Data mining" for myoglobin helices. The calculated spectra contributions to the amide I contour for the eight helical regions are plotted. The two short helices (3,4) have the lowest intensity amide I contours. The triple humped contour of helix 3 may reflect "end effects" in short helical segments

spectral range for these helices far exceeds the narrow shapes assumed in band fitting analyses.

2.2.4 IRRAS Simulations

Thus far we have reported on simulations of the amide I contour of proteins and peptides in aqueous solution. Now we turn our attention to another important system, namely, proteins adsorbed as monolayers at the air–water interface. This experimental paradigm serves as a model for many physiologically relevant sampling conditions such as those occurring at the air–alveolar interface in a mammalian lung or at during the interaction of air-borne pathogens with the host defense system in the same environment.

Infrared reflectance-absorption spectroscopy (IRRAS) produces amide I contours that depend on the orientation of the protein on the water surface. This is so because the protein molecule is optically anisotropic; that is, its frequency-dependent IR extinction coefficients along three Cartesian coordinates in a molecular frame are all different. Then a monolayer of similarly oriented, adsorbed protein molecules will form an anisotropic film whose laboratory-frame, Cartesian extinction coefficients are all different.

The determination of the protein orientation at an interface is facilitated by reflecting light polarized parallel (p) and perpendicular (s) to the plane of incidence. A ratio of the IR absorptions in the two polarizations can be calculated. The absorption must be simulated from the extinction coefficients of the protein in its molecular frame translated into the laboratory frame via the Euler coordinate axes rotation matrix. However, only the relative magnitude of the extinction coefficients in the molecular frame is needed since a ratio of absorptions in the two polarizations is to be calculated. The Euler angles in the rotation matrix that enable agreement with the experimental dichroic ratio define the molecular orientation in the laboratory frame.

The computation of the ratio of absorptions of s- and p-polarized light uniquely samples the three different extinction coefficients of a biaxial film. If the normal to the surface is the z direction and the plane of incidence is the zy plane then s radiation interacts with k_x, the extinction coefficient in the x direction, and p radiation interacts with both k_y and k_z. The relative participation of k_y and k_z in the parallel absorption depends on the angle of incidence of the light. The simulation of the IRRAS contour in the two polarizations requires a combination of an optical theory of reflectance and a method of amide I reconstruction. Our amide I contour reconstruction method can provide relative values for the three extinction coefficients over the frequency range of the band. The optical theory of Kuzmin [30–32] can calculate the reflectance–absorption of a thin film between two semi-infinite phases such as air and water in s and p polarization. The theory requires as input the laboratory-frame Cartesian components of the extinction coefficient and the index of refraction of the film and the optical constants of the air and aqueous subphase.

An example of this combination of calculations is our investigation [33] of the interaction of the neck and carbohydrate recognition domain (NCRD) of recombinant surfactant protein D (SP-D) with a lipopolysaccharide (LPS) at the air–water interface. The intent was to determine the location on the surface of the protein fragment of binding to the LPS. This would be indicated by the tilt angle of the of the symmetry axis of the trimeric fragment with the normal to the water surface. The tilt angle corresponds to the polar Euler angle, Θ. SP-D kills gram-negative bacteria by attaching to the LPS in their cell walls and then breaking the wall down. This is an important mechanism of antimicrobial host defense.

The simulated (inverted) IRRAS spectrum for the neck and CRD regions of SP-D based on the X-ray coordinates of the protein structure with bound maltose is shown for p (parallel) and s (perpendicular) polarized radiation in Fig. 2.6a, b. The calculation assumed a 36° angle of incidence with the tilt angle of the neck axis aligned at 90° (solid line) and 0° (dashed line) with respect to the normal to the air/water interface. The most significant changes in the contour as the protein orientation is altered occur at 1,654 cm^{-1}. Thus, the dichroic ratios formed from the s and p absorption intensities in the amide I contour at 1,654 cm^{-1}, over a range of angles of incidence, were used to determine the orientation of the NCRD SP-D on the water surface in the presence of adsorbed LPS. Generally, band areas should be used in computing dichroic ratios, but in this case the use of areas proved to be insensitive to the tilt of the protein on the water surface because most of the molecule consists of randomly oriented peptide groups. Simulations showed that the absorption at 1,654 cm^{-1} due to the neck region

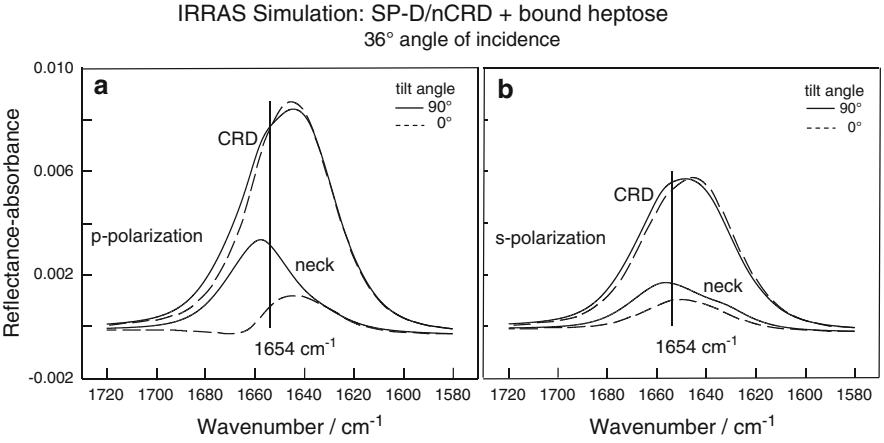

Fig. 2.6 Simulated IRRAS amide I subbands (*inverted*) for the neck and carbohydrate recognition domains of human SP-D based on the X-ray coordinates of the protein with bound L,D heptose. The peptide bond coordinates were downloaded from Protein Data Bank structure 2RIB. The calculation assumed an angle of incidence of 36° for (**a**) p-polarized and (**b**) s-polarized IR radiation with the tilt angle of the neck axis aligned at 90° (*solid line*) or 0°(*dashed line*) with respect to the normal to the air/water interface.The *vertical line* drawn at 1,654 cm^{-1} is intended to highlight the sensitivity of the intensity in the neck subband to orientation

of the molecule was the spectral feature most sensitive to protein orientation as indicated above.

Details of the dichroic ratio simulation follow. The molecular-frame Cartesian components of the resultant transition-dipole moment for each amide I normal mode of the SP-D fragment were calculated from a simulation of the amide I contour of the fragment. The PDB file 1PW9 supplied the atomic coordinates for the simulation. Before the simulation was undertaken the atomic coordinates were rotated to make the threefold symmetry axis the molecular z axis by using an Euler angle rotation matrix. The positive direction of the z axis extends up through the neck region and through the carbohydrate recognition domain.

The amide I contour of the rotated NCRD SP-D was simulated to reproduce the experimental contour of the molecule in aqueous solution. The molecular-frame Cartesian components of the transition-dipole moments for each amide I normal mode were imported into a reflectance-absorption program based on Kuzmin's equations. Here, the molecular-frame components were transformed into laboratory-frame Cartesian components by an Euler angle coordinate axes rotation matrix. The values of the Euler angles give the orientation of the protein in the laboratory frame. The transformed transition moment components were then squared to get the corresponding components of the extinction coefficient. The Cartesian component of an extinction coefficient is proportional to the square of the corresponding transition moment. The laboratory-frame extinction coefficient components were each broadened with a 70% Lorentzian and 30% Gaussian band shape function centered at the frequency of each normal mode and summed. The result is a band of amide I extinction coefficients at $1\,\mathrm{cm}^{-1}$ intervals. The extinction coefficient components were each multiplied with a common strength factor adjusted to make the simulated bands in the s and p polarizations approximately fit the experimental bands. The value of the strength factor is not critical in computing a dichroic ratio because the same value is used for both polarizations. The Cartesian components of the refractive index of the film which exhibits anomalous dispersion were related to the corresponding extinction coefficient by an equation of the form

$$nx_i(\nu) = 1.41 - 2s\,kx_i(\nu)(\nu - \nu_{i,0})/[4(\nu - \nu_{i,0})^2 + \sigma^2].$$

ν is the frequency of the infrared light in cm^{-1}, $\nu_{i,0}$ is the frequency of mode i, $kx_i(\nu)$ is the x direction, frequency-dependent extinction coefficient of the ith mode, s is a strength factor, and σ is a width parameter. Subbands of this shape were summed over all the normal modes to obtain an anomalous-dispersion, amide I refractive index band.

In the Euler rotation matrix, the polar angle, θ, represents the tilt of the molecular z axis from the laboratory Z axis which is perpendicular to the water surface and positive upward. This angle was varied in calculating IRRAS bands in s and p polarization. The other two Euler angles Φ and χ were kept constant at 45°.

In the presence of the LPS monolayer, the tilt angle needed to reproduce the experimental dichroic ratios as the function of angle of incidence was in the range of 0–20° as shown in Fig. 2.7. On a clean water surface, the tilt angle was 90°. These

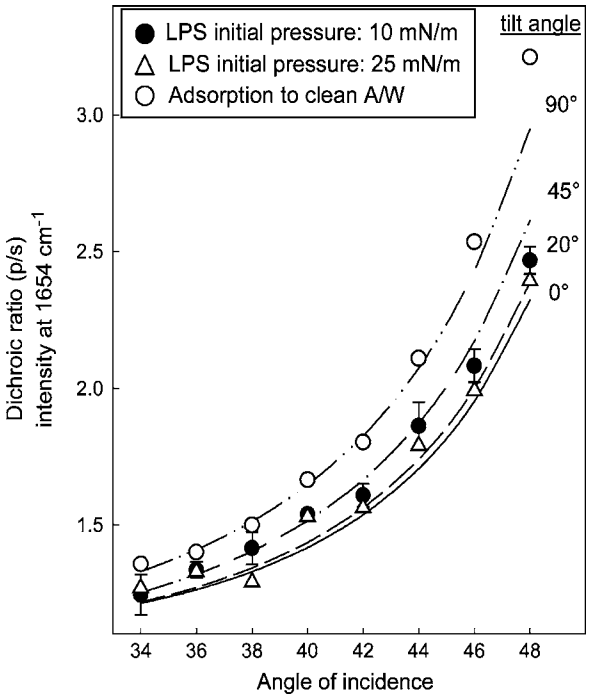

Fig. 2.7 Determination of SP-D (neck + carbohydrate recognition domain) orientation at the A/W interface in the presence and absence of lipopolysaccharide (Rd1) monolayers. Dichroic ratios calculated from simulated amide I band intensities (peak heights at 1,654 cm^{-1} over a range of incident angles are compared with experimentally measured dichroic ratios. (symbols: (*open circle*) clean A/W interface; (*filled circle*), Rd1 aqueous monolayer folm compressed to 10 mN/m, and (*open triangle*) to 25 mN/m. The mean and standard deviations for four separate experiments are shown for monolayers compressed to 10 mN/m, while averages for two experiments are shown for the remaining two pressures

results suggest that all three CRDs contact the LPS monolayer and that the coiled neck region extends into the subphase approximately normal to the water surface. The 90° tilt in the absence of the LPS indicates the equilibrium orientation of the fragment on a water surface in the absence of any specific binding.

It is possible to "push the envelope" of the amide I simulation procedures by use of data mining to attempt to identify the structural origin of small shifts in the amide I contour upon ligand binding to proteins. The approach is shown in Fig. 2.8a–d.

In Fig. 2.8a, b, the *simulated* IRRAS contours of SP-D in the presence and absence of bound maltose are compared with the *experimental* IRRAS contours in the presence and absence of a physiological LPS ligand (R7). The similarity between the simulations and the experiments in each of these physically different situations provides some validity for assuming that the comparison of amide I changes induced upon binding the soluble (maltose) and the native ligand (R7) provides a valid means to track protein conformational alterations upon the binding

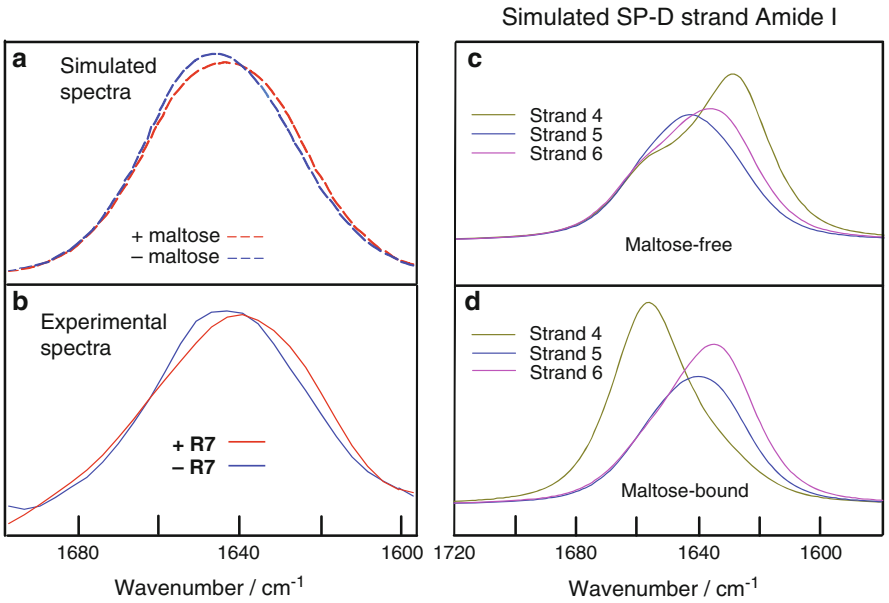

Fig. 2.8 Extracting conformational information from small regions of a large peptide structure. (**a**) Simulated amide I spectra for two different crystal forms of the SP-D (neck + carbohydrate recognition domain), one with bound Ca^{2+} and maltose (*dashed red line*, simulated from Protein Data Bank structure 1PWB) and one with bound Ca^{2+} only, no maltose (*dashed blue line*, simulated from Protein Data Bank structure 3DBZ). (**b**) Experimental IRRAS spectra from SP-D (neck + CRD) color-coded as *solid lines*; these spectra were acquired in the presence (*red solid line*) and absence (*blue solid line*) of R7, an Rd form of LPS from *Salmonella*. Small shifts in the experimental amide I contour upon R7 binding are almost quantitatively reproduced in the simulated spectra. (**c**) Simulated amide I spectra of three strands close to the Ca^{2+} binding site in the absence of maltose are depicted. (**d**) Simulated amide I spectra of three strand regions close to the Ca^{2+} binding site in the presence of maltose are depicted.

of native ligands. It is emphasized that crystals could not be obtained from SP-D/R7 preparations, so that no other means to acquire information about protein conformational changes on physiological ligand binding are available.

The major change occurs specifically in the spectrum derived from strand 4 (green line in Fig. 2.8c, d), whose location in the structure is shown in Fig. 2.9. The other two strands near the Ca^{2+} binding sites are altered much less by maltose interaction. Thus, our simulation methods can begin to extract some three-dimensional information about the location of structural changes upon ligation.

Figure 2.8c, d depicts the results of an initial attempt to interpret the very small spectral shifts seen in Fig. 2.8a, b. The origin of the spectral changes was traced specifically to a tetrapeptide fragment (strand 4 in the SP-D structure). This strand is a tetrapeptide: A_{290} A_{291} F_{292} L_{293} with three peptide bonds. The simulated amide I contour of strands 4, 5, and 6 in the presence and absence of maltose is shown in Fig. 2.8c, d, respectively. The location of the strand 4 relative to the sugar binding

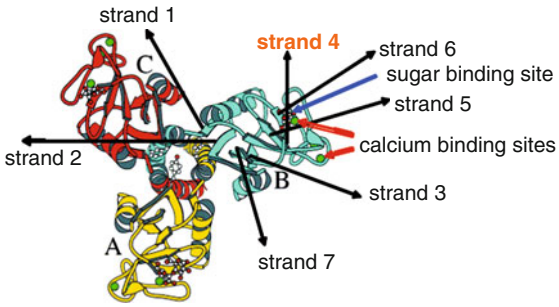

Fig. 2.9 The structure of the trimer containing the neck + carbohydrate recognition domains of SP-D is shown. Strand 4 is the tetrapeptide: A_{290} A_{291} F_{292} L_{293} whose structure change upon maltose binding is suggested in the simulations. See text for details. Professor Barbara Seaton of Boston University is thanked for her generosity in providing and permitting us to use the Figure

site is depicted in Fig. 2.9. The simulated contour of strand 4 changed dramatically on maltose binding, while the other two sheets near the calcium binding sites are altered much less by maltose interaction. Thus, our simulations can begin to extract some three-dimensional information about the location of structural changes upon ligand binding.

We are well aware that many assumptions have to be correct for the simulation of these structural changes in small peptide fragments of a large protein structure to be meaningful. We are moderately hopeful that with improved algorithms, the approach outlined above may provide a useful means to determine subtle structural changes upon ligand binding.

2.3 Conclusions and Future Prospects

We believe that the current approach presents a useful balance between exact theoretical computation of the amide I contour, which is not feasible and in any case would be quite difficult to generalize to a wide variety of proteins, and the empirical methods currently in vogue involving deconvolution/band fitting, which utilize unjustified and, in many cases, wrong assignments of particular spectral features. A general conclusion from the current study is that the widely used correlations between the amide I contour and the secondary structure content of a globular protein are at best approximate. Interactions between oscillators are generally not uniform within a given secondary structure and they extend beyond that structure. For example, it is clear from our "data mining" of the horse heart myoglobin contour that the IR amide I contour of short α-helical segments bears little resemblance to the contours observed for long homopolypeptide sequences. In addition, the amide I contours of extended β-sheets as calculated both from the quantum mechanical approach of the Keiderling group and from our methods again reveal the inadequacy of simply transferring experimental frequency assignments

from β-sheet homopolypeptides to the much narrower segments found in globular proteins.

On the positive side, the sensitivity of the spectral signatures of secondary structures to tertiary structure, as is indicated by the substructure band broadening shown for ribonuclease in Fig. 2.5a, may be used for investigating tertiary structure changes in localized regions of a protein molecule in those cases where a new structure can be reasonably proposed. An example of this is shown in Fig. 2.8c, d where the beta-strand in SP-D most actively involved in the binding of maltose is apparent as was discussed above.

Future progress in simulation may be advanced by experiments in which the magnitude of the off-diagonal terms in the **F** matrix can be established and compared with those determined in the model. The obvious candidate for these experiments is nonlinear 2D-IR, which has become fairly widely available [34, 35] over the past decade. The technique, which bears similarity to 2D-NMR spectroscopy (with, of course, completely different technologies in the two cases) but at vibrational timescales, directly probes the magnitude of interactions between particular vibrations in a protein and can track protein structural changes.

Acknowledgements We thank Prof. Tim Keiderling for productive discussions about the effects of extended β-sheet structures on the amide I contour. We thank Dr. Carol Flach for a critical reading of the MS. The writing of this chapter was supported in part from the US Public Health Service NIH/NIAID program project award AI083222 Program Project "Collectins and Innate Defense against Inhaled Pathogens" (9/1/09-8/31/11), PI, Barbara Seaton, Boston University; sub award to RM. We thank Prof. B. Seaton for permission to use Fig. 2.9.

References

1. T. Miyazawa, T. Shimanouchi, T. Mizushima, J. Chem. Phys. **24**, 408 (1956)
2. T. Miyazawa, T. Shimanouchi, T. Mizushima, J. Chem. Phys. **29**, 611 (1958)
3. T. Miyazawa, J. Mol. Spectros. **4**, 168 (1960)
4. T. Miyazawa, Bull. Chem. Soc. Jpn. **34**, 691 (1961)
5. A. Elliot, E.J Ambrose, Nature **165**, 921 (1950)
6. A. Elliot, Proc. R. Soc. Lond. A **226**, 408 (1954)
7. E.J. Ambrose, A. Elliot, Proc. R. Soc. Lond. A **208**, 75 (1951)
8. T. Miyazawa, J. Chem. Phys. **32**, 1647 (1960)
9. S. Krimm, Y. Abe, Proc. Natl. Acad. Sci. U.S.A. **69**, 2788 (1972)
10. E.B. Wilson, J.C. Decius, P.C. Cross, *Molecular Vibrations* (McGraw-Hill Book Co., New York, 1955)
11. J. Kubelka, T.A. Keiderling, J. Am. Chem. Soc. **123**, 6142 (2001)
12. J. Kubelka, T.A. Keiderling, J. Am. Chem. Soc. **123**, 12048 (2001)
13. B.L. Crawford Jr., W.H. Fletcher, J. Chem. Phys. **19**, 141 (1951)
14. H. Torii, M. Tasumi, J. Chem. Phys. **96**, 3379 (1992)
15. R.G. Snyder, J.H. Schachtschneider, Spectrochim. Acta **19**, 85 (1963)
16. J.W. Brauner, C. Dugan, R. Mendelsohn, J. Am. Chem. Soc. **122**, 677 (2000)
17. J.W. Brauner, C.R. Flach, R. Mendelsohn, J. Am. Chem. Soc. **127**, 100 (2005)
18. M.A. Bryan, J.W. Brauner, G. Anderle, C.R. Flach, B. Brodsky, R. Mendelsohn, J. Am. Chem. Soc. **129**, 7877 (2007)

19. K.R. Symon, *Mechanics* (Addison-Wesley, Reading, 1960)
20. H. Torii, M. Tasumi, in *Infrared Spectroscopy of Biomolecules*, ed. by H.H. Mantsch, D. Chapman (Wiley-Liss, Inc., New York, 1996), p. 1
21. L. Eyges, *The Classical Electromagnetic Field* (Dover Publications, New York, 1972)
22. G.A. Jeffrey, *An Introduction to Hydrogen Bonding* (Oxford University Press, New York, 1997)
23. R.A. Gangani, D. Silva, W. Barber-Armstrong, S.M. Decatur, J. Am. Chem. Soc. **125**, 13674 (2003)
24. R.D.B. Fraser, T.P. MacRae, *Conformation of Fibrous Proteins* (Academic Press, New York, 1973)
25. K.J. Halverson, I. Sucholeiki, T.T. Ashburn, P.T. Lansbury Jr., J. Am. Chem. Soc. **113**, 6701 (1991)
26. C. Paul, P.H. Axelsen, J. Am. Chem. Soc. **127**, 5754 (2005)
27. H. Torii, J. Phys. Chem. B **112**, 8737 (2008)
28. O.N. Antzutkin, J.J. Balbach, R.D. Leapman, N.W. Rizzo, J. Reed, R. Tycko, Proc. Natl. Acad. Sci. U.S.A. **97**, 13045 (2000)
29. Y.A. Lazarev, B.A. Grishkovsky, T.B. Khromova, Biopolymers **24**, 1449 (1985)
30. V.L. Kuzmin, A.V. Michailov, Opt. Spectrosc. (USSR) **51**, 383 (1981)
31. V.L. Kuzmin, V.P. Romanov, A.V. Michailov, Opt. Spectrosc. (USSR) **73**, 1 (1992)
32. A. Gericke, A.V. Michailov, H. Hühnerfuss, Vibr. Spectrosc. **4**, 335 (1993)
33. L. Wang, J.W. Brauner, G. Mao, E. Crouch, B. Seaton, J. Head, K. Smith, C.R. Flach, R. Mendelsohn, Biochemistry **47**, 8103 (2008)
34. C. Fang, A. Senes, L. Cristian, W.F. DeGrado, R.M. Hochstrasser, Proc. Natl. Acad. Sci. U.S.A. **103**, 16740 (2006)
35. L.P. DeFlores, Z. Ganim, R.A. Nicodemus, A. Tokmakoff, J. Am. Chem. Soc. **131**, 3385 (2009)

Chapter 3
Millisecond-to-Minute Protein Folding/Misfolding Events Monitored by FTIR Spectroscopy

Heinz Fabian and Dieter Naumann

Abstract The adaptation of conventional mixing or temperature-jump technologies to the specific requirements of time-resolved FTIR spectroscopy enables to explore protein folding/misfolding events on the millisecond to minute timescale by monitoring spectral changes in several spectral windows simultaneously. We present a description of experimental setups and representative results on different proteins, which illustrate the advantages and limitations of the FTIR approaches.

3.1 General Considerations

IR spectroscopy is known since 1950 to be potentially useful in providing information on structural features of peptides and proteins [1]. Later, these experimental observations were refined by making detailed vibrational analyses of the structure-sensitive protein backbone absorption bands in order to establish a correlation between the frequencies of these bands and the type of secondary structure. Nine such IR bands do exist, which are termed amide A, amide B, and amides I–VII, in the order of decreasing frequency [2–4], and three of them are particularly useful as structural probes (Table 3.1).

The best characterized in this respect is the amide I band, which is dominated by the C = O stretching vibration of the amide groups coupled to in-plane bending of the N–H and stretching of the C–N bonds and occurs in the region 1,600–1,700 cm^{-1}. Typically, the amide I band of proteins consists of a series of overlapping component bands which result from the various secondary structures present in such molecules [5–7]. As a consequence, the individual component bands, representing different structural elements, such as α-helices, β-sheets, turns, and

D. Naumann (✉)
Robert Koch-Institute, Biomedical Spectroscopy, Nordufer 20, 13353 Berlin, Germany
e-mail: NaumannD@rki.de

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*,
Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_3,

Table 3.1 Characteristic infrared bands of the peptide linkage in the range 1,000–4,000 cm^{-1}

Nomenclature (amide)	Approximate frequency (cm^{-1})	Vibrational modes
A	~3,300	N–H stretching
I	1,610–1,695	C = O stretching
II	1,480–1,575	N–H bending and C–N stretching
III	1,220–1,320	C–N stretching and N–H bending

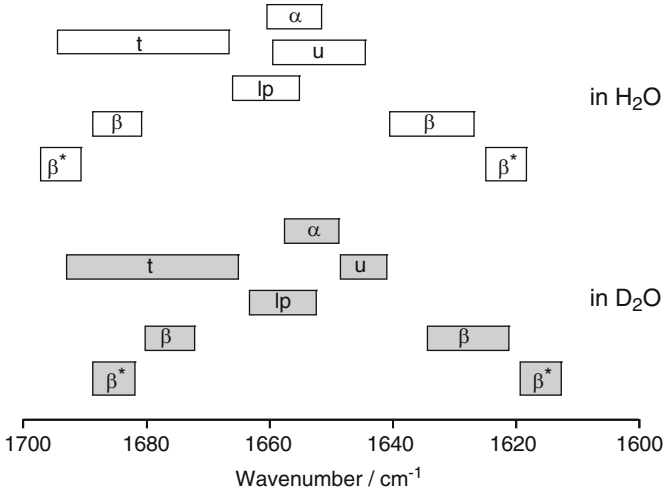

Fig. 3.1 Guidelines for protein secondary structures based on their amide I/I′ frequencies. Amide I (in H_2O): *blank frames*; amide I′ (in D_2O): *shaded frames*; α : α-helices. β : β-sheet structures. For proteins often more than one band component is observed. This reflects differences in hydrogen bonding (the stronger and shorter the H-bond, the lower the frequency) as well as differences in transition dipole coupling in different β-strands. β* : β-strands in ordered or amorphous aggregates, often characterized by a significant increase in splitting between the low- and high-frequency β-components in comparison to that observed for β-sheets in natively folded proteins; u: unordered parts of the polypeptide backbone; lp: loops; and t: turns.

irregular structures, are often not resolved in the broad amide I band contours of the experimentally obtained spectra. There are empirically determined guidelines for protein secondary structure based on their characteristic amide I frequencies (Fig. 3.1). They illustrate that α-helical and irregular structures show bands very close together (1,645–1,660 cm^{-1}), which may cause complications in the analysis. On the other hand, amide groups in β-sheet structures give rise to major diagnostic bands between 1,615 and 1,640 cm^{-1} and weaker bands near 1,670–1,695 cm^{-1}. This makes FTIR spectroscopy, in contrast to CD spectroscopy, very useful to indicate the presence of, and to monitor changes in, β-sheet structures. The secondary structure dependence of the amide I band arises from couplings among neighboring amide groups and intramolecular hydrogen bonding [see Chaps. 2 and 9

for more details]. Furthermore, hydrogen bonding involving solvent molecules can also have an impact on the amide I frequency [8].

The structure–spectrum correlation for the amide II and amide III bands is less well understood, but the former band may provide useful information on the solvent accessibility of the protein backbone from hydrogen/deuterium (H/D) exchange experiments, as this holds true even more for the amide A band (see below).

3.2 FTIR Spectroscopy, Experimental Aspects

3.2.1 Proteins in Aqueous Solutions

IR experiments with proteins in water are complicated by the fact that the H–O–H bending vibration of H_2O absorbs very strongly near 1,640 cm^{-1}. As a consequence, very short pathlength cells of 6–8 μm are needed for transmission measurements in the amide I region to prevent total IR absorption in the spectral regions of the water. Such short pathlengths limit the intensities of the IR bands and the signal-to-noise ratio at a given sample concentration. Consequently, relatively high sample concentrations, in the order of >10 mg/ml, are required for the measurement. Figure 3.2a (solid line) shows the IR spectrum of the protein β_2-microglobulin (β_2m) in H_2O buffer. For these transmission measurements, a drop of 3 μl of the proten solution was placed between a pair of calcium fluoride (CaF_2) windows separated by a pathlength of 8 μm. The solvent spectrum (Fig. 3.2a, dashed line) was measured in a matched second cell of slightly reduced pathlength, which takes into account the slightly lower water concentration in the protein sample measured. To obtain the spectrum of the protein (Fig. 3.2b), digital subtraction of solvent/buffer absorptions from the spectrum of the sample is required. For appropriate subtraction, the spectrum of the solvent/buffer should be recorded under identical physicochemical parameters (such as temperature, ionic strength, and pH) because variations will cause changes in the spectral features of the H_2O bands, thereby preventing an ideal subtraction of the buffer contribution. For example, the temperature of the sample in aqueous solution and that of the reference buffer should match within 0.1°C in order to avoid artefacts caused by temperature differences. The subtraction of water from a protein spectrum requires a reference water band that does not overlap with that of the sample. Often, the weak combination band of water around 2,128 cm^{-1} (which is more than six times less intense than the water band of interest near 1,640 cm^{-1}) is used to ensure proper subtraction of water, which then results ideally in a straight baseline between 1,900 and 2,400 cm^{-1}. If water subtraction results in derivative-type artifacts in this region, this implies that the spectroscopic characteristics of the water in the sample water are not identical with those of the reference water. The latter artefact may often be observed at very high protein concentrations, since the interaction of a protein with water

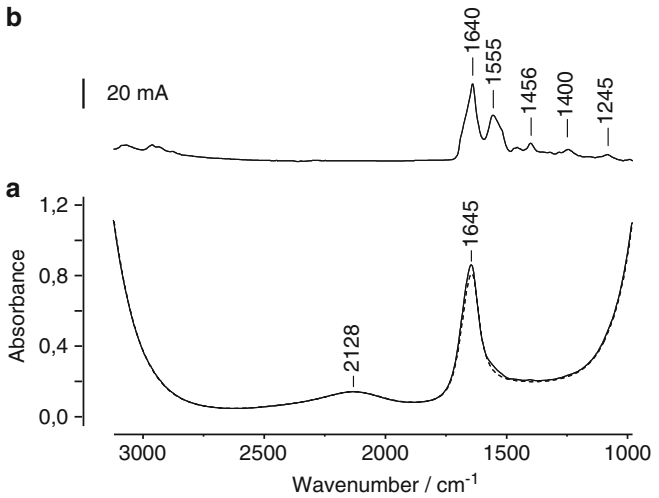

Fig. 3.2 (**a**) IR spectrum of the protein β_2-microglobulin (β_2m) in H_2O buffer (*solid line*) at a protein concentration of 28 mg/ml placed between a pair of CaF_2 windows separated by a pathlength of 8 μm, together with the buffer spectrum (*dashed line*) measured in a matched second cell of slightly reduced pathlength. (**b**) IR spectrum of β_2m after subtraction of the buffer spectrum [ordinate scale expanded by a factor of 5 compared to the scale in (**a**)]

causes the creation of hydration shells around the protein and therefore modifies the water vibrations and hence the shape of the water bands in the IR spectrum. Consequently, the reference spectrum of the H_2O buffer cannot completely match the spectral conditions of the water in the protein solution (containing bulk water plus protein-bound water). The weak water band at 2,128 cm^{-1} may serve as a good approximation to interactively subtract the water features. The final water subtraction, however, should be performed using a different spectral region with much stronger H_2O absorption, such as on the slope of the O–H stretching band in the vicinity of ~3,150 cm^{-1} [9].

3.2.2 Measurements in D_2O

Experimentally it is simpler to obtain protein spectra in deuterium oxide (D_2O) solution than in H_2O solution. The IR bands of D_2O occur at lower wavenumbers than those of H_2O, creating a spectral window of relatively low absorbance between 1,300 and 1,800 cm^{-1}, an ideal region to observe the weak IR bands of the dissolved or solubilized protein. Much longer pathlengths of 40–80 μm may then be used; hence, much lower sample concentrations are required to obtain high-quality spectra. Despite these positive practical aspects, D_2O as a solvent needs to be considered carefully because the exchange of D for H also impacts on certain IR bands of the protein under study.

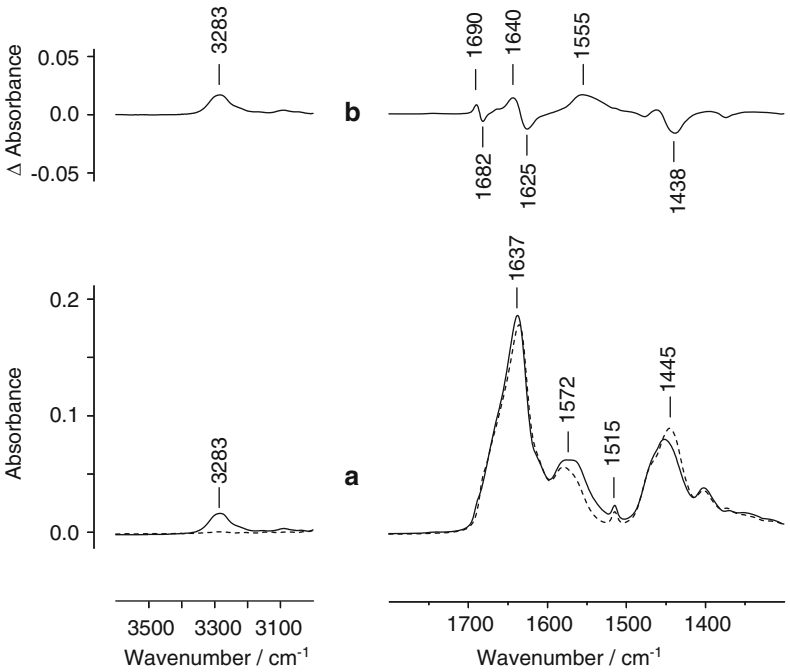

Fig. 3.3 IR spectra of β_2m in D_2O buffer at a protein concentration of 10 mg/ml and placed between a pair of CaF_2 windows separated by a pathlength of 50 μm. (**a**) IR spectrum of fully exchanged β_2m after subtraction of the buffer spectrum (*dashed line*), together with the spectrum of β_2m in a partially exchanged state (*solid line*). Complete deuteration of β_2m was achieved by keeping the protein solution overnight at 37°C before cooling back down to room temperature. (**b**) IR difference spectrum obtained by subtracting the spectrum of fully exchanged β_2m from the spectrum of the partially exchanged protein

Figure 3.3 summarizes spectral effects of partial and complete H/D exchange on the amide bands of a protein spectrum. The dashed line in Fig. 3.3a shows the IR spectrum of β_2m in D_2O buffer after complete H/D exchange, which was achieved by keeping the protein solution overnight at 37°C [10]. The solid line in Fig. 3.3a represents the IR spectrum of the protein β_2m that has been allowed to exchange in D_2O buffer for 1 h at room temperature. The residual intensity in the amide II region at about 1,555 cm^{-1}, together with the presence of the amide A band centered at 3,283 cm^{-1}, indicates that a significant number of the amide protons are not exchanged after 1 h of exposure to D_2O, a situation common in many proteins. The amide A band is the best indicator for residual non-exchanged N–H groups due to the lack of other protein absorptions in the range 3,200–3,300 cm^{-1}. The same assessment cannot easily be made based on the residual intensity in the amide II region, since IR bands arising from amino acid side-chain groups (Table 3.2) overlap with the remaining amide II band, e.g., in β_2m the two bands are at 1,572 and 1,515 cm^{-1}. Nevertheless, the disappearance of the amide II band near 1,555 cm^{-1}

Table 3.2 Approximated frequencies of characteristic amino acid side-chain absorption bands in the range 1,480–1,750 cm^{-1}

Side chain/assignment	Absorbance maximum (cm^{-1})	
	In H_2O	In D_2O
Asn (C = O)(ν)	1,678 (s)	1,646 (s)
Asn (NH_2)(δ)	1,622 (m)	
Asp (COOH/COOD) (ν)	1,716 (m)	1,720 (m)
Asp (COO^-)(ν_{as})	1,574 (s)	1,585 (s)
Arg ($CN_3H_5^+$) (ν_{as})	1,673 (s)	1,605 (s)
Arg ($CN_3H_5^+$) (ν_s)	1,633 (s)	1,586 (s)
Glu (COOH/COOD)(ν)	1,712 (m)	1,706 (s)
Glu (COO^-)(ν_{as})	1,560 (s)	1,570 (s)
Gln (C = O)(ν)	1,670 (s)	1,635 (s)
Gln (NH_2)(δ)	1,610 (m)	
Lys (NH_3)(δ_{as})	1,629 (m)	
Lys (NH_3)(δ_s)	1,526 (m)	
Phe (Ring)	1,494 (w)	1,498 (w)
Tyr (Ring)	1,614 (m)	1,615 (m)
Tyr (Ring)	1,518 (s)	1,515 (s)

s strong, *m* medium, *w* weak

and the appearance of the N–D absorption near 1,438 cm^{-1} (labelled amide II′ by convention) (Fig. 3.3b) is often used to monitor the time course and extent of hydrogen–deuterium exchange [11], and can provide valuable information on the structure and flexibility of proteins [12–18].

Compared to the large effects on the amide II band, the shift of the amide I band upon deuteration of the backbone hydrogens (amide I′) is relatively small (5–10 cm^{-1}). However, individual spectral components of the amide I band of a protein often reveal different exchange kinetics. This greatly assists the assignment of absorption bands arising from different secondary structural classes. Despite the positive aspects, it can also complicate the interpretation of the amide I′ region if a protein cannot be completely exchanged, because partial deuteration leads to band components with frequencies in between those of fully protonated and fully deuterated amide groups.

Figure 3.4 shows the absorbance and second-derivative spectra of six proteins with differing secondary structures in D_2O buffer after H/D exchange and subtraction of the buffer: Two predominantly α-helical proteins (bovine hemoglobin and calmodulin, Fig. 3.4a, b), a protein composed of four α-helices and a three-stranded parallel β-sheet (barstar, Fig. 3.4c), two proteins with a mixture of α-helical and antiparallel β-sheet structure (λ-Cro repressor and ribonuclease T1, Fig. 3.4d, e), and the β-sandwich protein β_2-microglobulin (Fig. 3.4f), where the two β-sheets are formed entirely from antiparallel β-strands.

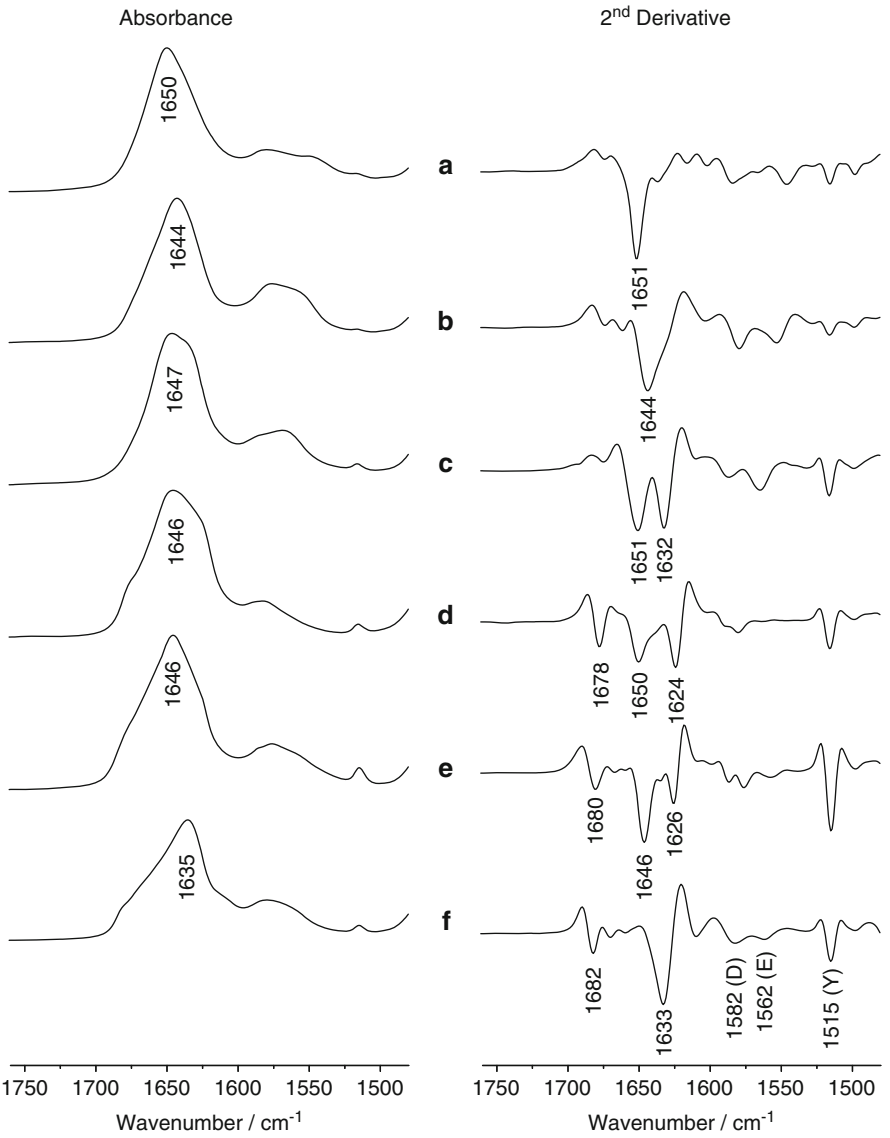

Fig. 3.4 IR absorbance and second-derivative spectra of two predominantly α-helical proteins bovine hemoglobin (**a**) and calmodulin (**b**), the protein barstar with a mixture of α-helical and parallel β-sheet structure (**c**), two α/β proteins λ-Cro repressor (**d**) and ribonuclease T1 (**e**), and the purely β-sheet protein β_2m (**f**). All spectra are shown after complete exchange of the amide protons with deuterons, except for hemoglobin which still contained traces of amide NH groups. The infrared bands that remain between 1,480 and 1,600 cm^{-1} in the spectra after complete H/D exchange are entirely due to amino acid side-chain absorptions (e.g., tyrosine, aspartate, and glutamate in the spectrum of β_2m)

3.2.3 FTIR Spectra of Chemical Denaturants

Obtaining IR spectra of proteins in the presence of the most commonly used denaturating agents, urea and guanidinium chloride (GdmCl), is not simple. First, the denaturant has very strong IR bands of its own. Therefore, IR cells with very short pathlengths of ~6 μm are required to prevent saturation of the IR detector by absorption of the chemical denaturant. Second, the major IR bands of urea or GdmCl (Fig. 3.5) mask the much weaker protein backbone bands of interest and thus may prevent their analysis. Isotopic labeling of the denaturant (e.g., the use of $^{13}C = O$ labeled urea) helps to circumvent this problem by shifting the intense urea band from 1,607 to 1,568 cm^{-1} (compare the dashed and solid lines in Fig. 3.5d), thus creating a window in the IR spectrum for analysis of the protein amide I′ bands [19].

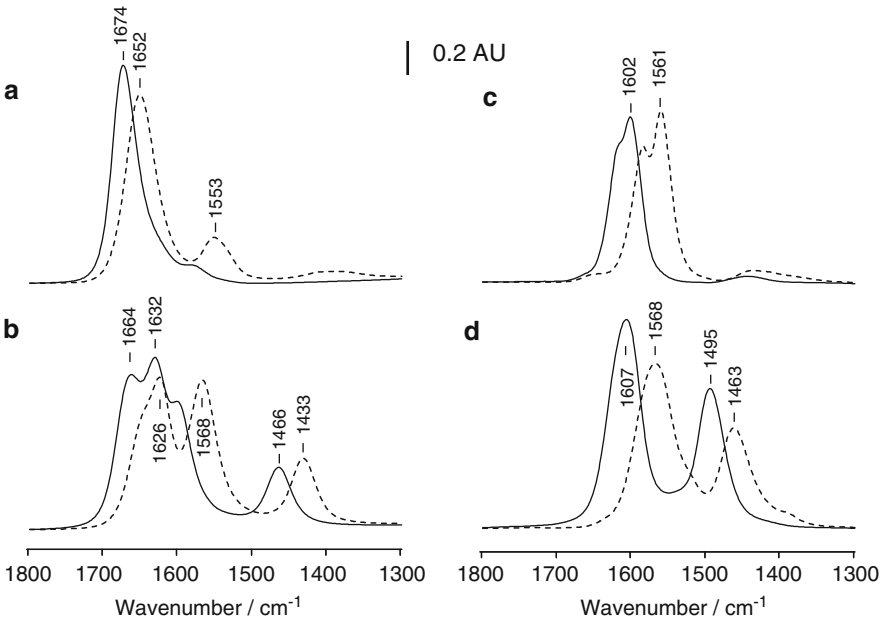

Fig. 3.5 IR spectra of chemical denaturants in H_2O (*left side*) and in D_2O (*right side*). The samples were placed between a pair of CaF2 windows separated by a pathlength of ~10 μm, the concentration of the denaturants was in the range of 2.3–3 M. All spectra are shown after subtraction of the corresponding solvent. (**a**, **c**) GdmHCl (*solid line*) and ^{13}C-GdmHCl (*dashed line*). (**b**, **d**) Urea (*solid line*) and ^{13}C-urea (*dashed line*)

3.3 Kinetic FTIR Experiments Applying Rapid Mixing and Temperature-Jump Approaches

3.3.1 Rapid-Scan FTIR Spectroscopy: Advantages and Limitations

Already the simplest FTIR spectrometers only require less than a second for collecting a single interferogram. Thus, even without sophisticated equipment, it is feasible to probe protein unfolding/folding events in the second to minute time domain by a series of FTIR spectra. With the rapid-scan option available, the observation time can be reduced to milliseconds. Moreover, these options allow the interferogram to be rapidly stored together with the parameters needed for the transformation in order to allow transformation afterward, thus reducing the intervals for spectral recording and allowing more interferograms to be coadded for a better signal-to-noise ratio. The main limiting factors are the scan rate of the interferometer (limited by the mechanical movement) and the throughput of data acquisition (limited by the speed of the analogue-to-digital converter). In practice, about 100 scans/s at 4 or 8 cm^{-1} spectral resolution can be reached.

Figure 3.6 shows the scheme of a typical time-resolved experiment performed on a Bruker IFS-66 rapid-scan FTIR spectrometer. Due to the parameters optimized for this particular experiment (scanner velocity 280 kHz, nominal resolution 4 cm^{-1}, and single-sided/forward–backward acquisition mode), one scan takes 37 ms. The experimental dead time, in this case, was estimated to be ~59 ms [19, 20], which allows to monitor events occurring over a time window from ~60 ms to minutes. Consequently, protein unfolding/folding events occurring on very fast timescales cannot be probed by this experimental setup. Nevertheless, the FTIR approach is of particular interest, because a complete spectrum is available for each time point of the measurement. In this way, several spectral windows are accessible simultaneously for the observation of the unfolding or the formation of different

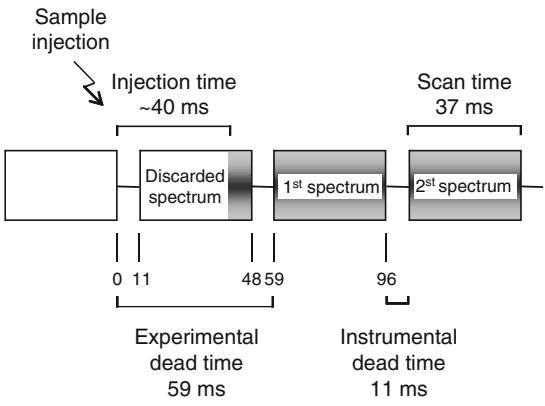

Fig. 3.6 Scheme of data acquisition of a typical time-resolved rapid-scan FTIR experiment. The *arrow* symbolizes the electric trigger signal from the spectrometer that initiates the pneumatic sample injection into the corresponding experimental setup

secondary structure elements and also events that can be attributed to changes in tertiary structure. Thus, perfect time correlation is ensured between events that can otherwise only be observed with different techniques, which often require different experimental conditions.

3.3.2 *Design and Operation of a Stopped-Flow Apparatus for Measurements in Heavy Water*

A schematic diagram of our stopped-flow system used to investigate events in D_2O solution is shown in Fig. 3.7. Pneumatically driven gas-tight Hamilton syringes deliver the two reactants through flexible tubes with small internal diameter into the rapid mixing attachment. The pneumatic drive is triggered by an electric signal from the FTIR spectrometer, thus permitting precise time correlation between data acquisition and mixing time [19, 20]. The drive system is vibrationally isolated from the spectrometer and the mixing system. The mixing attachment consists of a six-jet unit, which is mounted into the front plate of the IR flow-through cell. The cuvette is very robust in that a 9 mm thick front plate is used together with six bolts to clamp the assembly together. Teflon gaskets are used to seal the CaF_2 windows to the front and back plates and a 50 μm Teflon spacer separates the windows. The mixing attachment and the IR cuvette, fabricated in a workshop at the University of Birmingham, UK, have been provided to us by Dr. Christopher W. Wharton from Birmingham, who has designed the mixing system [21]. The output of the cuvette is connected to a stop syringe to arrest the flow after the drive pulse ceases. To reduce the dead volume of the protein solution, an additional three-way valve can be integrated into the protein flow line, which allows one to push the protein solution directly in front of the mixing attachment. In this way, e.g., 50 μl of the protein solution can be preinjected into the flow line before the pneumatic

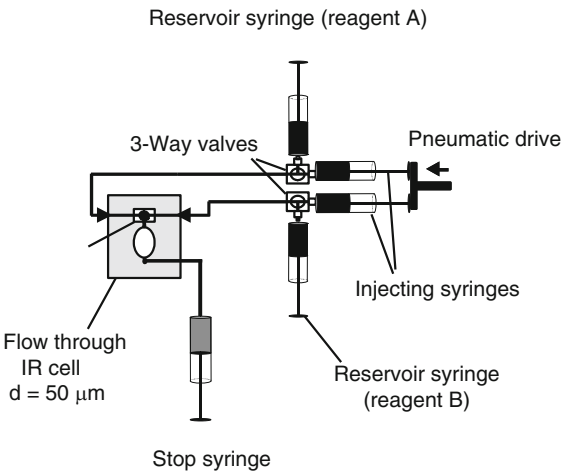

Fig. 3.7 Scheme of a stopped-flow and 50 μm IR cuvette system developed to monitor events in D_2O solution by time-resolved FTIR spectroscopy

drive additionally injects 30 μl of the protein solution and a larger volume of buffer solution into the mixer.

Other apparatus for time-resolved studies of conformational events in protein, either based on a commercial SFM-300 stopped-flow module (Bio-logic, Claix, France) equipped with an IR observation cell [22] or a home-made flow-through IR cell with a T-shaped mixer built into part of the Teflon spacer sandwiched between two CaF_2 plates [23], were described.

3.3.3 *A Stopped-Flow Apparatus for Measurements of H_2O-Protein Solutions*

For stopped-flow experiments using H_2O/protein solutions very short optical pathlengths are needed (~5–8 μm), which necessitates the use of high-pressure pumping systems to push the solutions through the cuvette. High-pressure induces optical pathlength changes which – though only in the order of a few nanometers – make it impossible to monitor reaction-induced changes quantitatively. To make protein folding studies feasible in aqueous solution by stopped-flow FTIR, several modifications of the conventional stopped-flow system described above are required [24]. One step is the replacement of the classical demountable optical flow-through cell equipped with teflon spacers by a novel flow-through cell with a pathlength of less than 8 μm (AquaSpec, micro-biolytics GmbH, Germany). In this AquaSpec cell (see Fig. 3.8, insert top left), the two CaF_2 windows are kept together by a special

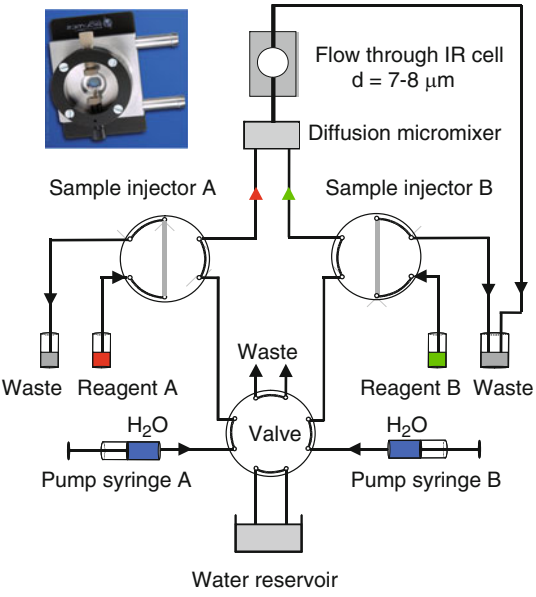

Fig. 3.8 Schematic diagram of the stopped-flow system developed to monitor protein folding events in H_2O buffer solutions by infrared spectroscopy (after Masuch et al. [24]). *Upper left panel* shows an AquaSpec cell from micro-biolytics GmbH, Germany

glue, which ensures a very fast relaxation to a constant pathlength after stopping the flow. Additionally, the new apparatus contains a miniaturized diffusion micromixer enabling a rapid and effective laminar mixing of the two reactants [24].

Figure 3.8 shows the scheme of the experimental setup including also the picture of an AquaSpec cell (see inset). The system consists of a two-channel high-pressure syringe pump (Harvard Apparatus, USA), two sample loops, HPLC tubing and fittings throughout, the diffusion micromixer, the AquaSpec flow-through cell, and three computer-controlled analytical HPLC valves (Valco Instruments, USA). The sample volume needed for a single experiment is of the order of 10–15 μl. At the beginning of each experiment, a constant water flow of pure water (e.g., 3 ml/min) is generated by the pump syringes A and B. The two sample loops containing the two reagents are switched by the computer-controlled sample injection valves A and B into the flow stream, and the two solutions are synchronously transported through the diffusion micromixer toward the sealed IR cell. After a defined time interval, depending upon the flow rate of the system, the IR cell is filled and the constant flow is stopped by the stopped-flow valve. At the same time, a trigger signal is sent to the FTIR spectrometer, which initiates the collection of the interferograms. The dead time of the experiment is defined by the time required for mixing the samples and for collecting the first interferogram, which is of the order of 150–200 ms with an instrumental scan rate of 180 kHz. This is considerably longer than the relaxation time of the flow-through cell after stopping the water flow, which is in the order of 70 ms.

Figure 3.9 shows original absorbance spectra from a series of independent shots mixing a solution of the Syrian hamster prion protein with buffer solution to yield an end concentration of approximately 4 mg/ml, which is rather low for IR measurements in H_2O. The spectra are nearly perfectly overlaid proving an excellent reproducibility of the system, which in turn allows the consistent use of difference spectroscopy for monitoring time-dependent processes after mixing the reagents. An application example of this methodology is given in Sect. 3.5.3.

3.3.4 T-Jump Experiments in Heavy Water

For experiments applying T-jumps, the solutions are injected into a thermostatted flow-through CaF_2 IR cell (home-made cell with Teflon spacer or an adapted micro-biolytics AquaSpec cell, each with an optical pathlength of ~50 μm). The flow-through cell is connected by a small diameter flexible Tefzel tube with a gas-tight Hamilton syringe localized outside the sample chamber (Fig. 3.10). The injecting syringe is equipped with a water jacket which allows adjustment of the temperature of the sample to be injected. The temperatures of the IR cell and the injecting syringe are controlled by different temperature baths. For unfolding experiments, the temperature of the injecting syringe is kept at low temperatures, while the IR cell is kept at the desired unfolding temperature. For technical reasons (leaking), the flow-through IR cell is typically not kept above 65°C. For refolding

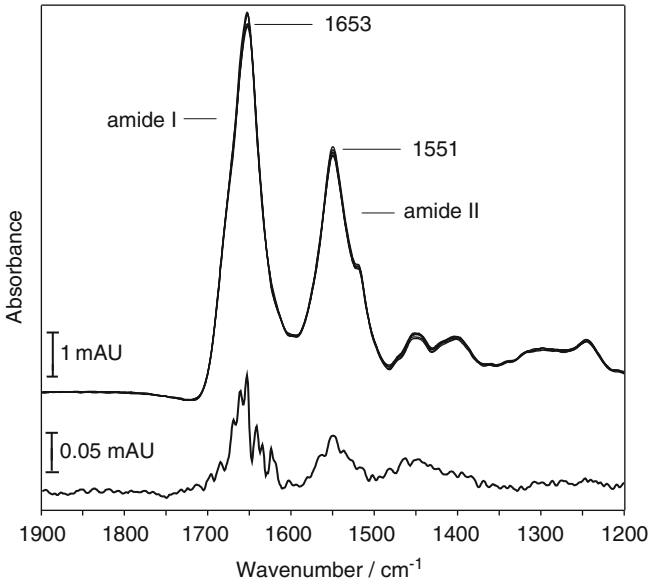

Fig. 3.9 *Upper traces*: Six overlaid independently measured absorbance spectra of the prion protein $SHaPrP^{90-232}$ after H_2O and buffer subtraction using the stopped-flow device shown in Fig. 3.8. The initial PrP concentration was 8 mg/ml, all PrP solutions were diluted 1:1 in the stopped-flow device yielding a final concentration of 4 mg/ml (~0.24 mM). Buffered solutions were used to ensure a proper pH value of 7. Steady-state spectra of either buffer or PrP + buffer were measured independently 5 times with 256 scans in each case and finally averaged. Buffer absorbance spectra were subtracted from PrP + buffer spectra. *Lower trace*: Standard deviation curve of the six spectra to their average

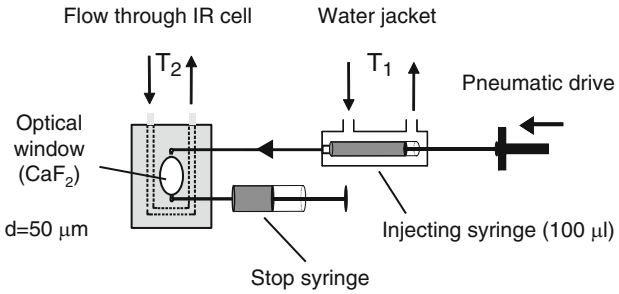

Fig. 3.10 Scheme of T-jump setup developed to monitor protein unfolding/folding events in D_2O solution by time-resolved FTIR spectroscopy

experiments, the solution is kept in the injecting syringe at high temperatures (up to 90°C feasible) for a certain time interval to ensure complete H/D exchange of all amide protons. Then, the solution is injected into the flow-through IR cell kept at the desired refolding temperature. The injection of the solution is accomplished using a pneumatic drive, which is triggered by an electric signal supplied from

the spectrometer, such as described in Fig. 3.10 for the stopped-flow apparatus. T-jump experiments with D_2O-containing buffer are performed under the same experimental conditions as the protein measurements. Each buffer spectrum is subtracted from the protein spectrum taken at the corresponding time point after initiation of the T-jump.

3.4 Examples of Applying T-Jumps onto a Protein Solution

3.4.1 Refolding of Wild-Type Ribonuclease T1 and Some of Its Mutants

Ribonuclease T_1 (RNase T_1) is a small globular, single domain protein with 104 amino acids. Its three-dimensional architecture, well characterized by several X-ray crystal structures, consists of an α-helix, extended antiparallel β-sheet with three long and two short β-strands, a short two-stranded antiparallel β-sheet, four wide loops, several turns, and two disulfide bonds connecting the C- and N-terminal regions [25]. Because of its small size, thermal stability, and the availability of efficient expression systems, RNase T_1 has often been used as a model system for protein folding studies. Moreover, its thermal unfolding is fully reversible even at the relatively high protein concentration of ∼10 mg/ml typically required for IR measurements. Native RNase T_1 has four proline residues with P55 and P39 in *cis* and P60 and P73 in *trans* conformation [25, 26].

It has been shown by several studies that certain variants where one or two amino acid residues had been exchanged (e.g., wild-type RNase T_1 and W59Y variant RNase T_1) altered the refolding kinetics of the protein coupled with the *trans–cis* isomerization reaction of P39 [27]. In another study, the role of the proline bond between S54-P55 that adopts a *cis* conformation in the native state on the refolding process was investigated. This *cis* bond was substituted by a *trans* G-54-N-55 bond by exchanging proline in position 55 by an asparagine and serine in position 54 by a glycine yielding the variant protein S54G/P55N. This RNase T_1 variant was known to refold after initial dilution of the denaturant by simplified kinetics. After a very rapid collapse to intermediates exhibiting already a native-like structure with the proline in position 39 still in the *trans* conformation. The rate-limiting *trans–cis* isomerization reaction then synchronizes the following refolding processes in all regions of the protein [28].

3.4.1.1 Structure Formation in RNase T1 Wild-Type and RNase T1 Variant S54G/P55N Under Destabilizing and Stabilizing Conditions

Infrared spectroscopy turned out to be ideal for studying the refolding processes of the wild type and certain variants of RNase T_1 after a rapid temperature jump using

the techniques and devices described above [29, 30]. In the following, an example of a study on folding events in the millisecond to minutes time range is given.

Before initiating the T-jump, the protein D_2O/buffer solutions were kept for 10° min at 65°C in the injecting syringes (see Fig. 3.10). This temperature is well above the thermal denaturation temperature ($T_m \sim 54$°C) of RNase T1 [31], which ensured their unfolding and avoided complications due to *trans–cis* isomerization. The pneumatic drive used, triggering, spectrometer settings, data acquisition, and evaluation were as described above and in [29, 30]. To obtain difference spectra for discrete incremental time intervals, single-beam spectra were ratioed as described in [29]. The kinetic data were obtained by selecting spectral bands specific for certain structural elements.

Figure 3.11 displays the spectral changes observed for the RNase T1 variant S54G/P55N in discrete time intervals after inducing the refolding process by a temperature jump. In these difference spectra positive bands are due to arising structures, whereas negative intensities are due to vanishing structures.

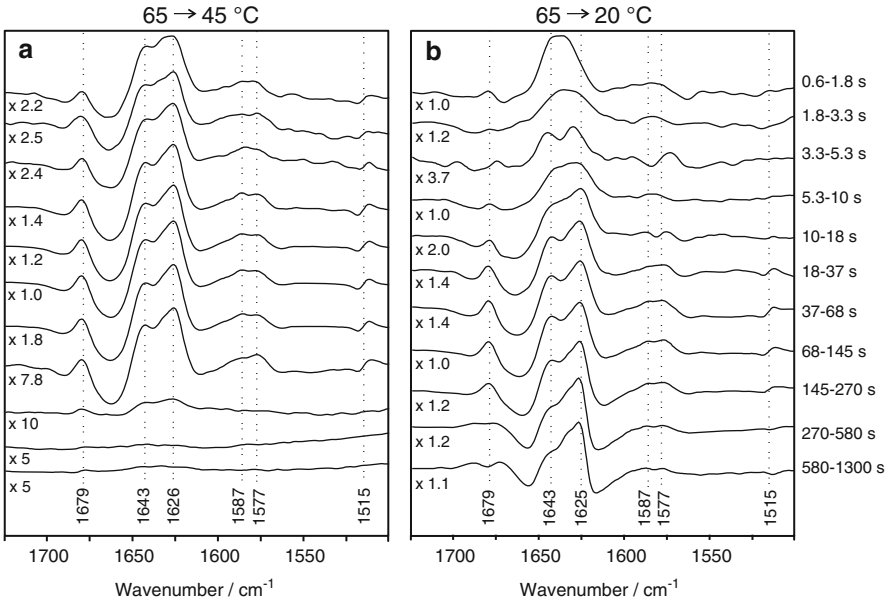

Fig. 3.11 FTIR difference spectra observed during the refolding of the RNase T1 variant S54G/P55N in discrete time intervals after a T-jump (**a**) *under destabilizing conditions* (T-jump from 65°C to 45°C) and (**b**) *under stabilizing conditions* (T-jump from 65°C to 20°C). The difference spectra were calculated according to $\log(S^{t_1}/S^{t_2})$, where S^{t_1} is the single-beam spectrum at a selected time t_1 and S^{t_2} the spectrum at a later refolding time t_2. To achieve a better signal to noise ratio, time corresponding absorbance spectra from three independent experiments were averaged over various time intervals at the expense of time resolution. The corresponding time intervals are given at the *right margin* for the averaged spectra. All difference spectra were expanded according to the cited scaling factors at the *left margin* in order to obtain comparable intensities

To discuss the ongoing structural changes in the RNase T1 variant after a rapid T-jump, some band assignments are necessary: The low-frequency band at 1,626 cm^{-1} and the high-frequency band at 1,679 cm^{-1}, both resulting mainly from transition dipole coupling, have been assigned to antiparallel β-sheets [7,31,32]. The band component at 1,643 cm^{-1} has been tentatively assigned to overlapping bands of irregular structures, turns and loops, or α-helices. Spectral contributions between 1,620 and 1,500 cm^{-1} of fully H/D-exchanged proteins are due to amino acid side-chain vibrations. The bands at 1,587 and 1,577 cm^{-1} arise from asymmetric carboxylate stretching vibrations of the amino acid side-chain groups of aspartate and glutamate, respectively, whereas the band at 1,515 cm^{-1} results from the aromatic ring stretching vibration of tyrosine [31, 32]. These amino acid side-chain marker bands, the tyrosine band in particular, are very useful in monitoring the formation of tertiary structure in proteins by means of FTIR spectroscopy [20, 29].

A comparison of the difference spectroscopic features observed in the various time intervals after a temperature jump from 65°C to 45°C (destabilizing conditions) reveals practically no qualitative differences for the various IR marker bands during the whole time course of refolding (Fig. 3.11a). This indicates a highly cooperative two-state refolding process which involves all structural elements simultaneously and is not complicated by the accumulation of intermediates.

Under stabilizing conditions (T-jump from 65°C to 20°C, Fig. 3.11b) the temperature-induced refolding of the RNaseT1 variant S54G/P55N turned out to be more complex. Clearly, the difference spectra observed within the first 18 s of the refolding process are different from each other, indicating that the renaturation process under stabilizing conditions is not as cooperative as at 45°C. In the time range between 18 and 270 s, the folding events are more cooperative and resemble closely the corresponding difference spectra shown in Fig. 3.11a. While at destabilizing conditions (T-jump from 65°C to 45°C) the refolding of the protein is already practically finished within 270 s, the difference spectra obtained at a refolding temperature of 20°C still show significant amplitudes particularly at 1,625 and 1,643 cm^{-1}. These bands generally indicate changes in β-sheet and irregular structures of proteins, respectively. Interestingly, after approximately 270 s only the low-frequency β-sheet, but not the corresponding high-frequency component near 1,679 cm^{-1}, exhibits positive difference signal intensity. It was suggested that the formation of ordered β-sheets is already completed and that only minor rearrangements in the outer regions of the β-strands and/or terminal β-strand segments are taking place in this late folding period [30].

The fitted kinetic refolding plots evaluated using two diagnostic IR marker bands of the S54G/P55N variant and – for comparison – the RNase T1 wild-type enzyme (corresponding difference spectra not shown here, see [30]) are displayed in Fig. 3.12. From these kinetic plots, it is evident that the protein variant refolds significantly faster to its native conformation than the wild type under both destabilizing (Fig. 3.12a) and stabilizing (Fig. 3.12b) conditions.

This statement is supported by the more quantitative data given in Table 3.3, which compare the kinetic time constants and amplitudes of the S54G/P55N variant and the wild-type RNase T1 using time-dependent changes of the same

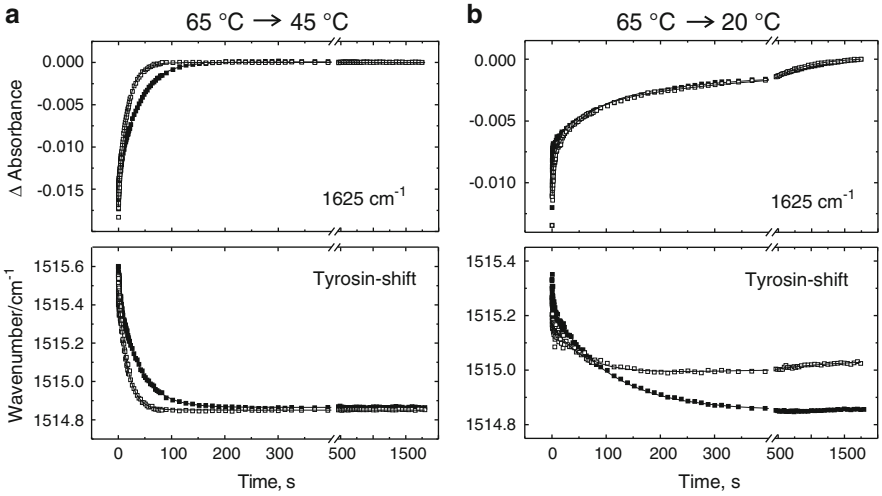

Fig. 3.12 Kinetic plots and their best fits of two infrared bands monitoring secondary or tertiary structure formation during the refolding of wild-type RNase T1 (*filled symbols*) and S54G/P55N RNase T1 (*open symbols*). (**a**) Shows the respective plots obtained after a T-jump from 65°C to 45°C, (**b**) demonstrates the corresponding plots for T-jumps from 65°C to 20°C. *Top*: Intensities of the low-frequency β-sheet band at 1,625 cm^{-1}. *Bottom*: Frequency of the tyrosine band at ~1,515 cm^{-1}. Data were evaluated from difference spectra (except the tyrosine shift) calculated from single-beam spectra according to $\log(S^{\infty}/S^{t_i})$, where S^{∞} is the last recorded single-beam spectrum (at $t = 30$ min) and S^{t_i} the single-beam spectrum at the time points t_i. For each plot the fits accord to a sum of exponential functions [29]. The tyrosine band shift was evaluated from the absorbance spectra

marker bands shown in Fig. 3.12. The data of Table 3.3 indicate that under destabilizing conditions (T-jump from 65°C to 45°C) the refolding process of the two proteins can be resolved into two phases based on the marker bands used. Furthermore, it is obvious that the S54G/P55N variant refolds to its native conformation approximately twice as fast as the wild type. This fact suggests that the *trans* → *cis* prolyl isomerization of Pro 55 must play an important role as a rate-limiting step in the refolding process of RNase T1 under destabilizing conditions [30]. At stabilizing conditions (T-jump from 65°C to 20°C) the refolding pathway of the S54G/P55N variant also proceeds faster and with simpler kinetics compared to the wild type (see time constants and amplitudes of Table 3.3).

3.4.1.2 Structure Formation Within the Experimental Dead Time

The FTIR technique used offers an unique approach for the quantitative detection also of those events occurring within the experimental dead time. Briefly, this is obtained by calculating the difference between the total spectral changes observed during a typical T-jump experiment and the spectral changes observed with the same sample under equilibrium conditions in the same temperature range [29, 30, 33].

Table 3.3 Time constants τ and relative amplitudes α of kinetic phases observed for RNase T_1 wildtype and variant S54G/P55N at different refolding temperatures using different structure sensitive IR marker bands

T-jump	Protein	Evaluated IR marker bands	$\tau 1$ [s]	$\tau 2$ [s]	$\tau 3$ [s]	$\alpha 1$ [%]	$\alpha 2$ [%]	$\alpha 3$ [%]
65°C → 45°C	Wild type	Tyrosine ($\sim$1,515 cm^{-1})	1.3 ± 0.8	39 ± 2	–	13 ± 5	83 ± 13	–
		Low-frequency β-sheet (1,625 cm^{-1})						
	S54G/P55N	Tyrosine ($\sim$1,515 cm^{-1})	0.5 ± 0.2	18 ± 1	–	14 ± 4	83 ± 9	–
		Low-frequency β-sheet (1,625 cm^{-1})						
65°C → 20°C	Wild type	Tyrosine ($\sim$1,515 cm^{-1})	0.7 ± 0.5	110 ± 20	–	15 ± 5	30 ± 3	–
		Low-frequency β-sheet (1,625 cm^{-1})	0.6 ± 0.2	77 ± 14	816 ± 50	16 ± 3	21 ± 2	14 ± 2
	S54G/P55N	Tyrosine ($\sim$1,515 cm^{-1})	–	90 ± 22	–	–	29 ± 4	–
		Low-frequency β-sheet (1,625 cm^{-1})	4.3 ± 1.8	72 ± 30	760 ± 320	13 ± 3	20 ± 2	22 ± 1

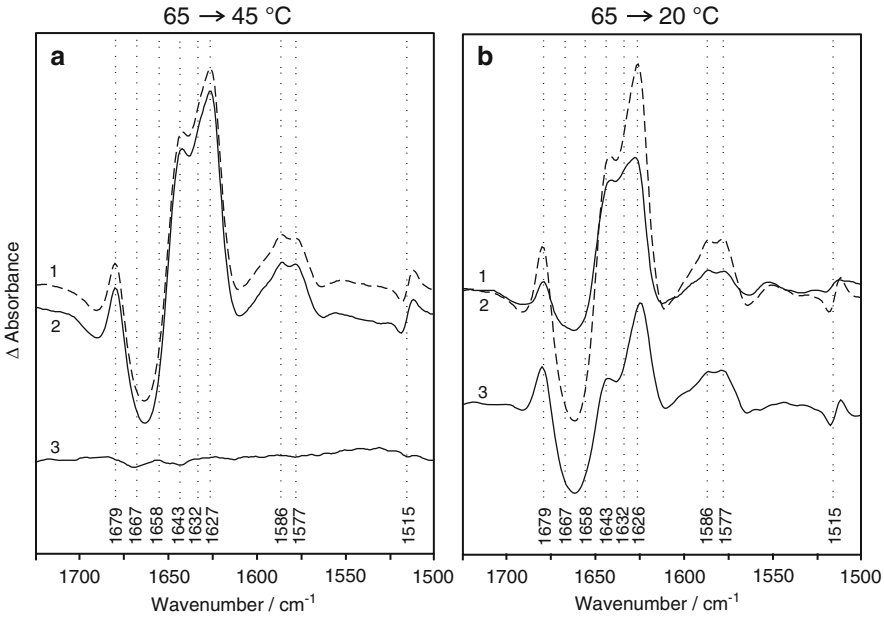

Fig. 3.13 Refolding events within the experimental dead time as determined for the RNase T1 variant S54G/P55N. Total changes (curves 1, *dashed lines*) observed during a refolding experiment under equilibrium conditions from (**a**) 65°C to 45°C and from (**b**) 65°C to 20°C. Difference spectra were calculated from the absorbance spectra recorded at the corresponding temperatures and corrected for changes due to unspecific temperature adaptation (for details, see [29, 30]). Total spectral changes (curves 2, *solid lines*) observed during the kinetic refolding experiment after T-jump from (**a**) 65°C to 45°C and (**b**) from 65°C to 20°C. The difference spectra were calculated between the corresponding averaged first five spectra which define a nominal dead time of ∼180 ms with the actual experimental dead time (the first single-beam spectrum defines the experimental dead time of ∼59 ms) and the last spectrum obtained after 30 min. Curves 3 (*solid lines*) at the *bottom* of (**a**, **b**): Spectral changes occurring during the dead time of the T-jump experiments. These double difference spectra were obtained by subtracting curves 2 from curve 1

Figure 3.13a shows the corresponding spectral changes observed for the S54G/P55N variant between 65°C and 45°C under destabilizing conditions, and Fig. 3.13b between 65°C and 20°C under stabilizing conditions. The difference spectra 1 and 2 in both figures show the changes under thermodynamic (equilibrium conditions) and kinetic (T-jump) control, respectively. The double difference spectrum obtained by calculating the difference between curve 1 (complete refolding events obtained from the unfolding/refolding experiment under thermodynamic control) and curve 2 (total events detected between 180 ms and 30 min after induction of the refolding) yields an almost flat line (curve 3 in Fig. 3.13a). This finding clearly suggests that essentially no secondary structure and tertiary contacts are formed within the experimental dead time (∼180 ms) under these conditions.

In contrast, under stabilizing conditions (T-jump from 65°C to 20°C, see Fig. 3.13b, curve 3) considerable secondary structure and also significant parts of the

tertiary structure are already formed within the experimental dead time as indicated by spectral features at 1,643, 1,627, 1,577, and 1,587 cm^{-1}, and the peak shift of the tyrosine band at 1,515 cm^{-1}. Since the latter three marker bands specifically monitor tertiary structure formation, it was concluded [30] that a significant amount of tertiary structure is already formed within the experimental dead time by rapidly "trapping" the proline peptide bond in its "incorrect" *trans* conformation. Thus, the species being generated within the experimental dead time can be considered to form a folding intermediate, as it has already been stated by other authors using different experimental conditions and techniques [27, 28, 34–38].

3.4.2 Unfolding of the λ-Cro Repressor

The λ-Cro repressor, a small dimeric protein consisting of two identical polypeptide chains, each containing 66 amino acids, is known to unfold via highly populated intermediate states under certain conditions [39]. Figure 3.14 (bottom panels) shows the IR spectra of the wild-type protein (left side) and an engineered λ-Cro repressor variant, in which Val-55 is replaced by Cys, designated as Cro-V55C (right side), measured at 20°C (solid lines) and 55°C (dashed lines). The spectra of the folded proteins at 20°C are dominated by three bands in the amide I′ region: bands at 1,624 and 1,678 cm^{-1} that confirm the presence of antiparallel β-strands and a band at 1,650 cm^{-1} that indicates the presence of α-helical structures (see also Fig. 3.4d). The interplay between changes in secondary structure and the state of association upon unfolding of the wild-type protein (Cro-WT) and a number of engineered variants was studied by FTIR spectroscopy and dynamic light scattering (DLS) [40–42]. The combined approach revealed that the first thermal transition of the Cro-V55C dimer involves the melting of the α-helices and only parts of its native β-sheet structure and that this event is accompanied by the formation of tetramers. The wild-type protein, on the other side, was found to have lost all of its α-helical and also most of its native β-sheet structure at intermediate temperatures of 50–60°C. Thermal unfolding initiated by a rapid T-jump from 20°C to 55°C and monitored by time-resolved FTIR spectroscopy permitted us to follow unfolding events on the timescale of 10^{-1} to 10^{3} s [43].

A quantitative comparison of the total spectral changes observed during the T-jump experiments from low to intermediate temperatures (Fig. 3.14b, left side) with those observed within the same temperature range for a corresponding experiment performed under equilibrium conditions (Fig. 3.14a, left side) reveals that a significant amount of the secondary structure of Cro-WT unfolds within the dead time of the experiment. Moreover, the spectral changes observed at 1,624 and 1,650 cm^{-1} are slightly different from each other (Fig. 3.15a, b), indicating that unfolding of the α-helices of Cro-WT is completed prior to the changes of its β-sheet structure [43]. Stabilization of the two protein subunits by a disulfide bridge in the variant Cro-V55C drastically changes its unfolding behaviour. Practically no changes at 1,650 cm^{-1}, but clear spectral changes associate with the high- and

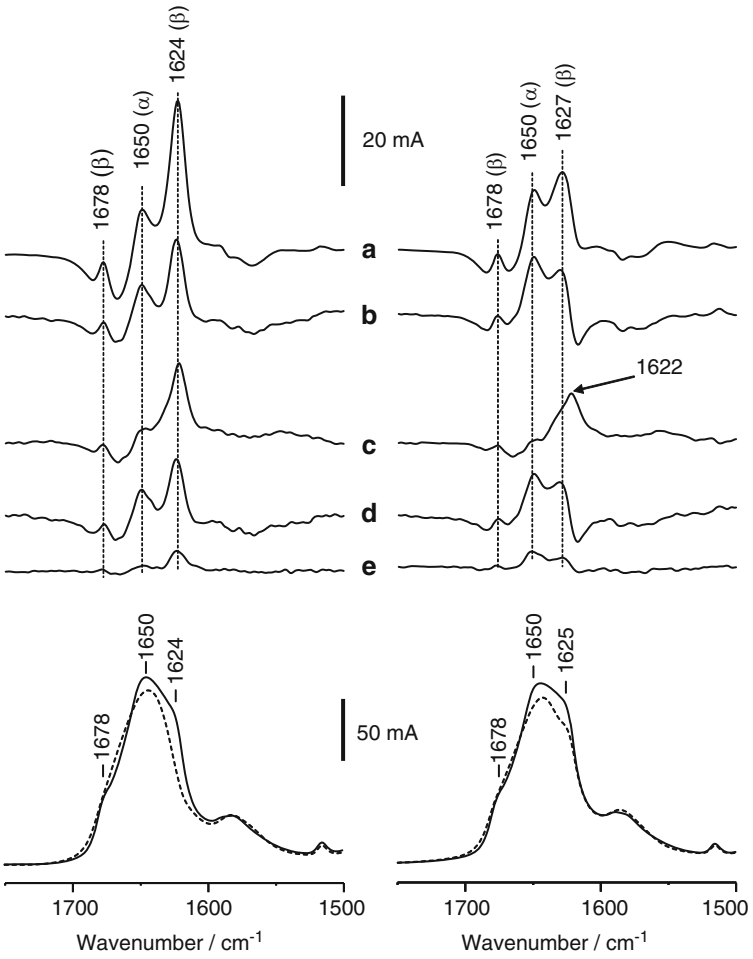

Fig. 3.14 *Bottom panels*: IR spectra of Cro-WT (*left side*) and Cro-V55C (*right side*) measured at 20°C (*solid lines*) and 55°C (*dashed lines*). *Top panels*: IR difference spectra observed during formation of the intermediate states of Cro-WT (*left side*) and Cro-V55C (*right side*). (**a**) Total changes observed during an unfolding experiment under equilibrium conditions between 20°C and 55°C. (**b**) Total changes observed during a T-jump experiment from 20°C to 55°C. (**c**) Double difference spectra (**a** minus **b**), representing the spectral changes occurring within the dead time of the T-jump experiment. (**d**) Spectral changes observed between 79 ms and 1 s of the T-jump experiment. (**e**) Spectral changes observed between 1 and 55 s of the T-jump experiment

low-frequency β-sheet band components are observed immediately after the T-jump (Fig. 3.14c, right side; Fig. 3.15c, d). This indicates that unfolding of the β-sheet structure of Cro-V55C temporally proceeds before changes in regions of its native α-helical structure occur. Moreover, quite different spectral changes in the region of the β-sheet band components are observed for the wild-type protein and Cro-V55C

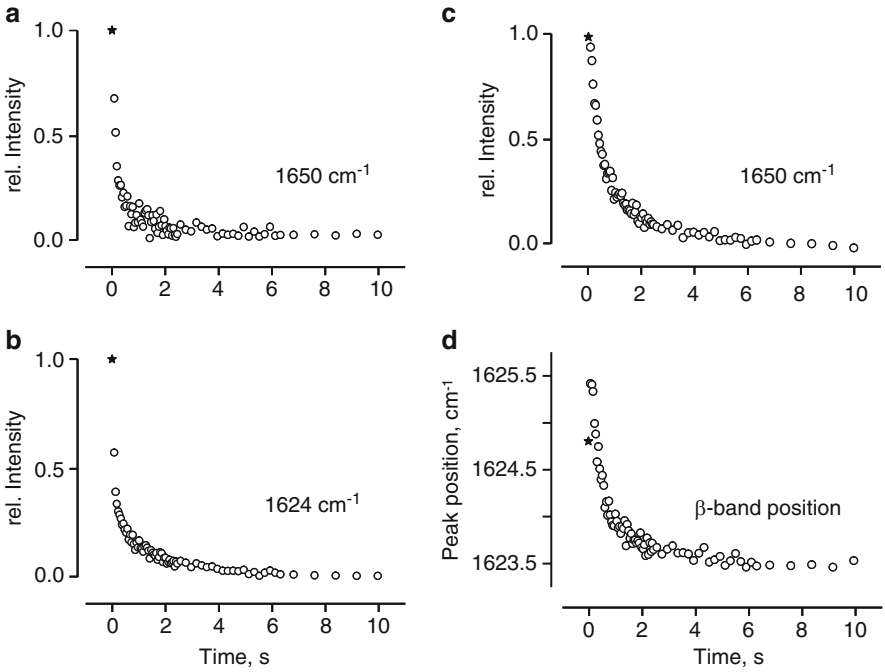

Fig. 3.15 Kinetic plots (*open circles*) of the IR bands monitoring (**a**, **c**) unfolding of the α-helices (1,650 cm^{-1}), (**b**) the disruption of β-sheet structure (1,624 cm^{-1}), and (**d**) the changes in β-sheet structure (β-band position) during the T-jumps with Cro-WT (*left side*) and Cro-V55C (*right side*). The plots are normalized to the total changes observed during unfolding experiments under equilibrium conditions (*filled stars*)

within the first seconds after the T-jump (compare Fig. 3.14c, d, left and right side). Most striking is a shift in peak position of the low-frequency β-band by 2 cm^{-1} for the variant (Fig. 3.15d), which is not observed for Cro-WT. This suggests that unfolding of the α-helices and disruption of parts of the native β-sheet structure of Cro-V55C are accompanied by a reorganization of the β-strands that remain in the tetrameric intermediate state of Cro-V55C at 55°C [41].

3.5 Examples Making Use of Rapid-Mixing Methods

3.5.1 Refolding of α-Lactalbumin Studied by Stopped-Flow Infrared Spectroscopy After a pH-Jump

α-Lactalbumin, the regulatory subunit of lactose synthase, is a 123 amino acids containing protein with an α-domain including four helices and a β-domain containing a small β-sheet and a 3_{10}-helix that has frequently been used as a model

system to study protein folding [44–46]. A high-affinity binding site for Ca^{2+}-ions is located in a loop between the 3_{10} helix (β-domain) and an α-helix (α-domain). The calcium-bound native form (so-called holo-form) is more stable and less flexible than the calcium free native conformation (apo-form). At acidic pH, α-lactalbumin forms the so-called partially folded A-state, also called a molten-globule state [47]. In the presence of calcium, the transition from the partially folded A-state at acidic pH to the holo-form is known to proceed fast and cooperative. Moreover, the presence of calcium during refolding leads to a nucleation mechanism after Ca^{2+}-binding and an early intermediate. Without calcium refolding is more complex and nonnative secondary structures are formed transiently [45, 48, 49]. Stopped-flow pH-jump FTIR experiments allowed to follow the refolding of α-lactalbumin from the A-state to the holo- and apo-form [50].

3.5.1.1 IR Spectra of α-Lactalbumin Under Equilibrium Conditions

Figure 3.16a shows the absorbance spectra for the A (molten globule)-, holo-, and apo-forms of α-lactalbumin under equilibrium conditions. Obviously, the spectral contours of the two native forms (holo and apo) are very similar, suggesting that the presence of Ca^{2+} has no major impact on the structure of the protein, which is verified by their difference spectrum (Fig. 3.16b, red trace). Figure 3.16b shows three difference spectra between the three different α-lactalbumin forms shown in Fig. 3.16a. These difference spectra clearly prove that the two native states are indeed very similar. The two native conformations exhibit significant differences only near 1,515 and 1,577 cm^{-1}, indicating mainly differences in tertiary structure, while minor differences in the amide I region not exceed the amount of 1–2% of the total absorbance of the amide I band, suggesting only small differences in secondary structure. The shift of the aspartate band from ~1,588 to 1,577 cm^{-1} can be explained by the interaction of aspartate carboxylate side-chain groups with Ca^{2+}-ions in the holo-form. The positive and negative peaks appearing in the difference spectra between the A-state and the two native states (Fig. 3.16b, black and blue traces), respectively, correspond to approximately 5% of the total absorbance of the amide I band, showing that the amount of secondary structure of the A-state is already close to that of the two native forms of α-lactalbumin. This is in agreement with X-ray scattering studies which revealed that α-lactalbumin at acetic pH is rather compact with a radius of gyration only approx. 10% larger than in the native state [51].

3.5.1.2 Refolding Kinetics of α-Lactalbumin Monitored by Stopped-Flow FTIR Difference Spectroscopy

Stopped-flow pH-jump FTIR experiments were performed to unravel differences between the refolding process from the A-state to the holo-form, and the A-state to the apo-form, respectively [50]. Refolding of the bovine α-lactalbumin was initiated

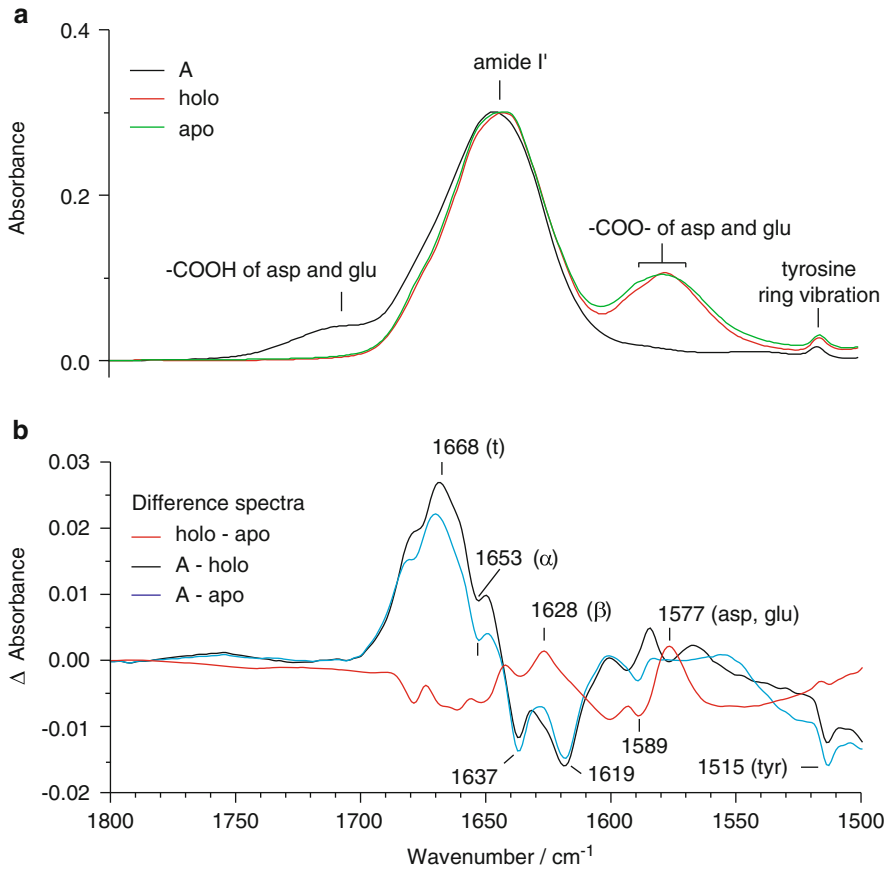

Fig. 3.16 IR spectra of α-lactalbumin in its folded states ("apo" and "holo") and in the partially folded A-state under equilibrium conditions. (**a**) Normalized absorbance spectra: holo-form (*red*), apo-form (*green*), and A-state (*black*). (**b**) Difference spectra obtained from the absorbance spectra in (**a**), *red*: "holo" minus "apo", *black*: A-state minus native "holo-state", *blue*: A-state minus native "apo-state". Difference peaks appearing at 1,668, 1,650, 1,627, 1,577, and 1,515 cm^{-1} stand for turn, α-helical, β-sheet, side-chain asp/glu, and side-chain tyrosine, respectively. Note that the Y-scale of the difference spectra was expanded by a factor of 10 compared to the scale of the absorbance spectra

by mixing the protein solution at pH 2 (A-state, molten globule) with a basic solution adjusted such that a final pD of ∼7.5 after mixing was obtained. 240 µl of protein solution containing 20 mg/ml of the protein in D_2O/DCl was mixed with 24 µl of a basic solution containing 460 mM Tris and either 10 mM EDTA (apo) or 100 mM $CaCl_2$ (holo).

For rapid-mixing stopped-flow experiments, the device shown in Fig. 3.7 of this chapter has been used to keep the temperature of the FTIR flow-through cell at 4°C to slow down the reaction rates (see also [50, 52]).

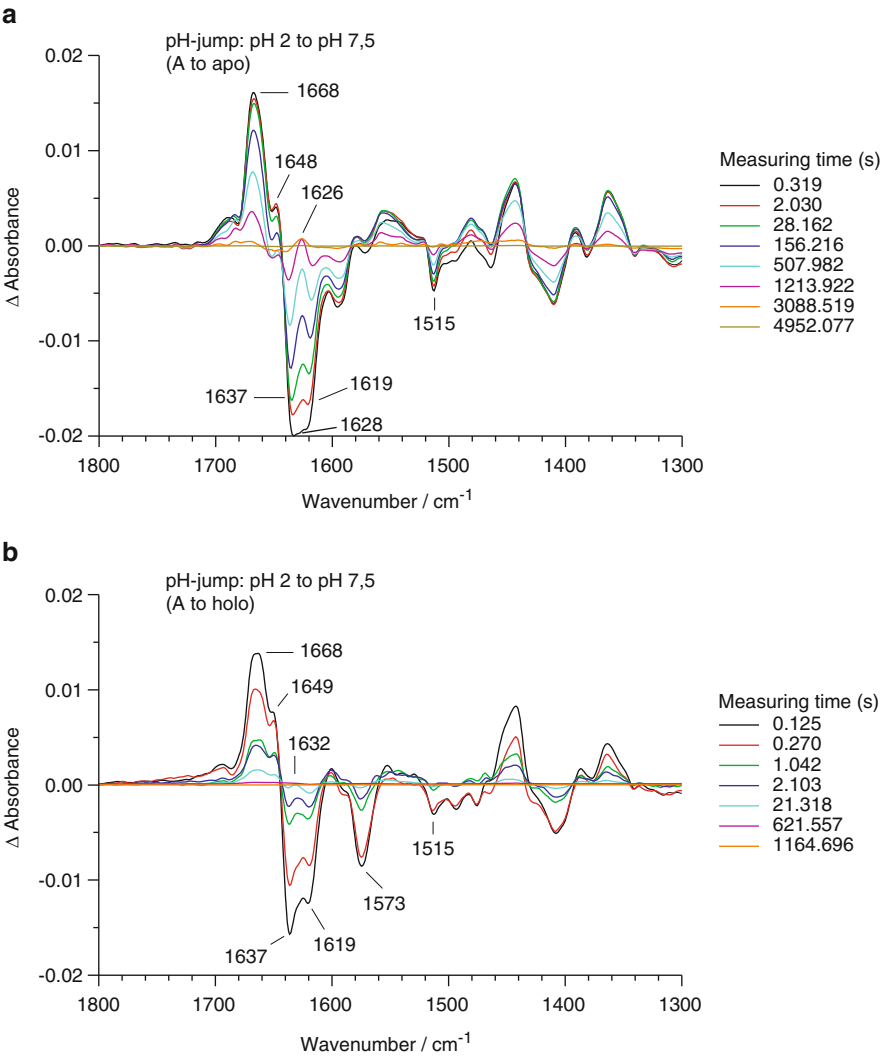

Fig. 3.17 Difference spectroscopic features time dependently observed for α-lactalbumin after a pH-jump from pH 2 to 7.5. (**a**) Difference spectra calculated for the transition from the A-state to the native apo-form. (**b**) Difference spectra calculated for the transition from the A-state to the native holo-form

Spectra were collected as a function of reaction time after mixing. The spectra of Fig. 3.17a, b have been obtained by calculating the difference between the last and any preceding time incremental spectrum. Thus, emerging structures produce negative bands, whereas disappearing structures result in positive bands. Positive and negative features converge to a zero line as the refolding process comes to its end.

Table 3.4 Rate constants for the apo and holoform of α-lactalbumin as deduced from the various difference peaks of Fig. 3.17 using a three-exponential fit [50]

	$k_1[s^{-1}]$	$k_2[s^{-1}]$	$k_3[s^{-1}]$
apo (without Ca^{2+})	0.046	0.0018	0.00094
holo (with Ca^{2+})	0.9	0.08	0.004

Both series of difference spectra shown in Fig. 3.17 indicate that the presence or absence of Ca^{2+}-ions has a strong influence not only on the general reaction rates but also on the time-dependent appearance of structural elements that can be identified by IR difference spectroscopy. This can be seen for the various secondary structure-specific difference peaks in the amide I region and the difference features attributed to various amino acid side-chain absorptions near 1,515 cm^{-1} (tyrosine) and between 1,577 and 1,588 cm^{-1} ($-COO^-$ of aspartate). As expected, refolding of the holo-form is faster than refolding of the apo-form and the time-dependent appearance of the various structures is more cooperative for the holo- than for the apo-form, as can be deduced from the qualitative structure of the difference spectra. In the case of the apo-form, the structural elements appear at a slower rate and are more complex (Fig. 3.17a). Different kinetic plots were constructed using the different absorbance bands responsible for the various structural elements present in α-lactalbumin. Three rate constants for transient intermediates turned out to be required to adequately fit the time-resolved FTIR spectroscopic data for the formation of the holo- and the apo-form (Table 3.4) [50].

Apart from the fact that refolding of the apo-form to the native state was found to be much quicker than for the holo-form, the kinetic IR data indicated a fairly complex refolding process where the protein refolds via several intermediate states after the dead time of the experiment [50].

3.5.2 Misfolding of β_2-Microglobulin

Native β_2m adopts a seven-stranded β-sandwich structure organized into two antiparallel β-sheets linked by an intersheet disulfide bond and represents the noncovalently bound light chain of major histocompatibility complex class I molecules [53]. β_2m is also the major component of the amyloid deposits that are found in the musculoskeletal system in patients suffering from dialysis-related amyloidosis [54]. Although one of the most extensively studied amyloidogenic proteins [55], the mechanisms by which soluble β_2m is converted into β_2m-amyloid fibrils in vivo are still largely unknown. In vitro, β_2m has been shown to be stably folded at physiological pH and intransigent to assembly into amyloid fibrils, and it is assumed that partial unfolding of the native protein is required for fibrillation to occur. Amyloid-like fibrillar structures of wild-type β_2m can form easily and with high yield under acidic conditions, when the protein is initially partially or fully unfolded. Depending on the growth conditions used, fibrillar structures

of different morphologies are formed [56, 57]. For example, long straight (LS) fibrils can form spontaneously upon agitation at pH 2.5 and low ionic strength [55]. Without agitation, β_2m seems not to be fibrillogenic below pH 2.5 at low ionic strength [58], but can form curved worm-like (WL) assemblies, also called protofilaments. Although the amyloid fibrils formed under acidic conditions are not stable at physiological pH [59, 60], and their relevance in vivo is consequently unclear, these conditions are ideal for an analysis of the process of β_2m assembly under different kinetic regimes.

We have performed a series of experiments using IR spectroscopy and static (SLS) as well as dynamic light scattering (DLS) to identify and characterize products emerging during assembly of β_2m. Specifically, filamentous species formed after adding sodium chloride to the acid-unfolded state of β_2m were compared, and their presence and kinetics of assembly are validated directly using the combined IR and SDS/DLS approach. Moreover, we studied the impact of moderate and extended heat treatment on the salt-induced filaments. The association of β_2m molecules was induced by adding a stock solution of 1 M NaCl in 10 mM DCl to the acid-denatured protein sample outside the spectrometer. After mixing, 15 µl of the sample mixture was filled into a home-made demountable IR cell [61] with an optical pathlength of 50 µm, and IR spectra were collected in the normal scan mode of the spectrometer. This simple hand mixing approach allowed us to measure the first IR spectrum ~2 min after initiation of the association process. In parallel, SLS and DLS experiments were performed, providing information concerning changes in molecular mass and hydrodynamic Stokes radii of the molecules at protein concentrations and over time periods practically identical to those used for the IR experiments.

The IR spectrum of native β_2m in D_2O buffer at pH 7.5 (Fig. 3.4f) is characterized by a major amide I band at 1,633 cm^{-1} and a weaker band component at 1,682 cm^{-1}, both together indicating the presence of intramolecular antiparallel β-sheet structures, which disappear in the acid-denatured state in 2 mM DCl at pH 2.4 [10]. The addition of NaCl to the acid-unfolded protein induces the formation of secondary structure, as indicated by the weak features at 1,631 and 1,682 cm^{-1} in the spectrum of β_2m 3 min after addition of NaCl (Fig. 3.18a, red trace). These spectral features of the initially salt-induced state of β_2m suggest the presence of β-strands organized in an intramolecular-like fashion, which later disappear. On hour after initiation of the association process, a new band component at ~1,616 cm^{-1} is observed (Fig. 3.18a, blue trace), whose intensity increases over the time. The strong low-frequency band component at ~1,616 cm^{-1} and a weaker band component at ~1,684 cm^{-1} indicate the presence of antiparallel intermolecular β-sheet structure in the associates of β_2m. DLS and IR experiments carried out over 2 days under the same conditions reveal that the kinetics of intermolecular β-sheet formation (Fig. 3.18b, red dots) and aggregation of the protein (Fig. 3.18b, open black squares) are complex. A plot of the temporal changes of the intensity of the IR band at 1,616 cm^{-1} versus the changes of the relative apparent mass (Fig. 3.18c) reveals two different phases of synchronous events of structure conversion and association. A change in these processes is observed at the stage of octamers to decamers

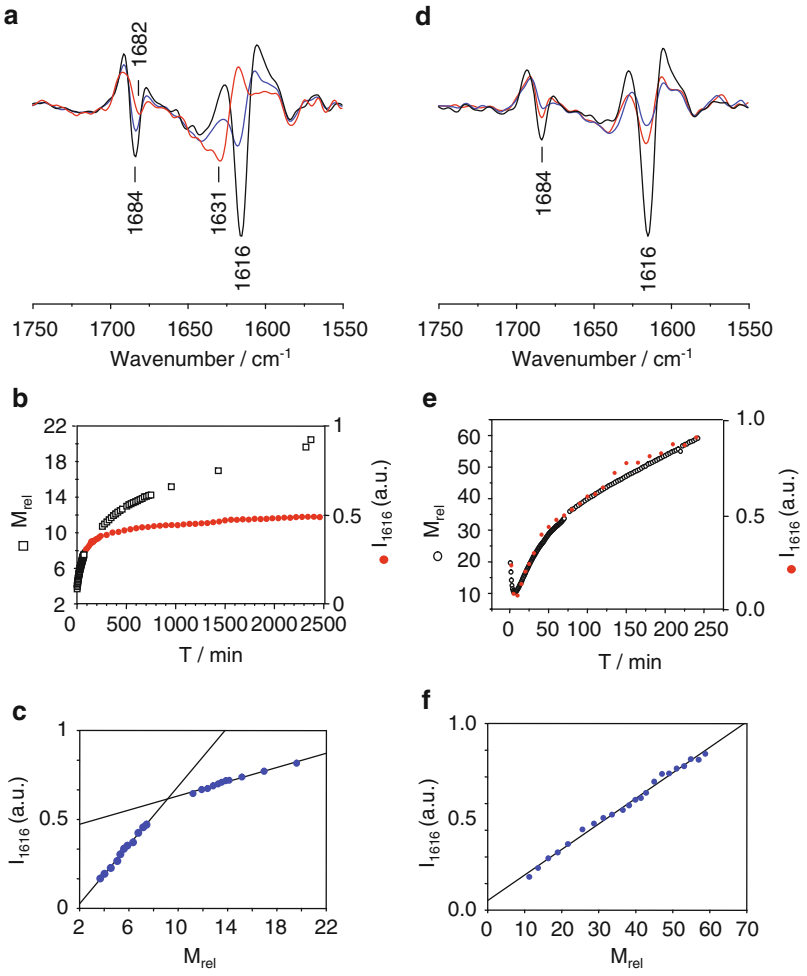

Fig. 3.18 Association-induced structural changes of β_2m at pH 2.1 in the presence of 0.2 M NaCl. (*Left side*) Formation of short WL assemblies at 25°C. (**a**) IR spectra (second derivatives) of β_2m 3 min (*red*), 1 h (*blue*), and 40 h (*black*) after addition of NaCl. (**b**) Comparison of the changes in intensity of the IR band at 1,616 cm^{-1} (*red circles*) as monitor of the formation of intermolecular β-sheet structure, with the changes in the relative mass of the aggregates (*squares*) as determined from the DLS data. (**c**) Temporal changes of the IR band at 1,616 cm^{-1} (I_{1616}) versus changes of the relative molecular mass (M_{rel}) observed over a time period of 40 h. (*Right side*) Heat-induced disaggregation/reassociation of WL β_2m assemblies into longer filamentous structures after a T-jump from 25°C to 50°C. (**d**) IR spectra (second derivatives) of β_2m at 50°C 3 min (*red*), 15 min (*blue*), and 18 h (*black*) after the T-jump. (**e**) Comparison of the changes in intensity of the IR band at 1,616 cm^{-1} (*red circles*) with the changes in aggregate size (*open circles*). (**f**) Temporal changes of the IR band at 1,616 cm^{-1} versus changes of the relative molecular mass observed over a time period of 4 h

(M_{rel} 8–10). Electron micrographs of the β_2m assemblies formed at pH 2.1 in the presence of 0.2 M NaCl revealed short WL structures [10], such as obtained by other groups under similar conditions [57, 62]. An increase in temperature of the corresponding β_2m samples in the IR or DLS cells from 25°C to 50°C within 2–3 min was able to disaggregate the WL assemblies, followed by the formation of longer filamentous structures [10]. The IR spectra of the β_2m sample exhibit initially a loss of band intensity at 1,616 cm^{-1}, followed by a pronounced increase in intensity of this band as a function of incubation time at 50°C (Fig. 3.18d). Practically identical kinetics for unfolding/re-formation of β-sheet structure on one side and disaggregation/reassociation of the β_2m molecules are observed (Fig. 3.18e). The linear correlation between change in IR signal and mass increase over the entire time range (Fig. 3.18f) is clearly different from the salt-induced changes described before (Fig. 3.18c). The long filamentous assemblies obtained after incubation for days at 50°C were not amenable to further structural changes by heating up to 90°C [10]. In contrast to that, extended heat treatment of the short WL particles may induce major structural changes, leading to a species with an IR spectrum different from that of the initial state (compare the blue and red traces in Fig. 3.19a, b). Most strikingly is the lack of a high-frequency β-sheet component at 1,684 cm^{-1} in the spectrum after heat treatment, indicating that the intermolecular β-sheet structure of the new state must clearly be different from that of the initial state (Fig. 3.19c). The electron micrographs of the heat-treated assemblies, characteristic for LS amyloid fibrils, are shown in Fig. 3.19d. Independent support for a significant structural reorganization upon heat treatment comes from the 1 cm^{-1} difference in frequency of the tyrosine band for the two states (Fig. 3.19b, blue and red traces).

Following theoretical calculations [4], we took the absence of the spectral component at ∼1,684 cm^{-1} in the IR spectra of the LS fibrils as indicative for a parallel organization of the β-strands in β_2m fibrils [10]. In contrast, the IR spectrum of the WL assemblies is dominated by a component at 1,616 cm^{-1} together with a weaker component at 1,684 cm^{-1}, both together suggesting that this nonnative intermolecular structure of β_2m involves an antiparallel arrangement of the β-strands. Very recently, direct proof for a parallel in-register arrangement of β-strands in LS β_2m fibrils, which is drastically distinct from that in native β_2m, came from magic angle spinning NMR [63] and electron paramagnetic resonance experiments [62]. The latter report also provided evidence that the WL particles of β_2m do not adopt a parallel arrangement, such as observed for the LS fibrils. It is interesting to note that the high-frequency β-band component is only very weak or even absent in the IR spectra of LS fibrils of other proteins, such as insulin [64], the SH3 domain [65, 66], lysozyme [67], transthyretin [68], α-synuclein [69], amyloid beta peptides [70], or peptide models [71, 72], and in the spectrum of natively folded barstar as well (Fig. 3.4c), a protein composed of four helices and a three-stranded parallel β-sheet. At the same time, different reports have shown that the high-frequency component at ∼1,684 cm^{-1} is well expressed in the spectra of other protein assemblies, such as oligomeric [70] or WL particles [71], and amorphous aggregates [59, 64, 65]. Apart from uncertainties in correlating IR spectral features

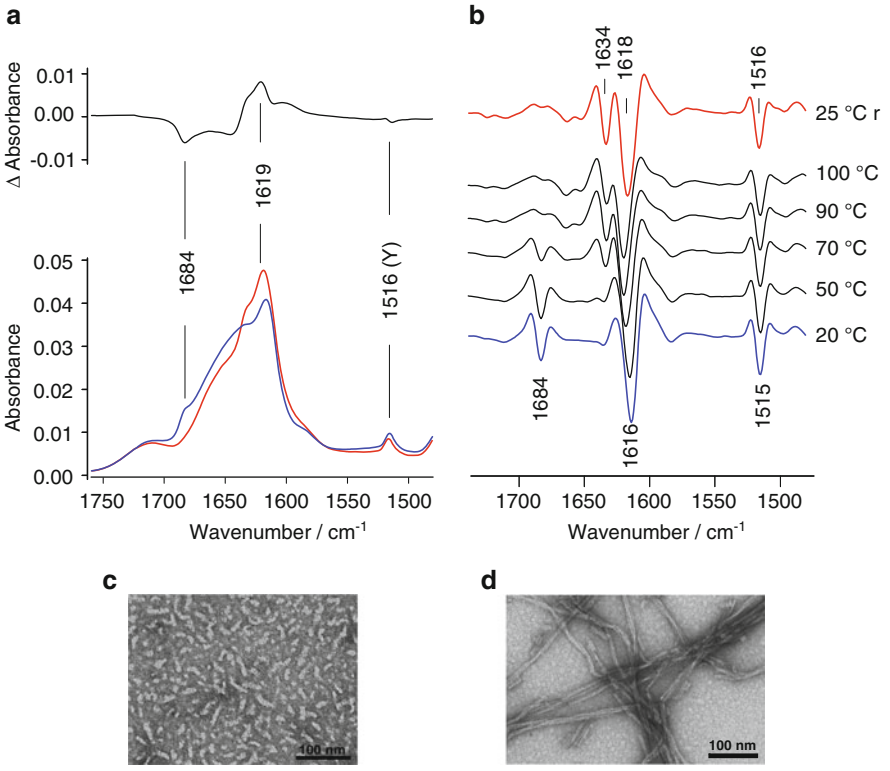

Fig. 3.19 (**a**) IR absorbance spectra of WL assemblies of β_2m at 25°C before heating (*blue line*) and of β_2m fibrils obtained by a rinse in temperature up to 100°C and cooling back to 25°C (*red line*). The *upper* trace shows the corresponding IR difference spectrum (after heating minus before heating). (**b**) Stacked representation of the second derivatives of the IR spectra of the assemblies at the indicated temperatures upon heating and after cooling from 100°C to 25°C (25°C r). (**c**, **d**) Corresponding electron micrographs for the short WL particles (**c**) and LS fibrils of β_2m (**d**)

with three-dimensional structural details (Chaps. 2 and 9 offer a detailed discussion), the observations described above support the assumption that a parallel organization of β-strands might be a characteristic of well-organized LS fibrils and that such an arrangement can be identified based on its IR spectroscopic characteristics. Further experiments on other well-characterized amyloid fibrils, together with theoretical considerations, should shed light on whether this assumption can be generalized.

Examples of IR spectra of different forms of β_2m are shown in Figs. 3.20 (deuterated samples) and 21 (protonated samples). The spectrum of the heat-induced amorphous aggregates (Fig. 3.20b) is clearly different from that of the native protein (Fig. 3.20a) and those of β_2m fibrils formed at 65°C by agitating the protein solutions at neutral pH and high ionic strength (Fig. 3.20c) or formed at 37°C by agitating the samples at acidic pH and low ionic strength (Fig. 3.20d). A high-frequency band component, attributed to the presence of an antiparallel arrangement

Fig. 3.20 IR absorbance spectra (*left side*) and the corresponding second-derivative spectra (*right side*) of different forms of fully deuterated β_2m. (**a**) native β_2m at neutral pH (**b**) amorphous aggregates of β_2m after heat treatment at neutral pH. (**c**) agitation-induced LS fibrils of β_2m in the presence of 1 M NaCl at neutral pH. (**d**) agitation-induced LS fibrils of β_2m in the presence of 0.05 M NaCl at pH 2.1. (**e**) acid-unfolded β_2m in 2 mM DCl at pH 2.4

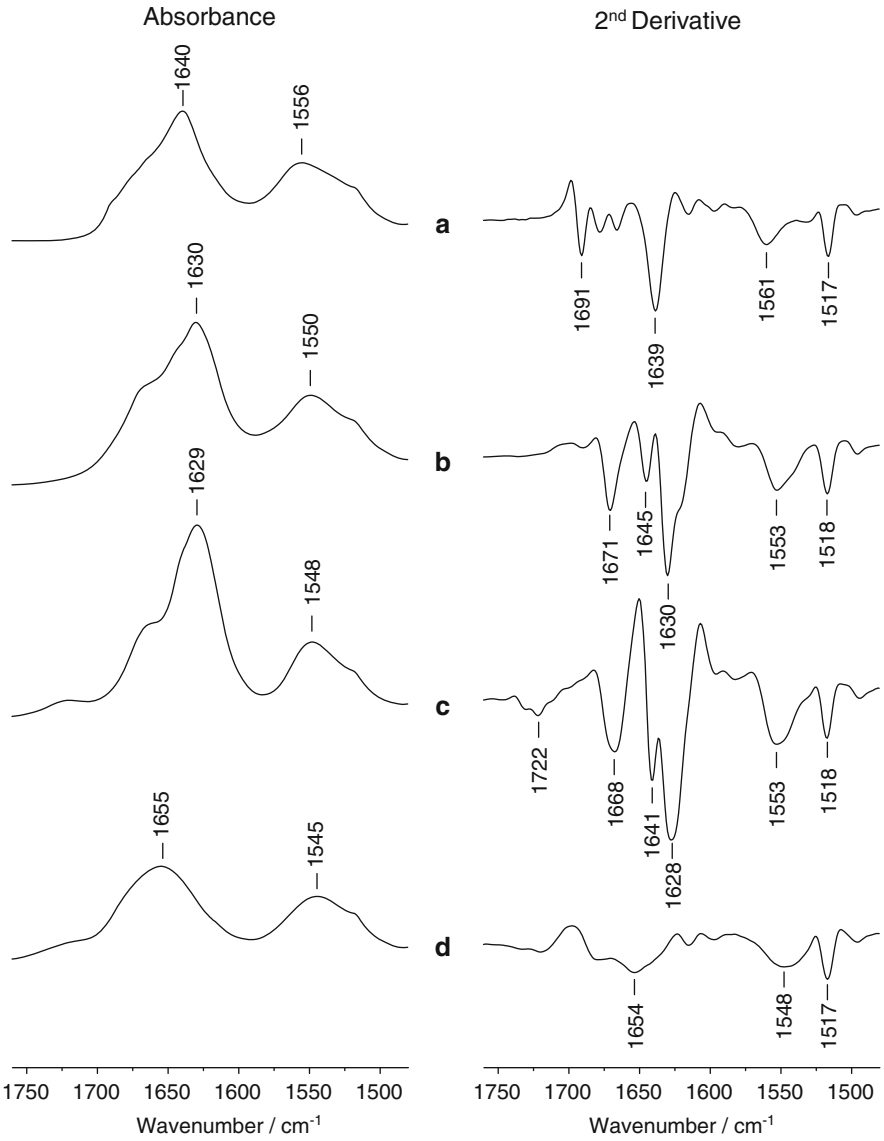

Fig. 3.21 IR absorbance spectra (*left side*) and the corresponding second-derivative spectra (*right side*) of different forms of β_2m in H_2O. (**a**) Native β_2m at neutral pH. (**b**) Agitation-induced LS fibrils of β_2m in the presence of 1 M NaCl at neutral pH. (**c**) Agitation-induced LS fibrils of β_2m in the presence of 0.05 M NaCl at pH 2.1. (**d**) Acid-unfolded β_2m in 2 mM HCl at pH 2.1

of β-strands, is missing in the spectra of the fibrils obtained at low and neutral pH values, both in D_2O (Fig. 3.20c, d) and in H_2O (Fig. 3.21b, c). The assignment of the spectral features at ~1,662 cm^{-1} and ~1,670 cm^{-1} in the spectra of the deuterated (Fig. 3.20c, d) and protonated (Fig. 3.21b, c) β_2m fibrils, respectively, is presently

uncertain. Possibly, these band components reflect a characteristic arrangements of β-strands in LS fibrils, which cannot be specified based on currently known structure—spectrum correlations for natively folded proteins. A pronounced feature at $\sim$1,662 cm^{-1} was also observed in the spectra of straight fibrils of the hamster prion protein, but not in the spectra of curvy S-like fibrils of the same protein obtained under different shaking conditions [73]. The authors suggested that the band component at 1,662 cm^{-1} is indicative of β-turns and loops, despite the fact that this spectral feature remained almost unaffected by heat treatment of the fibrils [74]. Although detailed molecular interpretations of such IR spectra currently are not possible, it is clear that the IR spectra are sensitive monitors of the arrangement of β-strands in various aggregate states.

3.5.3 The α-to-β Conversion Process of the Prion Protein

Following the protein only hypothesis, the prion protein (PrP) is the sole agent causing a group of transmissible neurodegenerative disorders, the so-called prion diseases or prionoses [75, 76]. The crucial step in transmission and manifestation of prion disease is the conversion of noninfectious cellular prion protein (PrP^c) exhibiting mainly an α-helical structure to the infectious multimeric prion protein (PrP^{Sc}), which is misfolded into predominantly aggregated β-sheet structures. The late products of this process are amyloid fibrils and amyloid plaques [77]. The in vitro conversion of PrP to amyloid forms with physical properties similar to PrP^{Sc} have been investigated by numerous groups, including some using FTIR spectroscopy in concert with other biophysical techniques [78–81]. To follow the kinetics of the conversion of soluble PrP^c into aggregated PrP^{Sc} species by FTIR techniques represents an experimental challenge and has been approached only seldom so far [78, 82].

We have performed stopped-flow FTIR experiments with the recombinant Syrian hamster prion protein comprising residues 90–232 ($SHaPrP^{90-232}$) [78] to study the kinetics of the conversion process of a prion protein in vitro. Since the amide I band components diagnostic for α-helices and β-sheets are well separated in the infrared spectra of proteins, changes in these types of structures can be monitored simultaneously and independently. Moreover, intramolecular β-sheets can be differentiated from intermolecularly aggregated β-sheet structures on the basis of their IR spectra [61].

Figure 3.22 shows a typical second-derivative spectrum of a $SHaPrP^{90-232}$ measured at pH 7 exhibiting relatively sharp band components assigned to amide and amino acid side-chain groups. The prominent bands at both 1,651 (amide I) and 1,551 cm^{-1} (amide II) are due to α-helical, and the amide I component at 1,678 cm^{-1} is due to turn and/or loop structures. Only a very weak amide I band at 1,628 cm^{-1} possibly due to the small β-sheet structure in $SHaPrP^{90-232}$ is visible. The band at 1,517 cm^{-1} is attributed to the amino acid side-chain absorbance of tyrosine [61].

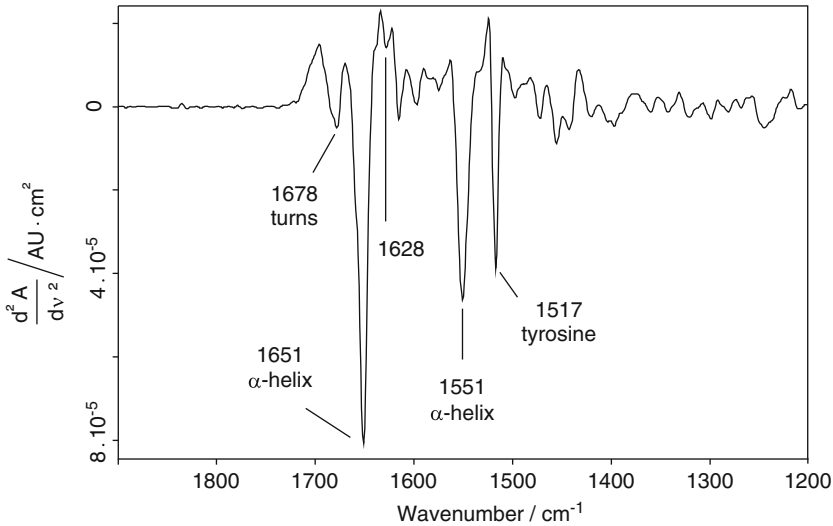

Fig. 3.22 Second-derivative spectrum of $SHaPrP^{90-232}$ at pH 7 calculated from a typical absorbance spectrum shown in Fig. 3.9

Under mildly acidic conditions (pH 4.2) and in the presence of low concentrations of GdmCl (∼1 M) the recombinant prion protein undergoes a transition from a monomeric α-helical form to a β-sheet rich oligomeric form [78]. The stopped-flow apparatus described in Fig. 3.8 permitted us to obtain highly reproducible IR spectra of the prion protein at fairly low concentrations in protein/H_2O solutions with excellent signal-to-noise ratio. The structural changes observed for $SHaPrP^{90-232}$ after a denaturant concentration jump from zero to 1 M GdmCl concentration were kinetically followed using FTIR difference spectroscopy. The reaction-induced difference spectra obtained at different times after this concentration jump of a protein/GdmCl solution at pH 4.2 are shown in Fig. 3.23a.

These data clearly revealed two main events: (a) the loss of α-helical structure, indicated by a decrease in intensity at 1,653 cm^{-1}, and (b) the formation of antiparallel β-sheet structures, indicated by the emerging band components at 1,621 and 1,691 cm^{-1}. The strong splitting between the low- and high-frequency amide I components is indicative of the presence of intermolecular β-sheet structures with very strong hydrogen bonds. Similar features were found in the IR spectra of other ordered protein aggregates [80].

The kinetic curves obtained from the time evolution of the difference peaks at 1,621 and 1,653 cm^{-1} (Fig. 3.23b) show the increase in intermolecular β-sheet structure synchronous with the decrease of α-helical structure. The formation of ordered aggregates was verified by dynamic light scattering and electron microscopy, which revealed that the α-to-β transition is accompanied by a defined oligomerization of the prion protein leading primarily to octameric aggregate structures [78]. Long-term measurements after this initial appearance of oligomeric

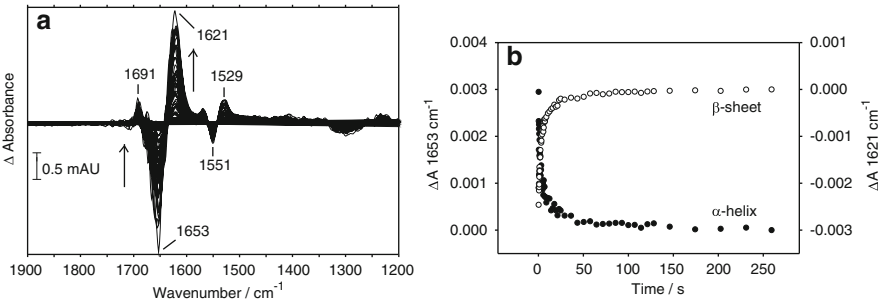

Fig. 3.23 The α-to-β conversion process of the prion protein $SHaPrP^{90–232}$ in H_2O solution monitored by stopped-flow FTIR difference spectroscopy in the time window from 190 ms to 260 s. (**a**) Sixty difference spectra obtained after 1:1 mixing of the prion protein with 2 M GdmCl at pH 4.2 (20 mM sodium acetate buffer). The final PrP and GdmCl concentration after mixing was 4 mg/ml and 1 M, respectively. All spectra were recorded continuously in the rapid-scan mode. The pathlength of the CaF_2 cell was approximately 8 μm. The difference spectra were calculated from single-beam spectra according to log (S_{last}/S_n), where S_n is the spectrum of the nth scan and S_{last} is the spectrum of the last scan. (**b**) The α-to-β conversion kinetics of the prion protein $SHaPrP^{90–232}$ were followed by using the amide I difference peak intensities at 1,653 cm^{-1} (α-helical structures) and at 1,621 cm^{-1} (intermolecular β-sheet structures). Changes are given with respect to the last spectrum obtained. Each data point represents the average of five individual measurements

structures over a time period of several days indicated that rearrangement of these oligomeric structures to protofibrillar structures is not accompanied by significant changes of secondary structure [78].

Summarizing, the data indicated that the use of FTIR stopped-flow difference spectroscopy is a very appropriate technical approach to elucidate in vitro the α-to-β conversion process of recombinant PrP. The IR spectroscopic approach is applicable to protein/H_2O solution at protein concentrations relatively low for FTIR experiments and should be useful in future for the investigations of other amyloidonic proteins as well. The use of the micro stopped-flow device shown in Fig. 3.8 allows to monitor the α-to-β conversion with sufficiently high time resolution and excellent reproducibility. The advantage of IR spectroscopy in following the α-to-β-conversion is that bands specific for α-helices, β-sheets, turns, and a number of amino acid side chains are well separated. Thus, changes in these types of structures can be detected independently.

Acknowledgements We thank our former PhD students Diane Reinstädler, Ralf Moritz, and Fabian Sokolowski for the data used in some of the figures herein, and Klaus Gast for longstanding fruitful collaboration. Moreover, financial support from the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

References

1. A. Elliot, E.J. Ambrose, Nature **165**, 921 (1950)
2. T. Miyazawa, in *Poly-α-Amino Acids*, ed. by G.D. Fasman (Marcel Dekker, New York, 1967) p. 69

3. H. Susi, Methods Enzymol. **22**, 455 (1972)
4. S. Krimm, J. Bandekar, Adv. Protein Chem. **38**, 181 (1986)
5. W.K. Surewicz, H.H. Mantsch, D. Chapman, Biochemistry **32**, 389 (1993)
6. J.L.R. Arrondo, A. Muga, J. Castresana, F.M. Goni, Prog. Biophys. Mol. Biol. **59**, 23 (1993)
7. M. Jackson, H.H. Mantsch, Crit. Rev. Biochem. Mol. Biol. **30**, 95 (1995)
8. S. Mukherjee, P. Chowdhury, F. Gai, J. Phys. Chem B **111**, 4596 (2007)
9. S.Yu. Venyaminov, F.G. Prendergast, Anal. Biochem. **248**, 234 (1997)
10. H. Fabian, K. Gast, M. Laue, R. Misselwitz, B. Uchanska-Ziegler, A. Ziegler, D. Naumann, Biochemistry **47**, 6895 (2008)
11. E. Goormaghtigh, V. Cabiaux, J.M. Ruysschaert, Subcell. Biochem. **23**, 329 (1994)
12. T. Heimburg, D. Marsh, Biophys. J. **65**, 2408 (1993)
13. N. Methot, C.N. Demers, J.E. Baenziger, Biochemistry **34**, 15142 (1995)
14. J. Backmann, C. Schultz, H. Fabian, U. Hahn, W. Saenger, D. Naumann, Proteins: Struct. Funct. Genet. **24**, 379 (1996)
15. H.H.J. de Jongh, E. Goormaghtigh, J.M. Ruysschaert, Biochemistry **36**, 13603 (1997)
16. W. Dzwolak, M. Kato, A. Shimizu, Y. Taniguchi, Biochim. Biophys. Acta **1443**, 45 (1999)
17. H. Fabian, H. Huser, D. Narzi, R. Misselwitz, B. Loll, A. Ziegler, R.A. Böckmann, B. Uchanska-Ziegler, D. Naumann, J. Mol. Biol. **376**, 798 (2008)
18. E. Dzafic, O. Klein, E. Screpanti, C. Hunte, W. Mäntele, Spectrochim. Acta A **72**, 102 (2009)
19. H. Fabian, D. Naumann, Methods **34**, 28 (2004)
20. D. Reinstädler, H. Fabian, D. Naumann, Biochemistry **35**, 15822 (1996)
21. A.J. White, K. Drabble, C. Wharton, Biochem. J. **306**, 843 (1995)
22. J. Tang, F. Gai, Appl. Spectrosc. **60**, 1477 (2006)
23. Q. Xu, T.A. Keiderling, Biochemistry **44**, 7976 (2005)
24. R. Masuch, D.A. Moss, Appl. Spectrosc. **57**, 1407 (2003)
25. B.A. Shirley, P. Stanssens, U. Hahn, C.N. Pace, Biochemistry **31**, 725 (1992)
26. J. Martinez-Oyanedel, C. Hui-Woog, U. Heinemann, W. Saenger, J. Mol. Biol. **222**, 335 (1991)
27. T. Kiefhaber, H.P. Grunert, U. Hahn, F.X. Schmid, Proteins **12**, 171 (1992)
28. J. Balbach, C. Steegborn, T. Schindler, F.X. Schmid, J. Mol. Biol. **285**, 829 (1999)
29. D. Reinstädler, H. Fabian, D. Naumann, Proteins **34**, 303 (1999)
30. R. Moritz, D. Reinstädler, H. Fabian, D. Naumann, Biopolymers (Biospectroscopy) **67**, 145 (2002)
31. H. Fabian, C. Schultz, D. Naumann, O. Landt, U. Hahn, W. Saenger, J. Mol. Biol. **232**, 967 (1993)
32. Y.N. Chirgadze, O.V. Fedorow, N.P. Trushina, Biopolymers **14**, 679 (1975)
33. D. Reinstädler, Dissertation, Freie Universität Berlin (1998)
34. T. Kiefhaber, R. Quaas, U. Hahn, F.X. Schmid, Biochemistry **29**, 3053 (1990)
35. T. Kiefhaber, R. Quaas, U. Hahn, F.X. Schmid, Biochemistry **29**, 3061 (1990)
36. T. Kiefhaber, H.P. Grunert, U. Hahn, F.X. Schmid, Biochemistry **29**, 6475 (1990)
37. T. Kiefhaber, H.-H. Kohler, F.X. Schmid, J. Mol. Biol. **224**, 217 (1992)
38. T. Kiefhaber, F.X. Schmid, K. Willaert, Y. Engelborghs, A. Chaffotte, Protein Sci. **1**, 1162 (1992)
39. V.V. Filimonov, V.V. Rogov, J. Mol. Biol. **255**, 767 (1996)
40. H. Fabian, K. Fälber, K. Gast, D. Reinstädler, V.V. Rogov, D. Naumann, D.F. Zamyatkin, V.V. Rogov, Biochemistry **38**, 5633 (1999)
41. H. Fabian, H.H. Mantsch, C. Schultz, Proc. Natl. Acad. Sci. U.S.A. **96**, 13153 (1999)
42. H. Fabian, K. Gast, V.V. Filimonov, D.F. Zamyatkin, V.V. Rogov, Vibr. Spectrosc. **38**, 45 (2005)
43. H. Fabian, Z. Wu, Y. Wang, D. Naumann, J. Mol. Struct. **974**, 203 (2010)
44. V. Forge, R.T. Wijesinha, J. Balbach, K. Brew, C.V. Robinson, C. Redfield, C.M. Dobson, J. Mol. Biol. **288**, 673 (1999)
45. B.A. Schulman, P.S. Kim, C.M. Dobson, C. Redfield, Nat. Struct. Biol. **4**, 630 (1997)
46. J. Balbach, V. Forge, N.A.J. van Nuland, S.L. Winder, P.J. Hore, C.M. Dobson, Nat. Struct. Biol. **2**, 865 (1995)
47. C.J. Tsai, B. Ma, R. Nussinov, Proc. Natl. Acad. Sci. U.S.A. **96**, 9970 (1999)

48. A.T. Alexandrescu, P.A. Evans, M. Pitkeathly, J. Baum, C.M. Dobson, Biochemistry **32**, 1701 (1993)
49. L.C. Wu, Z.Y. Peng, P.S. Kim, Nat. Struct. Biol. **2**, 281 (1995)
50. A. Troullier, D. Reinstädler, Y. Dupont, D. Naumann, V. Forge, Nat. Struct. Biol. **7**, 78 (2000)
51. M. Kataoka, K. Kuwajima, F. Tokunaga, Y. Goto, Protein Sci. **6**, 422 (1997)
52. D. Reinstädler, H. Fabian, J. Backmann, D. Naumann, Biochemistry **35**, 15822 (1996)
53. P.J. Bjorkman, M.A. Saper, B. Samraoui, W.S. Bennet, J.L. Strominger, D.C. Wiley, Nature **329**, 506 (1987)
54. F. Gejyo, T. Yamada, S. Odani, Y. Nakagawa, M. Arakawa, T. Kunitomo, H. Kataoka, M. Suzuki, Y. Hirasawa, T. Shirahama, A.S. Cohen, K. Schmid, Biochem. Biophys. Res. Commun. **129**, 701 (1985)
55. G.W. Platt, S.E. Radford, FEBS Lett. **583**, 2623 (2009)
56. N.M. Kad, N.H. Thomson, D.P. Smith, D.A. Smith, S.E. Radford, J. Mol. Biol. **313**, 559 (2001)
57. J. Kardos, D. Okuno, T. Kawai, Y. Hagihara, N. Yumpto, T. Kitagawa, P. Zavodszky, H. Naiki, Y. Goto, Biochim. Biophys. Acta **1753**, 108 (2005)
58. A.M. Smith, T.R. Jahn, A.E. Ashcroft, S.E. Radford, J. Mol. Biol. **364**, 9 (2006)
59. D.P. Hong, M. Gozu, K. Hasegawa, H. Naiki, Y. Goto, J. Biol. Chem. **277**, 21554 (2002)
60. S.L. Myers, S. Jones, T.R. Jahn, I.J. Morten, G.A. Tennent, E.W. Hewitt, S.E. Radford, Biochemistry **45**, 2311 (2006)
61. H. Fabian, W. Mäntele, in *Handbook of Vibrational Spectroscopy*, ed. by J.M. Chalmers, P.R. Griffiths (Wiley, Chichester, 2002), p. 3399
62. C.L. Ladner, M. Chen, D.P. Smith, G.W. Platt, S.E. Radford, R. Langen, J. Biol. Chem. **285**, 17137 (2010)
63. G.T. Debelouchina, G.W. Platt, M.J. Bayro, S.E. Radford, R.G. Griffin, J. Am. Chem. Soc. **132**, 17077 (2010)
64. M. Bouchard, J. Zurdo, E.I. Nettleton, C.M. Dobson, C.V. Robinson, Protein Sci. **9**, 1960 (2000)
65. J. Zurdo, J.I. Guijarro, J.L. Jiminez, H.R. Saibil, C.M. Dobson, J. Mol. Biol. **311**, 325 (2001)
66. J. Zurdo, J.I. Guijarro, C.M. Dobson, J. Am. Chem. Soc. **123**, 8141 (2001)
67. E. Frare, M.F. Mossuto, P. Polverino de Laureto, M. Dumoulin, C.M. Dobson, A. Fontana, J. Mol. Biol. **361**, 551 (2006)
68. Y. Cordeiro, J. Kraineva, M.C. Suarez, A.G. Tempesta, J.W. Kelly, J.L. Silva, R. Winter, D. Foguel, Biophys. J. **91**, 957 (2006)
69. V.N. Uversky, J. Li, A.L. Fink, J. Biol. Chem. **276**, 10737 (2001)
70. R. Sarroukh, E. Cerf, S. Derclaye, Y.F. Dufrene, E. Goormaghtigh, J.M. Ruysschaert, V. Raussens, Cell. Mol. Life Sci. **68**(8):1429–1438 (2011)
71. K. Janek, J. Behlke, J. Zipper, H. Fabian, Y. Georgalis, M. Beyermann, M. Bienert, E. Krause, Biochemistry **38**, 8246 (1999)
72. K. Yamaguchi, S. Takahashi, T. Kawai, H. Naiki, Y. Goto, J. Mol. Biol. **352**, 952 (2005)
73. N. Makarava, I.V. Baskakov, J. Biol. Chem. **283**, 15988 (2008)
74. O.V. Bocharova, L. Breydo, A.S. Parfenov, V.V. Salnikov, I.V. Baskakov, J. Mol. Biol. **346**, 645 (2005)
75. S.B. Prusiner, Science **216**, 136 (1982)
76. S.B. Prusiner, Proc. Natl. Acad. Sci. U.S.A. **95**, 13363 (1998)
77. T. Wisniewski, P. Aucouturier, C. Soto, B. Frangione, Amyloid **5**, 212 (1998)
78. F. Sokolowski, A.J. Modler, R. Masuch, D. Zirwer, M. Baier, G. Lutsch, D.A. Moss, K. Gast, D. Naumann, J. Biol. Chem. **278**, 40481 (2003)
79. A.J. Modler, H. Fabian, F. Sokolowski, G. Lutsch, K. Gast, G. Damaschun, Amyloid **11**, 215 (2004)
80. E.M. Jones, W.K. Surewicz, Cell **121**, 63 (2005)
81. N. Makarava, I.V. Baskakov, J. Biol. Chem. **283**, 15988 (2008)
82. J. Ollesch, E. Künnemann, R. Glockshuber, K. Gerwert, Appl. Spectrosc. **61**, 1025 (2007)

Chapter 4
Watching Dynamical Events in Protein Folding in the Time Domain from Submilliseconds to Seconds: Continuous-Flow Rapid-Mixing Infrared Spectroscopy

Satoshi Takahashi and Tetsunari Kimura

Abstract A continuous-flow rapid-mixing infrared spectrometer was developed to investigate the kinetic processes of protein folding in the time domain from 100 μs to seconds. The design of the developed devices and the practical issues in applying the devices were discussed. The kinetic investigation for the folding of apomyoglobin and single-chain monellin suggested that the collapsed folding intermediates possess a hydrated and fluctuating core.

4.1 Introduction

Many dynamical events are involved in the protein folding process, in which the native conformation of proteins is selected from the ensemble of unfolded conformations. The native conformation is specific and determined by the primary sequence of amino acid residues [1, 2]. In contrast, unfolded proteins formed in the presence of high concentrations of denaturant behave as random coils, in which the sampling of dihedral angles of one peptide bond is independent from the angles of the neighboring peptide bonds [3, 4]. This interpretation further suggests that the number of possible conformations of an unfolded polypeptide is astronomical and that the search process of the unfolded proteins for the native conformation, if it occurs randomly, might require an extremely long time [5]. However, actual proteins can fold to native conformations within biologically relevant timescales. This observation implies that the dynamics of protein folding is organized such that the unfolded proteins can find the native conformation efficiently without entrapment to misfolded or aggregated species. To understand the efficiency of the

S. Takahashi (✉)
Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan
e-mail: st@tagen.tohoku.ac.jp

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*, Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_4,

conformational search of actual proteins, it is important to characterize details of the structural events of protein folding [6].

Infrared absorption spectroscopy has been used to elucidate the conformational and dynamical properties of proteins [7–9]. In particular, the amide I band of the peptide bond appearing at around $1,660\,\mathrm{cm}^{-1}$ is related directly to dihedral angles, hydration status, and hydrogen bonding of the peptide bonds. It can reveal detailed structural and environmental information of proteins. Furthermore, the method is useful to characterize proteins in solution without extensive data averaging. Accordingly, the method is suitable for the characterization of transient phenomena occurring in solution such as protein folding. The most general strategy to conduct time-resolved observation of protein folding is the rapid mixing of the unfolded protein solution with an excessive amount of the refolding buffer. Consequently, infrared spectroscopy combined with the rapid-mixing method can reveal direct information related to the dynamics of protein folding [10, 11].

For more than a decade, we have investigated protein folding dynamics by developing a continuous-flow rapid-mixing apparatus and combining the apparatus with various spectroscopic and scattering methods [12–14]. The developed apparatus achieved time resolution on the order of a hundred microseconds. The combination of the mixing apparatus with infrared absorption spectroscopy is the most important among our attempts; it entails various technical improvements in the spectrometer, the mixing apparatus, and the sample preparations [15–17]. In this chapter, we summarize the efforts of our group and others that have enabled rapid-mixing infrared spectroscopy. In addition, we provide insight into protein folding obtained through those efforts.

4.2 The Collapse and Search Mechanism of Protein Folding

4.2.1 The Protein Folding Mechanism Depends on the Chain Length

The mechanism of protein folding is dependent on the chain length. Small proteins, possessing chain length between ca. 50 and 100 residues, demonstrate perfect cooperativity in protein folding, in which only the fully unfolded state and the native state are involved [18–20]. Extensive investigation of the structure of the folding transition states based on Φ-value analysis suggested that the folding of small proteins occurs in the smooth energy landscape without forming pseudo-stabilized conformations [21–23]. In contrast, intermediate proteins possessing chain length between ca. one hundred and several hundred residues usually possess several intermediates in both equilibrium and kinetic conditions [18]. Large proteins, whose chain length is greater than several hundred residues, frequently form irreversible aggregates during the folding process and require cellular devices such as molecular chaperons for correct folding.

The origin of the difference in the folding mechanisms among proteins with different chain lengths remains unclear. One explanation suggests that the intermediate proteins comprise several domains; the formation of the intermediates is ascribed to the organization of some domains before the folding of the entire structure [21]. Another explanation includes the assumption that the intermediate formation is associated with the dominance of a certain intramolecular interaction, such as hydrophobic interaction, in a particular phase of protein folding [24]. A time-resolved investigation of the folding of the intermediate proteins, that is the structural characterization of the successive intermediates appearing in time series, would reveal how the intramolecular interactions are organized to construct the overall structure.

4.2.2 Kinetic Investigation of Protein Folding for Intermediate Proteins

Early investigations of the kinetic process of protein folding used stopped-flow devices that can mix two solutions within several milliseconds [25]. These investigations demonstrated that many proteins form kinetic intermediates during the mixing dead time of the device. The intermediates are generally collapsed, and show large amounts of secondary structures and tertiary contacts mainly in the core domain. Structural and kinetic properties of the kinetic intermediates are frequently identical to those of the static intermediates, termed molten globule states [26–28]. The observations suggest that the molten globule state is a general intermediate of protein folding [26]. To elucidate how and why the intermediates are formed, it is necessary to resolve the formation process of the intermediates in the time domain that is inaccessible by the stopped-flow devices.

We developed a continuous-flow rapid-mixing device and combined it with various spectroscopic and scattering techniques. To detect overall dimensions of the transient intermediates, we used small angle X-ray scattering (SAXS) [13]. We used circular dichroism (CD) spectroscopy to obtain rough estimates of the secondary structure content [12]. We characterized how the dimension and the secondary structure content change in the folding process of three proteins, cytochrome *c*, apomyoglobin (apoMb), and single-chain monellin (SMN) [12, 13, 29, 30]. The results are presented in Fig. 4.1, in which the ordinate and abscissa, respectively, represent the secondary structure contents and the radius of gyration in the various kinetic states of these proteins.

Although the three proteins that were selected possess distinct secondary structures and overall shapes in the native state (Fig. 4.1), they showed a strikingly similar folding mechanism. In case of cyt *c*, we detected a collapse as the initial event occurring within several hundreds of microseconds, leading the unfolded protein to the initial intermediate with a small amount of the secondary structures. The second intermediate was subsequently formed with a time constant of several milliseconds, which is similarly collapsed and which possesses an increased amount

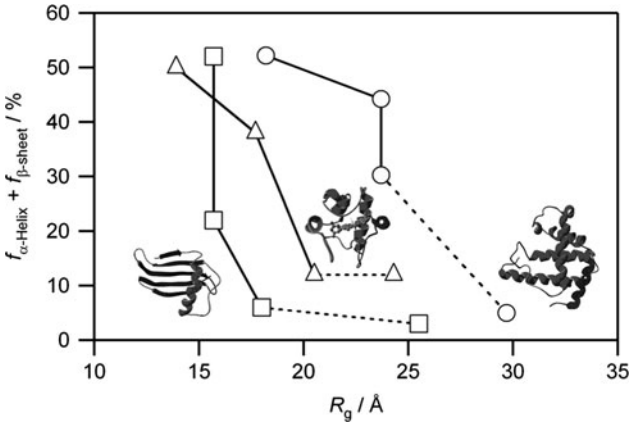

Fig. 4.1 Comparison of the folding trajectories for cytochrome *c* (cyt *c*), apomyoglobin (apoMb), and single-chain monellin (SMN) in the two-dimensional conformational space defined by secondary structural content (sum of the helix and β-sheet contents) deduced from the circular dichroism spectroscopy and the radius of gyration (R_g) obtained from the small angle X-ray scattering. The *circles, triangles*, and *squares*, respectively, correspond to the folding trajectories for apoMb, cyt *c*, and SMN [12–14, 30]. Modified from the original figure in [30]

of the secondary structures. Finally, the conversion of the second intermediate to the native state was observed. The mechanism, in which the rapid and significant collapse of the main chain is followed by the stepwise formation of the secondary and tertiary structures leading to the native conformations, was commonly observed for apoMb [29] and for SMN [30]. We designated these observations as the "collapse and search" mechanism of protein folding [29]. Although the time-resolved investigations described above clarified changes in the overall shape and secondary structure of the proteins, details of the events remain to be clarified. In particular, CD spectroscopy yielded rather good estimates for the secondary structure contents, but detailed information related to the conformations and environments of the protein backbone were lacking. We therefore initiated kinetic infrared investigations of protein folding.

4.3 Development of Continuous-Flow Time-Resolved Infrared Spectrometer

4.3.1 Comparison of Different Methods for Triggering Protein Folding Events

To detect kinetic changes in the conformations of proteins based on spectroscopic methods, it is necessary to initiate the structural changes of all molecules in the observation volume simultaneously. Various laser-triggering methods were

proposed to investigate the time domain shorter than several milliseconds, the limit attainable by the stopped-flow apparatus. The most fruitful among these methods is the laser-induced temperature jump [31–34] (see also Chap. 6). Using pulsed infrared lasers, a rapid increase in the sample temperature as large as ca. 10 °C can be induced within 10 ns. The sample response, such as helix melting, can be monitored after the jump. A disadvantage of the method is its difficulty in characterizing the folding intermediates that are usually populated in the forward folding process. In the unfolding process after the temperature rise, because of the faster unfolding rate of the intermediate than that of the native state, the intermediates do not usually accumulate.

Several other methods were similarly developed to initiate the forward folding process using pulsed lasers [35–38]. For example, a photocleavable polar group was labeled to hydrophobic residues of a protein to stabilize the unfolded state [37]. The folding processes were initiated by the photochemical disruption of the labeled groups. A disadvantage of the method is a moderate quantum yield of the triggering reactions, causing only a part of the molecules in the observation volume to be photocleaved. Accordingly, the response of proteins after laser excitation is frequently the changes in only a fractional population of the molecules. In addition, the labeled groups sometimes show spectral changes that overlap the region important for the interpretation of the structural changes. Here, it is necessary to use spectroscopic methods that can eliminate contributions from nonphotolyzed samples and from the labeled groups.

In contrast to the methods that depend on the laser triggering, solution mixing is superior in terms of several aspects. The latter method can initiate the forward folding reaction of all molecules in the observation volume. By selecting a proper method of protein unfolding, such as a pH jump, the spectral contribution from the components other than sample protein can be minimized. Accordingly, the method enables the detection of "pure" spectroscopic signals after rapid changes in the solution condition. A disadvantage of the solution mixing method was its difficulty in obtaining faster time resolutions. Several recent designs of devices now enable the mixing of two solutions within several tens of microseconds [39]. We describe the details of these mixing devices below.

4.3.2 Development of a Continuous-Flow Cell with a T-Shaped Flow Channel

The continuous-flow rapid-mixing method was an early experimental protocol used in the kinetic observation of chemical reactions. In 1923, the method was first used in the investigation of the binding process of oxygen to hemoglobin by Hartridge and Roughton [40]. The method was then replaced with the stopped-flow method, which can drastically reduce the amount of the sample consumption compared to that of the continuous-flow method [41]. However, the mixing dead time of the current stopped-flow devices, several milliseconds, is close to the practical

limit attainable by the device. A breakthrough in the time resolution of the rapid-mixing experiments was realized by Regenfuss et al. [42, 43]. They returned to the continuous-flow protocol and introduced a narrow gap (ca. 10 μm) by using a small ball placed in a glass capillary to create turbulence at a relatively low speed of flow (Fig. 4.2a). The device achieved mixing times of ca. 50 μs, which is more than two orders of magnitude faster than that of the stopped-flow device. The device, a radical departure from the stopped-flow protocol, became the prototype for all later designs of the turbulent mixing devices by other researchers.

It was almost a decade later that the turbulent mixing method was used by other researchers. Takahashi and Rousseau used a narrow mixing gap with diameter of

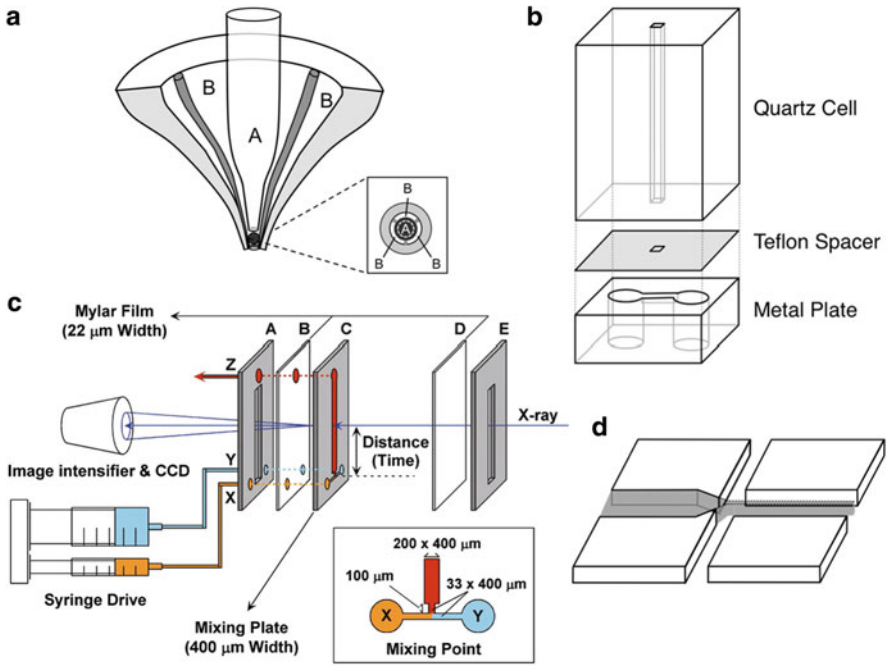

Fig. 4.2 Comparison of different designs of the continuous-flow rapid-mixing cells. (**a**) The original ball-type turbulent mixer developed by Regenfuss et al. [42]. Two solutions supplied from coaxial tubes (A and B) form turbulence at the interface between the ball and capillary wall. *Inset* is the slice view of the exit of the mixer. The ball with a diameter of 50–100 μm is held by three glass rods whose diameter is 10 μm. Redrawn from the original figure in [42]. (**b**) The T-shaped turbulent mixer connected to quartz cell for resonance Raman measurements [48]. Two solutions are supplied from two holes of a metal plate from the *bottom*, flow a thin channel (100 μm wide and 25 μm depth), and enter an observation cell. Redrawn from the original figure in [48]. (**c**) The T-shaped mixing channel sandwiching the mixing plate with window materials developed for small angle X-ray scattering measurements [13]. Adopted from [13]. (**d**) The diffusive mixing channel made on a silicon wafer [50]. A sample solution was supplied from the left channel, and the solutions to be mixed are supplied from the *top* and *bottom channels*. The thickness of the sample flow can be reduced to less than 1 μm by hydrodynamic focusing. The depth and width of the channels are both 10 μm. Redrawn from the original figure in [50]

less than 50 μm and length of 100 μm to mix two solutions and to create a free running jet, achieving mixing time of several tens of microseconds [44, 45]. The design of the mixing cell having no moving components was pursued further in the subsequent investigations by Rousseau and his collaborators [48, 49]. They used a "T-shaped" mixing channel and connected the exit of the channel to the entrance of the quartz flow cell (Fig. 4.2b). The design can mix two solutions within several hundreds of microseconds, and was used for resonance Raman investigation of the folding reaction of cytochrome *c* [48, 49]. In contrast, Roder and his collaborators used the ball-type mixer and connected it to a quartz observation cell [46, 47]. They used the device for investigation of the protein folding dynamics based on fluorescence detection. The ball-type design is effective for creating the turbulent flow and achieving the short mixing time. However, the performance of the mixer is critically dependent on the location of the ball in the channel, which must be adjusted carefully. In contrast, the design without moving components enables us the easy reproduction of the mixing time, although the mixing time is somewhat longer than that achievable by the ball-type mixer.

Takahashi et al. further improved the T-shaped design without moving components so that the device could be incorporated into various spectroscopic and scattering techniques [13]. The central component was a stainless mixing plate in which the T-shaped slit was carved (Fig. 4.2c). By sandwiching the plate between two windows, the mixing and observation channels can be constructed in a single device. For X-ray scattering measurements, mylar film was used as window material. For infrared absorption spectroscopy, calcium fluoride was used [15]. The mixing time of the device can be reduced to less than one hundred microseconds. Due to the reproducibility of the mixing time, the device can be easily disassembled, cleaned, and used after reassembly without a calibration of the mixing time.

Another strategy of mixing solutions within a short period is the diffusive mixing device [50]. In the first report, Knight et al. constructed a "crossroad-shaped" channel and injected a sample solution from the left inlet, and a solution to be mixed from the top and bottom inlets (Fig. 4.2d). The three streams of the solutions form a lamellar flow in the observation channel on the right. The sample flow thickness can be adjusted to less than a micrometer by hydrodynamic focusing, and the molecules from the side flows can diffuse into the sample flow with the shortest duration of ca. 10 μs. The device, first used for the time-resolved X-ray scattering experiments [51], was later applied for the kinetic infrared study of protein unfolding processes [52]. Currently, the diffusive mixer is frequently used in single-molecule fluorescence detection [53]. However, the diffusive mixer might not be best suited for kinetic infrared spectroscopic measurements, because the thin sample flow required for the faster mixing makes it difficult to acquire infrared high-quality spectra.

Several recent investigations have aimed at reducing the mixing time of the continuous-flow method. One design realized by our group used "premixing gaps" for the initiation of turbulence in the two parent solutions before their introduction to the mixing point, achieving a mixing time of ca. 11 μs [39]. The mixing channel was constructed on a stainless plate by using a laser ablation technique. The other

design used three repeats of an alcove and achieved a mixing time of ca. 22 μs [54]. Here, the channel was constructed on silicon by using photolithographic techniques. These results suggest that the mixing time can be further shortened by improving the channel design and by reducing the size of mixing channel based on the technologies to construct microfluidic devices.

In summary, the continuous-flow rapid-mixing method presents a practical approach for kinetic investigations of protein folding in the time domain of several tens of microseconds. In particular, the mixing device based on the T-shaped mixing channel sandwiching the mixing plate with window materials is of most practical use because of its reproducibility and ease of handling.

4.3.3 Construction of the Time-Resolved Spectrometer Based on Infrared Microscopy

Time-resolved spectrometers based on the continuous-flow rapid-mixing method are surprisingly simple compared to their stopped-flow counterparts. For the stopped-flow method, the synchronization of stopped-flow mixing and scanning of the interferometer requires a rather complex electronic configuration. Careful mechanical insulation of the mixing device from the spectrometer is also necessary. In contrast, the continuous-flow method can be constructed easily merely by installing the mixing device, which causes no mechanical vibration, inside conventional spectrometers with no additional modification of the electronic configuration.

The only practical consideration when installing the rapid-mixing device into the spectrometer is how to focus the infrared beam into the narrow flow channel. We used a Cassegrain-type infrared microscope (UMA-500; Bio-Rad Laboratories Inc.) connected to the Michelson interference spectrometer (FTS-575C; Bio-Rad Laboratories Inc.), and placed the T-shaped mixing cell on the sample stage of the microscope (Fig. 4.3). The use of the microscopic system is advantageous because the optical alignment can be confirmed easily by visualizing the image of the flow cell. Time-dependent changes in the folding reaction were detected by adjusting the distance between the mixing point of the mixer-flow cell assembly and the focusing point of the infrared light. We used a mercury cadmium telluride (MCT) detector for infrared detection. Although it is sensitive, the MCT detector has a nonlinear response to the infrared light intensity, which sometimes degrades the spectra after subtraction of the baseline and background lines of moisture. We introduced an optical filter before the sample stage, which cuts the infrared light over 2,300 cm^{-1} and maintains the intensity of the light within the linear response regime of the MCT detector.

The quality of the spectrum obtained by our kinetic spectrometer is sufficient for detailed investigations of protein folding. At the typical protein concentration of ca. 2 mg/ml dissolved in D_2O solution and with the optical path length of 100 μm, the optical density of the amide I band is about 0.08. With about 60 s of accumulated

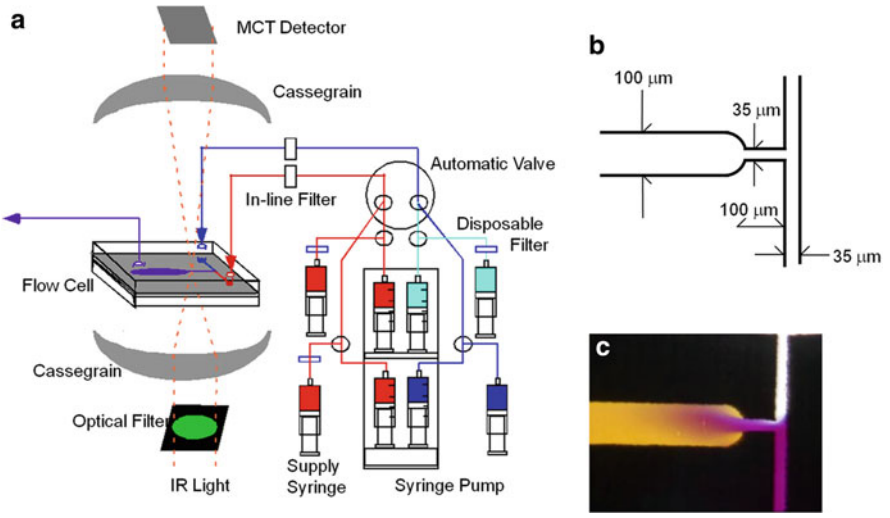

Fig. 4.3 (**a**) Schematic drawing of the system for time-resolved infrared measurements based on the rapid solution mixer. Solution mixing is performed using the T-shaped mixing plate in which the two solutions are introduced continuously and mixed. The infrared spectra are observed by focusing the infrared light on the observation channel of the mixing plate. The time-resolved measurements are performed by varying the distance between the mixing point and the observation point or by varying the flow velocity. (**b**) The detailed representation of the mixing point. The protein solution and refolding buffer are supplied from the *top* and *bottom* channels, and then flow through the observation channel. (**c**) The complete mixing was confirmed using the discoloration reaction of a dye solution. After mixing of sodium acetate (flowing from the *top*) and *bromocresol purple* (flowing from the *bottom*), the solution in the observation channel turns yellow within 100 μs

data, noise at around 1,600 cm^{-1} can be reduced to less than 0.0005 in optical density, which is sufficient for analysis of the main chain structures. To cancel the contributions of the water vapor lines, which may change sensitively according to small fluctuations of temperature and pressure, we used a tandem syringe pump that can switch the sample and reference flows without intervention (Fig. 4.3). Using the tandem pump, we observed the sample and buffer spectra almost simultaneously and obtained the "clean" difference spectra. Additionally, we covered the entire microscope with a glove box filled with dry air, and controlled the sample flow using an automated valve installed inside the glove box. We usually applied a solution flow of 0.17 ml/s to detect the fastest point that corresponds to 150 μs after the mixing. In this way, the sample consumption after 60 s of data accumulation is about 10 ml. Consequently, the current system can acquire time-resolved data with a high signal-to-noise ratio while consuming a reasonably small amount of sample.

It is worth noting here that the time range after several milliseconds cannot be followed by the above continuous-flow mixers because the turbulence required for the mixing cannot be generated at the low flow speed. Events in the millisecond

time range can be investigated using stopped-flow approaches, but the use of continuous-flow method is sometimes advantageous for infrared spectroscopy because the method generates no mechanical vibration during its operation. To cover a larger time range, several static mixers can be connected to the flow observation cell of the infrared spectrometer. In our group, we developed a continuous-flow mixing cell to cover the time range from 10 to 25 ms in collaboration with Unisoku (Hirakata, Japan) (Fig. 4.4) [16]. To cover the time range between several tens of milliseconds to more than seconds, we used a nanomixer designed by Upchurch Scientific Inc. (Oak Harbor, WA) [55] and connected it to the flow channel for the infrared detection. The use of several mixing devices enabled us to monitor the entire time range of protein folding from several tens of microseconds to several seconds, based on the continuous-flow protocol.

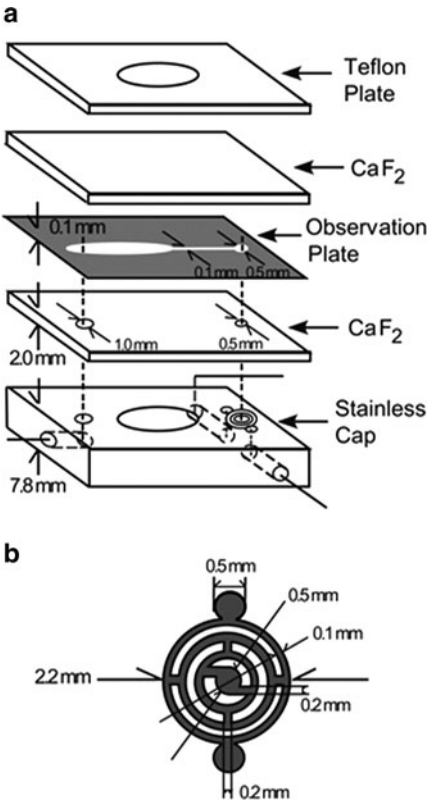

Fig. 4.4 (**a**) Schematic drawing of the continuous-flow mixing device for observations in the time domain from 10 to 25 ms. The mixing point was engraved at the stainless cap, which is pushed against the observation assembly by CaF_2 plates (thickness of 2 mm) and the stainless observation plate (thickness of 100 μm). IR light is focused on the flow channel (100 × 100 μm) of the observation plate. The time-resolved measurements are performed by varying the flow velocity. (**b**) Detailed representation of the mixing point: The two solutions of protein and refolding buffer are sent to the up and down channels. The two solutions are mixed at the *narrow* channels, and then sent to the observation plate though the center channel

4.4 Practical Issues for Kinetic Infrared Investigations of Protein Folding

4.4.1 Selection of the Initial Unfolded State

The initial state of proteins when performing kinetic folding experiments is frequently the unfolded state in the presence of high concentrations of denaturants [56]. The use of denaturants, however, is unsuitable for infrared measurements because common denaturants such as urea and guanidinium possess a strong absorbance in the amide I region. Although the absorption band can be shifted by isotopic substitution of the denaturants [57], it is not practical for continuous-flow experiments that consume a rather large amount of samples. Accordingly, we commonly use the pH-induced unfolded state as the initial state for kinetic infrared investigations.

However, not all proteins demonstrate pH unfolding [58]. In some cases, the conformational states formed at extremely low pH conditions are not the unfolded state but an intermediate state [59]. It is also worth noting that the structural property of the acid- or alkaline-unfolded states usually differs from that of the denaturant-unfolded states [60]. For example, the acid-unfolded state of apoMb is slightly structured compared to that of the urea-unfolded state [61]. Accordingly, in designing the kinetic infrared experiments, it is necessary to check carefully the reversible pH unfolding and the properties of the pH-unfolded state.

The use of rapid-mixing infrared spectroscopy for the investigation of protein folding is still limited. This can be attributed to the scarcity of proteins that demonstrate reversible pH unfolding transitions. The use of other unfolding methods that can unfold proteins without spectral contribution of the denaturant in the amide I region is strongly desired to expand the applicability of kinetic infrared spectroscopy. One of the candidates for such unfolding reagent is dimethylsulfoxide (DMSO), which is a strong denaturant and possesses no absorbance in the amide I region [62]. However, compared to urea or guanidinium, the property of DMSO as denaturant has not been fully understood. The development of a new unfolding technique is warranted for use in kinetic infrared investigation.

4.4.2 Suppression of the Aggregate Formation

Even if we can find a proper condition for the reversible pH unfolding, another requirement exists that the protein of interest must fulfill for the success of the kinetic infrared measurements. The infrared measurements require concentrated proteins, typically around 2 mg/ml. In such a condition, it is sometimes the case that the intermediates form dimeric or higher multimeric (aggregate) components during the kinetic folding process. In fact, all proteins we have examined to date formed aggregates during the initial trials of the pH-jump experiments; thus, a large

share of our experimental effort was directed to the survey of conditions that can avoid aggregate formation.

The order of the events in the process of the aggregate formation is the initial formation of the collapsed and monomeric intermediate from the unfolded state within several hundreds of microseconds, and its subsequent conversion to dimeric and multimeric conformations. Accordingly, a proper choice of the protein concentration, pH, and buffer composition in the solution after the mixing is important. The aggregates form more easily in the electronically neutral state. Consequently, pH near the isoelectric point should be avoided. The other choice is the use of a cosolvent that can attenuate the hydrophobic interactions. For example, the addition of 10% ethanol as a cosolvent increases the critical concentration of aggregate formation of the molten globule state of apoMb [63].

Regarding apoMb, another strategy was effective to eliminate the aggregate formation [29]. We presumed that a small amount of the damaged molecules present even in the purified sample might form a conformation that is prone to attract other molecules, and thereby becomes the "nucleus" of the aggregates. To prevent the formation of aggregates through this mechanism, we developed the double-jump technique. The apoMb solution was subjected to the first refolding pH jump without conducting spectroscopic measurements. The solution after the first jump was slightly turbid caused by the formation of a small amount of aggregates, which were removed by centrifugation. The major monomeric component was collected, concentrated, and then used in a second jump for the spectroscopic measurements, in which the aggregate formation is minimized. This method might be applicable for other protein systems as well.

4.4.3 Method of Spectral Analysis

Information related to the secondary structure contents and the environment around the amide backbone was obtained by analysis of the IR spectra in the amide I region. The method of data analysis obtained using the kinetic infrared spectroscopy is identical to that used for the other infrared applications. Determination of the number of the peaks, peak frequencies, widths, and relative area of components under the broad amide I band contour is the key procedure [7–9, 64–66]. Although several techniques such as resolution enhancement by Fourier self-deconvolution and Gaussian fitting have been widely accepted, the unambiguous determination of the fitting parameters always constitutes a difficult problem. Regarding kinetic spectroscopy, we can take advantage of the time-resolved spectra that are obtained at almost identical experimental conditions. Accordingly, analysis of all the time-resolved spectra based on consistent fitting parameters would reduce the ambiguity of the spectral analysis.

For analysis of the data obtained for apoMb [16], we first determined the number of the peaks buried in the amide I region of all the time-resolved spectra using a combination of Fourier self-deconvolution and second-derivative calculation.

Second, to fix the peak frequencies, the self-deconvoluted spectra were fitted with the Gaussians without restricting the peak frequencies. The most probable frequency was obtained by using averages of the frequencies of the corresponding peaks in the time-resolved spectra. Third, to determine the averaged width for each peak, the self-deconvoluted spectra were again fitted by Gaussians with fixed peak wavenumbers, but with unrestricted widths. Finally, the relative area of the peaks was estimated by fitting the original data using the Gaussians with fixed frequencies and widths.

For analysis of the data of single-chain monellin [17], we used a different approach. We first conducted a global fitting analysis of the entire time-resolved data and estimated the "pure" spectra for two kinetic intermediates. The resultant spectra were resolution-enhanced using Fourier self-deconvolution, and fitted by Gaussians, whose number and peak frequencies were determined using the fourth derivative of the original spectra.

The relative areas of the peaks estimated using the procedures described above should be considered as rough estimates of the secondary structure contents of proteins [67, 68]. Amide I mainly originates from the C=O stretching vibration of amides, whose intrinsic frequency is modulated by hydrogen bonding and an electrostatic potential [69–71]. Furthermore, the transition dipole coupling between the neighboring amides splits the amide I into several lines having different intensities [69, 72]. The assignments of the deconvoluted lines of amide I to the representative secondary structures are based on the accumulated knowledge of the spectra of proteins with known structures. With these limitations in mind, however, the band fitting analysis yields important information on protein folding dynamics.

4.5 Application to Protein Folding

4.5.1 Pioneering Investigations of Rapid-Mixing Infrared Spectroscopy

The application of the continuous-flow rapid-mixing method to kinetic infrared spectroscopy was first reported by Marinkovic et al. [73]. They designed a turbulent mixing cell with a dead time of ca. 0.5 ms, and combined it with a strong infrared beam generated at a synchrotron facility. They observed the conversion process of the acid-unfolded cyt *c* to the molten globule intermediate stabilized using 400-mM KCl. Kaufmann et al. developed a diffusive mixing cell and observed the unfolding process of β-lactoglobulin induced by mixing with trifluoroethanol (TFE) [52]. The characteristic time of the diffusion of TFE to the protein solution was 0.4 ms. A distinct intermediate with a loosely packed β-sheet structure was identified before the formation of the nonnative α-helix. Kimura et al. studied the helix formation process of polyglutamic acid after a rapid pH jump based on kinetic infrared and CD spectroscopic techniques [15]. They used the turbulent mixer with a T-shaped

mixing channel, and observed helix formation and elongation processes after the pH jump.

4.5.2 Apomyoglobin

4.5.2.1 Sequential Folding Scheme of Apomyoglobin

Apomyoglobin (apoMb), the apo form of myoglobin, comprises 153 residues. Myoglobin is a fully helical protein possessing eight helices termed A–H [74]. ApoMb retains the overall conformation of myoglobin except for a part of E helix and the terminal of H helix that are fluctuating in the native conformation [75]. ApoMb forms an acid-unfolded state at pH 2.0, which possesses ca. 5% of helices as judged from its CD spectrum in the far UV region. Some NMR investigations have shown that it is the A helix that forms a fluctuating helix [61, 76]. In contrast, the unfolded state prepared in the presence of a high concentration of urea possesses no secondary structures. The radius of gyration (R_g) of the acid-unfolded state is 29.7 Å, which is smaller than that of the urea-unfolded state (34.2 Å) [29, 77]. Consequently, the acid-unfolded state is slightly structured compared to the urea-unfolded state. In addition, apoMb forms an equilibrium intermediate at pH 4.2, which is one of the representative examples of the molten globule state [26]. Hydrogen–deuterium exchange experiments demonstrated that the intermediate possesses the partially formed helices A, G, and H [78]. Furthermore, there exist specific tertiary contacts among the three helices [79, 80]. A pulsed hydrogen–deuterium exchange experiment demonstrated that a conformation identical to that of the equilibrium intermediate at pH 4.2 is formed within the burst phase of the pH-jump experiments [81].

Several detailed investigations of the kinetic process of apoMb folding revealed that the folding intermediate is separable into two components [82, 83]. The first, formed at around pH 3.6, is the form that appears first after the pH jump from 2.0 to 6.0 [82]. The first intermediate is in equilibrium with the second intermediate, which is stable at pH 4.2 and which is the rate-limiting intermediate. The presence of those two intermediates is further supported in the kinetic CD and SAXS investigations, whose results are summarized in the two-dimensional plot for R_g and secondary structure content (Fig. 4.1) [29]. The first intermediate is formed within 300 μs after the pH jump. It is collapsed ($R_g = 23.7$ Å) and possesses partial helix contents (30%). The form converts to the next intermediate with a time constant of ca. 5 ms, and increases its secondary structure content to 44%.

Consequently, the folding of apoMb involves the acid-unfolded state, the two intermediates, and the native state. By conducting a time-resolved infrared investigation on the folding of apoMb by rapidly changing the solution condition from pD 2.2 to pD 6, we can expect to observe detailed changes in main chain conformations associated with the initial rapid collapse and the stepwise secondary structure formation.

4.5.2.2 Kinetic Infrared Spectra of Apomyoglobin Folding and Their Analysis Based on Band Fitting

We first examined the equilibrium unfolding transition of apoMb by obtaining amide I spectra at various conditions, and analyzed them based on fitting using seven Gauss functions assignable to turns (1,666/80/95 cm^{-1}), a buried helix (1,652 cm^{-1}), a disordered structure (1,644 cm^{-1}), a solvated helix (1,633 cm^{-1}), and a loop (1,619 cm^{-1}) (Fig. 4.5) [64–66]. The native protein at pD 6.0 possesses a sharp amide I centered at 1,650 cm^{-1} corresponding to the mostly helical conformation (Fig. 4.5a). The band fitting of the spectrum gave 39% of buried helix and 14% of solvated helix [84, 85], whose sum is in accord with the helix content estimated by CD spectroscopy (55%) [29]. The spectrum for the intermediate at pD 4.2 suggests that the amounts of buried and solvated helices are 20% and 24%, respectively, whose sum (44%) is somewhat larger than that estimated using CD spectroscopy (33%) (Fig. 4.5b) [29]. In contrast, the acid-unfolded state at pD 2.2 possesses a broad spectrum with a shoulder at 1,630–1, 640 cm^{-1}, which was shown

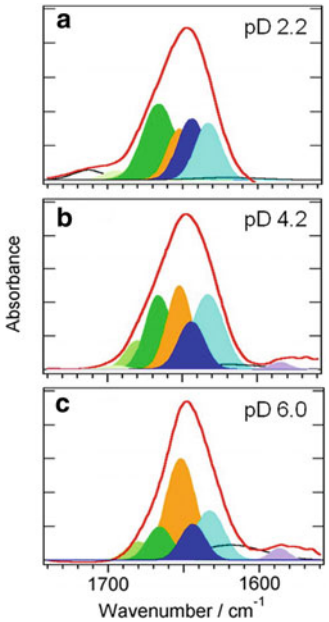

Fig. 4.5 Infrared absorption spectra in the amide I region for various equilibrium states of apoMb. The panels (**a**), (**b**), and (**c**) respectively present spectra obtained for samples in the acid-unfolded state at pD 2.2, in the intermediate state at pD 4.2, and in the native state at pD 6.0 (*red traces*). The spectrum at pD 2.2 was obtained in the absence of ethanol. The spectra at pD 4.2 and 6.0 were obtained in the presence of 10% (v/v) ethanol. The traces were fitted by Gauss functions for turns at 1,619 cm^{-1} (*black trace*), solvated helix at 1,633 cm^{-1} (*light blue*), disordered structure at 1,644 cm^{-1} (*navy blue*), buried helix at 1,652 cm^{-1} (*orange*), and turns at 1,666, 80, and 95 cm^{-1} (*dark green*, *green*, and *pale green*) [16]

to possess sizable amounts of buried (16%) and solvated (19%) helices (Fig. 4.5c). The IR spectrum of the urea-unfolded state (not shown), which is distinct from that of the acid-unfolded state, is broadly centered at around 1,645 cm^{-1}, suggesting the fluctuating and unfolded conformation.

The results presented above demonstrate the advantages and the possible shortcomings of the analyses of amide I based on the band fitting. The analysis largely reproduced the structural changes of apoMb obtained using different spectroscopic methods. For example, the content of the helical structures for the intermediate at pH 4.2 estimated from the CD spectroscopy, 33%, was ascribed to the helices A, G, H, and a part of helix B [78, 86], which is consistent with the current results (44%). In addition, the current results provided insight into the distinction between solvated and buried helices. In contrast, the analyses of the amide I spectra did not agree with the results of the other spectroscopic methods for the acid-unfolded state. As we discussed above, band fitting is a qualitative method of spectral analyses, established by comparing amide I spectra and three-dimensional structures of the folded state of many proteins. Accordingly, the results of band fitting used to analyze the conformations in the unfolded states should be interpreted with care.

We next conducted kinetic pD-jump experiments by changing the solution pD from 2.2 to 6.0 (Fig. 4.6), and analyzed the time-resolved spectra based on the fitting procedure described above. After changes in pD, the amide I spectra demonstrated a rapid and drastic change showing an increase in the amount of helix after the pD jump (Fig. 4.6a). There exists a shoulder at around 1,630–1,640 cm^{-1} for components obtained immediately after the pD jump, demonstrating the presence of a substantial amount of solvated helices. As time proceeds, the amide I shifts to higher wavenumbers, corresponding to the formation of the buried helix. These trends are reflected in the subtraction spectra (Fig. 4.6b) and in the amide I band fitting results (Fig. 4.7). The amount of the solvated helix increased significantly from 19% of the acid-unfolded state to 25% of the transient conformation at 100 μs. This amount remains almost constant in the time range from 100 μs to 5 ms, in which the first intermediate is mainly populated, and decreases slightly in the time range from 10 to 23 ms, in which the second intermediate is mainly populated. In contrast to the solvated helix, the content of the buried helix is low in the conformation immediately after the pD jump (19%). In the final phase occurring in the time domain longer than 10 ms, the increase in the buried helix and the decrease in the solvated helix were observed.

The kinetic infrared investigation of the folding of apoMb demonstrated clearly that a significant amount of solvated helix is accumulated immediately after the pD jump. The kinetic SAXS experiments for the apoMb folding demonstrated that the initial event after the pD jump is the collapse of the main chain [29]. The temporal coincidence of the two kinetic events suggests that the helix formation is associated with the collapse of the main chain, which should be promoted by hydrophobic interactions. However, the current result further demonstrated that the environments around the backbone amides remain hydrophilic after the collapse. Desolvation, the process that converts solvated helix to buried helix, occurs at the slowest kinetics of

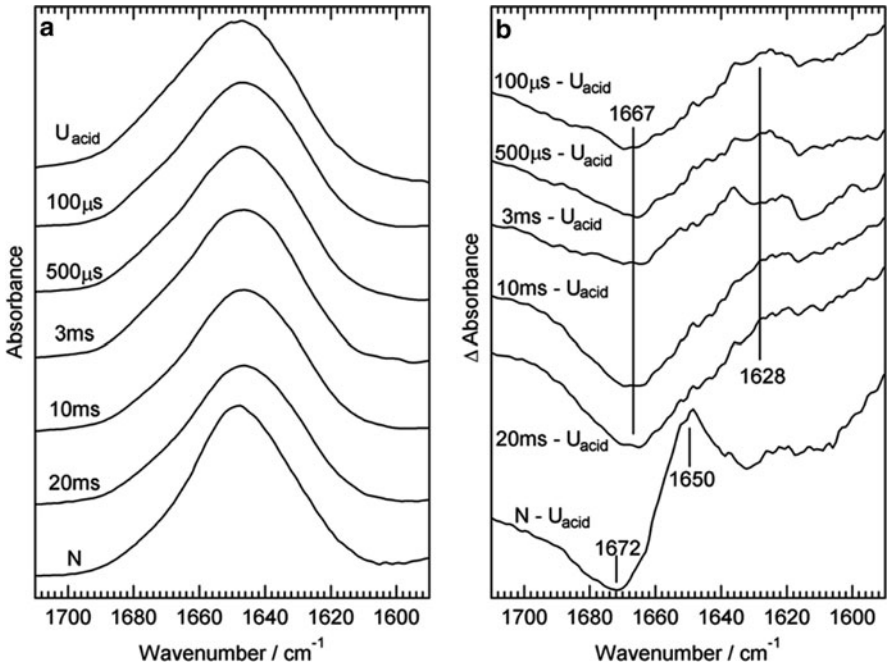

Fig. 4.6 Time-resolved infrared absorption spectra for the folding process of apoMb. (**a**) The representative spectra are presented as obtained at 100 μs, 500 μs, 3 ms, 10 ms, and 20 ms after initiating folding by the pD jump from 2.0 to 6.0. Static spectra corresponding to U_{acid} and N are also presented. (**b**) The kinetic difference spectra for the folding process of apoMb initiated by a pD jump from 2.0 to 6.0. The differences between the time-resolved data and the spectrum for the initial state was calculated. The difference spectrum between the spectra for the native state and the acid-unfolded state was also presented for comparison. Adapted from [16]

apoMb folding, suggesting that dehydration of the hydrophilic amide group might be the rate-determining step in protein folding.

4.5.3 Single-Chain Monellin

4.5.3.1 Folding Intermediates of Single-Chain Monellin

Single-chain monellin (SMN) is a β-sheet protein with 94 residues [87] comprising a five-stranded β-meander and an α-helix [88, 89]. SMN shows reversible alkaline unfolding. The folding dynamics of SMN initiated using a pH jump from 13.0 to 9.4 involves two kinetic intermediates, I_1 and I_2, which were proposed as connected between the unfolded state (U) and the native state (N) sequentially [30]. The kinetic CD and SAXS measurements demonstrated that I_1, formed within 300 μs after the pH jump, possesses a small amount of secondary structures with a collapsed oblate

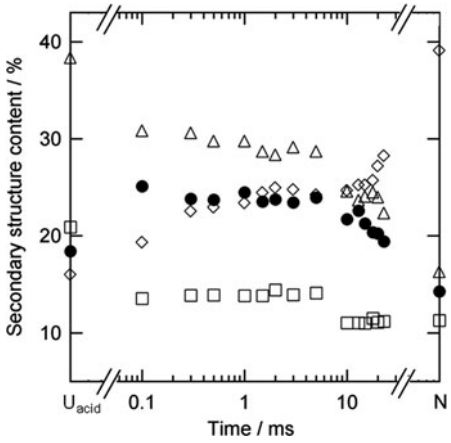

Fig. 4.7 Kinetic changes in the secondary structure contents of apoMb in its folding process after the pD jump. The time-resolved data were analyzed based on the band fitting procedure. *Filled circles, open squares, open diamonds*, and *open triangles* respectively correspond to the contents for the solvated helix at 1,633 cm^{-1}, the disordered structure at 1,644 cm^{-1}, the buried helix at 1,652 cm^{-1}, and the turns at 1,666, 80, and 95 cm^{-1}. Adapted from [16]

shape. I_1 converts to I_2 possessing an increased amount of secondary structures and a dimension that is indistinguishable from N with a time constant of 14 ms. The conversion of I_2 to N occurs with a time constant of 1.2 s. To explain the rapid construction of the oblate shape with a minimum amount of a detectable β-sheet, we proposed that the formation of β-turns and the hydrophobic collapse drive the rapid construction of I_1. Detailed analyses of the secondary structure content in each intermediate based on the kinetic infrared spectroscopy are expected to provide insight into the organization process of the β-sheet.

4.5.3.2 Kinetic Infrared Investigation of the Folding of Single-Chain Monellin

We conducted time-resolved infrared measurements based on the continuous-flow mixers to follow the folding of SMN upon a pD jump from 13.0 to 9.4 and obtained time-resolved spectra in the time domain from 100 μs to 10 s [17]. The time-resolved data possess no isosbestic points, suggesting the formation of the kinetic intermediates (Fig. 4.8a). To analyze the kinetic data quantitatively, we performed global fitting of the entire set of the kinetic IR spectra based on the sequential folding model, in which the intermediates are formed in the order of I_1, I_2, and N. The calculated spectra and the second derivatives for each kinetic component are presented respectively in Fig. 4.8b, c. The time-dependent changes at several wavenumbers are plotted in Fig. 4.8d. We obtained two rate constants of $10.0 \pm 0.7\ s^{-1}$ and $0.91 \pm 0.15\ s^{-1}$, respectively, for the first and second transitions.

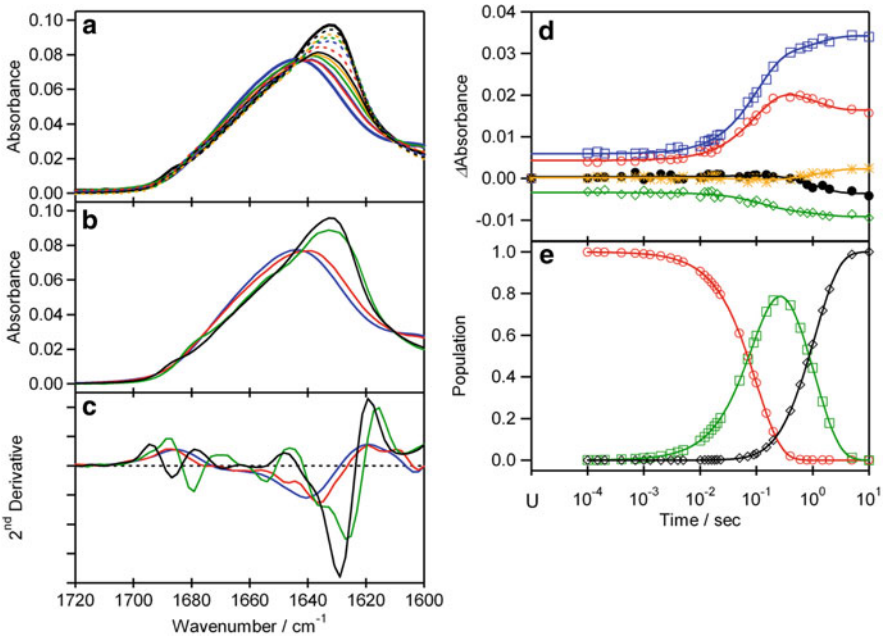

Fig. 4.8 (**a**) Time-resolved infrared spectra during the folding of SMN. Only 12 of 34 spectra are shown for clarity, corresponding to the initial unfolded state at pD 13.0 (*bold blue*), kinetic data at 150 μs (*red*), 12 ms (*blue*), 50 ms (*green*), 90 ms (*orange*), 100 ms (*black*), 200 ms (*dotted red*), 600 ms (*dotted blue*), 800 ms (*dotted green*), 2 s (*dotted orange*), 10 s (*dotted black*) after the pD jump, and the native state at pD 9.4 (*bold black*). (**b**) Infrared spectra for the kinetic components analyzed using the global fitting procedure based on a sequential folding scheme. Extracted spectra correspond to I_1 (*red*), I_2 (*green*), and N (*black*). The spectrum of the unfolded state at pD 13.0 (*blue*) is shown for comparison. (**c**) The second-derivative spectra of U (*blue*), I_1 (*red*), I_2 (*green*), and N (*black*). (**d**) Time courses of the changes in the difference absorbance at 1,621 (*red*), 1,629 (*blue*), 1,654 (*green*), 1,677 (*black*), and 1,687 cm^{-1} (*orange*). Continuous lines are obtained by fitting based on the double exponential with the rate constants of 10.0 and 0.91 s^{-1}. (**e**) Time courses of the population of each component; I_1, *red*; I_2, *green*; N, *black*. Modified from the original figures in [17]

The populations of the kinetic components are plotted in Fig. 4.8e. The success of the data analysis verified the sequential folding scheme with two kinetic components established earlier [30].

The spectrum of I_1 that accumulates from 100 μs to 10 ms is distinct from that of U. The conversion of U to I_1 corresponds to the burst phase, which is similarly observed in the kinetic SAXS and CD measurements. In the second-derivative spectrum of I_1, we identified a component assignable to a solvated helix (1,637 cm^{-1}; ca. 11% content) and weak negative features at 1,666 and 1,675 cm^{-1} assignable to turns. The presence of turns and absence of β-sheet in I_1 support our earlier proposal that the formation of I_1 might be driven by the turn formation. In addition, the presence of a significant amount of solvated helix suggests that

the N-terminal helix might be formed and become a part of the loosely assembled core in I_1. The second-derivative spectrum of I_2 shows a feature of solvated helix at 1,637 cm^{-1} (12%) and a split pattern of β-sheet at 1,625 cm^{-1} (34%) and 1,680 cm^{-1} (3%). Although the total content of solvated helix and α-helix (17%) is comparable to that of N (17%), the total content for β-sheet (37%) for I_2 is smaller than that of N (43%), indicating that the formation of β-sheet is not completed in I_2 and that additional stabilization of secondary structures is necessary to form N.

The major and minor amide I band components characteristic of β-sheet for I_2 and N might be interpreted in terms of the hydration status of the main chain amides. The two components in I_2 (1,625 and 1,680 cm^{-1}) had both shifted to lower frequencies than those of N (1,629 and 1,687 cm^{-1}). The average and difference of the major and minor amide I components correspond to the intrinsic frequency and the twice the dipolar coupling constant between the nearest C = O groups in the adjacent strands. The intrinsic frequency for β-sheet in I_2 is shifted to lower frequency by 4.6 cm^{-1}. We assigned the shifts to result from the hydration of C = O groups in the β-sheet for reasons similar to those proposed for solvated α-helix [84, 85]. A similar low-frequency shift of the intrinsic frequency for β-sheet was observed in the equilibrium intermediates of SMN, in which the shift in the peak of tryptophan fluorescence to longer wavelength was observed. Together with the presence of the significant amount of solvated helix in I_2, we proposed that the core domain formed in the interface between N-terminal helix and β-sheet might be hydrated.

By comparing the kinetic infrared results with the time-resolved SAXS, CD, and fluorescence data [30], detailed structural information for the kinetic intermediates is obtainable. The collapsed I_1 consists of the loosely assembled core and some specific structures such as N-terminal solvated helix and turns. These structures are expected to fluctuate, however, because a hydrophobic extrinsic dye cannot bind to I_1. In contrast, in I_2, a more structured core is suggested by the binding of the hydrophobic dye, which is consistent with the formation of solvated helices and solvated β-sheets, and with an overall oblate shape that is indistinguishable from N. The final process of the SMN folding is the conversion of the water-penetrated core to the dehydrated core and the additional stabilization of secondary structures.

4.6 Summary and Perspective

Time-resolved infrared spectroscopy based on continuous-flow rapid-mixing devices revealed detailed information related to the mechanism of secondary structure formation and the protein main chain dehydration in protein folding (Fig. 4.9). Comparison of data obtained using infrared spectroscopy and other techniques suggests that the arrangement of the ordered secondary structures occurs slowly after the main chain collapse and formation of the overall shape. Furthermore, these results demonstrate that the folded structures of proteins are constructed in several steps.

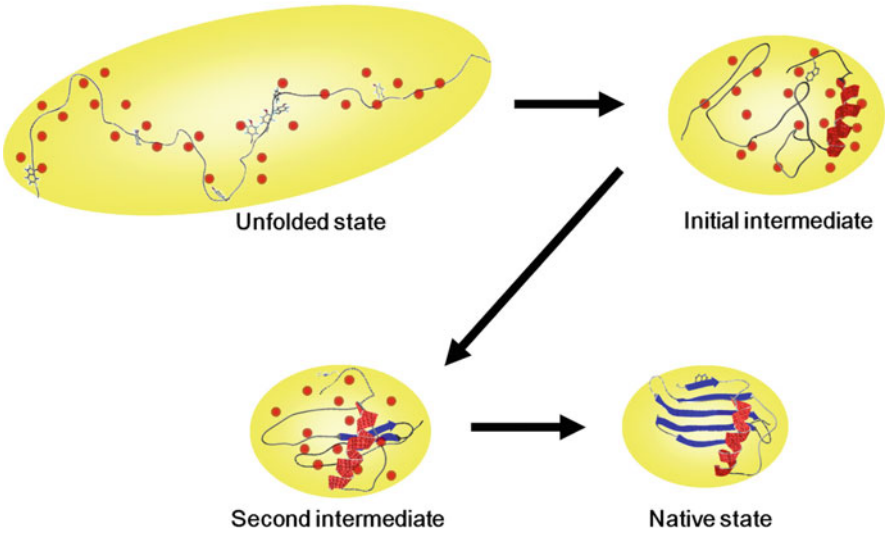

Fig. 4.9 Schematic representation of the mechanism of protein folding deduced from the time-resolved infrared spectroscopy based on rapid-mixing devices. The unfolded protein is expanded and fully hydrated. Red balls represent water molecules around the main chain. The unfolded state converts to the initial intermediate within several hundred of microseconds, which is collapsed and possesses partial secondary structures such as hydrated helix and turn. The second intermediate is more structured with significant amount of secondary and tertiary structures; however, a large fraction of the main chain is still hydrated. The formation of the native state from the second intermediate is associated with the desolvation of the main chain and is the rate-limiting step in protein folding

The initial event observed immediately after the initiation of the folding process of apoMb and SMN is the formation of solvated helices and turns (Fig. 4.9, upper right). For SMN, the split amide I lines characteristic of β-sheet were not detected in the initial intermediate. Both, turns and helices are stabilized by intramolecular interactions between the residues located within a short range along the polypeptide chains. In contrast, the formation of β-sheet structure requires hydrogen bonding between residues separated along the polypeptides, and might not be formed in the initial process. The time-resolved SAXS measurements showed that both apoMb and SMN are already collapsed in the initial intermediate [29, 30]. Consequently, the initial collapse, driven by hydrophobic interaction of the side chains, is concomitant with the formation of solvated helices and turns. We infer that the elements of the secondary structures stabilized by the local interactions might help the rapid collapse of the main chain and the formation of the core domain.

In contrast to the moderate content of the secondary structures in the initial intermediate, the secondary structure content increased significantly in the second intermediate to the level that resembles the static molten globule state (Fig. 4.9, lower left). In the case of apoMb, the conversion of the initial intermediate to the

second intermediate engenders the increase and decrease in the amounts of buried and solvated helices, respectively, suggesting partial desolvation of the interface of helices. For SMN, the conversion of I_1 to I_2 involves the appearance of the split amide I lines, demonstrating the formation of the hydrogen bonded network of the β-sheet. These observations support the earlier proposal that the kinetic phases occurring after the collapse of the main chain represent the search for the correct secondary and tertiary structures.

The noticeable difference between the second intermediate and the native state is in the solvation status of the main chain. The kinetic infrared spectroscopy demonstrated a low-frequency shift of amide I lines unequivocally for both helix and β-sheet in the second intermediate. The main chain amides are polar, forming a strong hydrogen bonding with water in the unfolded conformation. Accordingly, the embedding of the main chain amides inside the core of the native protein was explained to entail a large energetic cost [90,91]. The observations based on kinetic infrared spectroscopy underscore the importance of the energetic costs of main chain dehydration in the dynamics of protein folding.

A kinetic infrared investigation of the folding of the intermediate proteins demonstrated that the folding of this class of proteins occurs in multiple phases, each of which is likely dominated by intramolecular interactions of different types. We observed a rapid collapse that is concomitant with the formation of partial helices and turns, which is ascribed to the hydrophobic interactions and to the local interactions. The collapsed intermediate then turns to the second intermediate, in which the contacts between the side chains separated for longer distances along the primary sequences are formed. A partial expulsion of water is associated with conversion of the first intermediate to the second intermediate. Finally, the conversion of the second intermediate to the native structure occurs as a result of the complete expulsion of water from the core domain. Based on these observations, we suggest that the presence of folding intermediates in the intermediate proteins is associated with a larger contribution of the interactions promoting the collapse of protein main chain, such as the hydrophobic interaction and the turn formation, in the early folding phase.

Many questions must be addressed in future investigations of protein folding based on kinetic infrared spectroscopy. For example, it is now possible to decrease the mixing dead time of the rapid-mixing devices to 10–20 μs by designing microfluidic channels for efficient mixing [39,54]. With this time resolution, the processes involved in the burst phase that lead to the formation of the collapsed domain might be resolved. Moreover, the structural identification of specific locations and regions of the main chain is now possible by introducing isotopes site specifically. The use of isotope-edited samples for the rapid-mixing infrared investigation would add a new structural information. Development of a new unfolding method that do not interfere with the detection of amide I, such as the DMSO-induced unfolding, would expand the number of proteins that can be investigated by the kinetic infrared spectroscopy. Such investigations will give a clearer view on the molecular mechanism of protein folding.

Acknowledgements We sincerely thank the following collaborators for sharing our enthusiasm for protein folding research; Mr. S. Nishiguchi, Mr. A. Maeda, Dr. S. Akiyama, Dr. T. Uzawa, Prof. T. Konno, Prof. K. Ishimori, Prof. Y. Goto, and Prof. I. Morishima.

References

1. A.R. Fersht, *Structure and Mechanism in Protein Science* (W. H. Freeman and Company, New York, 1998)
2. A.V. Finkelstein, O.B. Ptitsyn, *Protein Physics* (Academic Press, London, 2002)
3. J.E. Kohn, I.S. Millett, J. Jacob, B. Zagrovic, T.M. Dillon, N. Cingel, R.S. Dothager, S. Seifert, P. Thiyagarajan, T.R. Sosnick, M.Z. Hasan, V.S. Pande, I. Ruczinski, S. Doniach, K.W. Plaxco, Proc. Natl. Acad. Sci. U.S.A. **101**, 12491 (2004)
4. P.J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, 1953)
5. C. Levinthal, J. Chim. Phys. **65**, 44 (1968)
6. R.L. Baldwin, Ann. Rev. Biophys. **37**, 1 (2008)
7. A. Barth, C. Zscherp, Q. Rev. Biophys. **35**, 369 (2002)
8. S. Krimm, J. Bandekar, Adv. Protein Chem. **38**, 181 (1986)
9. H. Torii, M. Tasumi, in *Infrared Spectroscopy of Biomolecules*, ed. by H.H. Mantsch, D. Chapman (Wiley-Liss, New York, 1996), p. 1
10. A. Troullier, D. Reinstädler, Y. Dupont, D. Naumann, V. Forge, Nat. Struct. Biol. **7**, 78 (2000)
11. E. Kauffmann, N.C. Darnton, R.H. Austin, C. Batt, K. Gerwert, Proc. Natl. Acad. Sci. U.S.A. **98**, 6646 (2001)
12. S. Akiyama, S. Takahashi, K. Ishimori, I. Morishima, Nat. Struct. Biol. **7**, 514 (2000)
13. S. Akiyama, S. Takahashi, T. Kimura, K. Ishimori, I. Morishima, Y. Nishikawa, T. Fujisawa, Proc. Natl. Acad. Sci. U.S.A. **99**, 1329 (2002)
14. T. Uzawa, C. Nishimura, S. Akiyama, K. Ishimori, S. Takahashi, H.J. Dyson, P.E. Wright, Proc. Natl. Acad. Sci. U.S.A. **105**, 13859 (2008)
15. T. Kimura, S. Takahashi, S. Akiyama, T. Uzawa, K. Ishimori, I. Morishima, J. Am. Chem. Soc. **124**, 11596 (2002)
16. S. Nishiguchi, Y. Goto, S. Takahashi, J. Mol. Biol. **373**, 491 (2007)
17. T. Kimura, A. Maeda, S. Nishiguchi, K. Ishimori, I. Morishima, T. Konno, Y. Goto, S. Takahashi, Proc. Natl. Acad. Sci. U.S.A. **105**, 13391 (2008)
18. S.E. Jackson, Fold. Des. **3**, R81 (1998)
19. D.J. Brockwell, S.E. Radford, Curr. Opin. Struct. Biol. **17**, 30 (2007)
20. A.R. Fersht, A. Matouschek, L. Serrano, J. Mol. Biol. **224**, 771 (1992)
21. D.E. Otzen, L.S. Itzhaki, N.F. El Marsy, S.E. Jackson, A.R. Fersht, Proc. Natl. Acad. Sci. U.S.A. **91**, 10422 (1994)
22. V. Muñoz, W.A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **96**, 11311 (1999)
23. J.N. Onuchic, P.G. Wolynes, Curr. Opin. Struct. Biol. **14**, 70 (2004)
24. T. Uzawa, T. Kimura, K. Ishimori, I. Morishima, T. Matsui, M. Ikeda-Saito, S. Takahashi, S. Akiyama, T. Fujisawa, J. Mol. Biol. **357**, 997 (2006)
25. M. Ikeguchi, K. Kuwajima, M. Mitani, S. Sugai, Biochemistry **25**, 6965 (1986)
26. M. Arai, K. Kuwajima, Adv. Protein Chem. **53**, 209 (2000)
27. P.A. Jennings, P.E. Wright, Science **262**, 892 (1993)
28. M. Ohgushi, A. Wada, FEBS Lett. **164**, 21 (1983)
29. T. Uzawa, S. Akiyama, T. Kimura, S. Takahashi, K. Ishimori, I. Morishima, T. Fujisawa, Proc. Natl. Acad. Sci. U.S.A. **101**, 1171 (2004)
30. T. Kimura, T. Uzawa, K. Ishimori, I. Morishima, S. Takahashi, T. Konno, S. Akiyama, T. Fujisawa, Proc. Natl. Acad. Sci. U.S.A. **102**, 2748 (2005)
31. C.M. Phillips, Y. Mizutani, R.M. Hochstrasser, Proc. Natl. Acad. Sci. U.S.A. **92**, 7292 (1995)
32. R.M. Ballew, J. Sabelko, M. Gruebele, Nat. Struct. Biol. **3**, 923(1996)

33. R. Gilmanshin, S. Williams, R.H. Callender, W.H. Woodruff, R.B. Dyer, Proc. Natl. Acad. Sci. U.S.A. **94**, 3709 (1997)
34. C.Y. Huang, Z. Getahun, Y. Zhu, J.W. Klemke, W.F. DeGrado, F. Gai, Proc. Natl. Acad. Sci. U.S.A. **99**, 2788 (2002)
35. C.M. Jones, E.R. Henry, Y. Hu, C.K. Chan, S.D. Luck, A. Bhuyan, H. Roder, J. Hofrichter, W.A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **90**, 11860 (1993)
36. T. Pascher, J.P. Chesick, J.R. Winkler, H.B. Gray, Science **271**, 1558 (1996)
37. S. Hirota, Y. Fujimoto, J. Choi, N. Baden, N. Katagiri, M. Akiyama, R. Hulsker, M. Ubbink, T. Okajima, T. Takabe, N. Funasaki, Y. Watanabe, M. Terazima, J. Am. Chem. Soc. **128**, 7551 (2006)
38. R.P. Chen, J.J. Huang, H.L. Chen, H. Jan, M. Velusamy, C.T. Lee, W. Fann, R.W. Larsen, S.I. Chan, Proc. Natl. Acad. Sci. U.S.A. **101**, 7305 (2004)
39. S. Matsumoto, A. Yane, S. Nakashima, M. Hashida, M. Fujita, Y. Goto, S. Takahashi, J. Am. Chem. Soc. **129**, 3840 (2007)
40. H. Hartridge, F.J.W. Roughton, Proc. R. Soc. Lond. A **104**, 395 (1923)
41. B. Chance, Q.H. Gibson, R.H. Eisenhardt, K.K. Lonberg-Holm, Science **146**, 1697 (1964)
42. P. Regenfuss, R.M. Clegg, M.J. Fulwyler, F.J. Barrantes, T.M. Jovin, Rev. Sci. Instrum. **56**, 283 (1985)
43. P. Regenfuss, R.M. Clegg, Biophys. Chem. **26**, 83 (1987)
44. S. Takahashi, Y.-C. Ching, J. Wang, D.L. Rousseau, J. Biol. Chem. **270**, 8405 (1995)
45. C.K. Chan, Y. Hu, S. Takahashi, D.L. Rousseau, W.A. Eaton, J. Hofrichter, Proc. Natl. Acad. Sci. U.S.A. **94**, 1779 (1997)
46. M.C. Shastry, H. Roder, Nat. Struct. Biol. **5**, 385 (1998)
47. H. Roder, K. Maki, H. Cheng, Chem. Rev. **106**, 1836 (2004)
48. S. Takahashi, S.R. Yeh, T.K. Das, C.K. Chan, D.S. Gottfried, D.L. Rousseau, Nat. Struct. Biol. **4**, 44 (1997)
49. S.R. Yeh, S. Takahashi, B. Fan, D.L. Rousseau, Nat. Struct. Biol. **4**, 51 (1997)
50. J.B. Knight, A. Vishwanath, J.P. Brody, R.H. Austin, Phys. Rev. Lett. **80**, 3863 (1998)
51. L. Pollack, M.W. Tate, N.C. Darnton, J.B. Knight, S.M. Gruner, W.A. Eaton, R.H. Austin, Proc. Natl. Acad. Sci. U.S.A. **96**, 10115 (1999)
52. E. Kaufmann, N.C. Darnton, R.H. Austin, C. Batt, K. Gerwert, Proc. Natl. Acad. Sci. U.S.A. **98**, 6646 (2001)
53. E.A. Lipman, B. Schuler, O. Bakajin, W.A. Eaton, Science **301**, 1233 (2003)
54. T. Egawa, J.L. Durand, E.Y. Hayden, D.L. Rousseau, S.R. Yeh, Anal. Chem. **81**, 1622 (2009)
55. M. Kakuta, P. Hinsmann, A. Manz, B. Lendl, Lab Chip **3**, 82 (2003)
56. B. Nölting, *Protein Folding Kinetics, Biophysical Methods*, 2nd edn. (Springer, Berlin, 2006)
57. H. Fabian, D. Naumann, Methods **34**, 128 (2004)
58. A.L. Fink, L.J. Calciano, Y. Goto, T. Kurotsu, D.R. Palleros, Biochemistry **33**, 12504 (1994)
59. Y. Goto, L.J. Calciano, A.L. Fink, Proc. Natl. Acad. Sci. U.S.A. **87**, 573 (1990)
60. I.S. Millett, S. Doniach, K.W. Plaxco, Adv. Protein Chem. **62**, 241 (2002)
61. J. Yao, J. Chung, D. Eliezer, P.E. Wright, H.J. Dyson, Biochemistry **40**, 3561 (2001)
62. M. Jackson, H.H. Mantsch, Biochim. Biophys. Acta **1078**, 231 (1991)
63. D. Eliezer, P.A. Jennings, H.J. Dyson, P.E. Wright, FEBS Lett. **417**, 92 (1997)
64. D.M. Byler, H. Susi, Biopolymers **25**, 469 (1986)
65. A. Dong, P. Huang, W.S. Caughey, Biochemistry **29**, 3303 (1990)
66. P.I. Haris, D. Chapman, Biopolymers **37**, 251 (1995)
67. A. Barth, C. Zscherp, Q. Rev. Biophys. **35**, 369 (2002)
68. M. Jackson, H.H. Mantsch, Crit. Rev. Biochem. Mol. Biol. **30**, 95 (1995)
69. S. Krimm, J. Bandekar, Adv. Prot. Chem. **38**, 181 (1986)
70. H. Torii, T. Tatsumi, M. Tatsumi, J. Raman Spectrosc. **29**, 537 (1998)
71. H. Torii, J. Phys. Chem. A, **108**, 7272 (2004)
72. H. Torii, M. Tasumi, J. Raman Spectrosc. **29**, 81 (1998)
73. N.S. Marinkovic, A.R. Adzic, M. Sullivan, K. Kovacs, L.M. Miller, D.L. Rousseau, S.-R. Yeh, M.R. Chance, Rev. Sci. Instrum. **71**, 4057 (2000)

74. J.C. Kendrew, R.E. Dickerson, B.E. Strandberg, R.G. Hart, D.R. Davies, D.C. Phillips, V.C. Shore, Nature **185**, 422 (1960)
75. D. Eliezer, P.E. Wright, J. Mol. Biol. **263**, 531 (1996)
76. R. Mohana-Borges, N.K. Goto, G.J.A. Kroon, H.J. Dyson, P.E. Wright, J. Mol. Biol. **340**, 1131 (2004)
77. M. Kataoka, I. Nishii, T. Fujisawa, T. Ueki, F. Tokunaga, Y. Goto, J. Mol. Biol. **249**, 215 (1995)
78. F.M. Hughson, P.E. Wright, R.L. Baldwin, Science **249**, 1544 (1990)
79. M.S. Kay, R.L. Baldwin, Nat. Struct. Biol. **3**, 439 (1996)
80. M.S. Kay, C.H.I. Ramos, R.L. Baldwin, Proc. Natl. Acad. Sci. U.S.A. **96**, 2007 (1999)
81. P.A. Jennings, P.E. Wright, Science **262**, 892 (1993)
82. M. Jamin, R.L. Baldwin, J. Mol. Biol. **276**, 491 (1998)
83. M. Jamin, S.-R. Yeh, D.L. Rousseau, R.L. Baldwin, J. Mol. Biol. **292**, 731 (1999)
84. E.S. Manas, Z. Getahun, W.W. Wright, W.F. DeGrado, J.M. Vanderkooi, J. Am. Chem. Soc. **122**, 9883 (2000)
85. S.T.R. Walsh, R.P. Cheng, W.W. Wright, D.O.V. Alonso, V. Daggett, J.M. Vanderkooi, W.F. Degrado, Protein Sci. **12**, 520 (2003)
86. D. Eliezer, J. Yao, H.J. Dyson, P.E. Wright, Nat. Struct. Biol. **5**, 148 (1998)
87. T. Konno, Protein Sci. **10**, 2093 (2001)
88. J.R. Somoza, F. Jiang, L. Tong, C.H. Kang, J.M. Cho, S.H. Kim, J. Mol. Biol. **234**, 390 (1993)
89. S.Y. Lee, J.H. Lee, H.J. Chang, J.M. Cho, J.W. Jung, W. Lee, Biochemistry **38**, 2340 (1999)
90. R.L. Baldwin, J. Mol. Biol. **371**, 283 (2007)
91. G.I. Makhatadze, P.L. Provalov, Adv. Prot. Chem. **47**, 307 (1995)

Chapter 5
High-Pressure Vibrational Spectroscopy Studies of the Folding, Misfolding and Amyloidogenesis of Proteins

Roland Winter, Matthias Pühse, and Jonas Markgraf

Abstract A review is given for high-pressure Fourier-transform infrared (FTIR) spectroscopy studies on biomolecular systems with a special focus on proteins. After reviewing the experimental techniques, we present data on the pressure-induced un- and refolding reaction of proteins as well as studies on enzyme reactions. Finally, studies on macromolecular protein mimetics and recent advances in using pressure for studying misfolding and aggregation/fibrillation of proteins are discussed.

5.1 Introduction to High-Pressure Bioscience

Interestingly, the greatest part of our biosphere on Earth is in the realm of environmental extremes, and for several decades now, the limits of the existence of life have been pushed to unexpected extremes of pressure, temperature, pH and salinity [1]. Although extreme temperatures and pressures significantly influence the structural properties and thus functional characteristics of cells, this has not prevented life from invading the cold and high-pressure habitats of marine depths (69% of the surface of the earth is covered by oceans, and the average pressure on the ocean floor is about 380 bar (0.1 MPa = 1 bar, 1 GPa = 10 kbar)). For example, deep-sea sediments and hydrothermal vents are densely crowded with barophilic–thermophilic (i.e. pressure- and heat-adapted species). Psychrophilic–barophilic (cold- and pressure-adapted) species, which live at ~2°C, are found on the deepest ocean floor (depth ~10,900 m) in the Mariana Trench and in deep-sea sediments [2–5]. The significance of extremophiles and extreme conditions of temperature, pressure, solvent conditions and water activity for biotechnology,

R. Winter (✉)
Faculty of Chemistry, Physical Chemistry I – Biophysical Chemistry, TU Dortmund University, Otto Hahn Str. 6, 44227 Dortmund, Germany
e-mail: roland.winter@tu-dortmund.de

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*, Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_5,

medicine and research into the origin of life and early evolution has been discussed in stimulating monographs [4,5]. One fascinating example is given by small animals called Tardigrades. They become immobile and shrink into an arrest state when the humidity of the surroundings decreases. In such a state, they can survive temperatures from −253 up to 151°C and pressures up to 6 kbar, whereas they are killed at about 2 kbar under normal conditions [6].

From a more physical–chemical point of view, in contrast to increasing temperature, which is characterized by energy and volume effects due to thermal expansivity, pressure effects are mainly due to volumetric aspects via the compressibility of the system. Interest in pressure as a thermodynamic and kinetic variable has been growing in studies of biological materials in recent years [1,7–15], and for the following reasons: (a) Changing temperature of a system at atmospheric pressure produces a simultaneous change in thermal energy and volume; therefore, to separate thermal and volume effects, one must carry out also pressure-dependent experiments. (b) Because noncovalent interactions play a primary role in the stabilization of biochemical systems, the use of pressure allows one to change the intermolecular interactions in a controlled way, without the major perturbations often produced by changes in temperature or cosolvent concentration. (c) Pressure affects chemical equilibria and reaction rates. Le Châtelier's principle predicts that the application of pressure shifts an equilibrium towards the state that occupies a smaller volume (standard reaction volume $\Delta V^{\circ} < 0$), and accelerates processes for which the transition state has a smaller volume than the ground state (activation volume $\Delta V^{\#} < 0$). Hence, with the knowledge of ΔV° and $\Delta V^{\#}$ values, one can then draw valuable conclusions about the nature of the reaction and its mechanism. (d) The viscosity of the solvent can be changed continuously by pressure. (e) One can extend the range of conditions and carry out experiments at subzero temperatures still within the liquid state. For example, protein solutions can be studied at subzero temperatures to investigate their cold-denaturation behaviour. Apart from equilibrium aspects, information is also obtained on the dynamics of the system in view of the relation between the isothermal compressibility and the volume fluctuations [7].

Pressures used to investigate biochemical systems generally range from 1 bar to 10 kbar. Such pressures only change intermolecular distances and affect conformations, but do not change covalent bond distances or bond angles. The covalent structure of low-molecular-mass biomolecules (peptides, lipids and saccharides), as well as the primary structure of macromolecules (proteins, nucleic acids and polysaccharides), is generally not perturbed by pressures up to about 30 kbar. Pressure acts predominantly on the conformation, quaternary and supramolecular structures of biomolecular systems.

In this chapter, we introduce Fourier-transform infrared (FTIR) spectroscopy as one of the major techniques for investigating pressure effects on biomolecules, such as proteins, and soft condensed matter in general, and we discuss selected results of pressure studies on the structure, phase behaviour and kinetic properties of proteins and synthetic macromolecular systems serving as simple protein mimetics. The experimental principles used in high-pressure FTIR spectroscopy are discussed as well.

5.2 Fundamental Concepts: Stability Diagram of Proteins

The protein folding reaction is one of the most crucial steps during the life of a protein. If some failure occurs in achieving the native functional conformation, this will generally render the protein totally inactive, or even worse, it can produce a misfolded molecule that can interfere or block components of the cellular machinery to the point of causing cell malfunction or even death. In recent years, it has become evident that a wide range of human diseases are associated with aberrations in the folding process [16]. These diseases, which are also coined "protein conformational diseases", include Alzheimer's disease (Aβ-peptide), Parkinsons's disease (α-synuclein), prion-protein-related encephalopathies and type II diabetes mellitus (islet amyloid peptide, IAPP). Since the discovery of high-pressure-induced protein unfolding and denaturation by nobel laureate P. W. Bridgman in 1914 [17], it has been shown in many studies now that hydrostatic pressure may also lead to disruption of the intra- and intermolecular interactions maintaining the native protein structure, which is accompanied by a decrease of the volume of the protein–water system, and simultaneous unfolding [7–15]. Subsequently, high hydrostatic pressures (HHP) have also been shown to be effective for disaggregation and refolding of proteins from insoluble aggregates prepared in vitro [18, 19].

The appropriate way of expressing the thermodynamic stability of a protein appears to be a multidimensional function of temperature, pressure and solution conditions, yielding an energy landscape as a multi-dimensional surface. When the solution conditions (pH, salt and cosolvent concentration) are kept constant, the stability of the protein is a simultaneous function of temperature and pressure, only. The Gibbs free energy difference $\Delta_u G$ between the denatured (unfolded, U) and native state (N), relative to some reference point $T_0,\ p_0$ (e.g. the unfolding temperature at ambient pressure), can be approximated – assuming a second-order Taylor series of $\Delta_u G(T,\ p)$ expanded with respect to T and p around $T_0,\ p_0$ – as [10, 20]:

$$\Delta_u G = \Delta_u G_0 + \frac{\Delta_u \kappa}{2}(p-p_0)^2 + \Delta_u \alpha (p-p_0)(T-T_0) - \Delta_u C_p \left[T\left(\ln \frac{T}{T_0} - 1 \right) + T_0 \right] + \Delta_u V_0 (p - p_0) - \Delta_u S_0 (T - T_0), \tag{5.1}$$

where $\Delta_u \kappa,\ \Delta_u \alpha$, and $\Delta_u C_p$ refer to the changes in compressibility, expansibility and heat capacity upon unfolding (index u). $\Delta_u V_0$ is the change in partial molar volume of the protein in state U relative to N, also at the reference point (T_0 and p_0). Generally, the volume change for unfolding is given by

$$\Delta_u V(p,T) = V_U(p,T) - V_N(p,T) = \Delta_u V_0 + \Delta_u \alpha (T - T_0) - \Delta_u \kappa (p - p_0). \tag{5.2}$$

The entropy and volume changes of unfolding may be decomposed into intrinsic contributions of the protein and those of the hydration shell. The transition line, where the protein unfolds upon a temperature of pressure change, is given by

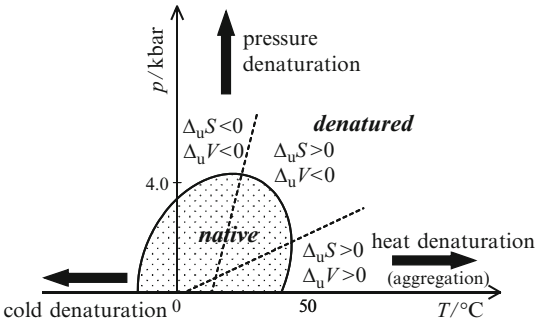

Fig. 5.1 Schematic pressure and temperature stability diagram of a typical monomeric protein [20]. Different routes of unfolding of the native protein as well as the corresponding thermodynamic changes are depicted. Heat denaturation is often accompanied by irreversible aggregation

$\Delta_u G = 0$. The physically relevant solution of the curve in the p, T-plane has an elliptic-like shape (Fig. 5.1). Exposure of nonpolar groups upon unfolding tends to bind and immobilize water molecules, thereby decreasing their entropy. This process is strongly temperature dependent. Its experimental signature is the positive heat capacity increment change at constant pressure, $\Delta_u C_p > 0$ (as has been determined by calorimetry).

According to the stability diagram shown in Fig. 5.1, both pressure and cold can also be used to destabilize proteins and have also been used to destabilize protein aggregate and fibrils, which is facilitated by a combination of factors: HHP leads to a weakening of hydrophobic interactions in general. Moreover, the presence of cavities within the folded proteins or in the interface of oligomers can favour the unfolding or dissociation of these structures. If the water molecules fill the cavities when they become accessible upon unfolding or dissociation of the proteins, this leads to a volume decrease and the event is hence favoured under high pressure. Also the dissociation of electrostatic interactions leads to a marked reduction in the volume caused by electrostrictive effects of water molecules around the unpaired charged residues exposing charged groups. In a similar way, solvation of polar groups results in a decrease in volume of the solution [7, 8, 18–20].

5.3 Experimental Methods

5.3.1 High-Pressure FTIR Spectroscopy

The variety of chemical groups present in amino acid side chains, the amide group itself, and the possible presence of prosthetic groups (e.g. in the case of hemoglobin) give rise to several vibrational bands that can be sensitive to structural changes induced in the protein molecule upon changes in environmental conditions, such as temperature and pressure. Out of seven existing infrared bands stemming from the peptide linkage, the amide-I band proved to bear most significant information for

a detailed characterization of protein's secondary structure [21–27]. The amide-I band appears in the wavenumber region 1,600–1,700 cm^{-1} and is due to the in-plane C = O stretching vibration (80%) weakly coupled with C–N stretching and in-plane N–H bending of the amide groups in proteins and polypeptides. As the energy of this vibration depends strongly on the presence and geometry of hydrogen bonding, the wavenumber of the amide-I band is subjected to changes in secondary structure, in which the particular amide bond is involved. As globular proteins generally contain segments of different substructures, the amide-I band usually appears as a broad band with several more or less pronounced maxima, and the changes in the relative intensities of these maxima are widely used for monitoring the changes in the conformational substructures of proteins. It should also be stressed that these wavenumbers might depend themselves on factors such as solvation of the polypeptide chain and possible distortions within helices [28]. Although similar to the case of the amide-I band, secondary structure assignments of the component frequencies to the amide-II band can be carried out. Generally high-pressure FTIR studies on proteins do not take advantage of the amide-II band other than monitoring of the pressure-enhanced H/D exchange, however [27]. The exchange rate is measured by following the decrease in the intensity of the amide-II band with time. This infrared band shifts to a wavenumber of 1,458 cm^{-1} when the protein atoms of the amide group are replaced by deuterium atoms. One of the advantages of IR spectroscopy over other methods for the H/D exchange study is that the conformational structure of the protein can be monitored simultaneously during the measurement of the H/D exchange rate constants.

As high pressure generally causes compression of chemical bonds equivalent to changes of the force constants between the involved atoms, pressure-induced wavenumber-shifts are frequently observed [27]. While this so-called elastic pressure effect is usually revealed by blue-shifting bands, hydrogen-bonded systems, such as proteins, often show a red shift of the IR bands, which is due to a strengthening of the O...H hydrogen bond at the expense of the C = O bond strength. The wavenumber versus pressure plots generally feature a negative slope for inter- or intrachain hydrogen-bound structures (as α-helices and β-sheets), and a positive slope for water-bound random coil [27]. Next to this elastic pressure effect, changes in secondary structure, e.g. upon unfolding of proteins, may be detected [27, 29–37]. Qualitative analysis of the secondary structure components can be carried out in a rather straightforward manner by conventional resolution enhancement methods (such as second derivative or Fourier self-deconvolution). It has been advocated that deconvolution coupled with multiband-fitting is a good option for quantitative analysis of protein high-pressure infrared spectra [31–37]. Quantitative analysis may be complicated by several factors, however. Firstly, the different elements of protein secondary structure may feature different molar absorption coefficients, i.e. transition dipole moments [27, 28, 38]. Also, IR protein studies often call for D_2O as the solvent, which is due to the strong absorption of H_2O in the amide-I spectral range. The isotopic substitution of solvent is followed by H/D exchange of exchangeable protein protons, which is revealed as a red shift of amide-I and amide-II bands (completely deuterated bands are marked as amide-I′

and amide-II′, respectively). The red shifts are rather insignificant for the former, but large for the latter band. Undeniably, H/D exchange complicates protein spectra, but, as its rate depends strictly on the accessibility of the polypeptide chain to solvent, it conveys important information about the protein compactness and its changes under high pressure [27]. Although IR studies of proteins dissolved in H_2O may be useful, the requirement for D_2O becomes generally mandatory in high-pressure FTIR spectroscopy, because, when a diamond anvil cell is used, the metal spacer (gasket) between the two diamonds has to be sufficiently thick (that is usually 0.05 mm, instead of only 0.006 mm for H_2O-based protein solutions at ambient pressure), which would drastically increase the intensity of the overlapping H_2O band.

5.3.2 Diamond Anvil Cell Technology

The most effective and convenient device for high hydrostatic pressure generation in FTIR spectroscopy is the diamond anvil cell (DAC) [39,40], which is schematically depicted in Fig. 5.2. The principal mode of operation is as follows. Force is generated via compression of a spring, which is then transmitted with the help of a metal lever on the upper diamond anvil. The diamond then presses on the sample, which is maintained by a metal gasket and the lower diamond [41]. This gasket normally consists of a circular plate of stainless steel with a drilling in the middle, where the sample is applied upon. The thickness of the gasket depends on the absorptivity of the sample as well as on the intended maximal pressure. Typically, values of 30–50 μm for biophysical applications with D_2O as solvent and pressures up to ∼10–15 kbar are used. One of the main advantages of this setup is the very low amount of sample required for the high-pressure measurement (several μL are

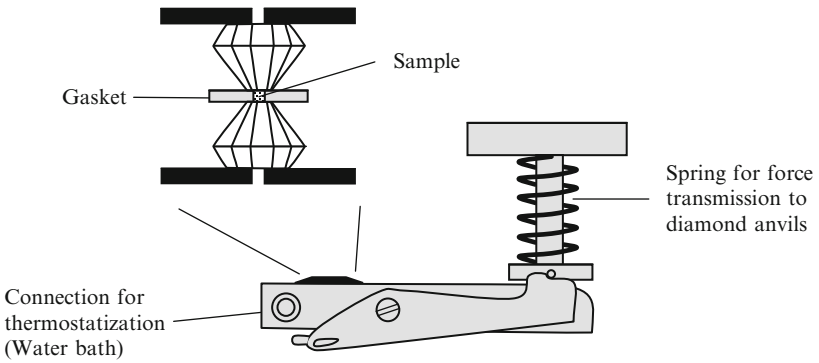

Fig. 5.2 Schematic drawing of a diamond anvil cell (DAC). The upper part shows the two diamond anvils with the gasket (*grey*) and the fluid sample in between. Thermostatization of the apparatus is possible via connection to a water bath

sufficient to fill the DAC). In fact, the effective pressurized initial volume is typically only around 10 nL, assuming a gasket thickness of 50 μm and a drilling diameter of 0.45 mm. The excess volume is squeezed out when closing the cell.

A further important issue is the choice of the diamonds. Normally, type IIa diamonds are used for the anvils, as they consist of pure carbon without spectroscopically unfavourable impurities, such as in type Ia and Ib diamonds, which are contaminated with nitrogen, and in type IIb diamonds that contain traces of boron [42]. With the diamond absorption between 2,650 and 1,800 cm^{-1}, there are free spectral windows for the investigation of biomolecules between 3,100 and 2,800 cm^{-1} as well as between 1,800 and 1,300 cm^{-1}, when using D_2O as solvent. As all kinds of diamonds absorb strongly around 2,000 cm^{-1}, one is left with sapphire windows as an option for high-pressure studies in this spectral range, e.g. for studies of the ligand conformation in carbon monoxide-bound proteins [43]. Sapphire, however, has mechanical properties inferior to diamond, thus limiting the range of accessible pressures to a few kbar.

5.3.3 Pressure Calibrants for Infrared Spectroscopy

When applying the DAC for high-pressure infrared studies, one has to rely on indirect methods for pressure determination, as the use of a manometer is not possible. In the last decades, three main calibrants were developed. First, the R_1 fluorescence line of ruby [44], second the phonon bands of α-quartz and finally the symmetric sulphate stretching mode of $BaSO_4$ [45, 46] (see Table 5.1).

The ruby method, prime in physics, has several disadvantages in biophysics. As it can be seen from Table 5.1, the fluorescence band is in the VIS- region of the electromagnetic spectrum, so it is necessary to exchange the spectrometer for pressure determination. Furthermore, as it was developed primarily for geophysical applications, the ruby fluorescence shift is not precise enough for pressures up to ∼10 kbar, the range being of prime interest in biophysics, and the strong laser irradiation needed for fluorescence excitation might damage sensitive biomolecules. α-Quartz or $BaSO_4$ is generally the calibrant of choice. $BaSO_4$ may be superior, as it exhibits a sharp and pronounced band, whose position changes linearly with pressure (blueshift of 0.41 $cm^{-1}\,kbar^{-1}$). In contrast, the phonon bands of α-quartz are significantly broader and might suffer from stronger spectral noise due to diffraction from coarse crystallites, which renders pressure determination less precise. It should be mentioned that the initial position of the $BaSO_4$ signal

Table 5.1 Pressure calibrants for high-pressure infrared spectroscopy

Pressure calibrant	Associated spectroscopic band
Ruby	14,405 cm^{-1} (694.2 nm)
α-Quartz	∼798/695 cm^{-1}
$BaSO_4$	∼983 cm^{-1}

is temperature dependent; it shows a red shift when temperature is elevated. The exact position also depends on the chemical environment, so the initial peak position should be determined at each temperature and for each new sample loaded. The general accuracy of pressure determination by the $BaSO_4$ method is about ±200–300 bar [46], which is sufficient for most biophysical applications.

As the diamond windows of the DAC absorb quite strongly in the Mid-IR region, the incident IR beam should have enough intensity and the detector sufficient sensitivity. The first requirement can be met by using lenses or mirror-based beam condensers, which focus the beam onto the pinhole of the DAC, whereas the second problem can be solved by implementing nitrogen-cooled mercury–cadmium–telluride (MCT) detectors. Especially the type MCT-A possesses the required sensitivity to generate spectra with a high signal-to-noise ratio.

There are only limited data available considering the pressure stability of buffers. High hydrostatic pressure is known to favour the dissociation of ion pairs due to the volume reduction associated with electrostriction of water. Compilations of pressure-stable buffer systems are available [47–49]. Tris and Bis-Tris buffers are the most common buffering agents for high-pressure spectroscopic studies on biomolecules.

5.4 Examples of Pressure Studies on Proteins and Polymers

5.4.1 Pressure-Induced Protein un- and Refolding Reactions

The pressure-induced folding reaction of proteins may conveniently be followed by FTIR spectroscopy using the diamond anvil cell technology. At pressures of 2–8 kbar, most small monomeric proteins unfold reversibly. As an example, we show the use of high-pressure FTIR spectroscopy to study the reversible pressure-induced unfolding and refolding of RNase A and compare it with the results obtained for the temperature-induced transition. RNase A is a single-domain protein, a pancreatic enzyme that catalyses the cleavage of single-stranded RNA. The crystalline state of the protein contains about 23% helices, 46% β-sheets, 21% turns and 10% unordered structures [50]. The protein has traditionally served as a model for protein folding because it is small and stable and has a well-known native structure. The protein was dissolved in D_2O at pH 2.5, and the solution was heated to 55°C for 20 min to allow for deuterium exchange of the labile amide protons of RNase A. The sample was then lyophilized from D_2O. The RNase A was dissolved to a concentration of 4 mM in 20 mM maleic acid buffer and 0.3 M NaCl in D_2O. For the pressure-dependent measurements, the infrared light was focused by a spectra bench onto the pinhole of the DAC with type IIa diamonds [33]. Each spectrum was obtained by coadding 512 scans at a spectral resolution of 2 cm^{-1} and was apodized with a Happ-Genzel function. The equilibration times to take one data point at each temperature and pressure were about 15 min. Fourier self-deconvolution of

the IR spectra was performed with a resolution enhancement factor of 1.8 and a bandwidth of 15 cm^{-1}. The fractional intensities of the secondary structure elements were calculated from a band-fitting procedure assuming a Gaussian–Lorentzian line shape function [33].

Figure 5.4a shows deconvoluted IR spectra of RNase A between 18 and 82°C at ambient pressure after H/D exchange at 55°C. At 20°C, the maximum of the amide-I′ band appears at 1,637 cm^{-1}. An increase of temperature to 82°C leads to a shift of the band maximum to larger wavenumbers and the absorption band becomes more symmetric. The evolution of the secondary structures with temperature is shown in Fig. 5.4a. The temperature-induced decrease of band intensities of β-sheets and α-helices between 50 and 70°C is accompanied by a concomitant increase in band intensities of non-periodic and turn structures. The non-monotonous changes in the component intensities as a function of temperature point to more steps in the unfolding process. The data also indicate that the level of ordered secondary structure elements at temperatures above 70°C is very low or almost negligible.

The deconvoluted FTIR spectra up to 10 kbar at $T = 20$°C are shown in Fig. 5.3b (for pressure calibration, the pressure-induced shift of the phonon bands of α-quartz have been used as depicted in Fig. 5.3c). The data reveal a conformational transition of RNase A in the pressure range around 6–7 kbar at this temperature, where the asymmetric amide-I′ band becomes more symmetric with a band maximum at 1,643 cm^{-1}. The changes of the fractional intensities of the different amide-I′ subbands are presented in Fig. 5.4b. The pressure-induced denaturation as judged by changes in the fractional band intensities begins at ∼5 kbar and the transition is complete at about 7.5 kbar. Although the fractional intensities of α-helices and β-sheets decrease significantly, these structures, in particular the β-sheet structures, are not completely disrupted by pressure. So far, all experimental evidence indicates that the pressure-denatured states of proteins have more ordered secondary structure than the temperature-denatured states [7–11].

In order to determine the standard Gibbs free energy change $\Delta_u G^o$ and the volume change $\Delta_u V^o = \mathrm{d}\Delta_u G^o(p)/\mathrm{d}p$ assuming a two-state pressure-induced unfolding process, the equilibrium profiles obtained from the FTIR spectra can be fitted to the relation $\Delta_u G^o(p) = -RT \ln K_{eq,u} = -RT \ln[(I_f - I_p)/(I_p - I_u)]$ using the data points at each pressure, I_p, and the asymptotic values of the fractional band intensities due to particular secondary structural elements for the folded and unfolded states, I_f and I_u, respectively. We obtain a standard Gibbs free energy change for unfolding of $\Delta_u G^o = 27$ kJ mol^{-1} and a volume change for unfolding of about $\Delta_u V^o \approx -45$ mL mol^{-1} at ambient temperature and pressure. For comparison, the corresponding data obtained for unfolding of SNase are 16 kJ mol^{-1} and −80 mL mol^{-1} [31, 32]. The volume change of SNase upon unfolding is thus about twice that of RNase A, indicating a larger void volume of native SNase. With increasing temperature, ΔV_u^o has been found to become less negative, giving a positive expansivity change of 1.33 mL $mol^{-1}K^{-1}$. Values of similar order of magnitude have also been recorded for other monomeric proteins, such as lysozyme $\Delta_u G^o = 9.5$ kJ mol^{-1}, $\Delta_u V^o \approx -40$ mL mol^{-1} at 25°C and pH 2 [51]).

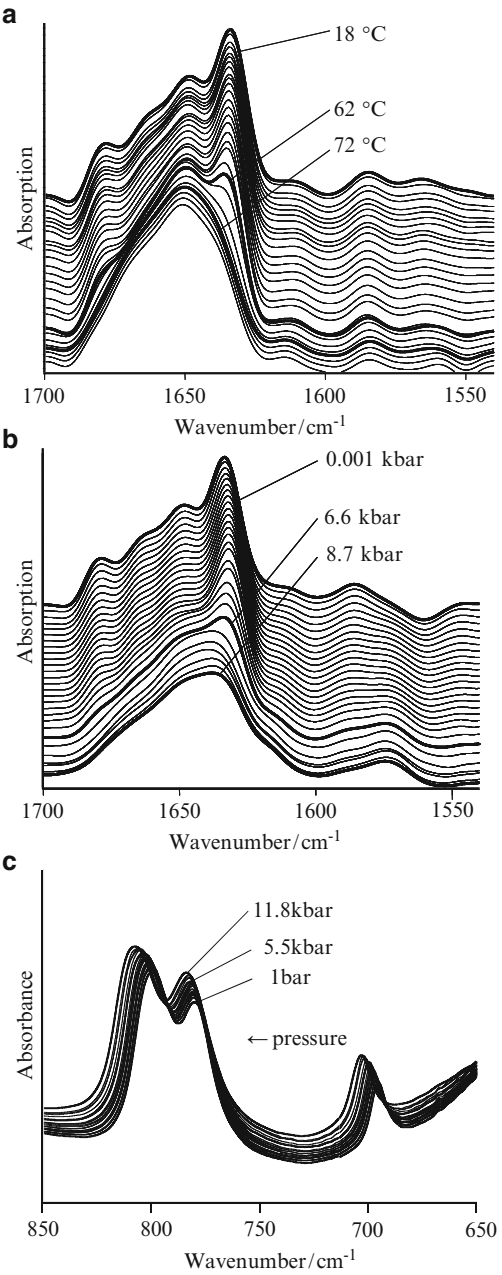

Fig. 5.3 Deconvoluted FTIR absorption spectra of RNase (5% (w/w), pH 2.5) (**a**) as a function of temperature at atmospheric pressure and (**b**) as a function of pressure at $T = 20°C$. (**c**) Pressure calibration by registration of the pressure-induced *blue* shift of the wavenumber of the phonon bands of α-quartz between 650 and 850 cm^{-1}. Bands appear at about 695, 780 and 801 cm^{-1}. The pressure shift of the band at 695 cm^{-1} is given by $p = 1.2062\Delta\tilde{\nu} + 0.015054(\Delta\tilde{\nu})^2$, and that at 801 cm^{-1} by $p = 1.168\Delta\tilde{\nu} + 0.0158(\Delta\tilde{\nu})^2 - 0.1660$; p is given in kbar [45, 46]

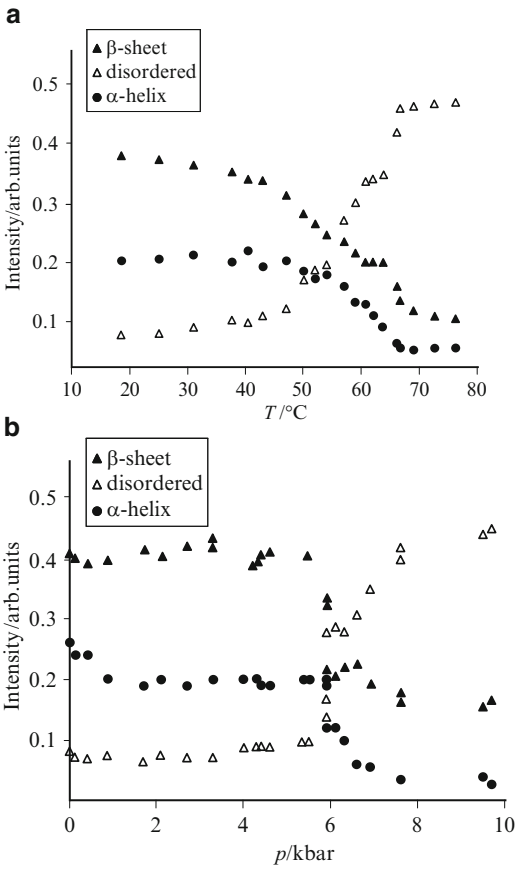

Fig. 5.4 (**a**) Temperature effect on the areas of the IR bands associated with β-sheets, α-helices and disordered structures of RNase A at pH 2.5 and ambient pressure. (**b**) Corresponding pressure-dependent data at $T = 20°C$

Figure 5.5 shows the p, T-phase diagram of RNase A and further monomeric proteins [7, 51]. For all of these proteins, an ellipsoid-type stability region has been observed, though the individual pattern is characteristic of each protein. Interestingly, some proteins such as chymotrypsin exhibit a small region of pressure-induced folding in the phase diagram at elevated temperatures. Differences in the orientation and inclination of the ellipse are largely due to differences in the heat capacity, $\Delta_u C_p$, and expansion coefficient, $\Delta_u \alpha$, upon unfolding.

5.4.2 *Protein Folding Kinetics*

Also with respect to the kinetics of the folding reaction, pressure studies are of particular use. Upon crossing the p, T-phase boundary of a protein through a p-jump,

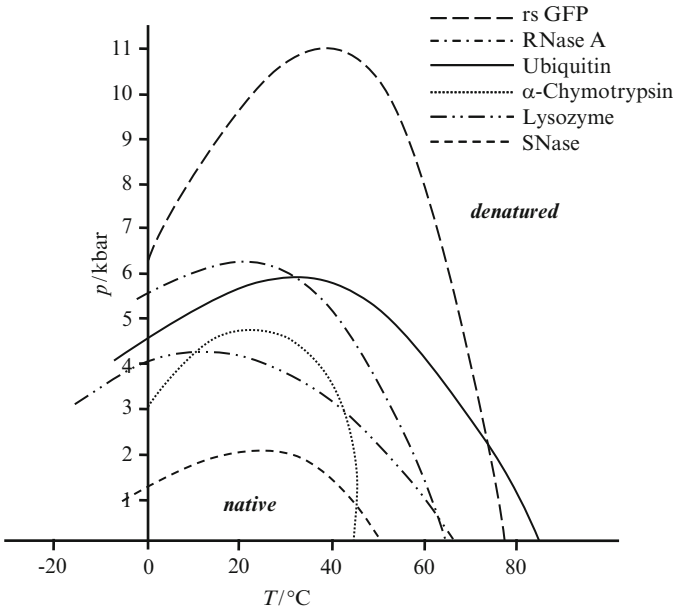

Fig. 5.5 The p, T-stability diagram of various monomeric proteins

the folding and refolding kinetics can be studied [52, 53]. Such kinetic data allow to evaluate the volume profile during the folding process and to characterize the nature of the barrier to folding or unfolding and the corresponding transition state. Moreover, pressure studies present an important advantage due to the generally observed positive activation volume for folding, the result of which is to slow down the folding reaction substantially, in turn allowing for relatively straightforward measurements of structural order parameters characteristic for folding intermediate states, which are difficult or even impossible to quantify on much faster timescales corresponding to ambient pressure conditions.

As an example, we discuss results on the folding reaction of SNase, which unfolds at ambient temperature at about 2 kbar (see Fig. 5.5). Figure 5.6a shows the deconvoluted infrared spectra of SNase after a pressure jump from 1 to 3 kbar at $T = 25°C$. The data reveal a significant change in the shape of the IR spectra during the first 15 min after the pressure jump. The band at a wavenumber of 1,627 cm^{-1}, which is assigned to β-sheets, dominates the FTIR spectra of the native protein. This peak gradually disappears and only a slight shoulder is evident after about 15 min at which point the spectrum exhibits a maximum around 1,641 cm^{-1}. Figure 5.6b presents the time dependence of the fractional intensities of secondary structure elements. The fractional band intensities of β-sheets and α-helices markedly decrease over the first 15 min after the pressure jump from 1 bar to 3 kbar, with a concomitant increase of the fractional intensity of the band at 1,641 cm^{-1}, which is due to disordered and turn structures. These data fit to a simple

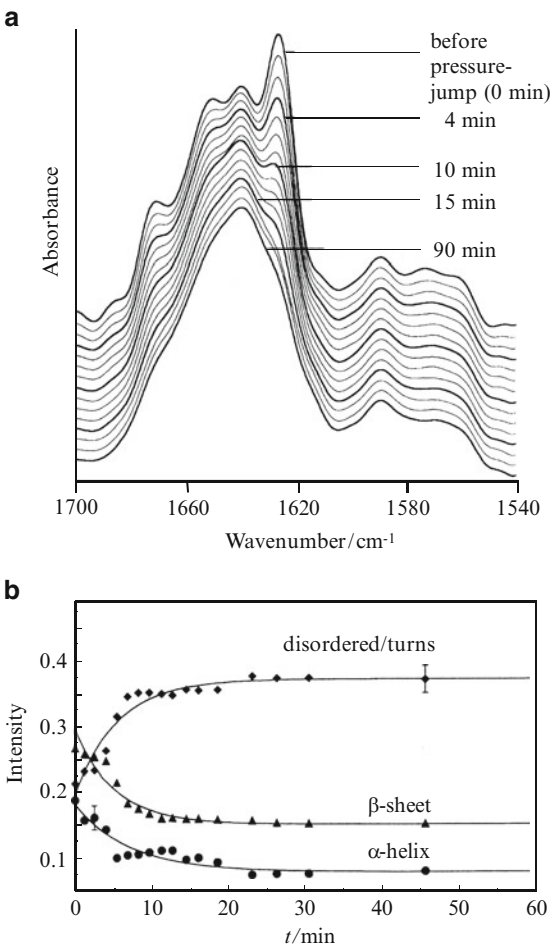

Fig. 5.6 (**a**) Deconvoluted FTIR spectra of SNase (5 wt%, pH 5.5, $T = 25°C$) after a ms-pressure jump from 1 to 3 kbar. (**b**) Time-dependent fractional intensities of secondary structure elements of SNase after a pressure jump from 1 to 3 kbar ($T = 25°C$). The lines represent a single-exponential fit to the data points (after [31])

two-state model assuming first-order kinetics for this relaxation process. Within the accuracy of the experiment, the calculated relaxation times τ at 3 kbar of the different secondary structure elements are similar and are on the order of $\tau \approx 5$ min. The refolding process proceeds much faster (data not shown). It is characterized by a comparable rapid formation of β-sheets and α-helices from disordered and turn protein segments as indicated by the increase of their band intensities within the first minute after the pressure jump from 3 kbar to 1 bar ($\tau \approx 5$ s) [32, 33].

Complementary pressure-jump small-angle X-ray scattering (SAXS) and fluorescence spectroscopy data provide strong evidence for a two-state folding/unfolding model for SNase under pressure, in that secondary structure, chain collapse and tertiary structure all exhibit relaxation profiles on similar timescales. Applying Eyring's transition state theory (i.e. assuming a single equilibrium transition state), we find that the activation volume for folding is large and

positive (57 mL mol^{-1}) and that for unfolding seems to be small and negative (−23 mL mol^{-1}). The volume of the protein solvent system in the transition state is thus significantly larger than in the unfolded state and somewhat smaller than in the folded state, so that the transition state lies closer to the folded than to the unfolded state in terms of system volumes. The positive activation volume of folding indicates that the transition state is accompanied by dehydration and chain collapse (with its accompanying packing defects). Compared to the urea and the temperature-induced unfolding/refolding processes at ambient pressure, the pressure-induced relaxation times of unfolding and refolding are much slower, in agreement with theoretical considerations by Hummer et al. [54]. Conversely, in RNase A, not only the temperature-induced but also the pressure-induced unfolding/refolding reaction is a multistep process, pointing to a more rugged folding funnel [33].

5.4.3 Pressure-Assisted Cold Denaturation of Proteins

The usually observed positive denaturational increment of the heat capacity, $\Delta_u C_p > 0$, of proteins implies that the enthalpy change of protein denaturation is a strong temperature-dependent function. Hence, the enthalpy of denaturation can, in principle, become zero and then even invert its sign at some low enough temperature, changing from the factor stabilizing the native protein structure into a factor destabilizing its structure. Therefore, protein denaturation can occur not only upon heating but also upon cooling. In contrast to heat denaturation, cold denaturation should then proceed with a release of heat, i.e. a decrease of enthalpy. While cold denaturation of proteins is now considered to be a very general phenomenon, its detection and characterization encounter numerous obstacles, mainly because it occurs for most proteins far below the freezing point of water. Many works attempted to overcome this problem by using cosolvents or codenaturants. However, pressure-assisted cold treatment makes a welcome alternative. This approach takes advantage of the fact that under the hydrostatic pressure of 2.2 kbar the freezing point of heavy water drops to about −19°C. By these means, the pressure-assisted cold-denatured state of many proteins is accessible. This method has been successfully applied in FTIR and NMR spectroscopic studies of the cold denaturation of proteins. FTIR spectroscopy was employed for the examination of pressure-assisted cold unfolding of a protein for the first time by Meersman et al. [55]. The study demonstrated that the cold-denatured state of metmyoglobin, just like the pressure-denatured state but opposite to the heat-denatured state, features some persistent secondary structure.

A study by Dzwolak et al. [56] revealed that α-lactalbumin undergoes cold unfolding only in the calcium-depleted form, while cooling of the native calcium-bound protein resulted in formation of non-native α-helices. Further experiments on SNase and other proteins revealed that this kind of unfolding scenario in the cold is also a “mild” one, leaving considerable residual structure of the protein intact. As an example, Fig. 5.7a shows the self-deconvoluted FTIR equilibrium spectra of

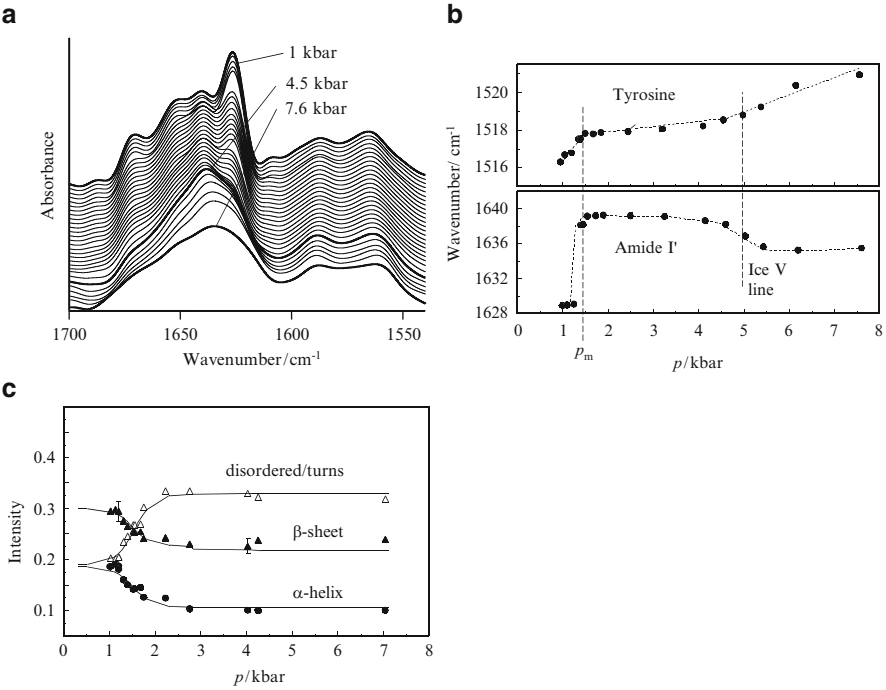

Fig. 5.7 (**a**) Deconvoluted FTIR spectra of SNase (5% (w/w), pD 5.5) at −4.5°C as a function of pressure (band assignment: 1,611 cm^{-1} side chains, 1,627/1,673 cm^{-1} β-sheets, 1,651 cm^{-1} α-helices, 1,641/1,659/1,666 cm^{-1} disordered structures/turns). (**b**) Pressure dependence of the position of the tyrosine band (*top*) and the maximum of the amide I′ band (*bottom*). (**c**) Pressure effect on the areas of the bands associated with β-sheets, disordered/turn structures and α-helices (after [57])

SNase at −4.5°C up to pressures of 7.5 kbar [57]. Figure 5.7b exhibits the pressure-dependent maximum wavenumbers of the amide I′ and tyrosine absorption band. The tyrosine groups of native SNase absorb at about 1, 515 cm^{-1}, and the location of the absorption maximum depends sensitively on its environment (exposure to water leads to an 1–2 cm^{-1} increase in wavenumber). Hence, following tyrosine band absorption reveals additional information about tertiary structural contacts. Clearly, a transition pressure from the native to a denatured state occurs around 1.4 kbar for that temperature. As the features of the tyrosine and amide I′ spectra occur concomitantly, we may assume that the changes of the secondary structure elements and tertiary contacts around the tyrosine residues occur at the same pressure, pointing to a two-state kind of transition for this pressure-assisted cold-denaturation process as well. Above about 4.5 kbar, the maxima of the bands shift again, which is due the pressure-induced formation of ice V. The content of secondary structure elements is not influenced by the freezing of the bulk liquid, however. The secondary structure of the cold-denatured state is similar to that of the pressure-unfolded state at room temperature [31, 32, 57].

5.4.4 Pressure Effects on Oligomeric Proteins and Chaperones

The stability diagrams and the folding reactions of oligomers and protein complexes are more complicated. Oligomeric proteins and multiprotein assemblies often dissociate into individual subunits already at pressures of 0.5–2 kbar [7–11, 58]. The effect of pressure in promoting dissociation may be explained by the imperfect packing of atoms at the subunit interface and by negative volume changes resulting from the disruption of polar and ionic bonds in the intersubunit region. After dissociation by pressure, subunits may undergo further conformational changes at higher pressures. For example, LADH (liver alcohol dehydrogenase), which consists of two 40 kDa subunits and binds the coenzyme nicotinamide adenine dinucleotide (NAD^+) and Zn^{2+} ions, exhibits dissociation of dimers above about 1 kbar, and the dimer ↔ monomer equilibrium is shifted to the right with increasing pressure. As expected, above a pressure of 1 kbar, also the activity of the enzyme ceases [7, 58].

As a further example, we discuss here the effect of pressure on α-crystallin, the major structural protein in mammalian eye lenses. α-Crystallin shares sequence similarity with many small heat-shock proteins (HSP), and its chaperone-like activity has been demonstrated for a number of systems. It can protect a variety of proteins against thermal and non-thermal aggregation in vitro, and this mechanism seems to function in vivo also in the eye lens [59, 60]. Hence, failure of α-crystallin to serve as a chaperone may play a role in the pathomechanism of cataract. The protective activity involves binding of denatured polypeptides to the hydrophobic α-crystallin to form larger soluble complexes (ATP is not required). The protein consists of two ∼20 kDa subunits A and B with high sequence similarity, and the secondary structure is about 17% α-helix, 33% β-sheets and ∼50% turns and coils. Although the protein is highly soluble, the subunits easily self-aggregate forming oligomers of about 800 kDa on average with high polydispersity [60]. Interestingly, pressure perturbation studies on α-crystallin revealed a significant increase in the chaperone activity in the lower pressure range [61, 62]. Figure 5.8 exhibits FTIR spectra of a 3 wt% α-crystallin solution at 25°C as a function of pressure. The IR bands and 1,617 and 1,682 cm^{-1} in the spectrum can be assigned to strong hydrogen bonds in intermolecular antiparallel β-sheets. With increasing pressure, this contribution decreases, indicating pressure-induced dissociation of the oligomers up to about 5 kbar, where no oligomers are detected anymore. At pressures above about 1–2 kbar, significant loss of ordered secondary structure is observed at the expense of formation of unordered structures, indicating reorganization of secondary structure elements and partial unfolding of the protein. This process is complete at about 6 kbar. Thus, in the range of 1–3 kbar, where the pressure treatment has been found effective in increasing the chaperone potency, loosening of the oligomeric organization occurs. After pressure release, reorganization of the oligomeric structure occurs again. Owing to these quaternary structural changes and the concomitant increase in the dynamics of the constituents, more hydrophobic patches are probably exposed, i.e. the amount of binding sites for damaged proteins increases. Such behaviour has been observed for other members of HSP as well [62].

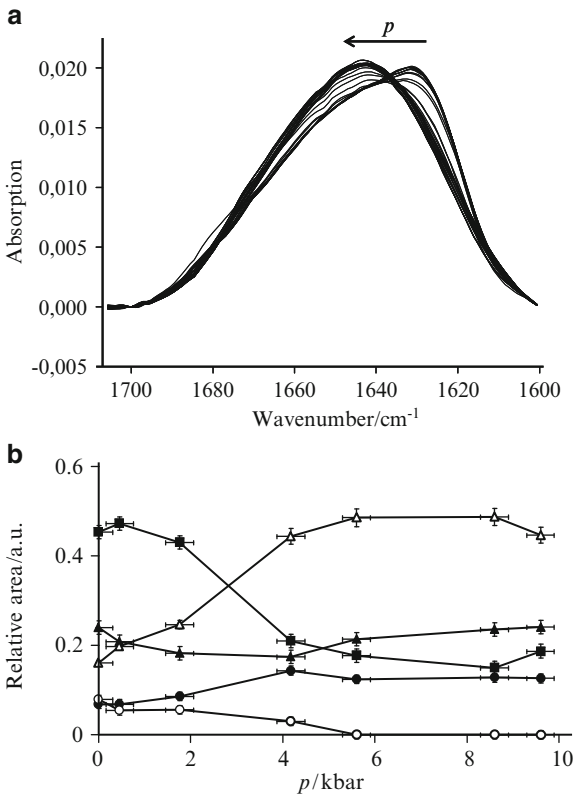

Fig. 5.8 (**a**) FTIR spectra of a 3 wt% α-crystallin solution in D_2O/Tris buffer, at pH 7, taken in 1,000 bar steps between 1 and 10,000 bar. The *arrow* indicates changes with increasing pressure. (**b**) Pressure dependence of the relative areas of the diagnostic amide-I′-bands of α-crystallin. Denotation: *open triangles* disordered structures; *black squares* intramolecular β-sheets ($1629\,cm^{-1}$); *black triangles* α-helical structures ($1657\,cm^{-1}$); *black circles* turns ($1676\,cm^{-1}$); *open circles* intermolecular β-sheets ($1682/1617\,cm^{-1}$)

5.4.5 Cosolvent Effects

It is well known that the cytoplasm of the cell comprises a complex aqueous solution of a variety of different salts and osmolytes and that the stability of proteins is modulated by the addition of these reagents. Some cosolvents, such as sugars, act as protein stabilizers, whereas others, including urea, denature proteins. Organisms living under extreme conditions, in order to protect proteins from denaturation, accumulate protein stabilizers in the cytoplasm. In addition, also the presence of other biomolecules in the cell affects the folding, stability and function of proteins (so-called molecular crowding effect). Hence, calculation of the stability diagrams of proteins also requires detailed knowledge of the effects of cosolutes and cosolvents on the thermodynamic properties of the proteins, which are largely determined by their solvational properties. We have explored the effect of various types of cosolutes and solvents on the temperature and pressure stability of SNase and other proteins [63, 64]. Changes in the denaturation temperature (T_m) and pressure (p_m), the volume change ($\Delta_u V^o$) and the standard Gibbs free energy change ($\Delta_u G^o$) of unfolding were obtained in the presence of different concentrations c of the

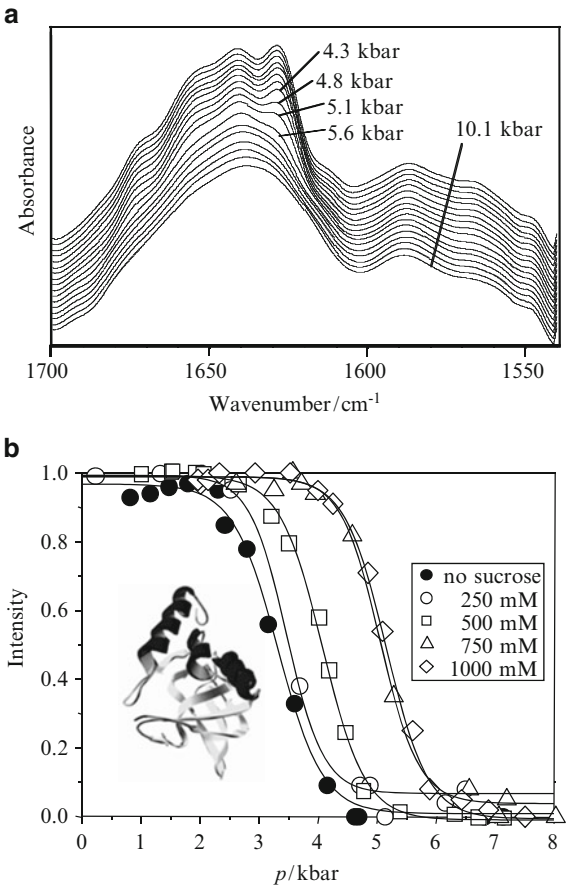

Fig. 5.9 Fourier self-deconvoluted infrared spectra of SNase (**a**) in the presence of 1 M sucrose at different pressures up to 10.1 kbar and (**b**) denaturation profiles of SNase in the presence of selected concentrations of sucrose at pD 5.5 and $T = 21°C$. The inset shows the structure of native SNase (after [64])

additives. Additionally, FTIR difference spectra were recorded and analysed in an effort to detect conformational changes in the native and unfolded state of the protein in the presence of the different additives.

As an example, Fig. 5.9a shows the FTIR spectra of SNase in the presence of 1 M sucrose. In Fig. 5.9b, the denaturation profiles of SNase in solution with different concentrations of sucrose are depicted. The volume change of unfolding, $\Delta_u V^o$, at the temperature T, and the standard free Gibbs energy change at pressure p, $\Delta G^o(p)$, were determined from the denaturation profiles as described above. The standard Gibbs free energy change $\Delta_u G^o$ at 1 bar is obtained by extrapolation of $\Delta_u G^o(p)$ to 1 bar. Up to 250 mM sucrose, there is no significant change in the denaturation pressure p_m. At higher concentrations, up to a concentration of 1 M sucrose, p_m increases roughly linearly at a rate of 2.3 ± 0.3 kbar M^{-1} with increasing sucrose concentration. The $\Delta_u V^o$ value of SNase with increasing concentration of sucrose appears to be only slightly smaller than that of SNase in pure buffer solution ($\Delta_u V^o \approx -65$ mL mol^{-1}) and is nearly independent of the concentration

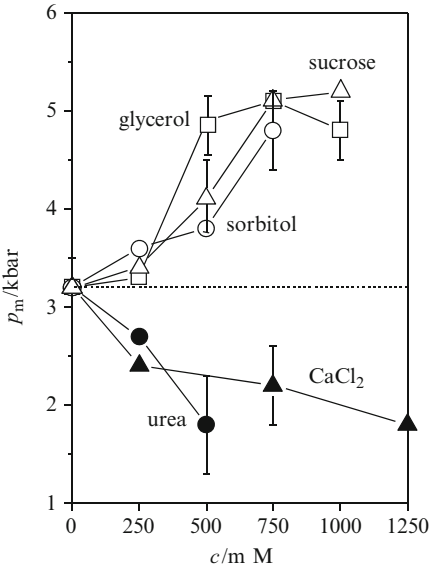

Fig. 5.10 Denaturation pressure p_m of SNase at different concentrations (c) of cosolvents (*open square* glycerol, *open circle* sorbitol, *open triangle* sucrose, *filled triangle* calcium chloride, and *filled circle* urea) (after [64])

of sucrose ($\Delta_u V^o \approx -70\,\text{mL mol}^{-1}$). No significant changes in the conformation of the denatured state are observed with respect to SNase in pure buffer solution. For higher sucrose concentrations, however, slightly more ordered secondary structure elements are retained in the pressure-denatured state [64]. Figure 5.10 depicts the comparison with further polyols and salts. With increasing sucrose concentration, the protein is preferentially hydrated; i.e. sucrose is largely excluded from the protein interface with the result of an increase in $\Delta_u G^o$. Since the unfolded state exposes more surface area than the folded state, the difference in preferential hydration between the two states is quite large, and this difference should become larger as the concentration of polyols increases.

With regard to temperature-induced denaturation, a different behaviour of $\Delta_u V^o(c)$ is observed, which is largely due to the strong temperature dependence of $\Delta_u V^o$ [63]. A more disordered structure in the pressure-denatured state, with regard to SNase without a cosolvent, is observed in the presence of chaotropic agents, such as urea or $CaCl_2$ (Fig. 5.10) [64]. The results for the various chaotropic and kosmotropic cosolvents clearly show that the cosolvents or salts not only markedly change the stability of proteins, but also their solvation, and they may also alter the conformation of the proteins in their unfolded, denatured states.

5.4.6 *Aggregation/Fibrillation Reactions of Proteins*

Protein folding is one of the most crucial steps during the life of a protein. If some failure occurs in achieving the native conformation, this will render the polypeptide inactive, or even worse, it can produce a misfolded molecule

that can interfere or block components of the cellular machinery to the point of causing cell malfunction or even death. In recent years, it has become evident that a wide range of human diseases are associated with aberrations in the folding process. These diseases, which are also coined "protein conformational diseases", include Alzheimer's disease (Aβ-peptide), Parkinsons's disease (α-synuclein), prion-protein-related encephalopathies, dialysis-related amyloidosis (β_2 microglobulin) and type II diabetes mellitus (islet amyloid peptide, IAPP) [16]. In an attempt to probe the stability and energetics of amyloid fibrils, besides temperature, pH and cosolvent dependent studies, also pressure perturbation has been in the focus, recently [34, 65–75]. High-pressure FTIR spectroscopy is apparently a well-suited method for this purpose, as the formation of amyloid fibrils is usually accompanied by the formation of parallel or antiparallel intermolecular β-sheets, which are easily distinguishable from IR spectra of native proteins.

As an example, Fig. 5.11 illustrates how high-pressure prevents aggregation of insulin upon prolonged heating at 70°C, which, under atmospheric pressure (Fig. 5.11a), causes infrared spectral changes indicative of the complete α/β-refolding and aggregation (spectral shift of the conformation sensitive amide I′ band from 1650.4 to 1624.7 cm^{-1}), preceded by the full solvent exchange (spectral shift of the amide II band from 1546.2 to 1445.8 cm^{-1}). If the 70°C incubation of insulin is carried out under high pressure, only solvent exchange takes place, while the absence of any significant spectral changes in the former band indicates no aggregation (Fig. 5.11b). This effect has been explained in pressure perturbation calorimetry studies on protein aggregation by a volumetric expansion that must accompany formation of the insulin intermediate, which is prevented under high pressure [76]. Uncommon to their precursors as discussed above, the mature insulin fibrils are pressure-insensitive, demonstrating that extensive hydrogen bonding and optimised side chain packing are crucial for their stability [7, 67].

Generally, the susceptibility of protein aggregates to pressure largely depends on the degree of the structural order of an aggregate. Fresh, amorphous aggregates are more sensitive to pressure and prone to refolding to the native state than mature amyloid fibrils. In the former case, the oligomers are stabilized mainly by electrostatic and hydrophobic interactions, yet they lack the precise packing of mature fibrils. This renders them susceptible to pressure-induced dissociation. In the latter case, effectiveness of pressure-induced dissociation depends on the particular mode of polypeptide backbone and side chain packing that allows reducing remaining void volumes. The pressure sensitivity of fresh aggregates and the virtual insensitivity of mature fibrils allows not only to differentiate between various stages of the amyloid formation, but also to obtain reliable thermodynamic data, such as Gibbs free energy and molar volume changes, of the early stage of the transition – the formation of the amorphous β-pleated sheet structure. This is only possible due to the reversibility of the process under high-pressure conditions.

One of the most important motives of high-pressure research on protein aggregates is rescuing – under high hydrostatic pressure – the native oligomeric and conformational state of a protein from aggregation. Some amyloid fibrils (in particular protofibrils) might exhibit a larger partial specific volume than that of

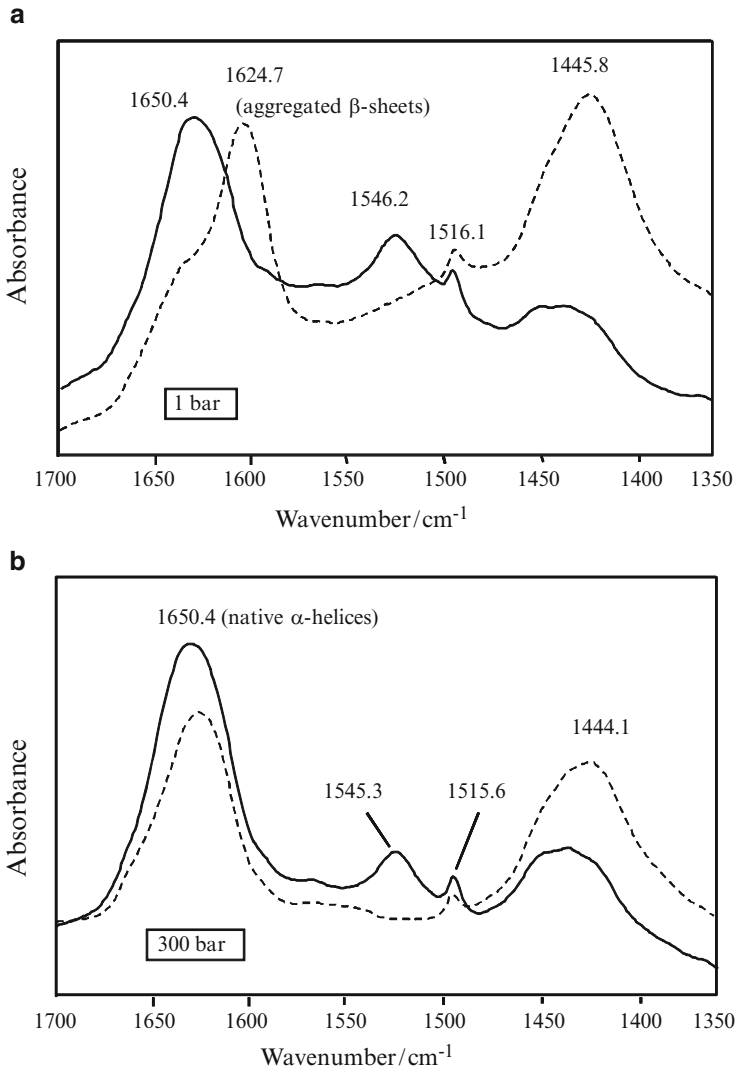

Fig. 5.11 FTIR spectra of bovine insulin (2 wt. % in D_2O, pD 1.9) fibrillating at 70°C at (**a**) atmospheric and (**b**) 300 bar pressure. *Solid lines* denote fresh, native protein, and the *dashed lines* correspond to the protein after 5 h of high-temperature incubation (after [76])

the composing proteins, because the assembly process, which involves various conformational transitions of native structures to non-native β-sheets, could create new water-excluded cavities and induce hydrophobic pockets. In these cases, HHP has the potential to dissociate also amyloid fibrils. For example, it has been shown that moderate pressures (1–3 kbar) can dissociate amyloid fibrils of lysozyme, transthyretin, β_2-microglobulin and α-synuclein. One of the most

thrilling prospects for application of high pressure in protein aggregation research is the idea of destroying prion infectivity through combined temperature–pressure treatment [74, 75]. The prion-type infectivity putatively stems from proliferation of the PrP^{Sc} protein template at the expense of its normal cellular isoform – the PrP^{C} monomer. The amyloidal nature of the PrP^{Sc} with its unmatched resistance to high temperature, proteases and many chemical denaturants has posed a significant challenge to medical and biotechnological protocols, leaving – for instance – no reliable tool of decontamination of neurosurgical equipment, or of ridding prions from beef (vide the BSE crisis). However, in light of the pronounced dissociative effect that pressure has unravelled to have on some amyloids, it was soon inquired whether the prion amyloid itself may be also irreversibly destroyed under pressure.

5.4.7 Enzymatic Reactions

In its simplest expression, an enzyme reaction comprises a series of elementary steps finally leading to the release of products. The effect of pressure on the overall reaction may be the consequence of its effects on the reaction itself, on enzyme conformational changes, on enzyme dissociation into subunits or on substrate and product levels [77–80]. In addition, the enzyme reaction may involve the interaction with the surrounding solvent, whose viscosity changes under pressure as well. Understanding of the principles of enzyme function under pressure is still in its infancy [77, 78], although experimental effects of pressure on enzyme activity have a long history, beginning with Eyring's study of invertase in 1946 [81]. Generally, kinetic models allow one to determine the activation volume $\Delta V^{\neq}$ of a reaction (i.e. the difference between the volume of the transition state and the reactant ground states) from the dependence of a rate constant, k, on pressure, according to:

$$k(p) = k_{\mathrm{o}} \mathrm{e}^{-p \cdot \Delta V^{\neq}/RT} \tag{5.3}$$

where k_{o} is the rate constant for the chemical transformation under atmospheric pressure conditions [78]. In a pioneering work, Wong et al. used high-pressure infrared spectroscopy in a series of enzyme kinetic studies [82]. Here we show one impressive example, the conversion of *p*-nitrophenol phosphate disodium into *p*-nitrophenol by alkaline phosphatase. The rate constant k of the enzyme conversion was determined at various pressures by monitoring the decrease with time in the intensity of the IR band of the substrate at 887.6 cm^{-1} at these pressures, which is depicted in Fig. 5.12. The rate constant increases almost linearly with increasing pressure up to 7.2 kbar and then decreases abruptly at ~8.3 kbar, a pressure, where drastic changes in the band shape and intensity in the amide-I region start to take place, i.e. the loss of enzyme activity of the alkaline phosphatase is the result of denaturation of the enzyme. According to transition state theory, an activation volume $\Delta V^{\#}$ of the enzyme reaction can be calculated from the slope of the $\ln k$ vs. p plot, yielding pressure-dependent values of about −9 mL mol^{-1} [82].

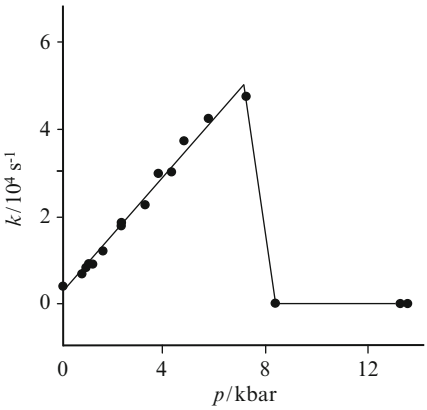

Fig. 5.12 Pressure dependence of the alkaline phosphatase reaction rate constant for the conversion from *p*-nitrophenol phosphate to *p*-nitrophenol (after [82])

5.4.8 Synthetic Polymers as Protein Mimetics

There are several known synthetic polymers that exhibit a more or less distinct change in their molecular level aggregational states in solution and volume phase transitions in their gel form (so-called thermoresponsive polymers) [83–88]. Poly(*N*-isopropylacrylamide) (PNIPAM) and other polyacrylamide derivatives with both hydrogen bonding properties and hydrophobic interactions in aqueous solution are some of the best studied examples. They show a reversible transition from their coiled (swollen) to collapsed (compact) states upon an increase in temperature (so-called inverse temperature or volume phase transition temperature). The swelling and shrinking of gels prepared from such polymers have also been extensively studied [83–88]. Besides temperature, these thermoresponsive polymers can undergo transitions by changing other factors, such as the salt or cosolvent concentration. Furthermore, PNIPAM and derivatives of this polymer undergo pressure-induced phase changes, both in solution and in their gel forms. Such changes are reminiscent of the behaviour of proteins. Actually, PNIPAM may be considered as a simple protein model and its thermodynamic properties have been investigated from such point of view [84–88]. The coil to collapse volume phase transition (which is also called cloud transition if accompanied by a drastic increase in turbidity) of these polymers upon increasing temperature corresponds to the transition from an unfolded protein to the folded structure across the cold-denaturation temperature, i.e. conveys a model of the process of cold denaturation of proteins. The temperature–pressure diagram for the cloud transition of PNIPAM was measured in aqueous medium and is depicted in Fig. 5.13a [84]. Near or at atmospheric pressure, the transition temperature increases with increasing pressure, but at higher pressures it decreases with increasing pressure. The experimentally obtained ΔC_p value (in the direction from the coil to the collapsed state) is negative, which is equivalent to a positive ΔC_p for cold denaturation (in the direction from the native to the denatured state).

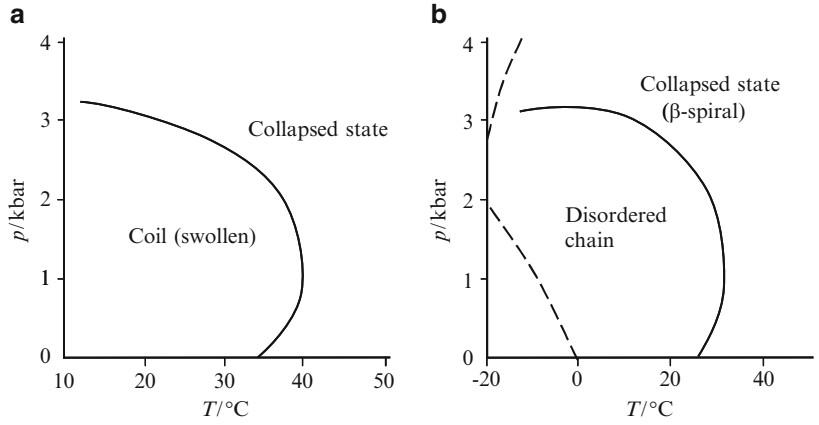

Fig. 5.13 (**a**) The p, T-phase diagram for the cloud point of a 0.5 (wt/v) % PNIPAM solution (after [85]). The exact location of the cloud point depends still on the molecular mass of the polymer, the pH and ionic strength. (**b**) The p, T-phase diagram of 0.05 (w/v) % $(VPGIG)_{40}$ in aqueous solution; *dashed line* freezing point of water under elevated pressure (modified after [86]). A roughly ellipsoidal cloud-point diagram is observed in the experimentally accessible temperature and pressure range, reminiscent of the denaturation contours of chymotrypsin (see Fig. 5.5)

Figure 5.14a illustrates the development of the amide-I′ spectra of PNIPAM upon heating, whereas the relative areas of the detected subbands, which represent the respective relative parts of the total hydrogen bonding pattern, in dependence of temperature at ambient pressure are displayed in Fig. 5.15a. The main diagnostic bands are located at ~1,652, ~1,626, ~1,606 and ~1,583 cm^{-1}, which can be assigned to intramolecular hydrogen bonding of carbonyl groups to amide deuterons, carbonyl groups in a medium hydrophobic environment and hydrogen bonding of carbonyl groups to D_2O (the latter both), respectively [83]. The bands at ~1,626 cm^{-1} and at ~1,606 cm^{-1} are dominant below 30°C, indicating a rather hydrophilic, coil-like state. Between 30 and 35°C, the band at 1,606 cm^{-1} shows a steep decrease and eventually vanishes at ~40°C. Also, the relative fraction of the band at ~1,626 cm^{-1} decreases, whereas the band at ~1,652 cm^{-1}, representing intramolecular hydrogen bonds, shows a steep increase. The midpoint of the transition appears at ~32°C, in good agreement with the location of the cloud point; i.e. the spectral data reflect nicely the coil-to-globule transition of PNIPAM [9, 27].

Upon application of high hydrostatic pressure at constant temperature, the polymer system shows an inverse behaviour when compared to temperature elevation. Figure 5.14b exhibits representative FTIR spectra of the pressure-dependent measurements of linear PNIPAM at 50°C. At 50°C, i.e. above the LCST, spectral changes start already between 1 and 500 bar. Up to ~5,400 bar, the relative fraction of the subband for intramolecular hydrogen bonds has decreased from ~50 to 17% and the relative fraction of the subband for carbonyl groups in medium hydrophobic environment has increased from ~50 to ~79%. Above ~5,400 bar, a new band for hydrogen bonding to D_2O emerges at 1,612 cm^{-1} and starts to increase together

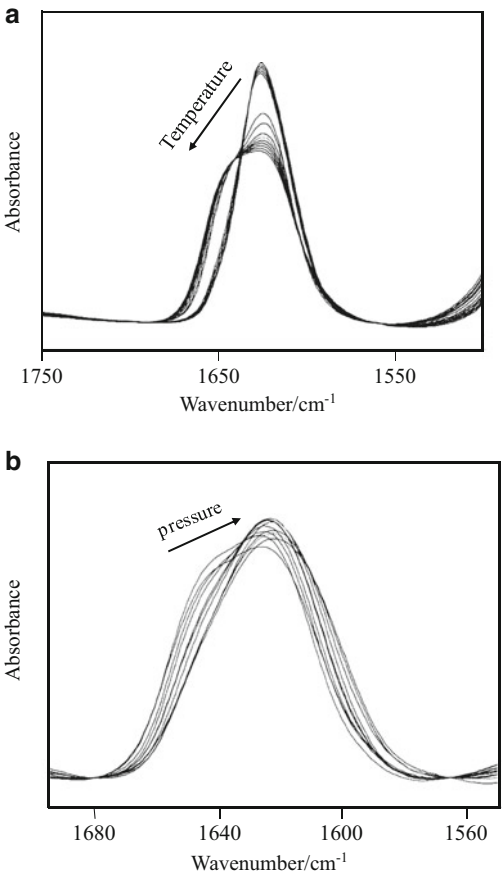

Fig. 5.14 (**a**) Normalized amide-I′–FTIR spectra of a solution of 2 (w/v) % PNIPAM in D_2O from 10 to 80°C at ambient pressure. The *black arrow* indicates the development of spectra with increasing temperature. (**b**) Normalized amide-I′–FTIR spectra of linear PNIPAM from 1 to ∼11,400 bar at 50°C. The *black arrow* indicates the direction of increasing pressure

with the other band for hydrogen bonding to the solvent (at $\sim$1,595 cm^{-1}) up to the maximal pressure of 11,400 bar. In this upper pressure range, a further slight pressure-induced increase in swelling of the polymer is observed, only.

These results show that the polymer is becoming more hydrated under high hydrostatic pressure conditions; i.e. pressure triggers the transition from the collapsed, globule-like state to a swollen coil-like state, which might be due to a weakening of hydrophobic interactions upon pressurization. The opposite effects of temperature and pressure are illustrated in Fig. 5.15c. Such results are in good agreement with high-pressure FTIR investigations of PNIPAM at 25°C, which also indicate the formation of more hydrated configurations of PNIPAM upon compression [87]. Assuming a two-state model for the transition, the associated free energy changes for the swelling process can be calculated. The calculated values for the globule-to-coil transformation are about $\Delta G^{o} = 4.2\,kJ\,mol^{-1}$ and $\Delta V^{o} = -29\,mL\,mol^{-1}$ at 50°C and ambient pressure, which is of the same order of magnitude as corresponding values for the pressure denaturation of proteins.

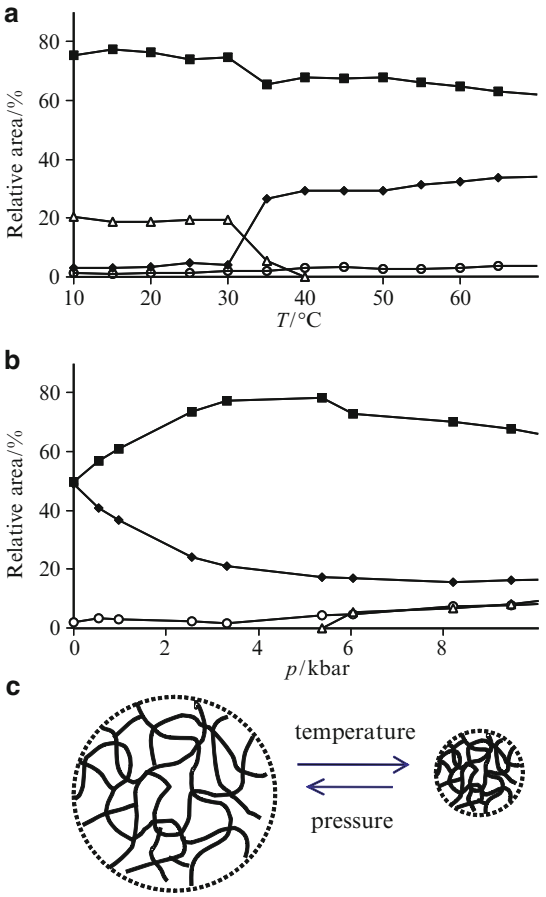

Fig. 5.15 (**a**) Temperature and (**b**) pressure dependence (at $T = 50°C$) of the relative areas of diagnostic amide-I′ bands of a 2 (w/v) % PNIPAM solution in D_2O. Symbol denotation: *filled diamond* intramolecular hydrogen bonds ($\sim$1652 cm^{-1}), *filled square* carbonyl groups in a medium hydrophobic environment ($\sim$1,626 cm^{-1}), *open triangle* hydrogen bonds to D_2O ($\sim$1,606 and $\sim$1,583 cm^{-1}). (**c**) Schematic drawing of the opposite effect of temperature and pressure on the swollen-to-collapsed transition of the PNIPAM polymer (after [89])

The negative ΔV^o value indicates the presence of significant void volume in the collapsed state. At sufficiently high pressures, continuous shrinking of the swollen state may occur again [88].

Apparently similar phase diagrams were obtained for the cloud-point transition of aqueous solutions of elastin-like polypeptides such as with the repeating pentapeptide Val–Pro–Gly–Ile–Gly sequence in $(VPGIG)_{40}$ (Fig. 5.13b) [85]. The $(VPGIG)_{40}$ solution exhibits a well-defined pressure-induced cloud point as well as a temperature-induced one, and this transition has been shown to occur through the collapse from an extended chain to a β-spiral-like structure with three VPGVG units per turn, each pentamer adopting a type II β-turn conformation. Calorimetric measurements on this polypeptide indicated that ΔH and ΔS of the transition are around 12 kJ (mol pentamer)$^{-1}$ and 30 J (mol pentamer)$^{-1}$ K^{-1}, respectively (in the direction from the coil to the collapse state), and ΔC_p was small but negative (-0.5 kJ mol^{-1} K^{-1}) [84]. The temperature- and pressure-dependent structure and phase behaviour of a shorter oligopeptide, GVG(VPGVG), which serves as a minimalistic elastin-like model system, has been studied as well [90]. This octapeptide

behaves as a two-state system which undergoes the inverse-temperature folding transition at ∼36°C, and, in addition, a slow trend reversal at higher temperatures ($T > 90$°C), finally leading to a reentrant unfolding close to the boiling point of water. The pressure dependence of the FTIR spectra of GVG(VPGVG) reveals that the fractional intensities of type II β-turn-and-loop structures decrease, whereas the population of disordered (non-periodic) structures increases with increasing pressure. The underlying driving force is probably the decrease in ΔV° as a result of the release of void volume in the folded state of the oligopeptide. Hence, these data suggest that elastins behave like other, ordinary proteins, with the crucial difference that the stability window of their folded (ordered) state approaches the boiling point of its solvent, water. This behaviour is a consequence of the essentially hydrophobic nature of its amino acid residues, which, as a consequence, places the hydrophobic hydration contribution to the protein's stability among the most prominent ones. The thermodynamic data obtained suggest that the folding is a consequence of hydrophobic interactions being largely driven by the increase in the entropy of water.

5.5 Conclusions and Outlook

We conclude that – besides by changing temperature and the chemical potential of additives – also pressure work on biomolecular systems can yield a wealth of enlightening new information on their structure, energetics, phase behaviour and on their dynamics and kinetic properties. For example, the extension of the thermodynamic characterization of globular proteins to the pressure axis increases our understanding of proteins as well as our ability to manipulate conformational states of proteins for practical purposes. It is clear, however, that the application of the pressure variable in this research area has only just started and many interesting results are expected in the near future. Besides using the pressure variable as thermodynamic and kinetic variable, pressure is also of significant biological and biotechnological relevance. Life on Earth can be traced back to as far as 3.8 billion years ago and may have originated in the depth of the proto-ocean of the Hadean Earth, under high hydrostatic pressure [1]. The large body of water could have filtered harmful radiation and buffered drastic physical–chemical variations, and therefore, could have provided a more convenient environment for pre-biotic biochemistry. Interestingly, whereas only a few families of organisms live in hot, acidic, halophilic, etc. environments, most, if not all, organisms can live under a large range of hydrostatic pressures. To yield a more fundamental understanding of these processes connected to the development of life requires more and extensive high-pressure studies on biological systems at different levels of complexity (proteome, genome, metabolome, etc.), however. Finally, pressure studies are also important for the development of new technological and pharmaceutical applications, such as high-pressure food processing, inactivation of food allergens, separation of protein complexes (e.g. inclusion bodies), supercritical biocatalysis and baroenzymology.

Acknowledgements Financial support from the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

References

1. I. Daniel, P. Oger, R. Winter, Chem. Soc. Rev. **35**, 858 (2006)
2. A.A. Yayanos, Proc. Natl. Acad. Sci. U.S.A. **83**, 9542 (1986)
3. D.H. Bartlett, Biochim. Biophys. Acta **1595**, 367 (2002)
4. M. Gross, *Life on the Edge* (Plenum Trade, New York, 1998)
5. D.A. Wharton, *Life at the Limits. Organisms in Extreme Environments* (Cambridge University Press, Cambridge, 2002)
6. K. Seki, M. Toyoshima, Nature **395**, 853 (1998)
7. R. Winter, D. Lopes, S. Grudzielanek, K. Vogtt, J. Non-Equil. Thermody. **32**, 41 (2007)
8. F. Meersman, C.M. Dobson, K. Heremans, Chem. Soc. Rev. **35**, 908 (2006)
9. J.L. Silva, D. Foguel, C.A. Royer, Trends Biochem. Sci. **26**, 612 (2001)
10. K. Heremans, L. Smeller, Biochim. Biophys. Acta **1386**, 353 (1998)
11. R. Winter, Biochim. Biophys. Acta **1595**, 160 (2002)
12. M.R. Arnold, R.H. Kalbitzer, W. Kremer, J. Magn. Reson. **161**, 127 (2003)
13. K. Akasaka, Chem. Rev. **106**, 1814 (2006)
14. G. Weber, H.G. Drickamer, Q. Rev. Biophys. **16**, 89 (1983)
15. J. Jonas (ed.), *High Pressure NMR* (Springer, Berlin, 1991)
16. F. Chiti, C.M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006)
17. P.W. Bridgman, J. Biol. Chem. **19**, 511 (1914)
18. R.J. St. John, J.F. Carpenter, T.W. Randolph, Proc. Natl. Acad. Sci. U.S.A. **96**, 13029 (1999)
19. R. Mishra, R. Winter, Angew. Chem. Int. Ed. **47**, 6518 (2008)
20. R. Ravindra, R. Winter, ChemPhysChem **4**, 359 (2003)
21. D.M. Byler, H. Susi, Biopolymers **25**, 469 (1986)
22. J. Bandekar, Biochim. Biophys. Acta **1120**, 123 (1992)
23. J.L.R. Arrondo, A. Muga, J. Castresana, F.M. Goi, Prog. Biophys. Mol. Biol. **59**, 23 (1994)
24. L.K. Tamm, S.A. Tatulian, Q. Rev. Biophys. **30**, 365 (1997)
25. M. Jackson, H.H. Mantsch, Crit. Rev. Biochem. Mol. Biol. **30**, 95 (1995)
26. H. Fabian, H.H. Mantsch, Biochemistry **34**, 13651 (1995)
27. W. Dzwolak, M. Kato, Y. Taniguchi, Biochim. Biophys. Acta **1595**, 131 (2002)
28. M. Jackson, P.I. Haris, D. Chapman, Biochim. Biophys. Acta **998**, 75 (1989)
29. L. Smeller, K. Heremans, in *High Pressure Research in the Biosciences and Biotechnology*, ed. by K. Heremans (Leuven University Press, Leuven, 1997), p. 131
30. L. Smeller, K. Goossens, K. Heremans, Vibr. Spectrosc. **8**, 199 (1995)
31. G. Panick, R. Malessa, R. Winter, G. Rapp, K.J. Frye, C. Royer, J. Mol. Biol. **275**, 389 (1998)
32. G. Panik, G.J.A. Vidugiris, R. Malessa, G. Rapp, R. Winter, C.A. Royer, Biochemistry **13**, 4157 (1999)
33. G. Panick, R. Winter, Biochemistry **39**, 1862 (2000)
34. Y. Cordeiro, J. Kraineva, R. Ravindra, L.M.T.R. Lima, M.P.B. Gomes, D. Foguel, R. Winter, J.L. Silva, J. Biol. Chem. **279**, 32354 (2004)
35. M. Zein, R. Winter, Phys. Chem. Chem. Phys. **2**, 4545 (2000)
36. G. Desai, G. Panick, M. Zein, R. Winter, C.A. Royer, J. Mol. Biol. **288**, 461 (1999)
37. H. Herberhold, S. Marchal, R. Lange, C.H. Scheyhing, R.F. Vogel, R. Winter, J. Mol. Biol. **330**, 1153 (2003)
38. H.H.J. de Jongh, E. Goormaghtigh, J.-M. Ruysschaert, Anal. Biochem. **242**, 95 (1996)
39. C.E. Weir, E.R. Lippincott, A. van Valkenburg, E.N. Bunting, J. Res. Natl. Bur. Stand. **63A**, 55 (1959)
40. A. Jayaraman, Rev. Mod. Phys. **55**, 65 (1983)

41. A. van Valkenburg, Rev. Sci. Instrum. **33**, 1462 (1962)
42. P.T.T. Wong, D.D. Klug, Appl. Spectrosc. **37**, 284 (1993)
43. C. Jung, O. Ristau, H. Schulze, S.G. Sligar, Eur. J. Biochem. **235**, 660 (1996)
44. G.J. Piermarini, S. Block, J.D. Barnett, R.A. Forman, J. Appl. Phys. **46**, 2774 (1975)
45. P.T.T. Wong, D.J. Moffatt, F.L. Baudais, Appl. Spectrosc. **39**, 733 (1985)
46. P.T.T. Wong, D.J. Moffatt, Appl. Spectrosc. **43**, 1279 (1989)
47. M. Tsuda, I. Shirotani, S. Minomura, Y. Teryama, Bull. Chem. Soc. Jpn. **49**, 2952 (1976)
48. Y. Kitamura, T. Itoh, J. Sol. Chem. **16**, 715 (1987)
49. R.C. Neumann, W. Kauzmann, A. Zipp, J. Phys. Chem. **77**, 2687 (1973)
50. A. Wlodawer, N. Borkakoti, D.S. Moss, B. Howlin, Acta Cryst. B **42**, 379 (1986)
51. A. Maeno, H. Matsuo, K. Akasaka, Biophysics **5**, 1 (2009)
52. J. Woenckhaus, R. Köhling, R. Winter, P. Thiyagarajan, S. Finet, Rev. Sci. Instrum. **71**, 3895 (2000)
53. J. Woenckhaus, R. Köhling, P. Thiyagarajan, K. Littrell, S. Seifert, C.A. Royer, R. Winter, Biophys. J. **80**, 1518 (2001)
54. G. Hummer, S. Garde, A.E. García, M.E. Paulaitis, L.R. Pratt, Proc. Natl. Acad. Sci. U.S.A. **95**, 1552 (1998)
55. F. Meersman, L. Smeller, K. Heremans, High Press. Res. **19**, 653 (2000)
56. W. Dzwolak, M. Kato, A. Shimizu, Y. Taniguchi, Biopolymers **62**, 29 (2001)
57. G. Panick, H. Herberhold, Z. Sun, R. Winter, Spectroscopy **17**, 367 (2003)
58. R. Winter (ed.), *High Pressure Bioscience and Biotechnology II* (Springer, Heidelberg, 2003)
59. R.C. Augusteyn, Clin. Exp. Optom. **87**, 356 (2004)
60. F. Skouri-Panet, S. Quevillon-Cheruel, M. Michiel, A. Tardieu, S. Finet, Biochim. Biophys. Acta **1764**, 372 (2006)
61. C. Böde, F.G. Tölgyesi, L. Smeller, K. Heremans, S.V. Avilov, J. Fidy, Biochem. J. **370**, 859 (2003)
62. F. Tölgyesi, C.S. Böde, L. Smeller, D.R. Kim, K.K. Kim, K. Heremans, J. Fidy, Cell. Mol. Biol. **50**, 361 (2004)
63. L. Mitra, N. Smolin, R. Ravindra, C. Royer, R. Winter, Phys. Chem. Chem. Phys. **8**, 1249 (2006)
64. H. Herberhold, C.A. Royer, R. Winter, Biochemistry **43**, 3336 (2004)
65. S. Marchal, E. Shehi, M.C. Harricane, P. Fusi, F. Heitz, P. Tortora, R. Lange, J. Biol. Chem. **278**, 31554 (2003)
66. F. Meersman, C.M. Dobson, K. Heremans, Chem. Soc. Rev. **35**, 908 (2006)
67. S. Grudzielanek, V. Smirnovas, R. Winter, J. Mol. Biol. **356**, 497 (2006)
68. K. Akasaka, A.R.A. Latif, A. Nakamura, K. Matsuo, H. Tachibana, K. Gekko, Biochemistry **46**, 10444 (2007)
69. E. Chatani, M. Kato, T. Kawai, H. Naiki, Y. Goto, J. Mol. Biol. **352**, 941 (2005)
70. A.D. Ferrao-Gonzales, L. Palmieri, M. Valory, J.L. Silva, H. Lashuel, J.W. Kelly, D. Foguel, J. Mol. Biol. **328**, 963 (2003)
71. R. Jansen, S. Grudzielanek, W. Dzwolak, R. Winter, J. Mol. Biol. **338**, 203 (2004)
72. G. Panick, R. Malessa, R. Winter, Biochemistry **38**, 6512 (1999)
73. D. Foguel, M.C. Suarez, A.D. Ferrao-Gonzales, T.C.R. Porto, L. Palmieri, C.M. Einsiedler, L.R. Andrade, H.A. Lashuel, P.T. Lansbury, J.W. Kelly, J. L. Silva, Proc. Natl. Acad. Sci. USA **100**, 9831 (2003)
74. J. Torrent, M.T. Alvarez-Martinez, M.C. Harricane, F. Heitz, J.P. Liautard, C. Balny, R. Lange, Biochemistry **43**, 7162 (2004)
75. P. Brown, R. Meyer, F. Cardone, M. Pocchiari, Proc. Natl. Acad. Sci. U.S.A. **100**, 6093 (2003)
76. W. Dzwolak, R. Ravindra, J. Lendermann, R. Winter, Biochemistry **42**, 11347 (2003)
77. C. Balny, in *Effects of High Pressure on Molecular Biology and Enzymology*, ed. by J.M. Markely, C.A. Royer, D.B. Northrop (Oxford University Press, New York, 1996), p. 196
78. D.B. Northrop, Biochim. Biophys. Acta **1595**, 71 (2002)
79. H. Ludwig (ed.), *Advances in High Pressure Bioscience and Biotechnology* (Springer, Berlin, 1999)

80. V.V. Mozhaev, K. Heremans, J. Frank, P. Masson, C. Balny, Proteins: Struct. Funct. Genet. **24**, 81 (1996)
81. E. Eyring, F.H. Johnson, R.L. Gensler, J. Phys. Chem. **50**, 453 (1946)
82. P.T.T. Wong, in *High Pressure Chemistry, Biochemistry and Materials Science*, ed. by R. Winter, J. Jonas (Kluwer Academic Publishers, Netherlands, 1993), p. 511
83. M. Keerl, V. Smirnovas, R. Winter, W. Richtering, Angew. Chem. Int. Ed. **47**, 338 (2008)
84. S. Kunugi, N. Tanaka, Biochim. Biophys. Acta **1595**, 329 (2002)
85. S. Kunugi, K. Kameyama, T. Tada, N. Tanaka, M. Shibayama, M. Akashi, Braz. J. Med. Biol. Res. **38**, 1233 (2005)
86. S. Kunugi, K. Takano, N. Tanaka, K. Suwa, M. Akashi, in *High Pressure Research in the Biosciences and Biotechnology*, ed. by K. Heremans (Leuven University Press, Leuven, 1997), p. 59
87. F. Meersman, J. Wang, Y. Wu, K. Heremans, Macromolecules **38**, 8923 (2005)
88. E. Kato, J. Chem. Phys. **106**, 3792 (1997)
89. M. Pühse, M. Keerl, C. Scherzinger, W. Richtering, R. Winter, Polymer, **51**, 3653–3659 (2010)
90. C. Nicolini, R. Ravindra, B. Ludolph, R. Winter, Biophys. J. **86**, 1385 (2004)

Chapter 6
Dynamics of α-Helix and β-Sheet Formation Studied by Laser-Induced Temperature-Jump IR Spectroscopy

Karin Hauser

Abstract Analysis of folding mechanisms requires a fast initiation method besides site-specific structural probes. Laser-induced temperature-jump IR spectroscopy is an ideally suited method to study α-helix and β-sheet formation in the nanosecond-to-microsecond time domain. The section gives an overview about the T-jump technique and its power to obtain site-specific dynamics by a combination with isotopic editing.

6.1 Peptide Folding Dynamics

6.1.1 Secondary-Structure Formation

Basic steps in protein folding such as the formation of turns, loops, and basic secondary-structure elements occur on the nanosecond to low microsecond timescale [1]. Stopped-flow techniques are limited to the millisecond time range due to their sample mixing times and are thus too slow, whereas pulsed laser techniques are capable to trigger and to analyze faster folding events. The complex folding process of a protein is most often studied with protein fragments or peptides serving as model systems. Infrared (IR) spectroscopy is a suitable method to monitor the molecular dynamics of secondary-structure formation [2–5] since it provides both structure sensitivity and the required time resolution.

K. Hauser (✉)
Biophysical Chemistry, Department of Chemistry, University of Konstanz, Universitätsstr. 10, 78464 Konstanz, Germany
e-mail: Karin.Hauser@uni-konstanz.de

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*, Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_6, © Springer-Verlag Berlin Heidelberg 2012

6.1.2 The Amide I Band as Structural Probe

The amide I vibration is commonly used as IR marker band for secondary-structure determination. It is a vibration of the polypeptide backbone and arises mainly (∼80%) from the CO stretching vibration of the peptide groups with minor contributions from the out-of-phase CN stretching vibration, the CCN deformation, and the NH in-plane bend. The amide I frequency is hardly affected by the nature of the side chains, but it significantly depends on the secondary structure of the backbone [3, 6]. Therefore, it is the most commonly used vibration for secondary-structure analysis. Experimental analysis of the secondary structure of proteins in their native aqueous environment has a long tradition and also theoretical studies of the amide I vibrations have become more and more intense during the last years. Reviews about experimental and theoretical amide I analysis can be found, e.g., in [4, 6–8].

Various secondary-structure elements of a protein absorb at different positions in the amide I region (1,700–1,600 cm^{-1}) of the IR spectrum. However, the frequency components largely overlap in a protein and a more or less featureless broad amide I band is observed in a protein absorption spectrum, in particular if proteins have various secondary-structure elements. Amide I band assignments to secondary-structure elements are commonly made either by band narrowing and curve-fitting for deconvolution [5] or by use of a calibration set of spectra from proteins and peptides with a known structure for pattern recognition [9–12]. Assignments have been collected and evaluated by Goormaghtigh et al. [7] and summarized by Barth and Zscherp [6]. Nevertheless, the amide I band assignment is not unique or straightforward [2]. For example, the amide I bands of α-helix and random coil overlap in protein absorption spectra. Hydrogen/deuterium exchange may partly help for assignment since the NH in-plane bending vibration of the amide I band is sensitive to the N-deuteration of the backbone and thus induces frequency shifts.

All peptide groups contribute to the amide I band. It arises by the same normal mode of the backbone repeat units, which is dominated by the C=O vibration. These polypeptide oscillators couple (Fig. 6.1) and it is the coupling of the amide I modes that is responsible for the secondary-structure dependence of the amide I band.

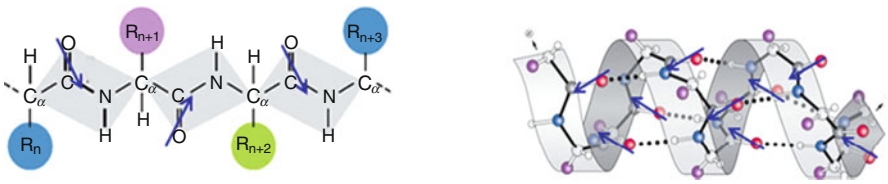

Fig. 6.1 The C=O vibrations dominate the amide I modes. The secondary-structure dependence of the amide I band arises from the coupling of the amide I modes. The coupling occurs via through-bond coupling along the chemical bonds, via the hydrogen bond network and via through-space transition dipole coupling which is the dominant factor that determines the amide I frequency position. *Left*: Polypeptide chain; *blue arrows* indicate the transition dipoles of the amide I modes. *Right*: Illustration of the hydrogen-bonded C=O oscillators within an α-helix

The coupling of the amide I modes occurs in different manners. One is the through-bond coupling, i.e., the coupling of the neighboring peptide group along the chemical bond. Since the amide I vibrations do not involve large displacements of the C_α atoms, one amide I mode will interact only slightly with the neighboring amide I mode, and in consequence, the effect on the vibration of the polypeptide system is minor. It cannot explain, e.g., the characteristic amide I band splitting of antiparallel β-sheet structures with a low- and high-frequency component [13]. Further coupling arises via the hydrogen bond network of the amide groups. Secondary-structure elements are stabilized by hydrogen bonds that connect non-neighbor peptide groups within the polypeptide chain. The frequency position of the amide I absorbance maximum of different secondary-structure elements correlates with the strength of hydrogen bonding. In general, the amide I frequency decreases with increasing hydrogen bond strength [6]. The solvent exposure and the resulting hydrogen bonds between the amide groups and the solvent molecules have also an impact on the position of the amide I frequency. For example, the amide I frequency of solvated α-helices is approximately $20\,\mathrm{cm}^{-1}$ lower than that of non-solvated α-helices [14, 15]. However, the dominant coupling mechanism that makes the amide I vibration sensitive to secondary structure and determines its frequency is the through-space transition dipole coupling (TDC) [13, 16–18]. It is a resonance interaction between the oscillating dipoles of the amide groups and the coupling depends on the relative orientations of, and the distance between, the dipoles. The coupling is strongest when the coupled dipoles vibrate with the same frequency. Only with the introduction of the TDC mechanism, the large splitting of the amide I band of β-sheet structures in a strong band component $\sim 1.630\,\mathrm{cm}^{-1}$ and a weak band component $\sim 1{,}690\,\mathrm{cm}^{-1}$ can be explained [6].

Structural IR studies of peptides and proteins are usually performed in the solvent D_2O instead of H_2O since H_2O has a strong absorption of the HO bending vibration at $\sim 1{,}650\,\mathrm{cm}^{-1}$, which overlays the amide I region. Due to the mass effect on the vibrational frequency, this vibration is shifted to $\sim 1{,}200\,\mathrm{cm}^{-1}$ in D_2O and thus out of the amide I region. The amide I band is termed as amide I′ if the protein is diluted in D_2O.

6.1.3 Equilibrium vs. Kinetic Data

Amide I spectra measured in thermal equilibrium provide information about the thermal stability of peptides. The change of the amide I band (peak intensity or frequency) in dependence of the temperature can be evaluated to determine thermodynamic parameters such as the transition temperature T_m of an unfolding process or the equilibrium constant K_eq [19]. The equilibrium constant is directly related to the free energy ΔG^0 (with the universal gas constant R):

$$K_\mathrm{eq}(T) = \mathrm{e}^{-\frac{\Delta G^0}{RT}}. \tag{6.1}$$

However, the thermodynamic characterization provides no information about folding pathways. Time-resolved techniques and kinetic data are required to get insights into the mechanisms of folding and folding rates.

6.1.4 Rate Constants

Folding mechanisms are usually studied by measuring relaxation times after triggering a folding or unfolding reaction. The aim of kinetics studies in general is to find a kinetic model that is capable to describe the experimental data. It is assumed that the mechanism involves a finite number of kinetic species that are separated by energy barriers significantly larger than the thermal energy, but each kinetic species may consist of different conformations in rapid equilibrium. Kinetic data analyzing protein folding mechanisms often focus on monomeric polypeptides. This has the advantage that all observed reactions are of first order. Consequently the time-dependent change in absorption $\Delta A(T)$ of an observed signal can be represented as a sum of n exponentials with relaxation times τ_i (or apparent rate constants $k_i = 1/\tau_i$) and corresponding amplitudes A_i:

$$\Delta A\,(T) = \sum_{i=1}^{n} A_i \cdot e^{-\frac{t}{\tau_1}} = \sum_{i=1}^{n} A_i \cdot e^{-k_i\,t}. \tag{6.2}$$

There is a general relationship between the number of apparent rate constants and the number of kinetic species. Any kinetic mechanism with n different species connected by first-order reactions results in $n-1$ observable rate constants. Applied to protein folding with initial and final states, the observation of $n-1$ apparent rate constants thus indicates the presence of $n-2$ transiently populated intermediate states.

In the case of a simple two-state folding process without detectable intermediate states, folding (k_f) and unfolding (k_u) rates can be derived from the observed relaxation rate (k_obs) by use of the equilibrium constant K_eq. A two-state transition can simply be described with:

$$U \underset{k_\mathrm{u}}{\overset{k_\mathrm{f}}{\rightleftharpoons}} N, \tag{6.3}$$

where the equilibrium of the reaction is determined by the flows from the unfolded state U to the native state N and vice versa:

$$k_\mathrm{f}\,[U]_\mathrm{eq} = k_\mathrm{u}\,[N]_\mathrm{eq}\,. \tag{6.4}$$

The equilibrium constant (K_eq) for folding is as follows:

$$K_\mathrm{eq} = \frac{[N]_\mathrm{eq}}{[U]_\mathrm{eq}} = \frac{k_\mathrm{f}}{k_\mathrm{u}}. \tag{6.5}$$

The relaxation rate constant (k_{obs}) for a two-state folding process is as follows:

$$k_{obs} = k_f + k_u. \tag{6.6}$$

Thus, if k_{obs} and K_{eq} are accessible from experiments, k_f and k_u can be derived from (6.5) and (6.6).

Stability and folding rates are often analyzed in dependence of temperature using the Arrhenius equation. In this treatment, the rate of escape from a metastable state in a chemical reaction, k_{obs}, is the product of a temperature-independent pre-exponential factor A and an exponential contribution, which depends on the inverse temperature and on the activation energy, E_a. The equation is usually written in the form (with the universal gas constant R):

$$\ln k_{obs} = \ln A - \frac{E_a}{RT}. \tag{6.7}$$

However, folding processes are often more complex with transient population of partially folded intermediate states. Transition landscapes may involve sequential and parallel barrier-crossing events, even in the absence of populated intermediates. In this case, analysis of data requires much more differentiated theories. An overview of applied theories to protein folding can be found in Bieri and Kiefhaber [20].

6.2 Laser-Induced T-Jump Technique

Beginning with the work of Eigen and coworkers [21], it has been recognized that temperature changes provide a means of disturbing a molecular system in order to perform rapid kinetic studies. In the study of protein folding kinetics, temperature jumps can be used both to study the rapid unfolding to thermally denatured states [15, 22] and the rapid folding of proteins from cold-denatured states [23, 24]. The first work with temperature jumps was carried out with electrical current pulses that were generated by discharging a capacitor to obtain rapid heating, however samples were required with sufficient conductivity and high ionic strength [21]. The heating rise time of the resistive T-jump method is determined by $\tau = RC$ (R: resistance and C: capacitance) of the capacitor circuit and is typically in the order of microseconds for a temperature jump of a few degrees [25]. Higher temperature jumps up to 10 K can be produced using a higher capacitance C, but at the cost of longer rise times.

Pulsed lasers provide a means for faster T-jumps without the necessity of working at high salt concentrations as required for electrical T-jumps. In a laser-induced T-jump experiment, the laser pulse energy is absorbed by a dye or by the solvent undergoing a radiationless decay to produce heat. While dye-mediated T-jumps face the principal problem that the dye may interfere with the studied chemistry and/or its photophysics may overlap the probe signals, the direct solvent heating is a trigger

without need of any mediator or molecular switch. The generation of a temperature jump by directly exciting water (H_2O or D_2O) requires near-infrared pulses with several mJ at wavelengths either between 1,300 and 2,100 nm for the excitation of the OH stretch overtone of H_2O or between 1,900 and 2,200 nm for excitation of the OD stretch overtone of D_2O [26–28]. The first experiments to directly heat the water were performed in 1970 and utilized the stimulated Raman effect in liquid nitrogen to shift a 1,060 nm fundamental of a Q-switched Nd:glass laser to 1,410 nm where the water absorption coefficient is $\sim 10\,cm^{-1}$ [29]. High-pressure gases, as e.g., hydrogen (H_2) [30, 31] or methane (CH_4) [32], were used as a more practicable alternative. Hydrogen shifts the Nd:YAG pulse from 1,064 to 1,909 nm where experiments with samples dissolved in both H_2O or D_2O are possible. However, the absorption coefficient of H_2O at 1,909 nm is much larger than of D_2O and thus a smaller T-jump is produced and the usable optical pathlengths are limited (see Sect. 6.2.1). Methane shifts the Nd:YAG fundamental to 1,542 nm where the absorption coefficient of H_2O is much smaller than at 1,909 nm; therefore, this medium would be more appropriate for experiments in H_2O. Absorption coefficients for H_2O and D_2O at various commonly used gases, as well as an historical overview about the lasers used to generate suitable IR pump pulses for the T-jump technique, can be found in reviews about the T-jump method [27, 28]. Alternative approaches for generating heating pulses have also been utilized, as e.g., the coherent mixing of the Nd:YAG fundamental with an output of a dye laser [33], or Nd:YAG pumped optical parametric oscillators (OPO's) [34, 35].

Overall, laser-induced T-jump provides a suitable method to study elementary processes in protein folding, such as α-helix and β-hairpin formation that typically occurs on the nanosecond-to-microsecond timescale [1]. The use of IR radiation for the T-jump minimizes interference between the pump pulse and optical detection of the sample response. Thus, a wide variety of probe signals, including visible and IR absorption, fluorescence, and Raman scattering, can be used to probe the dynamic response of the heated sample.

6.2.1 Generation of the Heating Pulse

The stimulated Raman effect provides a useful and relatively simple means to shift Q-switched Nd:YAG laser pulses to the near IR. This approach for generating the heating pulse is described here. A Raman effect is the inelastic light scattering, where the energy of the scattered photon is shifted with respect to that of the incident photon. One observes frequency shifts of the scattered light to lower energies (Stokes lines) and to higher energies (Anti-Stokes lines). The frequency shift is a property of the Raman active medium whose molecules vibrate with the eigen frequency ν_{vib}. The Stokes lines refer to transitions from a lower vibrational state v (e.g., the ground state $v = 0$) to a higher vibrational state v' (e.g., first excited state $v' = 1$), whereas the scattering process in the Anti-Stokes process starts from an excited vibrational state and ends in a lower vibrational state (Fig. 6.2). In the

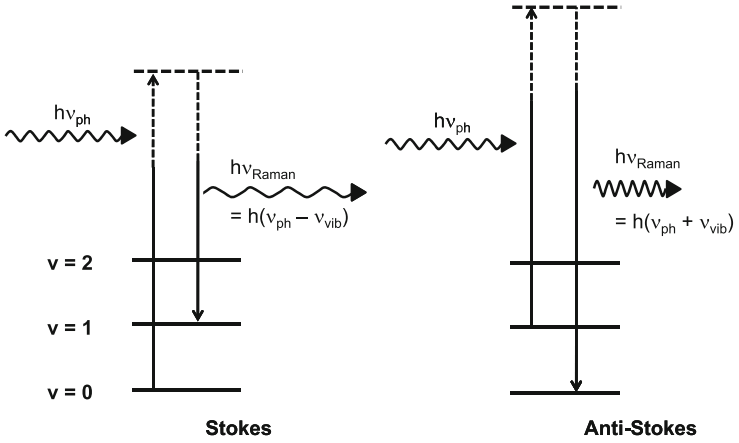

Fig. 6.2 Illustration of vibrational Raman scattering. In the Stokes process, the molecule takes up energy and undergoes a transition from a lower to a higher vibrational state (as shown here for $v = 0 \rightarrow v' = 1$), whereas in the Anti-Stokes process, the molecule releases energy and undergoes a transition from a higher to a lower vibrational state (as shown here for $v = 1 \rightarrow v' = 0$). The incident photon ν_{ph} excites the molecule from its initial energy level to a virtual energy level. The scattered Raman light ($E = h\nu_{\mathrm{Raman}}$) has less (Stokes lines) or more (Anti-Stokes lines) energy than the incident beam ($E = h\nu_{\mathrm{ph}}$)

Stokes process, the molecule takes up energy from the photon and the molecule releases energy in the Anti-Stokes process.

Without considering rotational states, the frequency of the scattered Raman-shifted light is given by:

$$\nu_{\mathrm{Stokes}} = \nu_{\mathrm{ph}} - n\nu_{\mathrm{vib}}, \tag{6.8}$$

for the red-shifted Stokes lines and

$$\nu_{\mathrm{Anti-Stokes}} = \nu_{\mathrm{ph}} + n\nu_{\mathrm{vib}}, \tag{6.9}$$

for the blue-shifted Anti-Stokes lines (with ν_{ph}: frequency of the excitation photons, ν_{vib}: vibrational frequency of the Raman medium molecules, n: order of the overtone). If H_2 gas is used as Raman medium, its eigen frequency ν_{H_2} of the hydrogen molecule is $\sim$4,160 cm^{-1} [36]. The photon frequency of the Nd:YAG laser beam is $\nu_{\mathrm{ph}} = 9{,}398\,\mathrm{cm}^{-1}$ (corresponding to $\lambda_{\mathrm{ph}} = 1{,}064$ nm). Thus, the red-shifted frequency of the first Stokes line is $\nu_{\mathrm{Stokes}} = \nu_{\mathrm{ph}} - \nu_{H_2} = 5{,}238\,\mathrm{cm}^{-1}$ (corresponding to $\lambda_{\mathrm{Stokes}} = 1{,}909$ nm) and the blue-shifted frequency of the first AntiStokes line is $\nu_{\mathrm{Anti\text{-}Stokes}} = \nu_{\mathrm{ph}} + \nu_{H_2} = 13{,}558\,\mathrm{cm}^{-1}$ (corresponding to $\lambda_{\mathrm{Anti\text{-}Stokes}} = 738$ nm in the visible spectral region). Overtones of the H_2 eigen frequency generate further Anti-Stokes lines at 564 nm and 457 nm, which are nicely observable since they are frequency-shifted to the visible spectral region. The spontaneous Raman scattering is nondirectional with a very small cross section: the intensity of the scattered radiation is on the order of 10^{-4}–10^{-5} of that of the incident beam and the

Anti-Stokes intensity is further attenuated by the lower population of the excited vibrational state [28]. At sufficiently high incident beam powers, the intensity of the Stokes radiation reaches the threshold where the Raman effect becomes stimulated which is practically achieved by focusing the laser pulses into the Raman conversion medium. High pressure of the Raman gas also contributes to a higher efficiency of the Raman effect. Besides H_2, also CH_4, D_2, and N_2 are other commonly used gases for Raman conversions [27]. The stimulated Stokes beam that is used as pump pulse has the properties of a laser beam, as it is coherent, propagates in the direction of the pump beam, and is very intense. Conversion efficiencies can reach up to 30% [28]. There is an optimum pumping energy at which the maximum Stokes output is produced. In general, the energy of the Stokes line increases with increasing pump power, until the threshold is exceeded for higher order Stokes and Anti-Stokes processes where the intensity of the Stokes line starts to decrease [28]. Further increase beyond the optimum also results in increasingly worse beam quality and shot-to-shot reproducibility. High repetition rate of cycling the Raman conversion can also decrease the conversion efficiency and output energy of the Raman cell as well as the beam quality and shot-to-shot stability [28]. However, the repetition rate of the T-jump experiment is usually limited by the rate of the thermal diffusion of the heated volume (cooling the sample).

The temperature jump (T-jump) perturbs the thermal equilibrium of the studied system and the kinetics of relaxation to the new equilibrium is spectroscopically followed. The temperature change ∂T produced by the laser pulse (size of the T-jump) determines the degree of equilibrium perturbation and therefore the amplitude of the relaxation signal. The rise time and duration of the T-jump limit the timescales of the processes that can be studied. In the absence of thermal diffusion, ∂T is given by [27]:

$$\partial T\,(r,\,z,\,t) = \frac{k}{\rho C_V \int_0^t I(r,z,t')dt'}, \tag{6.10}$$

where k is the absorption coefficient, ρ is the solution density, C_V is its heat capacity, and $I(r, z, t)$ is the instantaneous laser fluence in $\mathrm{W/cm^2}$ at time t and axial position z in which the pulse propagates. The intensity dependence of an absorbing medium on the spatial coordinate z is given by Lambert–Beer's law:

$$I\,(r,\,z) = I_0\,(r,\,z)\exp\left(-kz\right). \tag{6.11}$$

The size of the T-jump therefore depends on the incident beam intensity, beam diameter, absorption coefficient, and path through the sample. The absorption coefficient k depends on the excitation wavelength, the used solvent, and the temperature. It can be determined by the temperature-dependent near-IR spectra of the respective solvent. For example, at 1,909 nm, the Stokes line of a H_2-Raman shifter, the absorption of the solvent D_2O increases with temperature and the absorption coefficients are $k = 9.5\,\mathrm{cm^{-1}}$ at 273 K and $k = 13.5\,\mathrm{cm^{-1}}$ at 353 K [27]. As a consequence, the T-jumps at 353 K will be about 30% larger than those

at 273 K for the same heating pulse energy density. The duration of the T-jump is limited by the thermal diffusion. Analytical expressions can be derived in two limiting cases: first, for the pathlength significantly longer than the beam diameter as it is applicable to most laser T-jump experiments that use fluorescence or UV absorption as probe and second, for a significantly shorter pathlength compared to the beam diameter as it is applicable to T-jump experiments that probe protein IR amide I′ in D_2O [32]. In the latter case, the duration of the T-jump is essentially independent on the heated volume, but depends on the pathlength; thus, there is only minimal advantage associated with increasing the beam diameter. The heat transfer occurs almost entirely through the windows. Commonly used cell windows in IR experiments are CaF_2. Dyer and coworkers carried out numerical calculations of the heat transfer from D_2O between CaF_2 windows [37] leading to a T-jump duration (using a definition of 95% of the initial temperature) of ∼0.02 ms for a 100 μm pathlength. The temporal decay of the pulse depends on the real experimental geometries, the homogeneity of the sample, and the uniform heating. But the T-jump duration and temperature stability are crucial for the limitation in the study of peptide dynamics. The longer the jump is stable in temperature, the less the water dynamics (cooling) will interfere the peptide dynamics.

We measured the D_2O temperature stability with our experimental geometries: 100 μm pathlength of a CaF_2 cell, probe wavelength in the amide I′ region (1,638 cm^{-1}), and D_2O sample temperature of 20 °C. Figure 6.3 shows the temperature-dependent absorption change representing the water dynamics of the D_2O after the T-jump. At $t = 0$ s, an absorption change of $\Delta A = -41.1$ mOD can be observed corresponding to a T-jump of $\Delta T = 13$ °C. The change in absorption is converted to the change in temperature by a calibration set of temperature-dependent

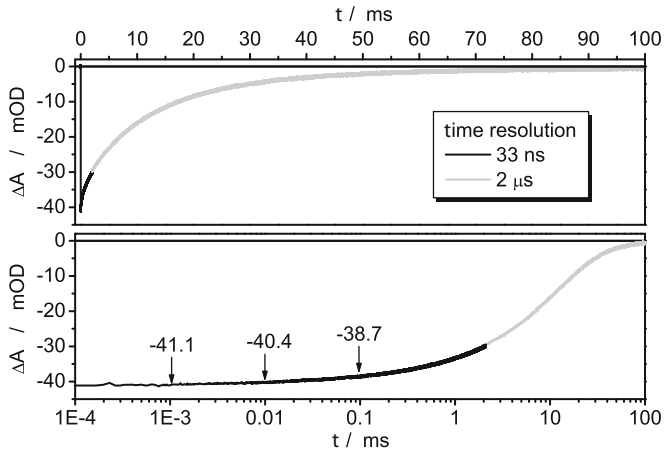

Fig. 6.3 D_2O dynamics at 1,638 cm^{-1} after a T-jump of 13 °C corresponding to ΔA of −41.1 mOD. The transient was measured with a time resolution of 33 ns (*black curve*) and 2 μs (*gray curve*). *Upper panel*: linear scale; *lower panel*: logarithmic scale

D_2O absorption spectra measured with Fourier transform infrared (FTIR) spectroscopy in thermal equilibrium. Thus, the solvent itself serves as intrinsic spectral thermometer. The excitation of the overtone vibrations as well as the thermal diffusion occurs on the picosecond timescale [33, 38, 39] and lies below our time resolution of 33 ns, but the temperature stability and cooling process can be monitored. The stability of the D_2O temperature becomes obvious in logarithmic scaling (Fig. 6.3, lower panel): the absorption decreases about 1.7% after 10 µs corresponding to $\Delta A = -40.4$ mOD and a temperature change of 0.2 °C. After 100 µs the change is 5.8%. Thus, our experimental geometries reproduce well the above-mentioned calculations for the T-jump duration [37]. The cooling of the water occurs in the ms time range and therefore interferes only minor in the µs time domain.

6.2.2 Photo-Acoustic Effects, Cavitation and Thermal Lensing

Laser-induced T-jumps are associated with the rapid expansion of the heated volume which generates photo-acoustic effects, cavitation, and thermal lensing [27, 28, 37]. A rapid temperature rise produces a large pressure increase within the heated volume. The shock wave subsequently propagates outward, reflects back from the cell walls, and produces large oscillations of the density and refractive index in the probed volume. The timescale is given by the speed of sound and the cell geometry. For most cell geometries and in aqueous solutions, these effects occur on the µs timescale. The photo-acoustic wave interferences are typically periodic oscillations observable in the measured intensity signal.

A stronger disturbance arises from cavitation effects that were experimentally and theoretically studied by Dyer and coworkers [37]. As the pressurized liquid expands outward, the pressure within the heated volume drops to negative values. When the tensile pressure exceeds a certain threshold value that depends on the concentration of impurities present, tensile failure occurs and bubble growth, or cavitation, is initiated. Cavitation is observed at tensions as low as 1 bar, however, indicating the existence of gas or vapor filled nuclei within the liquid or on the surfaces on bounding solids [37]. The bubbles perturb the probe beam by deflecting or defocusing the light, which causes a dramatic decrease in the detected signal. Cavitation is typically observed as large unperiodical drops in the signal intensity ranging from several hundred nanoseconds up to several microseconds [37]. Since cavitation and photo-acoustic effects in general arise from the volume change, they will be eliminated at temperatures when the solvent thermal expansion is minimal, i.e., at ~4 °C for H_2O and at ~12 °C for D_2O. Of course, studies in folding dynamics cannot be restricted to these temperatures. Filtration, distillation, UV exposure, and degassing are all effective in reducing size and number of undissolved impurities in water and, thereby, increasing its tensile strength.

Expansions of the heated volume produce a refractive index gradient that can focus or defocus the probe beam. This thermal lensing can affect the beam direction

or distribution over the detector, which can produce apparent changes in the signal if the response is not perfectly homogeneous [28]. The effect of thermal lensing depends on the experiment geometry. The temperature gradient sensed by the probe beam can be minimized by a careful alignment of the pump and probe beams.

6.2.3 *Experimental Setup*

There are various realizations of spectrometer setups using a laser-induced temperature jump. An overview can be found in [27,28]. The following describes the basics of a typical setup, illustrated in Fig. 6.4, which we have developed to study protein folding dynamics [40–42]. The temperature jump is generated by a Q-switched Nd:YAG laser pulse focused into a Raman shifter filled with H_2 (30 bar) that shifts the laser fundamental from 1,064 to 1,909 nm. In the described setup, the energy of the incident Nd:YAG fundamental is ∼560 mJ and the Raman-shifted Stokes line at 1,909 nm has an energy of ∼80 mJ; thus, the conversion efficiency is about 15%. The near-infrared pulse excites an overtone vibration of the solvent D_2O and thus initiates the temperature jump. For more homogeneous sample heating, the pulse is divided by a beam splitter into two counterpropagating pulses of equal energy in order to excite the solvent from both front and back. This results in a temperature gradient of less than 5% through the cell [36, 43]. Furthermore, the focus of each beam is adjusted to lie beyond the sample windows for both front and back excitation to avoid high power density within the sample. The area exposed to the heat inducing beams is approximately 1 mm in diameter, while that of the probe

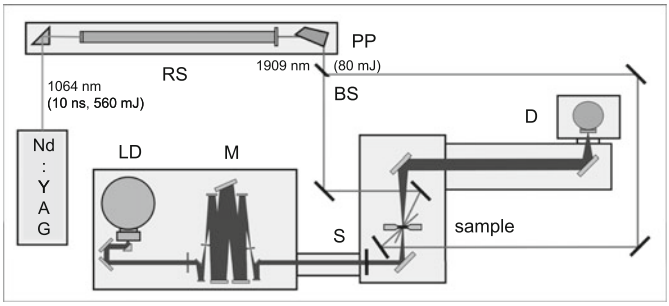

Fig. 6.4 Schematic layout of the T-Jump IR spectrometer, illustrating all principal components, including the Q-switched pulsed Nd:YAG laser (Nd:YAG), the Raman shifter (RS), the Pellin Brocca Prism (PP) for separating the 1,909 nm excitation wavelength, the beam splitter (BS), and dielectric mirrors. With ∼560 mJ pulse energy of the Nd:YAG fundamental, the Raman-shifted Stokes line at 1,909 nm has an energy of ∼80 mJ. The beamsplitter generates two counterpropagating pulses of approx. equal energy in order to excite the solvent of the sample and to induce the T-jump. For a T-jump of ∼10 °C, we use about 40 mJ excitation energy at the sample position. The IR spectrometer comprise cryogenically cooled and tunable lead salt laser diodes (LD) used as IR source in continuous-wave (cw) mode, the monochromator (M), the shutter (S), the thermostat-controlled sample holder (sample), and the HgCdTe detector (D)

beam is smaller (0.5 mm). These modifications permit more homogenous sample heating and reduce cavitation effects. With a sample pathlength of 100 μm, we use energies of ∼20 mJ for each of the two excitation pulses at the sample position in order to increase the temperature by ∼10 °C within 10 ns.

A continuous-wave (cw) lead salt infrared diode laser functions as IR source that is tunable in a spectral region from 1,600 to 1,750 cm^{-1}. Variation of temperature and current is used to tune the frequency and intensity characteristics of the laser diode over the emitted radiation modes. Since the laser diodes used here are multimode emitters, the output beam is directed through a monochromator in order to separate the desired probe wavelengths from the other spectral modes. The monochromator ensures single mode operation, but reduces the probe wavelength intensity up to 60%. Transient absorption changes of the probe beam through the sample are detected by a photovoltaic HdCdTe detector and are digitized by a transient recorder board. Signals originating from the black body radiation of the sample were minimized by a germanium band pass filter with a transmission range of 1,550–1,750 cm^{-1} positioned directly in front of the detector window. Further reduction was achieved by a shutter device that enables the alternative acquisition of any thermally induced black body radiation for one pump pulse without the probe light and then the transmission signal of the sample with the probe light on the next pulse. These two time-resolved intensity signals are used to evaluate the transients. Several thousand single transients are averaged for an improved signal-to-noise ratio. Since the sample transmission signals are monitored within a shutter sequence of two pulses, any detector drifts are minimized, which might be noticeable as an intensity offset occurring over several minutes. The sample cell consists of two CaF_2 windows that are separated with a spacer of typically 50–100 μm. The sample cell holder is temperature controlled by a connected water bath.

6.3 T-Jump Relaxation Kinetics

6.3.1 Two-State and Multistate Folders

Two-state processes in folding studies refer to the assumption that the protein has two defined states, a folded (native) and an unfolded state. In this case, folding and unfolding rates can be extracted from the observed rate of the T-jump studies by using the equilibrium constant as pointed out in Sect. 6.1.4. The equilibrium constant can be derived from FTIR measurements in thermal equilibrium and the temperature dependence of the amide I′ band intensity or frequency. An example is given here for the α-helical peptide polyglutamic acid, a peptide whose helix-to-coil transition has several kinetic steps, but a slow phase can be appropriately analyzed according to a two-state model. Figure 6.5 shows the transition from α-helix to random coil measured in thermal equilibrium. The spectral changes of the amide I′ absorption band clearly indicate the structural transition. The maximum of the amide I′ band is ∼1,640 cm^{-1}, a quite low amide I′ frequency, but typical for solvated,

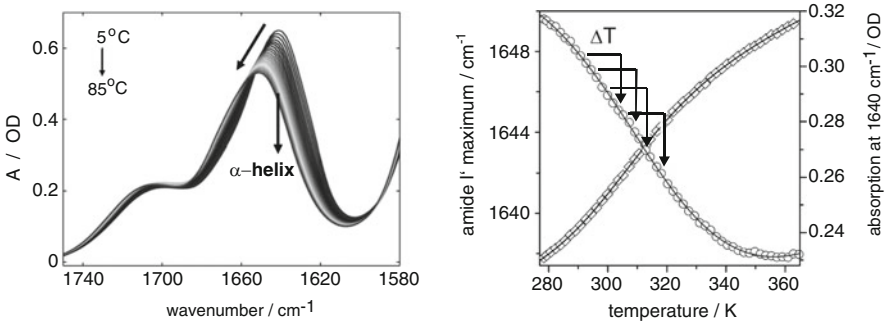

Fig. 6.5 Spectral changes of the temperature-induced helix-to-coil transition of polyglutamic acid at pD 5.4. FTIR measurements were performed in thermal equilibrium (from 5 to 85°C resp. 278–359 K). Low-temperature spectra represent the folded state and high-temperature spectra the unfolded state. *Left*: Amide I′ absorption spectra recorded at different temperatures. *Right*: Frequency position of the amide I′ maximum (*left y*-axis, *open rectangle* symbols), and amide I′ intensity change at 1,640 cm^{-1} (*right y*-axis, *open circle* symbols), both in dependence of the peptide temperature (*x*-axis). *Black arrows* illustrate the laser-excited temperature jumps ΔT at different peptide temperatures inducing thermal unfolding in a few nanoseconds

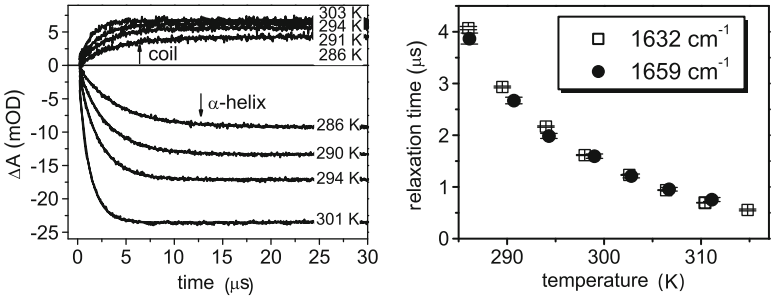

Fig. 6.6 T-jump kinetics of the helix-to-coil transition of polyglutamic acid at pD 5.4. *Left*: Transients of the α-helix decay (probed at 1,632 cm^{-1}) and coil rise (probed at 1,659 cm^{-1}) are shown exemplarily for some peptide temperatures. *Right*: Relaxation times in dependence of the peptide temperature (*open rectangle*, α-helix decay; *open circle*, coil rise)

hydrogen-bonded helices in contrast to the amide I′ frequency of shielded helices in proteins with higher frequencies ∼1,650 cm^{-1} [5, 6, 44–46].

T-jump kinetics provide the relaxation times of a peptide after initiation of a fast ns-temperature jump. The transients in Fig. 6.6 represent the decay of the α-helix probed at 1,632 cm^{-1} and the rise of the coil structure probed at 1,660 cm^{-1} shown exemplarily for some peptide temperatures. The temperature refers to the final peptide temperature after the T-jump. Evaluation of the transients yields the relaxation times of the peptide [19]. The relaxation time τ depends on the peptide temperature and it decreases with increasing temperature. It becomes obvious that the relaxation times for the α-helix decay and those for the rise of the coil match well in the considered temperature and time domain. Thus, a two-state model is

used to describe this transition. If intermediates occur on the transition pathways and the peptide unfolds in a multistate manner, the relaxation times will differ, as e.g., demonstrated for the Trpzip β-hairpin peptide in Fig. 6.8 (Sect. 6.4.2).

In case of a two-state folder, folding and unfolding rate constants can be extracted from the observed relaxation rates by use of the equilibrium constant K_{eq} as pointed out in Sect. 6.1.4. The Arrhenius plots of the relaxation, folding, and unfolding rates show a monotonic temperature dependence (Fig. 6.7). The transition rate k is directly correlated to the transition time $\tau(k^{-1} = \tau)$. At the transition temperature T_m (307 K), folding and unfolding times are the same ($\tau_u = \tau_f = 1.84\,\mu s$). Below the transition temperature T_m, the folding is faster than the unfolding, whereas above the T_m the behavior is vice versa (Table 6.1).

In case of a multistate folder, it is difficult to clearly define a folding and unfolding time. The kinetic data need to be evaluated with more complex models

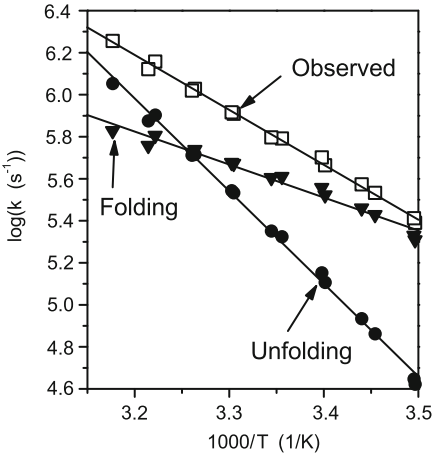

Fig. 6.7 Arrhenius plot of polyglutamic acid. Rates for folding k_f (*filled inverted triangle*) and unfolding k_u (*filled circle*) are derived from the relaxation rates k_{obs} obtained from kinetic T-jump studies (*open rectangle*) and the equilibrium constant K_{eq} obtained from equilibrium FTIR measurements. k_f and k_u are identical at the transition temperature T_m

Table 6.1 Time constants of polyglutamic acid relaxation for different peptide temperatures

T/K	$\tau_{obs}/\mu s$	$\tau_f/\mu s$	$\tau_u/\mu s$
290	2.86	3.65	12.80
300	1.44	2.42	4.03
307	0.91	1.84	1.84
310	0.76	1.64	1.35
320	0.41	1.14	0.49

τ_{obs} : observed time constant derived from the T-jump data; τ_f : time constant for folding; τ_u : time constant for unfolding. The temperature T refers to the final peptide temperature after the T-jump

that comprise intermediate states. Three of the most important classical models are the off-pathway model, the on-pathway model, and the sequential model. Different models imply different interrelationships between relaxation times and measured amplitudes. Models are usually chosen based on which one gives the best fit to the experimental rates and amplitudes. Newer models replace the pathway concept of sequential events with the folding funnel concept of parallel events. Energy landscapes are discussed where each protein conformation is represented by a point on a multidimensional energy surface [47].

6.3.2 Helix Dynamics

Several T-jump studies have been performed with α-helical peptides to analyze the helix-to-coil dynamics. The folding/unfolding transition has been reported to occur on the 100 ns timescale [33, 48–54]. The helix-to-coil relaxation time has been shown to be affected by the peptide length [55], temperature [19, 49, 55, 56], solvent viscosity [56], and ionic strength [53]. Alanine-based model peptides are often used to study helix formation and transitions since they form helical structures in aqueous solution and due to the high helix propensity of alanine. Experimental studies on short α-helical alanine-based peptides (<25 aa) have shown that the relaxation dynamics observed on the 100 ns timescale encompass both helix propagation (elongation of existing helical segments) and nucleation (formation of the first helical turn) and their reverse processes, with different activation energies and entropies [33,50,51,57]. Thus, the helix-to-coil relaxation is a multistep process and requires more complex models than two-state kinetics. However, as shown in Sect. 6.3.1, in the helix-to-coil transition of longer α-helical peptides, like the polyglutamic acid peptide that we studied (390–780 aa), there might be kinetic steps that can be well described by single-exponential behavior and a two-state model [19]. Other studies of model alanine-based peptides could show that the substitution of individual amino acid residues not only affects the helical stability of the peptide, but also the kinetics of its helix-to-coil transition [58]. Amino acids have significantly different helix propensities and thus the helicity of the peptide will depend on the amino acid sequence. T-jump studies of alanine-based peptides with isotopic editing indicate that the folding of an α-helix proceeds more rapidly at the helical C terminus than at the N terminus, presumably because of reduced steric restrictions due to the non-hydrogen-bonded carbonyl groups [59].

6.3.3 Hairpin Formation

Besides α-helices, β-sheets are one of the main extended structural elements in proteins. However, β-sheets pose difficulties for the study of folding mechanisms because of their variable structures and propensity for aggregation. A β-hairpin,

composed of two antiparallel strands connected by a turn, provides an ideal model system of a monomeric, well-defined structure containing a sheet segment. β-Hairpins are probably important structural elements in the initiation of sheet formation. It is still discussed whether hairpin folding itself is initiated by sequence propensities to form a turn and bring the strands together or by cross-strand interaction to stabilize a compact state and allow the turn to form [60–62]. The latter is normally attributed to a hydrophobic collapse, which can be aided by the sequence design with nonpolar residues on alternate positions in the cross-strand segments of the hairpin to interact and thereby excluding water. Cochran et al. developed the tryptophan zipper (Trpzip) β-hairpin sequences that are stabilized by Trp–Trp cross-strand interaction of four tryptophan residues forming a hydrophobic cluster [63, 64]. Spectroscopic studies on different Trpzip variants have shown changes in structure with variation in temperature in both equilibrium and kinetic measurements [62, 65–70]. T-jump measurements on Trpzip variants have been performed with infrared [41, 42, 62, 70–72] and fluorescence [68, 69, 73, 74] probes, detecting the backbone C=O groups or the tryptophan side chains, respectively. Also simulations have been employed to analyze the thermal stability and folding mechanism [69, 70, 75–77]. Although the understanding of the detailed mechanism of hairpin formation is still under study, there is agreement that hairpin formation is a multistate mechanism.

6.4 Site-Specific Dynamics with Isotopic Editing

6.4.1 Site-Specific Frequency Shifts

The secondary-structure dependence of the amide I band arises from the coupling of the polypeptide amide I modes. However, only average backbone conformational data can be determined from the amide I band. The delocalization of the vibrational modes prevents the use of the amide I band to get structure information on individual residues within the polypeptide. In order to elucidate structure and dynamics of polypeptides on the residue-level, isotopic labeling of individual carbonyl C=O groups of the polypeptide backbone provides a powerful means to isolate amide I modes of specific residues within the secondary structure [78, 79]. A backbone carbonyl that is labeled with ^{13}C is shifted to a lower frequency by $\sim$37 cm^{-1} if the mode is completely localized [78]. However, in a structured peptide, the amide I modes are coupled and the observable frequency shifts strongly depend on the coupling mechanisms between the ^{13}C-/^{13}C-carbonyls. The various coupling mechanisms between the C=O oscillators of the backbone (see Sect. 6.1.2) influence how far the ^{13}C amide mode frequency is shifted out of the overall ^{12}C amide band and how intense the frequency shifted mode is. Thus, the frequency shifts of ^{13}C labels are position-dependent and provide very sensitive probes for conformation. Isotope-editing permits structure resolution at the residue level at least for smaller polypeptides. Large proteins with hundreds of residues and several

secondary-structure elements will have a broad amide I band and the ^{13}C modes of individual labeled residues will not be frequency-shifted out of the ^{12}C amide band envelope. Reaction-induced difference spectra may provide the required spectral resolution to isolate individual amide modes and to monitor conformational changes on the residue level of larger proteins [80].

Several isotope-edited infrared studies of α-helical and β-sheet peptides have demonstrated the potential to analyze structural stability and conformational dynamics using the amide mode shifts [16, 81–89]. If peptides are isotopically labeled, certain amide modes become sensitive due to the vibrational coupling between the labeled residues. But also single-labeled residues may induce diagnostic frequency shifts independent from the coupling to other modes if the local mode is sufficiently sensitive to variations in the secondary structure as discussed in Sect. 6.4.3.

Studies with $^{13}C{=}O$ isotopically labeled α-helical peptides provided insights into helix conformation, stability, formation, and structural transitions [49, 50, 57, 81, 84, 87, 90, 91]. In the case of $^{13}C{=}O$ isotopically labeled peptides with β-sheet conformation, frequency shifts of amide I band components with very high intensity can be induced [16,82,83,85,86,88,89,92,93]. This anomalous intensity is attributed to a mixing of $^{13}C{=}O$ and $^{12}C{=}O$ modes between strands: There is interstrand coupling between the ^{13}C- and ^{12}C-carbonyls whose peptide groups are linked by a hydrogen bond, and as a result, the ^{13}C mode picks up substantial intensity [82]. This effect is strongest when ^{13}C-labeled residues are hydrogen-bonded to ^{12}C residues on adjacent strands. When the labeled groups are aligned between the strands, there is less $^{12}C/^{13}C$ mixing and therefore weaker intensity enhancement of the ^{13}C modes [89]. Also other mechanisms can contribute to the intensity enhancement, such as in-phase vibrations of interstrand-coupled ^{13}C modes, primarily from the central portions of the β-sheet [16].

6.4.2 Insights into Folding Mechanisms on the Residue Level

Site-specific T-jump IR kinetics provide insights into folding and unfolding processes on the residue level as shown here. Cross-strand isotopically labeled β-hairpins, derivatives of the tryptophan zipper hairpin Trpzip2 [63, 64], have been used in our study [41] with ^{13}C labels on selected amide C=O positions on opposite strands. The unlabeled variant is a mutant derivative of Trpzip2 and termed **Tripzip2C**. The labeled **Tripzip2C** variants are denoted as **A1A10, A3A10**, and **A3K8** with the designated residues corresponding to the labeling positions in the sequence (Fig. 6.8, upper panel). The absorption spectra of the various labeled variants demonstrate that the frequency shifts induced by the $^{13}C{=}O$ amide depend on the specific labeling position (Fig. 6.8, middle panel). As previously seen for other hairpin systems [83, 93–95], the double-labeled cross-strand interacting pair of amide C=O groups induces frequency shifts and their unique coupling results in a band pattern that is characteristic of their relative structure. Thus, site-specific

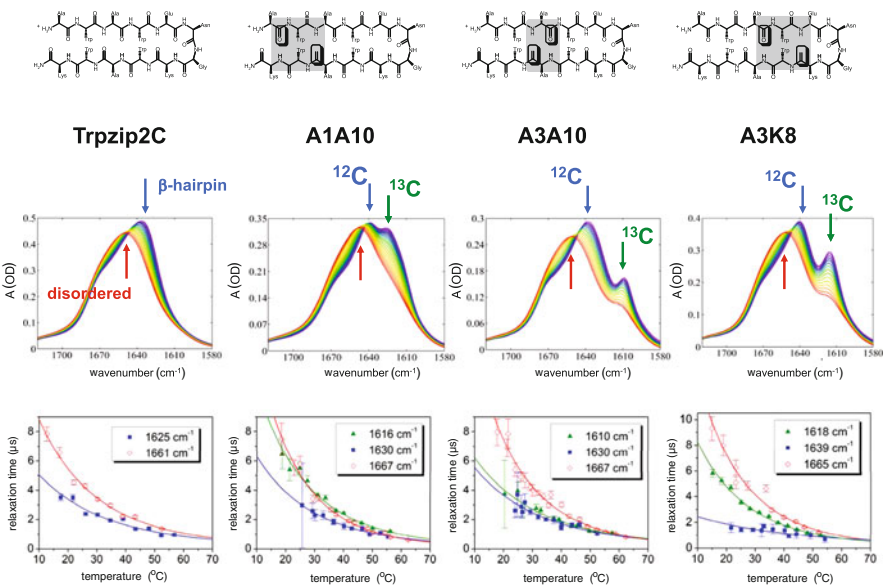

Fig. 6.8 Folding stability monitored with cross-strand isotopically labeled β-hairpin peptides (Trpzip2C: unlabeled variant, A1A10: labels at the hairpin termini, A1A10: labels in the hairpin core, A3K8: labels adjacent to the turn). *Upper panel*: Schemes of the hairpin peptide with various labeling positions. *Middle panel*: FTIR measurements in thermal equilibrium, heated up stepwise from 5 °C (*blue*, hairpin structure) to 85 °C (*red*, disordered structure). The frequency shift and the intensity of the ^{13}C-amide mode (*green arrow*) strongly depend on the labeling position within the hairpin structure. *Lower panel*: Relaxation times for different peptide temperatures derived from kinetic measurements at single wavelengths after a temperature jump of ∼10 °C. Shown are results for probe wavelengths representative for the rise of the disordered structure (*red*) and the decay of the β-hairpin, monitored at the ^{12}C- (*blue*) and the frequency-shifted ^{13}C-amide band component (*green*). Data points were fitted to the Arrhenius equation

frequency shifts can be used to monitor differences in structural stability within the secondary structure.

T-jump studies have been performed with labeled **Trpzip2C** variants at selected single wavelengths, which represent the rise of the disordered structure (∼1,665 cm^{-1}) and the β-hairpin decay upon thermal unfolding [41]. The β-hairpin decay is probed at two wavelengths, one chosen at the peak maximum of the ^{12}C=O amide band component (∼1,630 cm^{-1}, denoted as ^{12}C in Fig. 6.8) and one chosen at the maximum of the frequency-shifted ^{13}C=O amide band component (denoted as ^{13}C in Fig. 6.8). Wavelength selection was based on optimal overlap of the maximum absorbance change and intense modes of the lead salt diodes used as IR source. Temperature jumps of ∼10 °C have been applied at different peptide temperatures and relaxation times (τ_{obs}) have been determined. Figure 6.8 (lower panel) shows the results of the variation in dynamic behavior. The relaxation time τ_{obs} occurs in a few microseconds; τ_{obs} depends on the peptide temperature and decreases with increasing peptide temperature. The unlabeled **Trpzip2C** was

probed at 1,625 cm^{-1} representing the β-hairpin decay (Fig. 6.8, lower panel, blue trace) and at 1,661 cm^{-1} (Fig. 6.8, lower panel, red trace) representing the raise of the disordered structure. At temperatures >40 °C, there is little discrimination between the relaxation times for different probe wavelengths. However, at low temperatures, the relaxation times differ for the band components. The band related to the β-strand decays faster than the band related to the disordered structure appears. This behavior is definitely not consistent with a two-state behavior and the folding and unfolding rates cannot be derived as described in Sect. 6.1.4. The labeled variants show the same thermal variation of the relaxation times between strand-centered and disordered structure originating $^{12}C{=}O$ amide I′ modes: at low peptide temperatures (<40 °C), the $^{12}C{=}O$ β-hairpin band (Fig. 6.8, lower panel, blue traces) decays with faster time constants than the disordered band (Fig. 6.8, lower panel, red traces), whereas little discrimination among the time constants can be observed at high peptide temperatures. It is this low-temperature region where the distinct behavior of individual resolved $^{13}C{=}O$ amide I′ bands as site-specific probes becomes obvious (Fig. 6.8, lower panel, green traces), monitored at 1,616 cm^{-1} for **A1A10**, 1,610 cm^{-1} for **A3A10**, and 1,618 cm^{-1} for **A3K8**.

It is quite obvious that the different kinetics reflect stability variations at different positions of the hairpin. Our interpretation of the data is the following [41]: The decays of the β-strand $^{12}C{=}O$ modes represent the dynamics of disturbing the intact β-hairpin structure. In **A3A10**, where the $^{13}C{=}O$ groups form a small H-bonded ring in the middle of the hairpin, the $^{13}C{=}O$ amide I′ band has the same temperature-dependent dynamics as the $^{12}C{=}O$ band component. It reflects the stability of the center of the antiparallel strands due to the hydrophobic interactions of the tryptophans. In **A1A10**, the $^{13}C{=}O$ amide I′ band decay differs from the $^{12}C{=}O$ band component and coincides with the slower dynamics of the disordered structure. The labeled residues are close to the terminal groups and consequently structurally more disordered and less cross-strand coupled than **A3A10** and **A3K8**. Since the termini dynamics represent a mode of decay for the hairpin structure, even if only from a partially folded state, the $^{13}C{=}O$ amide I′ band in **A1A10** reflects the dynamics of the frayed ends upon unfolding. The labels in **A3K8** are adjacent to the turn and clearly show cross-strand coupling seen by the distinct and intense coupled $^{13}C{=}O$ amide I′ band in the absorption spectra (Fig. 6.8, middle panel). However, the turn residues stress the hairpin structure, resulting in a distortion of the H-bond between positions 5 and 8. This distortion may influence the local unfolding dynamics and reasonably reflect an intermediate kinetic behavior between the strand and the disordered components. Taken together, the labels of **A1A10** and **A3K8** sense more the dynamics of the unfolded state, whereas the **A3A10** labels sense the disruption of the fully intact segment of the β-hairpin backbone structure in the core. The faster, more β-strand-like $^{13}C{=}O$ kinetics of **A3A10** suggest that the central part of the strand, where the tryptophans are in the sequence, is the stable component and that unfolding progresses from the termini toward the center. By microscopic reversibility, this would support that folding is initiated by a hydrophobic collapse. Such reversibility would only be tenable for a two-state

model, but site-specificity of the relaxation times could only be obtained at low temperatures where the relaxation kinetics indicate the formation of intermediate states and thus a non-two-state folding process. Nevertheless, although the site-specificity of the relaxation times could only be derived from a partially folded state, the data are still supportive for the hypothesis of a hydrophobic collapse-mediated folding mechanism of the **Trpzip** hairpin.

6.4.3 Single and Multiple Isotope Labels

Labeling peptides with isotopes offer a means of enhancing the low spatial resolution IR technique by giving it site-specific sensitivity. Single labels offer a simpler means of identifying parts of a sequence than multiple labels, both from the synthetic and interpretive points of view, but there has to be a molecular mechanism that makes the single label sensitive to the conformational change of the peptide sequence. Double or multiple labels, if strongly coupled, offer more specific structural sensitivity. However, in some cases single labels yield the same information. We have compared the sensitivity of single versus cross-strand double labels of **Trpzip2C** that have been used as spectral probes for monitoring site-specific folding dynamics with T-jump IR spectroscopy [42]. In Fig. 6.9, IR-absorption spectra measured in thermal equilibrium as well as relaxation time constants determined with T-jump IR spectroscopy are shown for the single-labeled **Trpzip2C** variants **A1** and **A10**. Measuring conditions are the same as for the double-labeled variant **A1A10** that is discussed in Sect. 6.4.2. Position 10 of the **A10** variant is located between the four Trp residues, in the center of the hairpin β-strands that are expected to be the most stable part. Position 1 of **A1** is on the N terminus which is partially frayed. As shown in the absorption spectra (Fig. 6.9, middle panel), a well-defined $^{13}C{=}O$ amide I′ band component can be monitored for **A10**, even if weaker than in the double label variant **A1A10** (Fig. 6.8, middle panel), but a frequency-shifted band is not identifiable in **A1**.

The **A1** absorption spectrum looks similar to that of the unlabeled variant (Fig. 6.8, middle panel, Trpzip2C). However, there is a down shift of the ^{12}C-labeled amide I′ **A1** band maximum from that of the unlabeled **Trpzip2C**, confirming a difference between these two peptides. In contrast to **A1**, the double-labeled **A1A10** variant yields a complex and unresolved, but distinct amide I′ pattern for the ^{12}C and ^{13}C contributions (Fig. 6.8, middle panel, A1A10). T-jump dynamics were measured for **A1** and **A10** in the same manner used for the double-labeled variant **A1A10**. Relaxation times of **A10** showed that the ^{12}C- and ^{13}C-β-strand component values (blue trace) were faster (lower) than those of the disordered component (red trace), which means that the loss of sheet is faster than the gain of disordered structure. Thus, the dynamics of the single label variant **A10** behave similar as previously discussed for the double-labeled variant **A3A10** (Sect. 6.4.2). **A3A10**, **A10**, and **A3** are variants whose labels are located between the Trp's in the center of the strands and all show a slower relaxation time for the rise of the disordered component

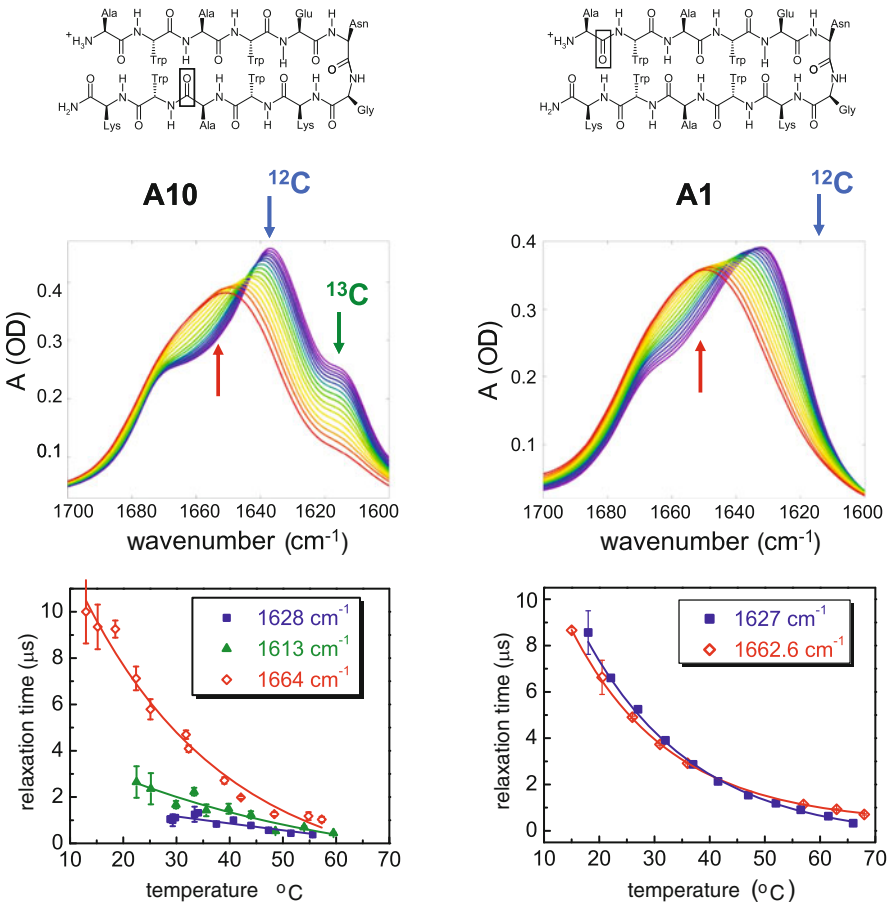

Fig. 6.9 Folding stability monitored with single isotopically labeled β-hairpin peptides (*left column*: **A10**; *right column*: **A1**). *Upper panel*: Schemes of the hairpin peptides with labeling position. *Middle panel*: FTIR measurements in thermal equilibrium, heated up stepwise from 5 °C (*blue*, hairpin structure) to 85 °C (*red*, disordered structure). There is no frequency-shifted ^{13}C-amide band in A1. *Lower panel*: Relaxation times for different peptide temperatures derived from kinetic measurements at single wavelengths after a temperature jump of ∼10 °C. Shown are results for probe wavelengths representative for the rise of the disordered structure (*red*) and the decay of the β-hairpin, monitored at the ^{12}C-(*blue*) and the frequency-shifted ^{13}C-amide band (*green*). The single label at the hairpin terminus (A1) shows different equilibrium and kinetic behavior. For further explanations, see text

than for the decay of the ^{12}C- and ^{13}C-β-strand components [41, 42]. The ^{13}C-band detected dynamics are, within error, quite similar to those detected with the ^{12}C-band. This reflects that the local loss of the central conformation occurs with the same dynamics as the loss of the overall strand structure. **A1** has very different equilibrium and kinetic spectra than **A10** and **A1A10**. The kinetics of **A1** probed at 1,627 cm^{-1} (Fig. 6.9, lower panel) have contributions from both the ^{12}C and

^{13}C components, and the relaxation times are about the same as for the disordered component. Comparing the **A1** dynamics to the **A1A10** dynamics, it seems that the ^{13}C contribution is quite dominant in what should be the β-sheet component, as the kinetics are slowed just as for the ^{13}C component of **A1A10** (Fig. 6.8, lower panel). This kinetic behavior of **A1A10** reflects the dynamics of the frayed ends upon unfolding and thus is also consistent for the relaxation data of the single variant **A1** [41, 42].

The sensitivity of single versus cross-strand double labels has been compared for the model β-hairpin **Trpzip2C**. Although single labels do not result in intensity enhancement, as seen for cross-strand labeling, the IR frequency shifts are still diagnostic for hairpin unfolding. If C=O carbonyls in the β-strand portion of the hairpin (between the Trp residues) are labeled, the dynamic behavior of the single labels is similar to the results obtained with double labels in terms of relaxation time and activation energy and closely tracks the kinetics of the β-strand components. This implies that either property, local secondary structure (change of ϕ, ψ), or cross-strand coupling enabled by strand formation and H-bonding relaxes with the same kinetic mechanism. Single-labeled residues on the terminal positions have a different behavior and are less able to be detected due to overlap with the ^{12}C components. In contrast, double labels involving terminal positions are enhanced due to coupling.

Acknowledgements The author thanks Dr. Carsten Krejtschi, former PhD student for the initial work on the spectrometer setup and various T-jump studies and Timothy A. Keiderling, University of Illinois at Chicago (USA), for inspiring collaboration and for generously providing us with isotopically labeled peptide variants.

References

1. O. Bieri, T. Kiefhaber, Biol. Chem. **380**, 923 (1999)
2. W.K. Surewicz, H.H. Mantsch, D. Chapman, Biochemistry **32**, 389 (1993)
3. S. Krimm, J. Bandekar, Adv. Protein Chem. **38**, 181 (1986)
4. J.L.R. Arrondo, A. Muga, J. Castresana, F.M. Goni, Prog. Biophys. Mol. Biol. **59**, 23 (1993)
5. D.M. Byler, H. Susi, Biopolymers **25**, 469 (1986)
6. A. Barth, C. Zscherp, Q. Rev. Biophys. **35**, 369 (2002)
7. E. Goormaghtigh, V. Cabiaux, J.-M. Ruysschaert, Subcell. Biochem. **23**, 405 (1994)
8. E. Goormaghtigh, V. Cabiaux, J.-M. Ruysschaert, Subcell. Biochem. **23**, 329 (1994)
9. F. Dousseau, M. Pezolet, Biochemistry **29**, 8771 (1990)
10. V. Baumruk, P. Pancoska, T.A. Keiderling, J. Mol. Biol. **259**, 774 (1996)
11. D.C. Lee, P.I. Haris, D. Chapman, R.C. Mitchell, Biochemistry **29**, 9185 (1990)
12. K. Rahmelow, W. Hübner, Anal. Biochem. **241**, 5 (1996)
13. S. Krimm, Y. Abe, Proc. Natl. Acad. Sci. U.S.A. **69**, 2788 (1972)
14. W.C. Reisdorf, S. Krimm, Biochemistry **35**, 1383 (1996)
15. R. Gilmanshin, S. Williams, R.H. Callender, W.H. Woodruff, R.B. Dyer, Proc. Natl. Acad. Sci. U.S.A. **94**, 3709 (1997)
16. J. Kubelka, K. Keiderling, J. Am. Chem. Soc. **123**, 6142 (2001)
17. N.A. Nevskaya, Yu.N. Chirgadze, Biopolymers **15**, 637 (1976)

18. Yu.N. Chirgadze, N.A. Nevskaya, Biopolymers **15**, 627 (1976)
19. C. Krejtschi, K. Hauser, Eur. Biophys. J. **40**(5), 673–685 (2011)
20. O. Bieri, T. Kiefhaber, in *Protein Folding: Frontiers in Molecular Biology*, ed. by R. Pain, (Oxford University Press, Oxford, UK, 2000), p. 34.
21. M. Eigen, L.de Maeyer, in *Techniques in Organic Chemistry*, ed. by A. Weisberger (Wiley Interscience, New York, NY, 1963), p. 895
22. R.B. Dyer, F. Gai, W.H. Woodruff, Acc. Chem. Res. **31**, 709 (1998)
23. R.M. Ballew, J. Sabelko, M. Gruebele, Nat. Struct. Biol. **3**, 923 (1996)
24. M. Gruebele, J. Sabelko, R.M. Ballew, J. Ervin, Acc. Chem. Res. **31**, 699 (1998)
25. B. Nölting, *Protein Folding Kinetics*, 2nd ed. (Springer, Heidelberg, 2006)
26. G.W. Flynn, N. Sutin, in *Chemical and Biochemical Applications of Lasers*, ed. by C.B. Moore (Academic Press, New York, 1974), p. 309
27. J. Hofrichter, Methods Mol. Biol. **168**, 159 (2001)
28. J. Kubelka, Photochem. Photobiol. Sci. **8**, 499 (2009)
29. J.V. Beitz, G.W. Flynn, D.H. Turner, N. Sutin, J. Am. Chem. Soc. **92**, 4130 (1970)
30. S. Ameen, L. Demaeyer, J. Am. Chem. Soc. **97**, 1590 (1975)
31. S. Ameen, Rev. Sci. Instrum. **46**, 1209 (1975)
32. A.P. Williams, C.E. Longfellow, S.M. Freier, R. Kierzek, D.H. Turner, Biochemistry **28**, 4283 (1989)
33. S. Williams, T.P. Causgrove, R. Gilmanshin, K.S. Fang, R.H. Callender, W.H. Woodruff, R.B. Dyer, Biochemistry **35**, 691 (1996)
34. G. Balakrishnan, Y. Hu, T.G. Spiro, Appl. Spectrosc. **60**, 347 (2006)
35. M. Sadqi, J.L. Lapidus, V. Munoz, Proc. Natl. Acad. Sci. U.S.A. **100**, 12117 (2003)
36. K. Yamamoto, Y. Mizutani, T. Kitagawa, Appl. Spectrosc. **54**, 1591 (2000)
37. W.O. Wray, T. Aida, R.B. Dyer, Appl. Phys. B **74**, 57 (2002)
38. P.A. Anfinrud, C. Han, R.M. Hochstrasser, Proc. Natl. Acad. Sci. U.S.A. **86**, 8387 (1989)
39. L. Genberg, F. Heisel, G. McLendon, R.J. Dwayne Miller, J. Phys. Chem. **91**, 5521 (1987)
40. C. Krejtschi, R. Huang, T.A. Keiderling, K. Hauser, Vibr. Spectrosc. **48**, 1 (2008)
41. K. Hauser, C. Krejtschi, R. Huang, L. Wu, T.A. Keiderling, J. Am. Chem. Soc. **130**, 2984 (2008)
42. K. Hauser, O. Ridderbusch, A. Roy, A. Hellerbach, R. Huang, T.A. Keiderling, J. Phys. Chem. B **114**, 11628 (2010)
43. R.M. Ballew, J. Sabelko, C. Reiner, M. Gruebele, Rev. Sci. Instrum. **67**, 3694 (1996)
44. S.Yu. Venyaminov, N.N. Kalnin, Biopolymers **30**, 1259 (1990)
45. G. Yoder, P. Pancoska, T.A. Keiderling, Biochemistry **36**, 15123 (1997)
46. S.T.R. Walsh, R. Walsh, P. Cheng, W.W. Wright, V. Daggett, J.M. Vanderkooi, W.F. Degrado, Protein Sci. **12**, 520 (2003)
47. K.A. Dill, H.S. Chan, Nat. Struct. Biol. **4**, 10 (1997)
48. W.L. Qian, S. Krimm, Biopolymers **34**, 1377 (1994)
49. C.-Y. Huang, Z. Getahun, Y. Zhu, J.W. Klemke, W.F. Degrado, F. Gai, Proc. Natl. Acad. Sci. U.S.A. **99**, 2788 (2002)
50. C.-Y. Huang, Z. Getahun, T. Wang, W.F. Degrado, F. Gai, J. Am. Chem. Soc. **123**, 12111 (2001)
51. P. Thompson, W.A. Eaton, J. Hofrichter, Biochemistry **36**, 9200 (1997)
52. A.P. Ramajo, S.A. Petty, M. Volk, Chem. Phys. **6**, 1022 (2004)
53. A.P. Ramajo, S.A. Petty, M. Volk, Chem. Phys. **323**, 11 (2006)
54. J. Bredenbeck, J. Helbing, J.R. Kumita, G.A. Woolley, P. Hamm, Proc. Natl. Acad. Sci. U.S.A. **102**, 2379 (2005)
55. T. Wang, Y. Zhu, Z. Getahun, D. Du, C.-Y. Huang, W.F. Degrado, F. Gai, J. Phys. Chem. B **108**, 15301 (2004)
56. G.S. Jas, W.A. Eaton, J. Hofrichter, J. Phys. Chem. B **105**, 261 (2001)
57. J.H. Werner, R.B. Dyer, R.B. Fesinmeyer, N.H. Andersen, J. Phys. Chem. B **106**, 487 (2002)
58. E.A. Gooding, A.P. Ramajo, J. Wang, C. Palmer, E. Fouts, M. Volk, Chem. Commun. **48**, 5985 (2005)

59. A.R. Ramajo, S.A. Petty, A. Starzyk, S.M. Decatur, M. Volk, J. Am. Chem. Soc. **127**, 13784 (2005)
60. S. Deechongkit, H. Nguyen, M. Jager, E.T. Powers, M. Gruebele, J.W. Kelly, Curr. Opin. Struct. Biol. **16**, 94 (2006)
61. M.S. Searle, B. Ciani, Curr. Opin. Struct. Biol. **14**, 458 (2004)
62. D. Du, Y. Zhu, C.-Y. Huang, F. Gai, Proc. Natl. Acad. Sci. U.S.A. **101**, 15915 (2004)
63. A.G. Cochran, N.J. Skelton, M.A. Starovasnik, Proc. Natl. Acad. Sci. U.S.A. **98**, 5578 (2001), Correction: Proc. Natl. Acad. Soc. U.S.A. **99**, 9081 (2002).
64. A.G. Cochran, N.J. Skelton, M.S. Starovasnik, Proc. Natl. Acad. Sci. U.S.A. **99**, 9081 (2002)
65. T. Wang, Y. Xu, D. Du, F. Gai, Biopolymers **75**, 163 (2004)
66. A.W. Smith, A. Tokmakoff, J. Chem. Phys. **126**, 045109 (2007)
67. A.W. Smith, H.S. Chung, Z. Ganim, A. Tokmakoff, J. Phys. Chem. B **109**, 17025 (2005)
68. W.Y. Yang, M. Gruebele, J. Am. Chem. Soc. **126**, 7758 (2004)
69. W.Y. Yang, J.W. Pitera, W.S. Swope, M. Gruebele, J. Mol. Biol. **336**, 241 (2004)
70. C.D. Snow, L. Qiu, D. Du, F. Gai, S.J. Hagen, V.S. Pande, Proc. Natl. Acad. Sci. U.S.A. **101**, 4077 (2004)
71. D. Du, M.J. Tucker, F. Gai, Biochemistry **45**, 2668 (2006)
72. D.C. Gadsby, Nature **450**, 957 (2007)
73. V. Munoz, P.A. Thompson, J. Hofrichter, W.A. Eaton, Nature **390**, 196 (1997)
74. V. Munoz, R. Ghirlando, F.J. Blanco, G.S. Jas, J. Hofrichter, W.A. Eaton, Biochemistry **45**, 7023 (2006)
75. J.W. Pitera, I. Haque, W.C. Swope, J. Chem. Phys. **124**, 141102 (2006)
76. B. Zagrovic, C.D. Snow, S. Khaliq, M.R. Shirts, V.S. Pande, J. Mol. Biol. **323**, 153 (2002)
77. J.P. Ulmschneider, W.L. Jorgensen, J. Am. Chem. Soc. **126**, 1849 (2004)
78. L. Tadesse, R. Nazarbaghi, L. Walters, J. Am. Chem. Soc. **113**, 7036 (1991)
79. S.M. Decatur, Acc. Chem. Res. **39**, 169 (2006)
80. K. Hauser, M. Engelhard, N. Friedman, M. Sheves, F. Siebert, J. Phys. Chem. A **106**, 3553 (2002)
81. R.A.G.D. Silva, J. Kubelka, P. Bour, S.M. Decatur, T.A. Keiderling, Proc. Natl. Acad. Sci. U.S.A. **97**, 8318 (2000)
82. J.W. Brauner, C. Dugan, R. Mendelsohn, J. Am. Chem. Soc. **122**, 677 (2000)
83. R. Huang, V. Setnicka, M.A. Etienne, J. Kim, J. Kubelka, R.P. Hammer, T.A. Keiderling, J. Am. Chem. Soc. **129**, 13592 (2007)
84. R. Huang, J. Kubelka, W. Barber-Armstrong, R.A.G.D. Silva, S.M. Decatur, T.A. Keiderling, J. Am. Chem. Soc. **126**, 2346 (2004)
85. A.W. Smith, A. Tokmakoff, Angew. Chem. Int. Ed. **46**, 7984 (2007)
86. S.H. Shim, R. Gupta, Y.L. Ling, D.B. Strasfeld, D.P. Raleigh, M.T. Zanni, Proc. Natl. Acad. Sci. U.S.A. **106**, 6614 (2009)
87. S.H. Brewer, B. Song, D.P. Raleigh, R.B. Dyer, Biochemistry **46**, 3279 (2007)
88. C. Paul, P.H. Axelsen, J. Am. Chem. Soc. **127**, 5754 (2005)
89. C. Paul, J. Wang, W.C. Wimley, R.M. Hochstrasser, P.H. Axelsen, J. Am. Chem. Soc. **126**, 5843 (2004)
90. S.M. Decatur, Biopolymers **54**, 180 (2000)
91. S.M. Decatur, J. Antonic, J. Am. Chem. Soc. **121**, 11914 (1999)
92. K.J. Halverson, I. Sucholeiki, T.T. Ashburn, P.T. Lansbury Jr., J. Am. Chem. Soc. **113**, 6701 (1991)
93. V. Setnicka, R. Huang, C.L. Thomas, M.A. Etienne, J. Kubelka, R.P. Hammer, T.A. Keiderling, J. Am. Chem. Soc. **127**, 4992 (2005)
94. J. Kim, R. Huang, J. Kubelka, P. Bour, T.A. Keiderling, J. Phys. Chem. B **110**, 23590 (2006)
95. P. Bour, T.A. Keiderling, J. Phys. Chem. B **109**, 23687 (2005)

Chapter 7
Light-Triggered Peptide Dynamics

Wolfgang Zinth and Josef Wachtveitl

Abstract Photoswitches suitably connected to model peptides allow to apply light as trigger for folding initiation. Time-resolved infrared spectroscopy provides a certain degree of structural information and allows to follow the kinetics of photoinduced conformational changes by investigating the transient, interconverting structures. Several secondary structure elements have been studied in detail with this approach. The recent development of light-responsive tertiary structures yields the next generation of model peptides on the way toward an understanding of protein folding.

7.1 Introduction

Most organisms contain chromo-proteins where the combination of a protein moiety with light absorbing prosthetic groups leads to new functionalities. Light absorption in visual proteins such as rhodopsins or sensory rhodopsins initiates a stimulus, which allows the organism to see the surroundings and/or to react on changes in the illumination conditions. Light may trigger or synchronize circadian cycles. Fluorescent chromo-proteins perform efficient wavelength conversion of light; photosynthetic proteins convert and store solar energy through rapid electron transfer processes. From the scientific point of view, the light absorbing groups of chromo-proteins allow to monitor function and dynamics of biological processes via

W. Zinth (✉)
Faculty of Physics, Department of Biomolecular Optics, Ludwig-Maximilians-University Munich, Oettingenstr. 67, 80538 Munich, Germany
e-mail: wolfgang.zinth@physik.uni-muenchen.de

J. Wachtveitl (✉)
Institute for Physical and Theoretical Chemistry, Goethe-University Frankfurt, Max von Laue-Straße 7, 60438 Frankfurt am Main, Germany
e-mail: wveitl@theochem.uni-frankfurt.de

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*, Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_7,

optical techniques. Time-resolved absorption and emission measurements characterize ultrafast processes in biological systems. Recently, fluorescing chromo-proteins have been used for subdiffraction-limited microscopic imaging. The application of chromo-proteins, however, is restricted due to their low abundance, the limited variety of functional properties and the rather large size of most native systems. The introduction of light switchable peptides combining light absorbing and photoreactive groups with small peptides of suitably designed amino acid sequence leads to molecular systems, which open numerous new applications. These light switchable peptides may act as light-activated and directed tools, may exhibit light-triggered enzymatic activity and can even lead to light controlled neuronal activity [1–3]. In addition, they could be used as biocompatible compounds for subwavelength microscopy or as switching elements for the study of structural dynamics in biomolecules. This novel field of research will be reviewed in this article.

7.2 Light-Triggered Peptides

A prototypical light switchable peptide (Fig. 7.1) is composed of three molecular parts, which are discussed in the following paragraphs: the photochromic switching unit (I) is attached via linking groups (II) to peptide moieties (III). In order to act as a successful light switchable peptide, the three units have to fulfill a number of requirements.

7.2.1 The Photochromic Switching Unit

Photochromic molecules are available with widely different properties. Here, we address only molecules that can be used to modify the structure of the complete

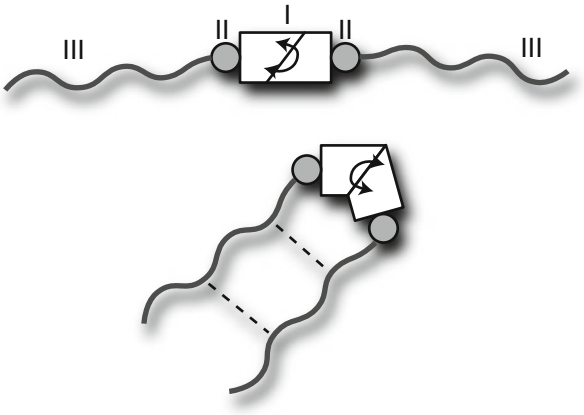

Fig. 7.1 Prototypical structure of a light-triggered peptide with the photochromic switching unit (I), the linkers (II) and the peptide moieties (III)

chromopeptide. Two types of photochromic switching molecules are suitable for these applications:

1. The switching molecule can access two stable states (conformers) with well-defined structures of different optical properties. Of special interest are photochromic molecules where both states are thermally stable (on the timescale of hours) and where both switching directions can be operated by light at different wavelengths. Such switching molecules are required if both directions have to be studied, when an enzymatic activity should be switched permanently or when the system should be controlled (e.g., activity on or off) by light.
2. In the second class of photochromic molecules, only one of the two states is stable and the reconversion occurs fast via a thermally driven back reaction. Here, only the optically controlled reaction can be studied directly. Since reformation of the initial state occurs automatically with a defined reconversion time, there is no need for the active reconversion by suitable illumination of the sample. An application of such a system is the control of enzymatic activity by intensity variations of the illumination.

A clear distinction between group one and two is not always possible. The energetic barriers in the electronic ground state which control thermal stability may strongly depend on specific substitutions of the switching chromophore. This can be used in the molecular design to adjust the switching molecule to the specific application.

For practical applications, switching molecules have to fulfill a number of additional requirements. (a) They should sustain many switching cycles (this requirement distinguishes them from caged compounds). (b) At least one (preferably both molecular states) can be prepared selectively. (c) The light-induced switching has reasonably high quantum efficiency. (d) The light-induced processes are fast (i.e., on the femtosecond and picosecond timescale) and should proceed over a small number of intermediate states. (e) Pronounced structural changes of the switching molecule, which are able to induce structural changes of the complete chromopeptide, occur upon light absorption. Here, we will not address switching processes that are restricted to the electronic properties of the molecular switch. These changes may be of importance for the transfer of energy and for applications as emitters in subdiffraction microscopy. (f) The molecular switch should be insensitive to specific surroundings encountered in the in vitro and in vivo experiments. It should have sufficiently high chemical stability. For in vivo applications, it should exhibit low toxicity.

Photoswitchable molecules have been synthesized in great varieties of structures (for a selection, see Fig. 7.2). Photochromism suitable for structural switches is advantageously based on chemical transformations such as pericyclic ring-opening or ring-closure reactions and Z/E isomerizations.

Isomerizations lead to structural changes of large amplitude, which may completely change the shape of the switch. For a suitably designed peptide moiety, this may change the structural arrangement of the whole system. When the photoisomerization is related with a downhill reaction path in the excited electronic state and with rapid interconversion to the ground state surface via a conical

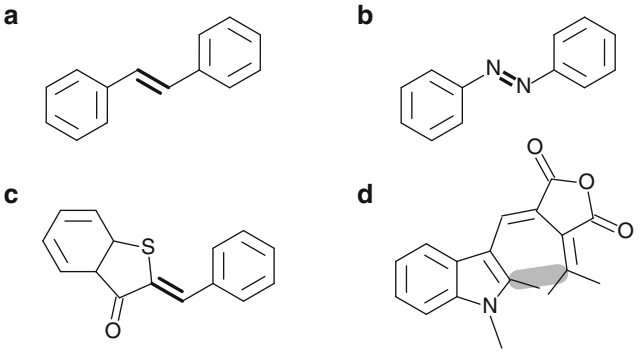

Fig. 7.2 Selection of photochromic molecules: stilbene (**a**), azobenzene (**b**), hemithioindigo (HTI) (**c**) with photoisomerization around the central double bond and an indolylfulgide (**d**) which undergoes light-induced ring-closure/opening reactions in the region highlighted in *grey*

intersection, an extremely fast formation of the product state is found. For steep potential surfaces large forces, which may induce strongly driven structural changes of the linked peptide moiety, result. An example of such strongly driven and ultrafast isomerization is the *cis*-to-*trans* isomerization of azobenzene after excitation in the blue via the $n\pi^*$ transition. For a recent review on azobenzene peptide switches, see [1–3]. This reaction occurs on the 300 fs timescale [4]. In other molecular switches acting via isomerization (e.g., hemithioindigo, HTI), the reaction in the excited electronic state proceeds via a minimum of the potential surface, separated from the reactive area (e.g., a conical intersection) for the internal conversion to the ground state product by an energetic barrier. For HTI, this barrier is sufficiently low and the reactions occur on the timescale of 10 ps [5, 6]. The high sensitivity of the barrier height on molecular surroundings and on substitution may strongly influence the reaction dynamics as could be demonstrated for HTI, where the Z to E reaction could be varied from 3 ps to 3 ns by appropriate substitutions at the stilbene part [7].

The photochromism of fulgides, a class of photoactive compounds extensively studied in the context of molecular memories, is based on ring-closure and ring-opening reactions. In recent studies on specific fulgides (indolylfulgides; see Fig. 7.2d) ring-closure reactions were found, which are extremely fast (several 100 fs) and largely insensitive to the molecular surroundings [8]. On the other hand, ring opening of the same molecules is somewhat slower (∼10 ps) and proceeds via a barrier on the excited state potential energy surface [9]. This leads to a pronounced temperature dependence of reaction time and quantum efficiency [10]. Recently, it could be demonstrated that the ring-opening reaction could even be modified by vibrational excess energy supplied by a pre-illumination light pulse [11, 12]. When ring-opening/closure reactions should be used for peptide switching, one has to consider that these reactions do not lead to large-scale force-driven structural changes. Here modifications of the complete construct will mainly occur from allosteric reactions where the peptide structure adjusts to the changed geometry and flexibility of the molecular switch by (slow) thermal motions.

7.2.2 The Linking Group

A very important function in a light switchable chromopeptide is executed by the linking groups. They connect the light operated switch to one or several peptide units. The requirements imposed by the intended properties can be fulfilled by different types of linking groups. When, for example, the light switch operates as a backbone element the linking group should provide the appropriate amino- and carboxy-groups. This masking as a pseudo-amino acid also facilitates its direct utilization in peptide synthesis. On the other hand, an optical switch designed for side-chain switching must contain reactive groups addressing the targeted side chains.

First experiments on the photocontrol of azobenzene containing peptides were performed on model peptides, where the photoswitch was introduced into the side chains of poly-lysines [13, 14]. Dependent on the linking group the two families of azo-modified poly-lysines exhibited completely different photoinduced conformational changes. Only when a sulfonamide group linked the azobenzene to the lysine unit, a reversible photoresponsive formation of a helical and random coil structure was observed.

When a rigid connection between switch and peptide is intended, the bridging group cannot be formed exclusively by single bonds. However, one has to consider that some electronic coupling occurs between the π-electron systems of the two linked molecular parts. In addition, the extension of the π-system of the switch into the bridging group and the peptide may influence not only the chemistry of the molecular parts, but also the electronic properties of the optical switching itself. This may result in changes of absorption properties and may influence potential energy surfaces in the electronic ground state as well as in the excited electronic state [15]. Severe modifications in the switching properties and in reaction dynamics may result. For instance, the azobenzene compound APB (see Fig. 7.3), where carboxy- and amino-groups are directly attached to the azobenzene part, exhibits a shift of the main $\pi\pi^*$-absorption peak of the *trans* form from ∼320 nm (unsubstituted azobenzene) to 420 nm [16]. In addition (in the ground state), the interconversion from the *cis* to the *trans* form is accelerated from hours in azobenzene to the range of a few minutes in APB. For the switch HTI, it has been shown that the directly attached groups strongly influence the dynamics also for light-induced reactions. The electronic coupling may also depend on the position with respect to the aromatic system of the optical switch where the linker is attached. For HTI, it could be shown that the effect of electron donating or electron withdrawing groups is reduced if the group is attached at the meta- instead of the para-position of the stilbene part of HTI [7]. An efficient decoupling of the two electron systems of switch and peptide may be obtained via saturated groups (e.g., CH_2) in the linker. However, the presence of the single bonds in the linker leads to a more flexible construct, with severe consequences for the generation of structural changes in the complete chromopeptide. A rigid coupling typically transmits strong forces or torques from the switching chromophore to the peptide moiety and the peptide will directly and

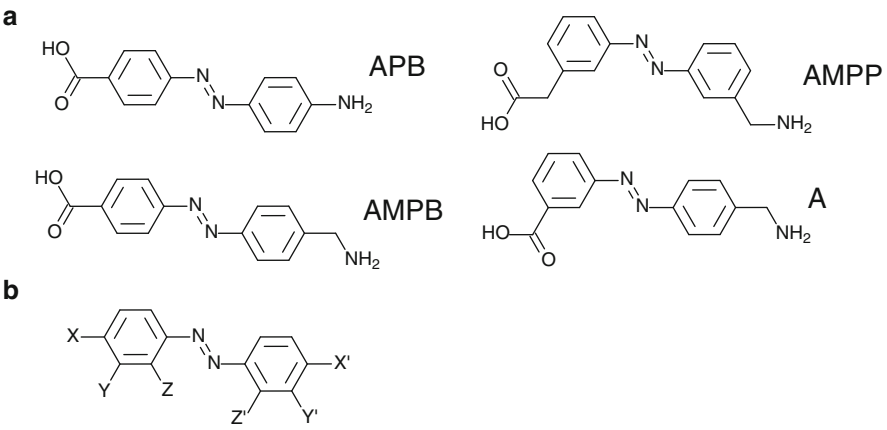

Fig. 7.3 (**a**) Azobenzene compounds used in the literature as backbone inserted photoswitches where different linking groups tune the chemical and structural properties. (**b**) Scheme showing different positions for substitutions

rapidly react on these forces. There will be strongly driven structural changes in the adjacent parts of the peptide that might even extend to more remote parts. Upon completion of the structural changes of the switch and rearrangements of the adjacent peptide groups further e.g., allosteric structural changes may occur. When a more flexible linker is used, the amount of force-driven structural changes is reduced and allosteric processes gain importance.

For the design of light switchable peptides, the structure of the switch itself, the position of the attachment, and the structure of the linker group are of major importance. The different constitutional isomers of azobenzene offer various positions where the two linking groups can be attached (Fig. 7.3b). There are a number of possibilities to connect peptide part I to positions X–Z and peptide part II to positions X′–Z′. The specific choice determines the structure of the construct before and after the switching photoisomerization and may strongly influence the geometric changes of the light-induced reaction. Four azobenzene constructs used for peptide switching are shown in Fig. 7.3a. In APB, the amino- and carboxy-groups are directly attached to the phenyl rings in para position. This switch-linker combination was used in cyclic and linear peptide constructs, showing pronounced and fast changes of the peptide in the vicinity of the switching chromophore group. When APB is used in a small cyclic system, the rigid construct strongly restricts the number of accessible states of the molecule. In simulations of the *cis*-to-*trans* reaction, it could be shown that even the thermally driven parts of the structural relaxation and reorganization are finished within 20 ns [17]. AMPP with amino- and carboxy-groups connected via methylene spacers to the azobenzene in meta position at both phenyl rings has been used as a switching element in a light-triggered hairpin model system. In the *cis* form of the AMPP the switch-linker construct acquires a geometry of a β-loop and the peptide folds into a β-hairpin, while in the *trans*

form the hairpin structure is destroyed [18]. Recently, a different azobenzene-based switch has been introduced (Fig. 7.3a, structure A) as a loop element in a cyclic model peptide forming a β-finger [19]. The selection of a specific molecular switch and the binding position for the selected linking groups in combination with the attached amino acid chain allows mimicking a wide range of structure elements found in native peptides. By this way, it becomes possible to produce switchable constructs for the different tasks reported above.

7.2.3 The Peptide Moiety

In order to obtain a light switchable peptide which can be used for the specific target application the peptide part has to meet important requirements. (a) The amino acid sequence has to be assigned in a way that the construct acquires intended molecular structures and/or function for both geometries of the molecular switch. This can be accomplished, e.g., by molecular dynamics simulations of the complete construct or by the use of sequences from known structural elements of proteins. The selection of a solvent with appropriate polarity may help to obtain stable structures of the dissolved chromopeptide [18]. A defined structure of a small peptide construct may be difficult to obtain. Here, additional constraints such as covalent bonds can help to reach the targeted structure. (b) The peptide moiety can be attached to the selected linking group with high efficiencies. (c) The amino acid sequence must be designed in a way that chemical reactions between the side chains and the optical switch or the linking groups are avoided. This can be accomplished by protecting groups on the side chains, which may be removed only after the assembly of the whole system. (d) Protecting groups used during peptide synthesis and the assembly of the complete construct have to be chosen in a way that they can be removed without damaging the construct. (e) The complete construct has to fit to the type of application and investigation. The solvent and the solubility of the light-triggered peptide have to be adapted to the techniques of investigation. For example, transient IR experiments often require high concentrations (in the mM range) and solvents with suitable IR transmission windows. In addition, aggregation of the construct should be negligible even at the high concentrations required for many ultrafast experiments.

7.3 Characterization of Light-Triggered Peptides by Stationary Spectroscopy

Optical spectroscopy in the visible and near-ultraviolet investigates the properties of the switching part of the light-triggered peptide. A comparison of the spectra of the isolated switch and the peptide construct yields information on the integrity of the switch after incorporation. By suitable illumination of the sample, one can

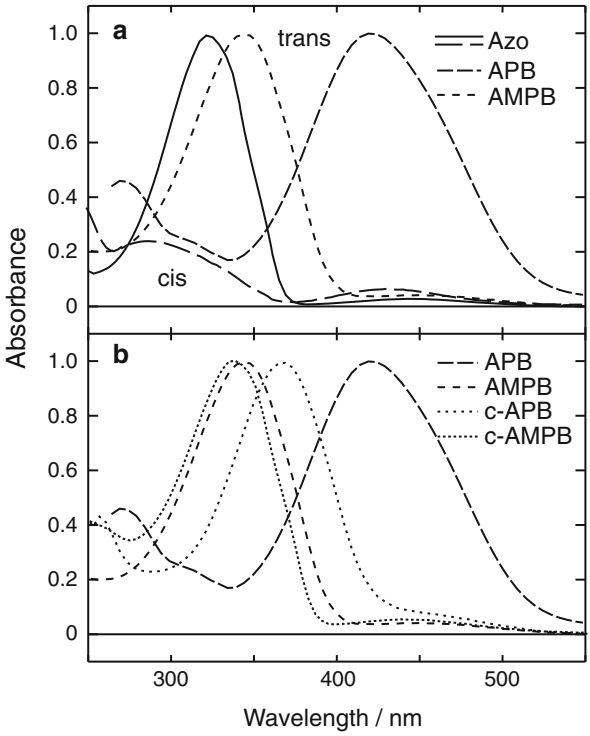

Fig. 7.4 Absorption spectra of (**a**) various azobenzene compounds azobenzene (Azo), APB, and AMPB and (**b**) the cyclic azobenzene peptides c-APB and c-AMPB

analyze if the incorporation into the peptide influences the switching properties or the equilibration between the two conformers. More detailed information may be obtained by a careful analysis of photostationary states reached by the illumination at specific wavelengths, by the determination of the quantum efficiency for light-induced switching or in quasistationary experiments, where the long time (> minutes) thermal reconversion of the sample is studied.

Figure 7.4 shows stationary absorption spectra of different azobenzene compounds. For the *trans* form of pure azobenzene (dissolved in DMSO), the absorption peak of the $\pi\pi^*$ transition is found at 320 nm and the weak $n\pi^*$ band at 450 nm. In the *cis* form, the absorption of the $n\pi^*$ band is increased and the $\pi\pi^*$ band shifts to shorter wavelengths ($\lambda_{max} = 275$ nm). For APB, i.e., azobenzene with directly attached amino and carboxyl groups, the charges at the terminal groups strongly shift the *trans* $\pi\pi^*$ absorption band to 420 nm. When APB is incorporated as a backbone switch into a peptide (see Fig. 7.4b, c-APB), the shift is reduced and the $\pi\pi^*$ *trans* absorption peaks now at ~370 nm [20]. These observations confirm that the rigid linking group used in APB strongly influences the electronic configuration of the system. In AMPB, where one methylene spacer is introduced the absorption spectrum of the isolated switch peaks at 345 nm, slightly red shifted as compared to pure azobenzene [21]. When AMPB is incorporated into the peptide, only a small additional blue shift of ~5 nm is found (see Fig. 7.4b, c-AMPB). The AMPP

switching unit which contains methylene spacers in both linking groups features the $\pi\pi^*$ *trans* absorption peak in the 325 nm range close to the absorption peak of unsubstituted azobenzene in solution [18, 22]. Thus, the methylene spacer – as expected – prevents the π-electron system of the switch to extend into the attached peptide.

Visible and ultraviolet spectroscopy of light-triggered peptides shows predominantly the few and broad bands of the molecular switch. The ultraviolet absorption bands of the peptide moieties (found below 300 nm) are quite weak and bear very limited information on the structural arrangement of the peptides. Thus, visible/near-ultraviolet spectroscopy can be used to yield information on the switching chromophore. On the other hand, infrared spectroscopy allows investigating *all* parts of the light switchable peptide. The advantage of infrared spectroscopy is evident when infrared difference absorption spectra are recorded. Of interest in light-triggered peptides are infrared absorption changes that occur upon operation of the molecular switch. Here, only molecular parts contribute which are influenced by the switching process. An example is given in Fig. 7.5 where infrared difference spectra of several azobenzene compounds are plotted. When azobenzene is switched by light at 370 nm from the *trans* to the *cis* form, distinct absorption changes are found at $\nu < 1{,}600\,\mathrm{cm}^{-1}$. This difference spectrum shows the disappearance of *trans*-azobenzene infrared bands as negative signals and the appearance of *cis* bands by a positive absorption change. For AMPB, strong absorption changes are found around $1{,}570\,\mathrm{cm}^{-1}$ and $1{,}600\,\mathrm{cm}^{-1}$. The difference spectrum of a bicyclic AMPB peptide, bc-AMPB, shows again these features related to the AMPB switch. However, much stronger absorption changes occur in the range of the amide I ($1{,}625$–$1{,}700\,\mathrm{cm}^{-1}$) and amide II ($1{,}500$–$1,570\,\mathrm{cm}^{-1}$) bands. They are connected

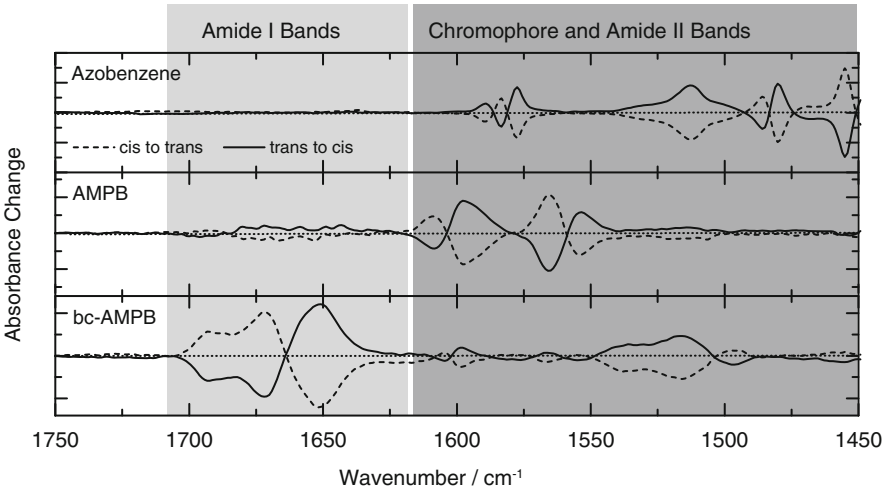

Fig. 7.5 Infrared difference spectra for azobenzene, AMPB, and the bicyclic azobenzene peptide bc-AMPB upon isomerization of the chromophore. Pronounced absorption changes from the chromophore and the peptide moiety are visible

with the spectral changes in the peptide moiety induced by the isomerization of the molecular switch. From the absorption difference spectrum (band positions and relative amplitudes), information on the structural changes can be obtained. The comparison with infrared spectra of model peptides and proteins can be used to obtain a first insight into the nature of these structural changes. Recently, theoretical modeling of infrared spectra of amide groups made considerable progress and the calculation of reliable infrared spectra of peptides of defined structure is foreseeable [23].

For proteins circular dichroism (CD) spectroscopy has proven to be an important technique. CD spectra in the 210 nm region contain signatures of the secondary structure. For light-triggered peptides CD spectroscopy gives a quick overview of changes in the arrangement between the two conformers. Most detailed structural information on the two isomers of the light switchable peptide is obtained from two-dimensional NMR. When applying 2D-NMR to small light switchable peptides, one has to consider the limitations of the technique: (a) A small peptide has many degrees of freedom and few through-space interactions. In most cases, a small peptide does not fold into a single structure. A well-defined structure may be reached if specific structural constraints are introduced, e.g., via covalent bonds or cyclization of the peptide. Often the peptide occupies an ensemble of molecular structures and the optical switch transforms one ensemble into another. (b) The structure determination in 2D-NMR relies on pairwise distance information between specific groups of the peptide. For small peptides, the limited number of distance information may not allow to define a single structure even if it would exist. When combined with high-quality molecular dynamic simulations, 2D-NMR data allow to characterize the ensembles connected with the two forms of the molecular switch. By this way, one obtains structural information for the two conformers, which is of major importance for the interpretation of the peptide dynamics.

7.4 Methods for the Study of Ultrafast Structural Dynamics

Folding reactions in proteins occur on a nested hierarchy of timescales and typically extends to the timescale of seconds (see Fig. 7.6). Molecular rearrangements of parts of the protein may occur much faster. For example, experiments on small peptides have shown that structural fluctuations of amino acids may occur on a timescale of picoseconds. Triplet–triplet energy transfer (TTET) experiments revealed that contact formation in short amino acid chains may occur down to the 100 ps time range [24]. Secondary structure elements are found to form within microseconds [25]. In the following, we present experimental techniques that allow studying even the fastest structural dynamics in the picosecond range.

We have seen above that the isomerization of the switching molecules may occur within a few 100 fs after the absorption of the exciting (pumping) light pulse. The excitation-induced changes in the optical properties of the light-triggered peptide can be recorded after a time delay t_D by a second, the probing light pulse of

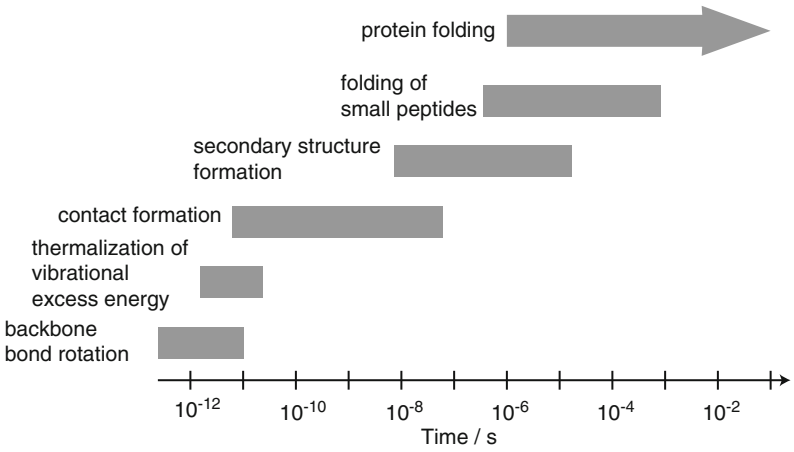

Fig. 7.6 Relevant timescales for folding dynamics of proteins and peptides

appropriate wavelength. Repeating the pump-probe experiment with different time delays, one can measure the light-induced absorption dynamics of the investigated peptide with high temporal resolution (see Fig. 7.7). During this experiment one has to assure that every pump-probe cycle investigates a fresh portion of the sample. This can be accomplished by a rapid exchange of the illuminated volume, e.g., via a peristaltic pump. The temporal resolution of the experiment is determined by the duration of the pump and probe pulses. Present experimental systems can generate pulses with durations well below 100 fs throughout the whole spectrum from the near ultraviolet to the mid infrared. The exact synchronization between pump and probe-pulse is necessary to maintain the high temporal resolution. On the timescale of femto- to nanoseconds this is accomplished by using a femtosecond laser system as the master light source. By means of a beam splitter a pair of exactly synchronized light pulses is generated from the original master pulse. Subsequent wavelength shifting via various nonlinear optical processes is used to generate pump and probe-pulses at the required spectral positions (see Fig. 7.7). The proper setting of the time delay between pump and probe-pulse is obtained by an optical delay line. For delay times t_D well above ca. 10 ns, an optical delay line becomes very difficult to operate since a pump probe experiment with the corresponding variable light path of (many) meters is difficult to realize. Therefore, one may use standard techniques of laser flash experiments for the investigation of slower processes. Recently, the high sensitivity of femtosecond IR experiments could be extended to longer time ranges by the electronic synchronization of two independent laser systems [26, 27].

By the specific choice of the probing wavelengths, the experiment may be tuned for the characterization of different parts of the light-triggered peptide. Experiments with probing in the visible and near-ultraviolet part of the spectrum record exclusively the dynamics of the light-operated switch in the peptide construct. The comparison of these results with those obtained for the isolated switch allows

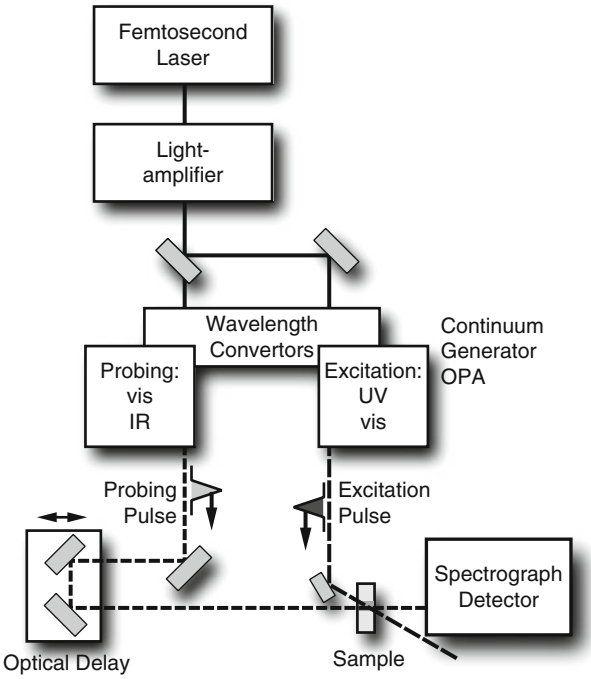

Fig. 7.7 Schematic of a typical experimental system (excite and probe setup) for ultrafast spectroscopy in the femto- to picosecond range

obtaining information on the interaction between switch and peptide. Experiments on azobenzene peptides have shown that the initial reactions dynamics of the switch are often slowed down as compared to pure azobenzene [21, 28, 29]. Apparently inertia and friction imposed by the attached peptide hinder the isomerizational motion of the switch. UV/visible experiments in the long wavelength part of the azobenzene absorption band revealed additional absorption transients on a much longer timescale where pure azobenzene did not show any transients (see Fig. 7.8a). These absorption changes could be assigned to the back action of the peptide onto the switch. On this timescale ($\approx$10 ps), the peptide has not yet reached its final arrangement and considerable strain between the isomerized switch and the peptide remains, which is released only on a longer timescale [21, 28, 29].

Probing in the mid infrared allows the direct access to structural changes of the peptide moiety. In most cases, amide I (1,620–1,700 cm^{-1}) and amide II bands (ca. 1,550 cm^{-1}) are investigated, which are very sensitive to structural rearrangements and changes in the hydrogen bonding pattern. Since the IR absorption changes induced by switching are very small (in most practical situations $\Delta OD < 10^{-3}$), elaborate optimization of infrared pulse generation and signal recording is necessary. Related experiments for different light-triggered peptides are presented below. Information on the vibrational system of a peptide may also be obtained by transient

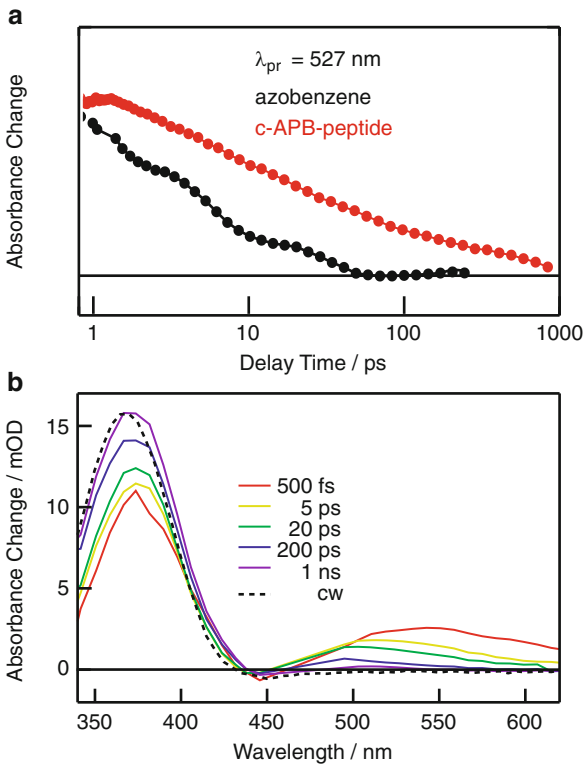

Fig. 7.8 (**a**) Comparison of the transient absorption changes of azobenzene and the cyclic azobenzene peptide c-APB for the cis to trans isomerization direction. (**b**) Absorption changes for c-APB recorded at different delay times showing dynamics with pronounced amplitudes in the time range of 50 ps due to the back action of the peptide on the isomerized chromophore

Raman techniques. However, up to date no transient Raman experiment has been published on the ultrafast dynamics of light-triggered peptides. Two-dimensional infrared (2D-IR) spectroscopy is another spectroscopic tool that combines structural information with ultrafast time resolution and is thus ideally suited to investigate structural dynamics of small peptides. This method has recently been reviewed [30, 31] and is thus not further discussed here.

7.5 Applications

7.5.1 Ultrafast Spectroscopy on Cyclic Azobenzene Peptides

In the first series of time-resolved experiments, light-triggered peptides have been investigated by ultrafast absorption spectroscopy in the visible and near-UV range. As discussed above, this type of experiment addresses the light switch itself and reveals its interactions with the peptide moiety. The related absorption transient for the *cis*-to-*trans* reaction of APB samples shows distinct differences between

the pure switch azobenzene, a linear azobenzene peptide where the strain between switch and peptide is released rapidly on the 10-ps timescale and the cyclic construct cAPB, where absorption transients extend out to 1 ns [21,29]. The transient spectra recorded at different delay times (see Fig. 7.8b) support this interpretation. Both the $n\pi^*$ absorption band around 450 nm and the $\pi\pi^*$ band around 370 nm are changed. Especially the $n\pi^*$ band is strongly shifted. The changes in absorption display dynamics in the 50 ps range together with slower reaction transients [21]. The simulation of the reaction dynamics with molecular dynamics techniques supports the given interpretation and finds a similar time constant for the relaxation of the strain. Very recent simulations have shown that the subsequent thermal processes establish the final *trans* arrangement of cAPB on the 20 ns timescale [17].

Further experiments on APB and AMPB-based azopeptides with visible fs spectroscopy have shown that the relaxation of the strain between switching chromophore and peptide moiety always occurs on the 10–100 ps timescale. This relaxation time depends on details of the linking groups, the peptide structure (linear, cyclic or bicyclic), and the viscosity of the solvent. The transfer of excess heat from the azobenzene to the surrounding solvent proceeds on the 10 ps timescale [4,21,32].

In the investigated peptides, the structural reorganization via force-driven processes and the complete thermal relaxation occur on the sub-100 ps timescale. Subsequent slower reactions will act on a molecule which is at the temperature of the solvent and contains the switch in the relaxed new conformation. As a consequence, the subsequent structural changes have to occur via thermal reaction steps bringing the system in an allosteric reaction on much longer timescales to its final arrangement [33].

Conformational transitions in the peptide moiety become clearly visible in time-resolved IR experiments [34]. For the bicyclic azobenzene peptide bc-AMPB, the transients revealed fast initial absorption changes representing bc-AMPB molecules with considerable vibrational excess excitation. It is on the 0–20 ps timescale that the spectra in the amide I region change considerably. These processes reflect the direct relaxation of the strain between chromophore and peptide and the release of vibrational energy. After this process, the IR difference spectrum already exhibits clear similarities to the stationary difference spectrum. However, a number of deviations are still present. On the timescale of sub-ns further relaxational processes occur; however, a complete agreement with the stationary spectrum is missing even after several ns. The slower processes observed in bc-AMPB indicate that even in peptides of small size the ensemble of molecules evolves on a multitude of timescales.

7.5.2 Unfolding and Folding of a Light Switchable Hairpin Model Compound

The azobenzene chromophore AMPP in its *cis* form can be used to mimic a β-loop and can serve as the central part of a β-hairpin structure [18]. When AMPP is in the

trans form, however, a β-hairpin structure cannot be formed. Thus, optical switching of the AMPP in the hairpin model peptide should transfer the system between states with different structural properties and the dynamics can be followed by ultrafast spectroscopy.

In the first study with stationary techniques (2D-NMR, CD spectroscopy, and IR spectroscopy), the ensembles of model peptides containing *cis*- and *trans*-AMPP, respectively, have been characterized [18]. The *cis*-AMPP hairpin peptide ensemble includes peptide structures with a larger amount of interstrand hydrogen bonds and with large probability to form the hairpin structure. This could be determined by NMR and CD spectroscopy. On the other hand, the peptides with *trans*-AMPP have a strongly reduced number of hydrogen bonds and no hairpin structure. Results from time-resolved IR experiments on the hairpin unfolding (*cis*-to-*trans* reaction of AMPP triggered with light at 400 nm) are shown in Fig. 7.9 [22]. At early times (within few ps, see Fig. 7.9b), one observes a first absorption decrease due to hydrogen bond breaking combined with the absorption changes expected in a vibrationally hot peptide. On the timescale of 20 ps, the vibrational excess energy relaxes and further hydrogen bonds break. In addition, one finds absorption increase around 1,690 cm^{-1}, which points to desolvated amide groups, i.e., to amide groups in a surrounding with reduced polarity. On the timescale of 50 ps resolvation occurs, which is visible by the disappearance of the 1,690 cm^{-1} band. However, no hydrogen bond reformation is found on this timescale. Finally with ∼500 ps, the system changes again its absorption by an absorption rise at 1,655 cm^{-1}. The absorption spectrum now closely resembles the stationary difference spectrum (see solid curve in Fig. 7.9c). Apparently the unfolding reaction of the AMPP hairpin model peptide is essentially completed on a timescale of 1 ns.

The inverse reaction, the folding reaction induced by the *trans*-to-*cis* isomerization of the AMPP photoswitch, shows again features related with vibrational heating (a broad absorption increase), desolvation, and subsequent resolvation processes on the sub-100 ps timescale. After these processes, the spectrum does not agree with the stationary difference spectrum. It is only on a much longer timescale that the stationary difference spectrum is reached (time constant 30 μs) in the region representing interstrand hydrogen bonds. Apparently the folding reaction is much slower than the photophysical processes occurring on the timescale of picoseconds. Entropic barriers have to be overcome before the right arrangement of the interchain hydrogen bonds of the β-hairpin is established.

7.5.3 Toward Light Switchable Tertiary Structures: (I) Azo-maquettes

The design of peptides with built-in chromophores that enable fast conformational changes by irradiation with monochromatic light is an approach that can be extended to the investigation of more complex structural transitions. Azobenzene derivatives

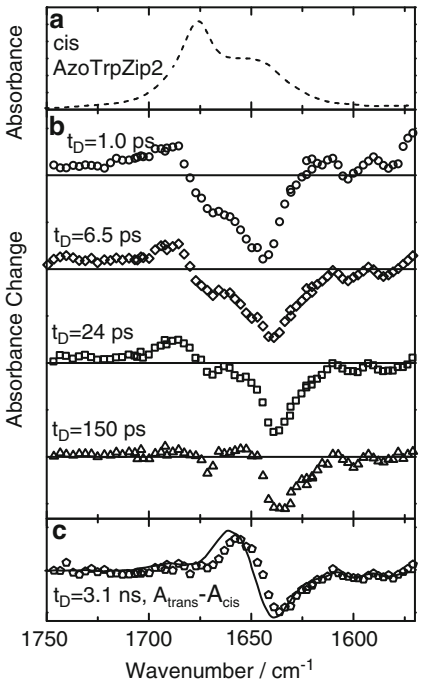

Fig. 7.9 Unfolding of the AMPP hairpin model peptide. (**a**) IR absorption spectrum recorded for the *cis* state of the AMPP chromophore. (**b**, **c**) Transient difference spectra taken at different delay times after switching the chromophore. *Solid line* in (**c**): Stationary absorption difference spectrum

can be used as photoswitches in larger model peptides with the potential to form stable tertiary structures. A 16mer α-helical system with a photoresponsive group has been described where the photoswitch is bound to the side chains of two cysteines [26]. Another approach is based on a photoswitchable α-helical peptide, composed of 30 amino acids (Fig. 7.10) [35].

The design of this peptide was based on an active site fragment that was already successfully used for the de novo synthesis of heme-binding maquettes [36]. Again, the photoresponsive group (AMPB) is directly incorporated into the peptide chain, which provides a direct link between switch and peptide moiety (Fig. 7.10, right). Second, this system can be extended for the construction of photoresponsive maquettes (similar to the maquette shown in Fig. 7.10, center) and then provide the possibility to monitor the dynamics of the chromopeptide assembly/disassembly via various spectroscopic assays. The behavior of AMPB and the hemes (incorporated into the peptide and solvent exposed, respectively) can be investigated by UV/vis spectroscopy and the behavior of the peptide as well as specific heme–peptide interactions could be monitored via IR spectroscopy. First time-resolved studies have demonstrated that it is possible to construct a photoresponsive helix of that size [35]. The data clearly indicate that although the photostability is slightly reduced compared to the smaller cyclic peptides, the system is stable enough to carry out the detailed spectroscopic investigations sketched above.

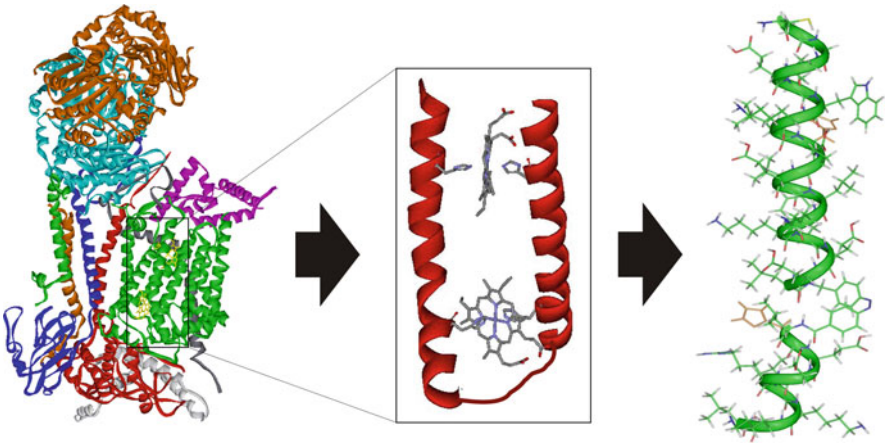

Fig. 7.10 The diheme cytochrome b subunit of the cytochrome bc_1 complex (*left*) served as a template for the synthesis of heme-binding proteins [36]. A di-helical peptide containing two histidines per helix was shown to bind two hemes (*middle*). The construct for photoswitchable maquettes is based on these of multi-heme proteins: two amino acids from the original sequence were replaced by an AMPB switch (*right*: molecular model of the *cis*-azopeptide obtained by INSIGHT II), amino acid sequence: Ac-CGGGELWKHEELLKKFEEL-AMPB-LHEERLKKL

7.5.4 *Toward Light Switchable Tertiary Structures: (II) Azo-collagens*

Collagen is the major stress-bearing component of connective tissue with a unique tertiary structure. Three parallel left-handed poly-Pro-II helices wind around a common axis to form a right-handed triple helix. This motif requires a characteristic primary structure of the individual strands with a Gly residue mandatory at every third position.

This Gly residue points inward and allows the tight packing of the triple helix. Collagen peptides consist of repeated (Gly-X-Y) triplets, with Gly-Pro-Hyp as most abundant and most stabilizing triplet. Local stability can be fine tuned by a variation of this motif. The triplets of the single strands are shifted by one amino acid in the triple helix, so that repetitive hydrogen bonds stabilizing the tertiary structure can be formed between the Gly (backbone amino group) and the Pro (backbone carbonyl group) residues.

Folding/unfolding of the collagen triple helix has been subject to extensive studies (for review see [37]). The currently accepted model for the formation of the triple helix is the association of the three collagen strands at the C-terminus and a subsequent zipper-like folding mechanism from the C- to the N-terminus with the slow *cis*/*trans* isomerization of Pro and Hyp as the rate-limiting steps. For a temporal resolution of the individual folding events, a system equipped with a fast conformational trigger is required.

Therefore, an azobenzene clamp was designed which when incorporated into the single collagen chains as side-chain-to-side-chain crossbridge can provide the required changes in the conformational space to trigger folding/unfolding of the collagen triple helix [38].

The model collagen is composed of three identical Ac-(Pro–Hyp–Gly)$_7$–Gly–Gly–NH_2 strands, which form a triple helix of relatively high thermal stability. As shown in Fig. 7.11 (top), a Pro and a Hyp residue were replaced in suitable positions with mercaptoproline, which formed the covalent bond with the azobenzene derivative. For an optimal transfer of the geometrical changes of the azobenzene switch to the collagen peptide, a rather rigid acetylene-type linker was chosen. Modeling studies predicted a self-association of the side-chain bridged collagen model peptide into a stable triple helix with the *trans*-azobenzene isomer, whereas its photoisomerization to the *cis* form leads to unfolding processes (Fig. 7.11, bottom) that can be directly assessed by NMR and IR spectroscopy [39]. Due to its regular composition, the IR spectra of triple-helical collagen peptides are characterized by well-defined amide bands [40, 41]. Time-resolved IR experiments can thus yield valuable information on the rate constants of formation for the periodic hydrogen bonding network. Although the isomerization of Pro is one of the rate-limiting steps in protein folding, it was observed that the fastest folding/unfolding processes of the triple-helical structure occur beyond the microsecond resolution of the stopped-flow techniques used in these studies [42]. In order to identify optimized conditions and to facilitate band assignments, temperature and light-induced unfolding was analyzed by FTIR. It turned out that unfolding of the tertiary structure can be induced both by heat and illumination. Temperature-dependent FTIR spectra of the azocollagen showed that the unfolding process is directly correlated to the H/D exchange at the glycine residues and that the amide II band is an excellent marker band to follow the destabilization and unfolding processes of the triple helix [43–45]. The situation is depicted in Fig. 7.12 a: the decreasing band around 1,555 cm^{-1} (amide II, N–H) and the increasing band around 1,475 cm^{-1} (amide II, N–D) are characteristic for the H–D exchange. It occurs, when the triple helix is destabilized or melts and the N–H protons, which are initially located on the inside of the triple helix and additionally stabilized by a hydrogen bond, are deprotected and become accessible to the solvent D_2O.

The azobenzene switch on the triple-helix is still able to photoisomerize, but the amplitude of the light-induced difference spectra, indicative for photomodulation of the tertiary structure, is clearly reduced at room temperature by the structural restraints of the triple helix. The photoinduced difference signal increases in amplitude with temperature (Fig. 7.12b), indicating that the conformational preferences of the azoswitch can only become effective upon weakening the interacting forces of the triple helix. At temperatures near the melting point, the maximum effects are reached and the azoswitch regains its full functionality. This change from a local disruption at room temperature to a complete unfolding at elevated temperatures upon *cis–trans* isomerization of the azobenzene switch is also clearly evidenced by the differences in the FTIR spectra shown in Fig. 7.12c.

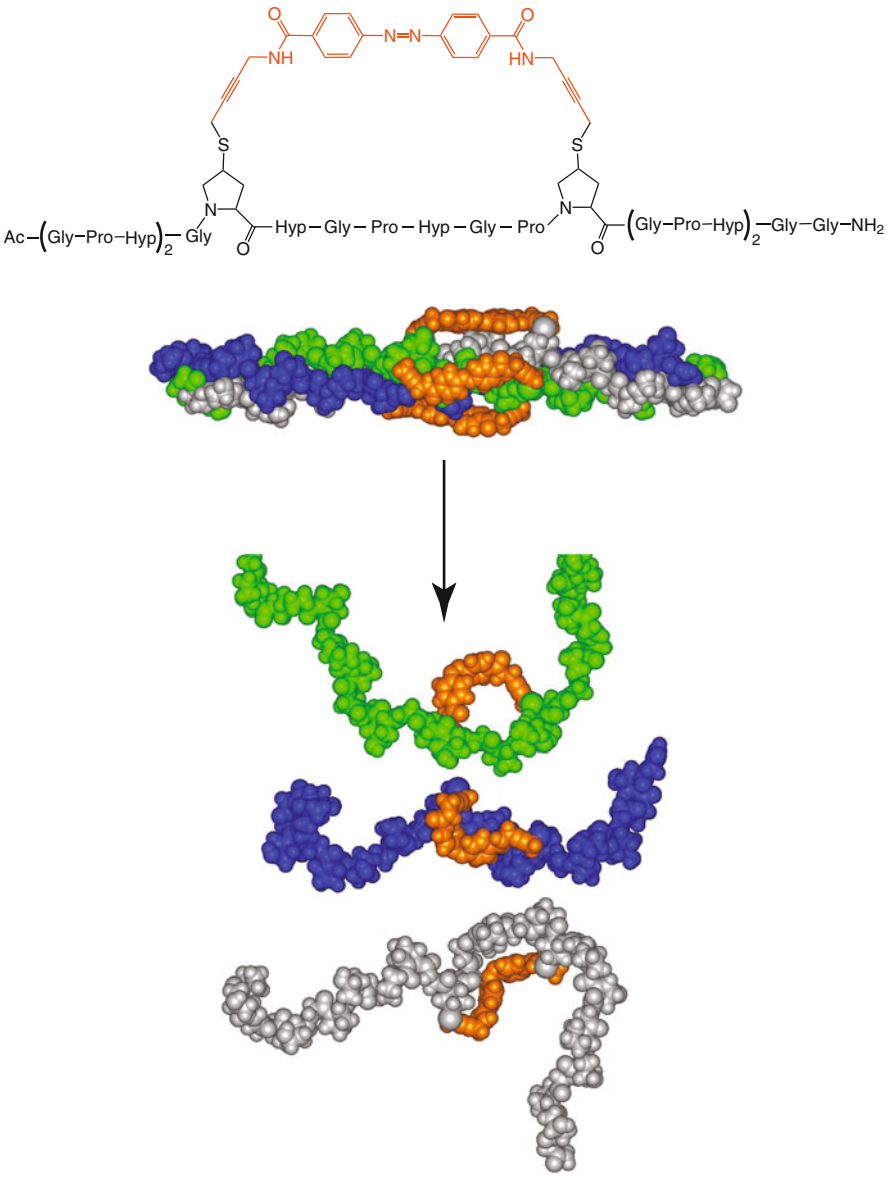

Fig. 7.11 Structure of a photoswitchable collagen peptide with an intramolecular side-chain-to-side-chain crossbridge containing the photoresponsive azobenzene moiety (*top*) and schematic illustration of the unfolding of the triple helix after *trans–cis* isomerization of the azoswitch. The azoswitch is depicted in orange, the three single strands in *blue, gray*, and *green*. This result of a molecular modeling was kindly provided by L. Moroder

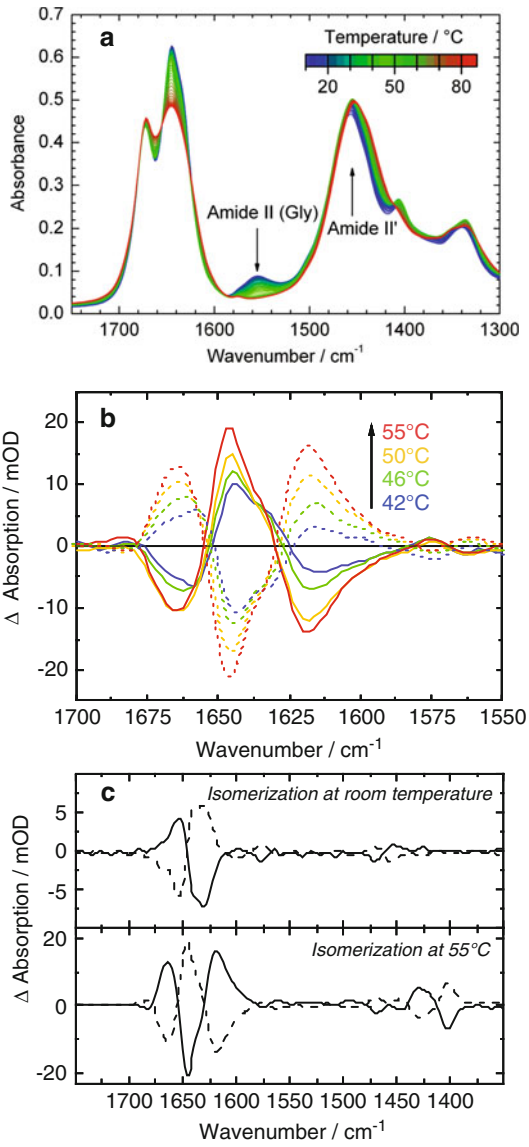

Fig. 7.12 (**a**) Temperature-dependent FTIR spectra of the azocollagen: The unfolding process is directly correlated to the H/D exchange at the glycine residue (amide II → amide II′). (**b, c**) Absorption changes induced by photoisomerization of the azobenzene moiety as a function of temperature; *trans*-to-*cis* (*dashed line*) and *cis*-to-*trans* (*dotted line*). Difference spectra of *trans*-to-*cis* (*dashed line*) and *cis*-to-*trans* (*dotted line*) isomerization at room temperature (*top*) and at the melting temperature (*bottom*)

7.6 Conclusion

The experiments on a number of light-triggered peptides with very different structure reveal certain common aspects of the structural dynamics. Direct force-induced structural changes of the peptide may occur ultrafast on the timescale of 10 ps. In these force-driven reactions predominantly amino acids in direct contact to the switching molecule are influenced. Apparently the flexibility of the amino acid

backbone prevents long-range ultrafast structural changes. On the same timescale, the excess energy released in the isomerization of the switch is dissipated to the surrounding solvent. On a longer timescale, the peptide moiety adjusts to the changed structure of the switch in thermal allosteric processes. Here, the nature of the structural changes and the restrictions imposed by enthalpic or entropic barriers determines the timescales. Even in the small peptides investigated up to now these processes may span the range between 100 ps and tens of μs. The investigations on light-triggered peptides have convincingly shown that ultrafast processes occur in folding reactions. It is evident that these processes have to be considered for a full description of protein folding.

References

1. A. Cattani-Scholz, C. Renner, C. Cabrele, R. Behrendt, D. Oesterhelt, L. Moroder, Angew. Chem. Int. Ed. **41**, 289 (2002)
2. R.H. Kramer, D.L. Fortin, D. Trauner, Curr. Opin. Neurobiol. **19**, 1 (2009)
3. C. Renner, L. Moroder, Chembiochem **7**, 869 (2006)
4. T. Nägele, R. Hoche, W. Zinth, J. Wachtveitl, Chem. Phys. Lett. **272**, 489 (1997)
5. T. Cordes, C. Elsner, T.T. Herzog, C. Hoppmann, T. Schadendorf, W. Summerer, K. Rück-Braun, W. Zinth, Chem. Phys. **358**, 103 (2009)
6. T. Cordes, D. Weinrich, S. Kempa, K. Riesselmann, S. Herre, C. Hoppmann, K. Rück-Braun, W. Zinth, Chem. Phys. Lett. **428**, 167 (2006)
7. T. Cordes, T. Schadendorf, B. Priewisch, K. Rück-Braun, W. Zinth, J. Phys. Chem. A **112**, 581 (2008)
8. T. Cordes, T.T. Herzog, S. Malkmus, S. Draxler, T. Brust, J.A. DiGirolamo, W.J. Lees, M. Braun, Photochem. Photobiol. Sci. **8**, 528 (2009)
9. H. Port, P. Gärtner, M. Hennrich, I. Ramsteiner, T. Schöck, Mol. Cryst. Liq. Cryst. **430**, 15 (2005)
10. T. Brust, S. Malkmus, S. Draxler, S.A. Ahmed, K. Rück-Braun, W. Zinth, M. Braun, J. Photochem. Photobiol. A **207**, 209 (2009)
11. T. Brust, S. Draxler, J. Eicher, W.J. Lees, K. Rück-Braun, W. Zinth, M. Braun, Chem. Phys. Lett. **489**, 175 (2010)
12. S. Draxler, T. Brust, S. Malkmus, J.A. Di Girolamo, W.J. Lees, W. Zinth, M. Braun, Phys. Chem. Chem. Phys. **11**, 5019 (2009)
13. A. Fissi, O. Pieroni, F. Ciardelli, Biopolymers **26**, 1993 (1987)
14. A. Fissi, O. Pieroni, E. Balestreri, C. Amato, Macromolecules **29**, 4680 (1996)
15. O. Sadovski, A.A. Beharry, F.Z. Zhang, G.A. Woolley, Angew. Chem. Int. Ed. **48**, 1484 (2009)
16. R. Behrendt, C. Renner, M. Schenk, F. Wang, J. Wachtveitl, D. Oesterhelt, L. Moroder, Angew. Chem. Int. Ed. **38**, 2771 (1999)
17. R. Denschlag, W. Schreier, B. Rieff, T. Schrader, F.O. Koller, L. Moroder, W. Zinth, P. Tavan, Phys. Chem. Chem. Phys. **12**, 6204 (2010)
18. S.-L. Dong, M. Löweneck, T.E. Schrader, W.J. Schreier, W. Zinth, L. Moroder, C. Renner, Chem. Eur. J. **12**, 1114 (2006)
19. C. Hoppmann, S. Seedorff, A. Richter, H. Fabian, P. Schmieder, K. Rück-Braun, M. Beyermann, Angew. Chem. Int. Ed. **48**, 6636 (2009)
20. R. Behrendt, C. Renner, M. Schenk, F.Q. Wang, J. Wachtveitl, D. Oesterhelt, L. Moroder, Angew. Chem. Int. Ed. **38**, 2771 (1999)
21. S. Spörlein, H. Carstens, H. Satzger, C. Renner, R. Behrendt, L. Moroder, P. Tavan, W. Zinth, J. Wachtveitl, Proc. Natl. Acad. Sci. U.S.A. **99**, 7998 (2002)

22. T.E. Schrader, W.J. Schreier, T. Cordes, F.O. Koller, G. Babitzki, R. Denschlag, C. Renner, M. Löweneck, S.-L. Dong, L. Moroder, P. Tavan, W. Zinth, Proc. Natl. Acad. Sci. U.S.A. **104**, 15729 (2007)
23. V. Schultheis, R. Reichold, B. Schropp, P. Tavan, J. Phys. Chem. B **112**, 12217 (2008)
24. B. Fierz, H. Satzger, C. Root, P. Gilch, W. Zinth, T. Kiefhaber, Proc. Natl. Acad. Sci. U.S.A. **104**, 2163 (2007)
25. O. Bieri, T. Kiefhaber, Biol. Chem. **380**, 923 (1999)
26. J. Bredenbeck, J. Helbing, J. Kumita, G. Woolley, P. Hamm, Proc. Natl. Acad. Sci. U.S.A. **102**, 2379 (2005)
27. M. Towrie, A. Gabrielsson, P. Matousek, A. Parker, A. Rodriguez, A. Vlcek, Appl. Spectrosc. **59**, 467 (2005)
28. H. Satzger, C. Root, C. Renner, R. Behrendt, L. Moroder, J. Wachtveitl, W. Zinth, Chem. Phys. Lett. **396**, 191 (2004)
29. J. Wachtveitl, S. Spörlein, H. Satzger, B. Fonrobert, C. Renner, R. Behrendt, D. Oesterhelt, L. Moroder, W. Zinth, Biophys. J. **86**, 2350 (2004)
30. J. Bredenbeck, J. Helbing, C. Kolano, P. Hamm, ChemPhysChem **8**, 1747 (2007)
31. P. Hamm, J. Helbing, J. Bredenbeck, Annu. Rev. Phys. Chem. **59**, 291 (2008)
32. V. Botan, E.H.G. Backus, R. Pfister, A. Moretto, M. Crisma, C. Toniolo, P.H. Nguyen, G. Stock, P. Hamm, Proc. Natl. Acad. Sci. U.S.A. **104**, 12749 (2007)
33. M. Löweneck, A.G. Milbradt, C. Root, H. Satzger, W. Zinth, L. Moroder, C. Renner, Biophys. J. **90**, 2099 (2006)
34. J. Bredenbeck, J. Helbing, A. Sieg, T. Schrader, W. Zinth, C. Renner, R. Behrendt, L. Moroder, J. Wachtveitl, P. Hamm, Proc. Natl. Acad. Sci. U.S.A. **100**, 6452 (2003)
35. S. Rehm, M.O. Lenz, S. Mensch, H. Schwalbe, J. Wachtveitl, Chem. Phys. **323**, 28 (2006)
36. D.E. Robertson, R.S. Farid, C.C. Moser, J.L. Urbauer, S.E. Mulholland, R. Pidikiti, J.D. Lear, A.J. Wand, W.F. DeGrado, P.L. Dutton, Nature **368**, 425 (1994)
37. H.P. Bächinger, J. Engel, in *Protein Folding Handbook*, vol Part II, ed. by J. Buchner, T. Kiefhaber (Wiley-VCH, Weinheim, 2005), p. 1059
38. U. Kusebauch, S.A. Cadamuro, H.J. Musiol, L. Moroder, C. Renner, Chem. Eur. J. **13**, 2966 (2007)
39. U. Kusebauch, S.A. Cadamuro, H.J. Musiol, M.O. Lenz, J. Wachtveitl, L. Moroder, C. Renner, Angew. Chem. Int. Ed. **45**, 7015 (2006)
40. Y.A. Lazarev, B.A. Grishkovsky, T.B. Khromova, Biopolymers **24**, 1449 (1985)
41. K.J. Payne, A. Veis, Biopolymers **27**, 1749 (1988)
42. A. Bachmann, T. Kiefhaber, S. Boudko, J. Engel, H.P. Bachinger, Proc. Natl. Acad. Sci. U.S.A. **102**, 13897 (2005)
43. S.W. Englander, N.R. Kallenbach, Q. Rev. Biophys. **16**, 521 (1983)
44. J. Backmann, C. Schultz, H. Fabian, U. Hahn, W. Saenger, D. Naumann, Proteins **24**, 379 (1996)
45. T.M. Raschke, S. Marqusee, Curr. Opin. Biotechnol. **9**, 80 (1998)

Chapter 8
Time-Resolved FTIR Spectroscopy of pH-Induced Aggregation of Peptides

John E.T. Corrie, Alex Perálvarez-Marín, and Andreas Barth

Abstract Reaction-induced infrared difference spectroscopy is a sensitive method to detect absorbance changes that accompany biomolecular reactions, even if they are very small. One of the ways to trigger reactions in the infrared cuvette is the use of caged compounds, photosensitive molecules that release a desired effector molecule when irradiated with near-UV light, and experiments with caged nucleotides and the sarcoplasmic reticulum Ca^{2+}-ATPase are used to introduce the methodology. A caged sulfate, which can be used to rapidly acidify protein or peptide samples in order to induce unfolding and misfolding, is discussed in detail. Applications described are the partial unfolding of myoglobin and the aggregation of the Alzheimer's peptide.

8.1 Introduction to Infrared Difference Spectroscopy

8.1.1 Principles

Infrared spectroscopy is one of the standard methods for the structural investigation of chemical compounds. When a molecule has a small number of atoms N, the $3N$- 6 normal modes of vibration lead to distinct absorption bands in the infrared spectrum. However, when the number of atoms becomes large, the spectrum is composed of many overlapping bands, so its information content is limited. The key to obtaining detailed structural information is to reduce the number of observed groups. This can be done by different techniques that are particularly suited for the study of biochemical reactions.

A. Barth (✉)
Stockholm University, Department of Biochemistry and Biophysics, Arrhenius Laboratories, 10691 Stockholm, Sweden
e-mail: barth@dbb.su.se

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*, Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_8,

In general, a difference spectrum is obtained by subtracting the absorption spectrum of a sample in a particular state A from a spectrum where it is in a state B. The resulting difference spectrum shows bands only from those groups that are affected by the change from state A to B. All "passive" residues are invisible in the difference spectrum, so the number of observed groups is dramatically reduced compared to the absorption spectrum. Therefore, a difference spectrum exhibits details of the reaction mechanism on the molecular level despite a large background absorption, and environmental changes around just a few atoms in a large protein can be detected [1].

In favourable cases, a change of state of a biomolecule can be observed in the infrared absorption spectrum. More usually, the effects are so small that subtracting spectra obtained from different samples is not sensitive enough. Instead, the transition between the two states of interest has to be initiated directly in the cuvette. This technique is termed *reaction-induced infrared difference spectroscopy* [2–5] and is illustrated in Fig. 8.1. In a typical experiment, the biomolecule is prepared in a stable state A and a spectrum of this state is measured. Then the reaction is triggered, the protein proceeds to state B and again the spectrum is recorded. From the spectrum recorded before the start of the reaction (state A) and the spectrum recorded after the reaction (state B) a difference spectrum is calculated. This reflects the changes in infrared absorption associated with the reaction. Instead of only one final state B, a sequence of transient states may be adopted in the course of the reaction. In this case, the interconversion between the product states B_1, B_2, etc. can be followed by time-resolved methods and is reflected in the difference spectra.

Reaction-induced difference spectroscopy combines several of infrared spectroscopy's advantages: (a) high time resolution ($\sim$10 ms for a complete spectrum, as discussed in this chapter, and $<$1 μs for single wavelength information in individual experiments. Full spectral information with $<$1 μs resolution can be obtained for systems that can be reproducibly excited several thousand times); (b) universal applicability from small soluble proteins to large membrane proteins; (c) the high molecular information content; and (d) a sensitivity high enough to detect a change in bond strength of one bond in a large protein.

8.1.2 Triggering Protein Reactions

A crucial problem in reaction-induced difference spectroscopy is how to trigger the protein reaction of interest. The number of methods for this has constantly increased in the last decade and the main approaches are:

1. Inducing photoreactions of biomolecules by illumination [4–11].
2. Releasing compounds from photosensitive and biologically inactive precursors by illumination [12–14]. The precursors are called caged compounds [15, 16] and the approach is discussed in more detail below.

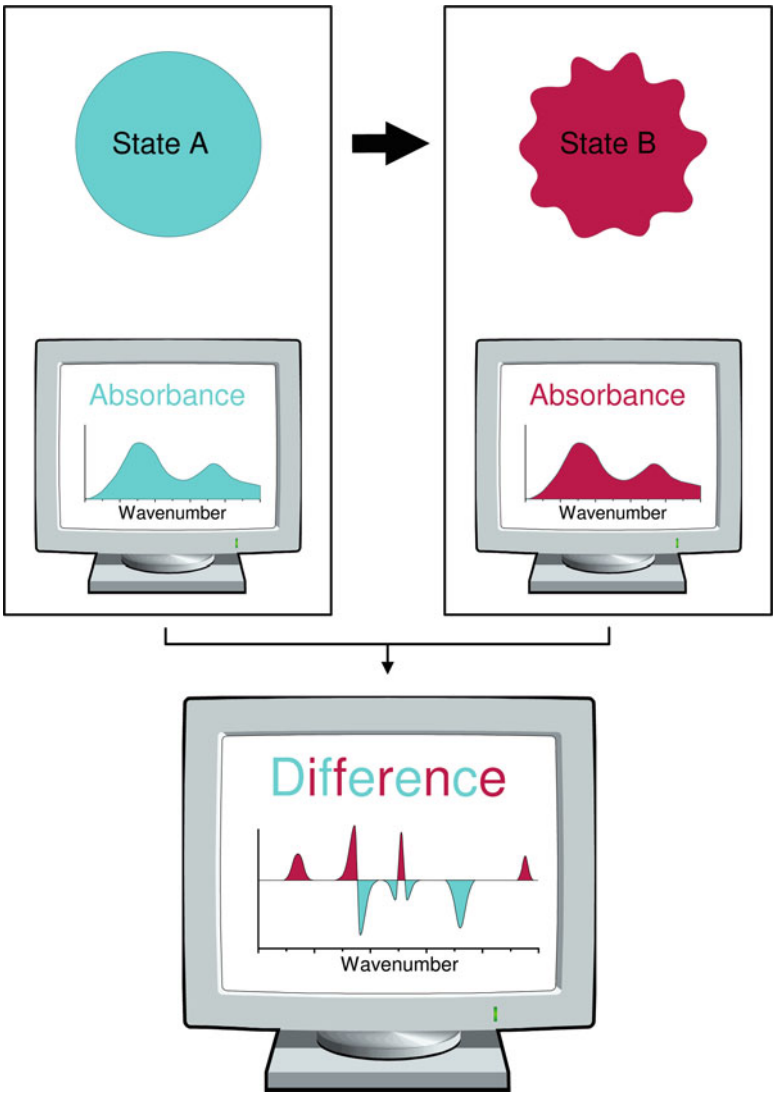

Fig. 8.1 The principle of infrared difference spectroscopy. See text for further explanation

3. Rapid mixing techniques. These are difficult to apply in infrared spectroscopy because of the viscous consistency of a typical sample and the small path length of less than 10 μm that is dictated by the strong $^{1}H_2O$ absorption ($\leq$50 μm for measurements in $^{2}H_2O$) [17, 18].
4. Attenuated total reflection (ATR) measurements [4, 5, 19, 20]. In an ATR experiment, a sample is placed on a crystal and infrared radiation is passed into the crystal at an angle such that the light undergoes total internal reflection at

the crystal–sample interface. The total reflection process generates an evanescent wave that extends about one wavelength away from the crystal surface and senses the absorption of the sample in this layer. The sample is usually a film, prepared by drying, with a buffer solution above it and the buffer can be exchanged to induce reactions in the sample. In a recent extension of the ATR technique, a dialysis membrane separates a sample compartment close to the ATR crystal and a reservoir. In this way, the medium in the reservoir can be altered without disturbing the protein sample in the sample compartment [21]. A film is not required, making the technique also applicable for soluble biomolecules.

5. Temperature and pressure jumps to study unfolding of proteins [22].
6. Equilibrium electrochemistry to initiate redox reactions using an ultra-thin-layer spectroelectrochemical cell [4, 5, 23].
7. Photoreduction by photoexcitable electron donors, so-called "caged electrons" [24], to induce redox reactions.

8.1.3 Interpreting Difference Spectra

Infrared spectra are sensitive to a number of factors, which can be exploited to obtain molecular information on the absorbing molecules and their environment:

The *chemical structure* of a molecule is the dominating effect that determines vibrational frequencies via the strengths of the vibrating bonds and the masses of the vibrating atoms. Because of this, changes of protonation state and the conversion of substrates to products in enzymatic reactions can be followed.

The *conformation* of a molecule affects the interaction of vibrating atoms with their environment and the coupling between vibrations. The latter depends on details of the molecular geometry, such as bond angles and proximity of the coupled groups. Since interaction and coupling alter vibrational frequencies, they can be detected in the infrared spectrum and provide insight into the three-dimensional structure of molecules.

Environmental effects such as hydrogen bonding and the electric field produced by the surroundings modify the electron density distribution of a given molecule and will thus affect the infrared spectrum.

Conformational freedom affects the width of infrared bands. Infrared spectroscopy provides a snapshot of the sample conformer population due to its short characteristic timescale (on the order of 10^{-13} s). As the band position for each conformer is usually slightly different, this heterogeneity results in band broadening. Flexible structures will thus give broader bands than rigid structures and the band width is a measure of conformational freedom.

The sensitivity of the infrared spectrum to many different factors is an advantage and a drawback at the same time: at first an infrared spectrum may seem hopelessly crowded with many overlapping bands. However, with increased understanding the spectrum gives detailed information about the sample since it allows one to monitor a variety of parameters besides the molecular events that are of prime interest.

Examples are protein concentration, enzyme activity, progress of caged compound photolysis and its yield, as well as buffer protonation or deprotonation.

The above factors are also those that cause bands in difference spectra, as illustrated in Fig. 8.2. Negative bands in a difference spectrum are characteristic of the state before the reaction and positive bands of the state(s) after or during the reaction. A change in chemical structure gives rise to a different absorption spectrum

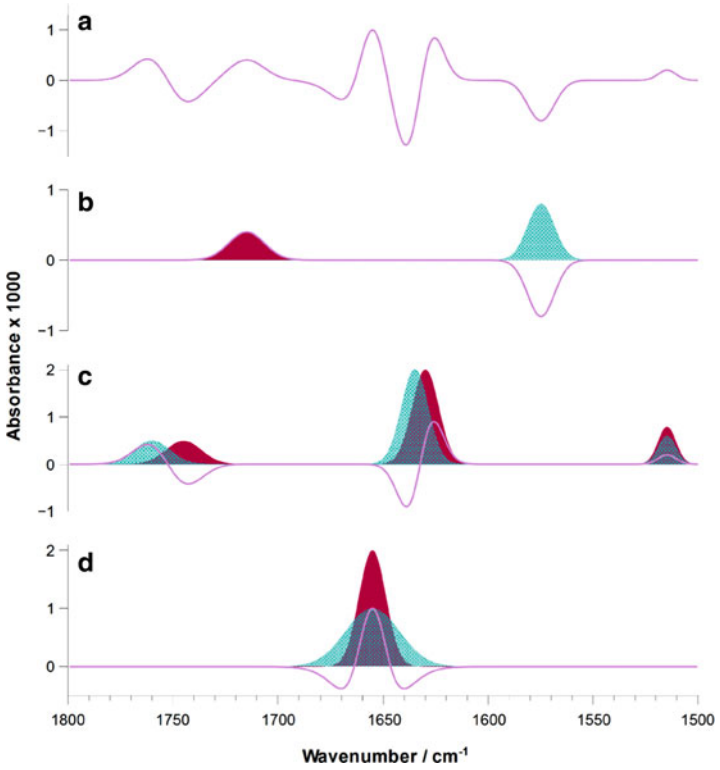

Fig. 8.2 The components of a difference spectrum. (**a**) A hypothetical difference spectrum obtained by calculating the absorbance of the final state B minus the absorbance of the initial state A. The difference spectrum is the composite of the spectra shown in panels (**b–d**). (**b**) Spectral effects of a chemical reaction, for example the protonation of a carboxyl group. The absorption of state A (*turquoise*) disappears and that of state B (*red*) appears. As a result, there are negative and positive bands in the spectrum which are often far apart. (**c**) Spectral effects of a change in conformation or environment. This often leads to small band shifts which result in paired positive and negative bands in the difference spectrum. The examples shown are for an environmental change around a protonated carboxyl group ($\sim$1,750 cm^{-1}) and for a conformational change of a β-sheet (1,635 cm^{-1}, for example a better alignment of the strands in the sheet). It may also occur that the predominant effect is a change in absorption coefficient and that a band shift is not obvious in the difference spectrum. In this case only a positive or a negative band is observed. This is illustrated at $\sim$1,515 cm^{-1}, characteristic for the side chain of Tyr. (**d**) Spectral effects of a flexibility change. In the example shown, an α-helix becomes more rigid, which makes the band of state B narrower than that of state A

before and after the reaction (Fig. 8.2b). Changes of conformation and environment cause band shifts and alterations of the infrared absorption coefficient (Fig. 8.2c). Changes in flexibility lead to a band that is flanked by two minor bands of opposite sign (Fig. 8.2d).

Not all features in a difference spectrum are readily interpreted and a first step is to regard the spectra as a fingerprint of the structural change. In order to extract more molecular information, additional experiments are often necessary. Common approaches for the molecular interpretation of spectra are modification of the sample (mutation, isotopic labelling and changes to substrate or protein cofactor), study of model compound spectra and comparison with computed spectra.

8.2 Caged Compounds

8.2.1 Introduction to Caged Compounds

A *photolytically induced concentration jump* can be achieved with photosensitive molecules that release a biologically active effector upon illumination in the UV spectral range (300–350 nm) [15,16]. These molecules are termed caged compounds and have been used for 20 years to study biological reactions with infrared spectroscopy [12]. In its caged form, the effector is modified such that it does not react with the biomolecule of interest. Flash photolysis of the caged compound leads on the μs to ms timescale to a concentration jump of the free effector, which initiates a biological reaction. A recent extension of the approach uses helper enzymes to induce a series of consecutive reactions [25]. In addition to protein and effector molecule bands, the photolysis reaction is reflected in the difference spectra [12–14, 26, 27].

In infrared studies of proteins, caged nucleotides, caged Ca^{2+} (Nitr-5 or DM-Nitrophen) and "caged electrons" have been used most often. The studies have dealt with two main aspects: substrate–enzyme recognition and the molecular basis of enzyme function. Most of these studies have been done on the sarcoplasmic reticulum Ca^{2+}-ATPase, which was the first enzyme to be studied with this technique [12].

In order to illustrate the approach and also different interpretation strategies, we describe two aspects of our work in more detail. More information can be found in a recent review [28]. The first study explored the use of infrared spectroscopy for the characterization of binding sites on proteins: it mapped the interactions between Ca^{2+}-ATPase and its substrate ATP. ATP binding induces a change in protein conformation which alters the amide I absorption (1,700–1,610 cm^{-1}) of the protein, as shown in Fig. 8.3. This absorption is mainly due to the stretching vibration of the backbone carbonyl groups and reflects protein structure. The signals near 1,693, 1,641 and 1,628 cm^{-1} are characteristic of β-sheets, those near 1,665 cm^{-1} are suggestive of turns and those near 1,653 cm^{-1} are indicative of

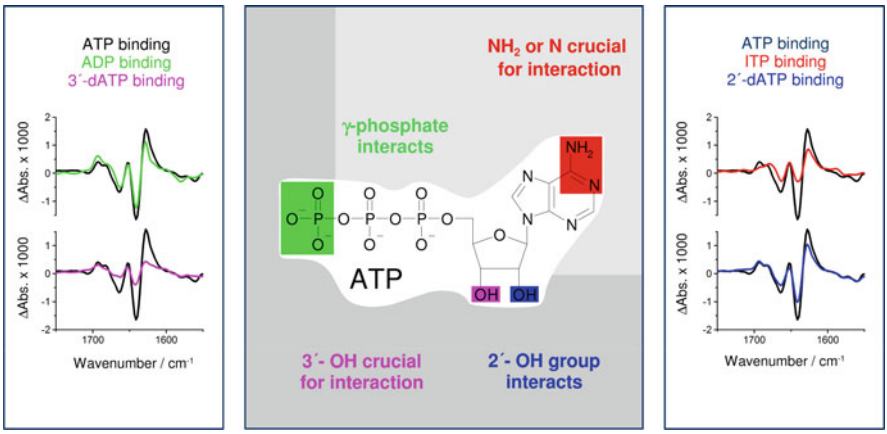

Fig. 8.3 Mapping the ATP-binding site of the Ca^{2+}-ATPase with infrared spectroscopy. *Left* and *right panel*: difference spectra caused by binding of ATP and ATP analogues: ATP (*black*), ADP (*green*), 3′-deoxyATP (purple), ITP (red) and 2′-deoxyATP (*blue*). *Middle panel*: Summary of the detected interactions. The protein environment is shown in *grey* and closeness between the *grey* area and the chemical structure of ATP indicates interaction between ATP and protein. The different shadings of *grey* symbolize different domains. The nucleotide-binding domain is represented by *light grey* and the phosphorylation domain by *dark grey*

α-helical structures. From the spectra, it can be concluded that α-helices, β-sheets and turns are affected by nucleotide binding [29].

Close ATP analogues produce nucleotide-binding spectra that are different from that obtained with ATP (see Fig. 8.3). Therefore, the conformational change upon nucleotide binding depends to a surprising degree on individual interactions between ATPase and nucleotide [29, 30]. The lack of individual interactions produces more than just local adjustments; it affects the entire conformation of the nucleotide–ATPase complex. Surprisingly, modification at different locations of the ATP molecule, interacting with different protein domains, produces similar effects. The spectra obtained with 3′-deoxyATP and ITP for example show a dramatic decrease of most bands in the amide I region, but their shape is still similar to the shape of the spectrum obtained with ATP. This suggests that all secondary structure elements involved in ATP binding are affected in a similar way by modifications at different sites of the ATP molecule. Thus, they do not adapt independently to the interaction sites provided by ATP, but move instead in a concerted way for which all interactions need to be in place [29]. As a consequence, the (average) structure of the nucleotide–ATPase complex is characteristic of the particular nucleotide bound.

From the sensitivity of the conformational change to individual interactions, it has been concluded that the ATPase interacts with the γ-phosphate [30], the ribose hydroxyls, and the amino function [29] of ATP. The interactions identified by infrared spectroscopy have later been confirmed by X-ray crystallography [31, 32]. The X-ray studies also provided an explanation for the drastic effect exerted by the modification of individual functional groups of ATP. Binding of ATP closes a

cleft between nucleotide-binding and phosphorylation domain of the ATPase. The two domains are bridged by ATP in the ATP–ATPase complex, which stabilizes the closed conformation like the filling of a sandwich that holds the two slices of bread together. When ATP is modified, the "glue" between the domains is partially lost and the conformation is more open than in the ATP complex.

Transfer of the nucleotidic γ-phosphate to Asp351 follows nucleotide binding. In this reaction, an interesting observation has been made with ITP: the phosphorylation spectrum obtained with ITP shows additional signals in the amide I region as compared to ATP, which are similar to nucleotide-binding signals. In contrast, the previous step of ITP binding gave much smaller signals than that of ATP binding. Thus in the phosphorylation reaction with ITP, the enzyme seems to catch up on a conformational change that could not be achieved in the binding step where the interactions between protein and base moiety were impaired [33].

The results discussed so far in this chapter were obtained by monitoring the conformational change of the peptide backbone. The signals in the amide I region were simply considered as a fingerprint of the conformational change of the protein and mainly their amplitude was evaluated. In spite of this simplistic interpretation, molecular information on the interacting groups of the substrate ATP has been obtained because the substrate could be modified. Small modifications are also the basis for other interpretation strategies, for example those based on mutation of protein residues and on isotopic labeling.

Isotopic labeling was used in the second aspect of our work presented here. The aim was to obtain information on the interactions between bound phosphate groups and ATPase. Since the vibrational frequency depends on the masses of the vibrating atoms, isotopic labeling shifts bands of the labeled group, which can then be identified in the spectrum. In studies of nucleotide binding to the Ca^{2+}-ATPase, β- and γ-phosphates were separately [^{18}O]-labeled. The spectral positions of the β- and γ-phosphate bands indicate that P–O bond strengths of bound β- and γ-phosphate are similar to those of ATP in aqueous solution; i.e. that hydrogen bonds of ATP to water are largely replaced by interactions with the protein for bound ATP [34] and that the protein environment makes little difference compared to water to the strength of the bonds within the bound phosphate groups.

In another study, the second ATPase phosphoenzyme intermediate E2P has been investigated to understand its high rate of hydrolysis of the aspartyl phosphate ester bond. In the experiment, ATP photoreleased from caged ATP transferred its γ-phosphate to Asp351 of the ATPase and led to the transient accumulation of E2P. This intermediate is long-lived under the conditions of the experiment and is in quasi-equilibrium with the unphosphorylated E2 state and non-covalently bound phosphate. Dephosphorylation occurs, but is followed by rapid rephosphorylation to regenerate E2P. This dephosphorylation/rephosphorylation cycle results in incorporation of oxygen from water into the phosphate of E2P. In the experiment, the oxygen isotopes of water and of ATP's γ-phosphate were different, which led to an oxygen isotope exchange at the phosphate group. This produced a difference spectrum, shown in Fig. 8.4a, to which only the isotope exchange reaction contributed [1]. In this way, the experiment selectively observed a group of only four

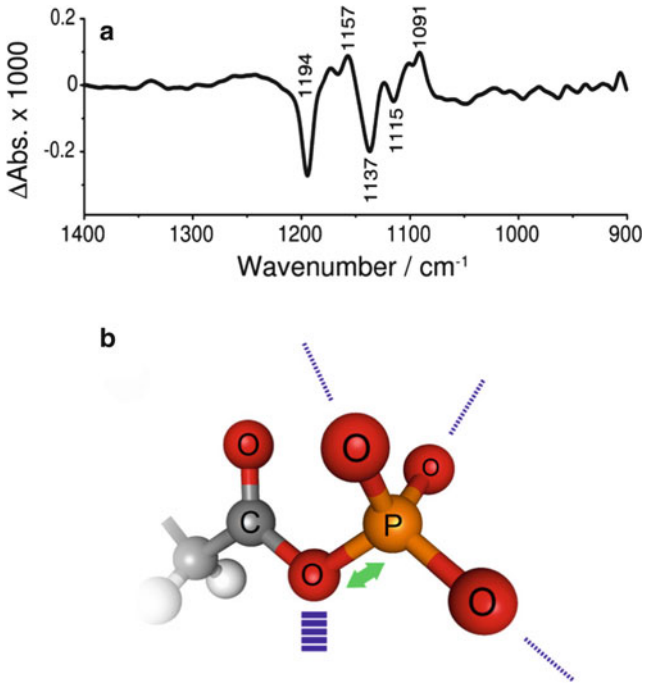

Fig. 8.4 Measuring bond lengths with infrared difference spectroscopy. (**a**) Difference spectrum from the experiment described in the text of the ^{16}O to ^{18}O isotope exchange at the phosphate group. Negative bands define the band positions of the ^{16}O isotopomer. The two main negative bands at 1,194 and 1,137 cm^{-1} are due to the second phosphoenzyme E2P; the minor negative bands near 1,170 and at 1,115 cm^{-1} are probably due to the first phosphoenzyme Ca_2E1P. The latter is a modification of our original interpretation [1], where also the 1,115 cm^{-1} band was attributed to E2P. However, occurrence of three P–O vibrations in a wavenumber interval of less than 100 cm^{-1} is incompatible with density functional theory calculations (M. Rudbeck et al., manuscript in preparation). Thus, we have to assume that the minor negative bands originate from a state that is different from the main E2P state and tentatively attribute them to Ca_2E1P in line with a previous study [34]. Positive bands are less obvious in the spectrum, likely because of overlap with the negative bands. (**b**) Molecular interpretation of the above data. The *green arrow* indicates the bond elongation induced by the enzyme environment. Bond elongation is caused by weaker interactions to the terminal phosphate oxygens and stronger interactions to the aspartyl oxygens [1] with the main effect at the bridging oxygen [35]. This is illustrated by the *dashed lines*: *narrow dashed lines* indicate reduced interactions in the protein environment compared to water. The *thick dashed line* illustrates stronger interactions

atoms in a protein with 15,000 atoms, or two phosphate vibrations out of ~50,000 protein vibrations. It exemplifies that difference spectroscopy drastically simplifies the spectrum and that this enables determination of the vibrational frequencies of particular groups. The latter is the prerequisite for a molecular interpretation, as discussed below for this example.

The phosphate frequencies obtained in the isotope exchange experiment revealed that the scissile bond between a covalently bound phosphate group and the protein is lengthened by ∼0.1 Å by the protein environment, which substantially weakens the bond and facilitates its hydrolysis as illustrated in Fig. 8.4b. Weakening and elongation of the P–O bond is not accomplished by external mechanical forces that pull the bond apart. Instead, it is an in-built response of aspartyl phosphate to a shift of interactions from phosphate to aspartyl oxygens, with only subtle changes in distances required. This provides an elegant "handle" for the enzyme to control hydrolysis [1, 35].

8.2.2 Caged Protons

The pH of a biological system is a crucial determinant of the structures and properties of its components, and studies of pH dependencies have a long tradition in the life sciences. Many perturbation studies have used rapid mixing techniques, but these are not always applicable, and time resolution below ∼1 ms is difficult to achieve. One alternative approach is to use caged protons, which are able to generate a pH jump much more rapidly than by conventional mixing techniques.

Photorelease from the majority of caged compounds is based on the well-known 2-nitrobenzyl type of rearrangement (Fig. 8.5). In all cases, a proton is liberated on the ns timescale [36] by ionization of the primary photochemical product, a nitronic acid **1**. Thus, the nitronic acid **1** is in rapid equilibrium with its conjugate *aci*-nitro anion **2**. The pK of the nitronic acid is ∼3.5 [36, 37] so the ionized species predominates in neutral solution. The lifetime of **2** varies over a range of a few microseconds to hundreds of milliseconds, depending on the nature of the attached OR group. Decay of **2** to the bicyclic intermediate **4** proceeds via the conjugate acid **1**, or more probably its isomer **3** protonated on the other oxygen of the nitronic acid, that is reached through rapid equilibration with **2** [36, 38]. Upon breakdown of **4**, the photolysis product RO^- is released and, to an extent dictated by its pK, will

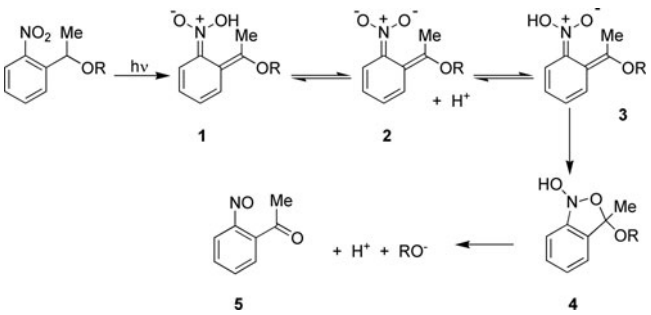

Fig. 8.5 Generalized reaction mechanism for photolysis of 1-(2-nitrophenyl)ethyl caged compounds. For caged sulfate, –OR is $–OSO_3^-$

neutralize the proton formed by the initial ionization of **1**. Hence, net proton release for the overall reaction may be total, partial or zero. If RO^- is a very weak base, full release of one proton per photolysed molecule will be observed, but if RO^- is a strong base the proton will be fully neutralized. The bicyclic intermediate **4** does not normally accumulate, as the rate-limiting step is considered to be reprotonation of the anion **2** [36, 38]. Therefore, the overall time course for approach to the post-photolysis pH value is a rapid acidification step, normally within the photolysis pulse, followed by an exponential approach to a new pH value as the *aci*-nitro anion **2** decays to the final products. However, if the pK of the photolysis product RO^- is substantially below the pH imposed in the initial jump, no decay of the initial rapid acidification will be observed.

A few compounds that function as caged protons capable of imposing rapid net acidification have been described. Among these is caged HPP [2-hydroxyphenyl 1-(2-nitrophenyl)ethyl phosphate] **6** (Fig. 8.6). The pK of the released 2-hydroxyphenyl phosphate is 5.3, so **6** cannot acidify solutions to values much below pH 5 [39]. Photochemical rearrangement of 2-nitrobenzaldehydes to 2-nitrosobenzoic acids has also been used as a source of caged protons, either with 2-nitrobenzaldehyde **7** itself [40, 41] or its water-soluble derivative 4-formyl-6-methoxy-3-nitrophenoxyacetic acid **8** [42]. The pK of 2-nitrosobenzoic acid does not appear to have been determined, but is evidently below 4 [41]. For the water-soluble derivative **8**, pK values of the photoproduct were reported as 0.75 and 2.76 [42]. The former value is a remarkably strong acidity for an aromatic carboxylic acid, but the second pK (for the oxyacetate side chain) would in any case exclude a pH excursion below pH ∼2.5 for an experiment that began near-neutral pH. This discussion excludes the elegant transient acidifications that can be imposed by pulse irradiation of phenolic compounds, for which the first excited state has very much higher acidity than the ground state [43]. These transients return to the initial level within a few microseconds, so are not relevant to the enduring acidifications considered here. An interesting recent study [44] has described a means to delay the proton reuptake after rapid ionization of a photoexcited naphthol derivative, but the acidification in this case still only lasts for ∼1 s.

A caged proton for the study of pH-dependent protein or peptide conformational changes with infrared spectroscopy should fulfil three main requirements: (a) the compound should be able to generate a large, permanent pH jump to values below pH 4; (b) proton release should be rapid and proceed with good photolysis efficiency; and (c) the infrared absorbance changes arising from photolysis of the

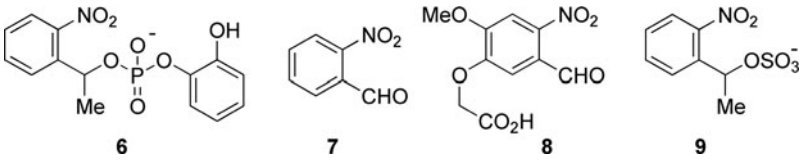

Fig. 8.6 Structures of caged proton reagents

reagent itself should be as few as possible. While 2-nitrobenzaldehydes **7** and **8** meet the first two criteria, they fail on the third. Photolysis of either compound converts an aldehyde to a carboxylate, both of which groups absorb strongly in regions where they would mask absorptions of Asp and Glu side chains and of amide vibrations that reflect changes in backbone conformation and hydrogen bonding [45]. Nevertheless, **7** has been used to study the coil-to-helix transition of poly-L-glutamate with time-resolved infrared spectroscopy on the ns timescale [45]. In spite of this successful application, the caged proton reagents described above do not fully meet the requirements for infrared spectroscopy and we describe here a reagent that was purpose-designed for infrared spectroscopy, although it may also have other applications as it can generate larger pH jumps than the previous caged protons **6–8**.

For this new caged proton, 1-(2-nitrophenyl)ethyl sulfate **9** (commercially available as Product No. 3512 from Tocris Biosciences, www.tocris.com), the released sulfate group has pK 1.92 at 20 °C [46]. Hence this caged proton, here called caged sulfate as it also releases a sulfate ion upon photolysis, is capable of generating sub-microsecond pH jumps down to pH ~2 [47]. The compound is highly water soluble and thermally stable and the absence of carboxylate groups in the reagent or its photoproducts makes it suitable to observe protonation changes of carboxylates with infrared spectroscopy. The quantum yield for photolysis of caged sulfate is 0.47. The *aci*-nitro decay rate constant in buffered solution at pH 7.0 is 34 s^{-1}, and 250 s^{-1} at pH 6.0 (both measurements at 20 °C). Proton release is very much faster and has been subject to a detailed study [48]. When the pH jump is initiated in a near-neutral solution, the proton is released by complete ionization of the photogenerated nitronic acid **1** (Fig. 8.5) with a rate constant of 1.58×10^7 s^{-1}. However, if the jump is initiated in solutions already at acidic pH (near to or less than the pK of 3.69 of the nitronic acid), a proportion of the nitronic acid does not ionize and its proton is released only as the sulfate ion is released via the process shown in Fig. 8.5. Under these circumstances, there is a biphasic release of protons: the initial phase with the rate constant given above, while the second phase is pH-dependent and in the region of 10^4 s^{-1}. Since most biological applications are likely to concern reactions starting at near-neutral pH, this biphasic release is unlikely to be of concern for most such studies.

8.2.3 Difference Spectrum of Caged Sulfate Photolysis

Figure 8.7 shows time-resolved infrared difference spectra of caged sulfate photolysis [47], with initial formation of the *aci*-nitro intermediate **2** (full line) and formation of the final products (dotted line). Negative bands arise from groups in caged sulfate that are modified in the photolysis reaction; positive bands are from groups formed upon photolysis in the *aci*-nitro intermediate or the final products. Both the intermediate and final spectra respectively have similarities with those of other 1-(2-nitrophenyl)ethyl esters such as P^3-[1-(2-nitrophenyl)ethyl]

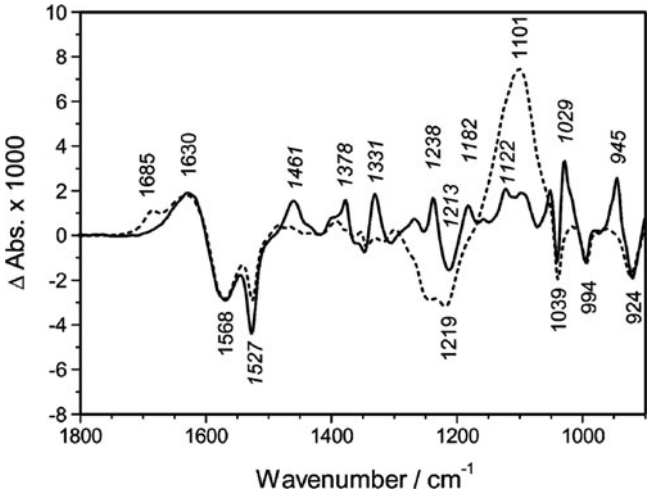

Fig. 8.7 Time-resolved spectra of caged sulfate photolysis in 400 mM Bicine buffer, pH 8.5 in ^1H_2O at 1 °C, recorded 0–0.32 s (*full line*, *italic labels*) and 16–29 s (*dotted line*, *roman labels*) after the photolysis flash

ATP (caged ATP) [26, 27] and 1-(2-nitrophenyl)ethyl methyl phosphate (caged methyl phosphate) [49]. To retard decay of the *aci*-nitro intermediate sufficiently to record its spectrum, it was necessary to work at pH 8.5 and 1 °C. The *aci*-nitro spectrum includes positive bands at 1,461, 1,378 and 1,331 cm^{-1}, which were assigned to vibrations of the nitronate group by analogy with bands at similar positions in the spectrum of the corresponding intermediate formed on photolysis of caged ATP [26, 27]. Both the intermediate and final spectra show strong signals for protonation of the Bicine buffer by the proton released on photolysis (negative band at 1,568 cm^{-1} and positive band at 1,630 cm^{-1}). These bands are missing when a different buffer is used because protonation-induced buffer signals are buffer specific.

The dotted line in Fig. 8.7 is the infrared difference spectrum of the overall photolysis reaction of caged sulfate at pH 8.5, i.e. after complete decay of the *aci*-nitro species **2**. The negative band at 1,527 cm^{-1} is from the antisymmetric stretching vibration of the nitro group [12, 27] of caged sulfate and the band at 1,219 cm^{-1} was assigned to its asymmetric SO_3^- stretching vibration [47, 50]. The negative bands below 1,050 cm^{-1} likely arise from C–O–S stretching vibrations. Positive bands were assigned to the products of photolysis as follows: the band at 1,685 cm^{-1} to the ketone group of 2-nitrosoacetophenone **5** (see Fig. 8.5) [27], and the intense broad band at 1,101 cm^{-1} to the released sulfate anion [50]. Further bands can be seen in the respective spectrum at pH 7 [47]: a negative band at 1,348 cm^{-1} due to the symmetric stretching vibration of the nitro group, positive bands at 1,424 and 1,378 cm^{-1} from the *cis*-nitroso dimer of the nitrosoketone and a small band at 1,269 cm^{-1} from the *trans*-nitroso dimer [14, 27].

As a direct infrared spectroscopic demonstration of a protonation reaction achieved by the photolysis of caged sulfate, spectra were also recorded in the presence of several carboxylate compounds. Full protonation of both acetate (pK 4.75) and methoxyacetate (pK 3.57) was achieved with a single light flash. 2-Nitrobenzoate (pK 2.21) required a second flash to release sufficient protons for full protonation, whereas no protonation occurred for trifluoroacetate (pK 0.52) because of the higher pK of the released sulfate [47].

8.3 Acidification-Induced Unfolding of Myoglobin

The first demonstration of the biological applicability of caged sulfate for pH jump experiments monitored the well-characterized acid-induced conformational change of metmyoglobin in the near-ultraviolet spectral range [47]. The transition from native state to a partially unfolded form proceeds near pH 4 [51–53] and leads to a broadening of the Soret band and a shift of its maximum from 409 to 363 nm. Figure 8.8 shows UV-visible spectra obtained upon sequential flashes on a solution containing metmyoglobin and caged sulfate. They were corrected for absorbance changes arising from photolysis of the caged sulfate, which are observed for all nitrobenzyl- and nitrophenylethyl-caged compounds. The spectra clearly demonstrate the shift of the Soret band described above. In this experiment, the caged sulfate had generated a pH drop to well below pH 4.0, thereby inducing the partial unfolding transition of myoglobin. This demonstrated its suitability for

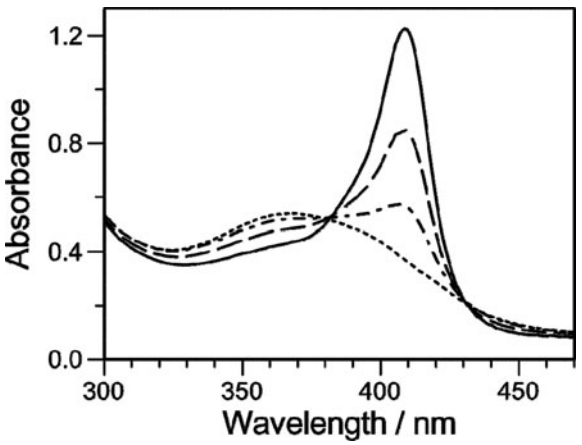

Fig. 8.8 UV-visible spectra of the partial unfolding transition of metmyoglobin induced by photolysis of caged sulfate. *Full line*: spectrum before the first photolysis flash, *broken line*: spectrum after 1 flash, *dash-dotted line*: spectrum after 2 flashes, and *dashed line*: spectrum after 5 flashes. The final pH was below 4. Spectra were corrected for the absorbance change of caged sulfate [47]

protein folding studies in that pH range, although under the particular experimental conditions, several flashes were required to cause the complete transition.

The full change can be achieved in a single flash as long as there is enough light energy and/or caged sulfate to convert a sufficient quantity of the reagent and achieve the necessary pH drop. This was demonstrated in a later study [54], which also showed that an alternative caged proton, 2-nitrobenzaldehyde **7**, was incapable of releasing sufficient protons to achieve unfolding because of buffering by the 2-nitrosobenzoate photoproduct formed from that compound. The use of caged sulfate to achieve a large pH drop also readily enabled time-resolved data for the absorbance change to be recorded.

8.4 Acidification-Induced Aggregation of the Alzheimer's Peptide

8.4.1 Introduction to the Alzheimer's Peptide

Misfolding of proteins can lead to their aggregation and subsequent onset of diseases such as Alzheimer's disease, transmissible spongiform encephalopathies and type II diabetes. In the case of Alzheimer's, the disease is associated with the deposition of so-called senile plaques in the brain, which have as major components the 40–42-residue β-amyloid peptides (Aβ). These peptides are the product of the processing of the amyloid precursor protein, a protein of unknown function. Mature amyloid precursor protein is bound to the plasma membrane via one transmembrane helix. Processing by β- and γ-secretases gives rise to the Aβ peptides which aggregate and form the insoluble fibrils that constitute the plaques ("amyloid") [55–57]. The aggregation and formation of fibrils are associated with the conversion of the peptide's secondary structure from predominantly random coil to β-sheet. Recent evidence has indicated that neurotoxicity is associated not so much with the insoluble plaques but with soluble oligomeric aggregates of Aβ [55, 57, 58], which represent an intermediate stage in fibril formation.

Among the many unresolved questions regarding the cellular events associated with Alzheimer's disease is where the crucial processing and early aggregation processes occur. Although the plaques are observed outside the cells, there is increasing evidence that Aβ processing, accumulation and aggregation also occur inside cells in the endosomal/lysosomal system [59, 60]. Other reports disagree with the endosomal/lysosomal origin of the large aggregates, since the morphology of large aggregates of the full-length $A\beta_{1-40}$ has shown significant differences between neutral and acidic pH [61].

Two regions of the peptide are considered to be the most important in the oligomerization/aggregation processes: a central hydrophobic segment, comprising approximately residues 16–22, and the hydrophobic C-terminal membrane-anchored segment, approximately residues 29–40 or 29–42. One of the most studied

fragments, the $A\beta_{1-28}$ peptide, retains the ability to form fibrils with a clear β-sheet architecture. It comprises the extramembranous part of the peptide with the central hydrophobic segment [62,63] and contains most of the main charged residues of the full-length sequence that are responsible for promoting or inhibiting aggregation via solvation, pH and metal ion interaction [64,65]. In vitro, the process leading to $A\beta_{1-28}$ fibril formation is slow (compared to the longer peptides), and is accelerated as the pH decreases (as for the longer peptides) [61, 63, 66]. Aggregates of this fragment seem to contain oligomeric species similar to those found in natural amyloid deposits [62].

Infrared spectroscopy has been used extensively to study aggregation processes, secondary structure transitions and folding of peptides and proteins [67, 68]. It has been very useful in the field of amyloid proteins, because the amide I band is sensitive to hydrogen bonding and backbone structure [69]. For the Aβ peptide, infrared spectroscopy has determined secondary structures and slow structural conversion kinetics for both the full-length peptide and shorter fragments [67, 70, 71]. Fibril formation and membrane interaction have also been studied by this technique [72, 73]. Fast events in the aggregation had not been studied with infrared spectroscopy before the use of photolytically induced acidification due to the problem of triggering the process.

8.4.2 Time-Resolved Infrared Difference Spectroscopy of the Aggregation of the Alzheimer's Peptide

8.4.2.1 Outline of the pH Jump Experiment

High time resolution and characterization of fast events in the aggregation process is possible with infrared spectroscopy if an appropriate trigger for aggregation can be employed. The sudden acidification induced by photolysis of caged sulfate is such a trigger. It models a process that may occur in the endosomal/lysosomal system and we describe our application of this approach [74] in the following. The sample consisted of 4 μl of 0.83 mM $A\beta_{1-28}$ peptide in D_2O (pD∼8.5), which was placed into a demountable 30 μm BaF_2 cuvette. After recording an absorbance spectrum of the unphotolysed sample and a series of control difference spectra to check the stability of the sample absorption, a reference single beam spectrum I_R was recorded against which the spectra after acidification $I_1 \ldots I_{40}$ were later ratioed for the calculation of time-resolved difference spectra (difference spectrum i is given by $-\log I_i/I_R$).

Acidification was induced by photolysis of caged sulfate using a xenon flash lamp, which illuminated the sample with ∼150 mJ near-UV light. The flash light was filtered by a Schott UG11 filter to reduce light below 280 nm and above 370 nm. Then time-resolved infrared single beam spectra were recorded for 4 min with a maximum time resolution of 70 ms and difference spectra calculated as described

above. This experiment allowed us to follow the secondary structure changes of the $A\beta_{1-28}$ peptide in real time [74]. The initial pH jump was sufficient to complete the aggregation process, since no structural changes of the peptide were observed upon subsequent flashes. After the experiment, the pH of the sample was measured with pH indicator paper, which indicated that the final pH was below pH 6. The kinetic evolution of selected bands in the difference spectra was analyzed using the integrated area of the bands. The integration was performed with respect to a baseline drawn between the average absorption in the 1,720–1,702 cm^{-1} range and the 1565–1531 cm^{-1} range. For the 1,643 cm^{-1} and 1,620 cm^{-1} bands of the $A\beta_{1-28}$ peptide, the integration limits were 1,656–1,641 cm^{-1} and 1,630–1,612 cm^{-1}. For the 1,687 and 1,601 cm^{-1} bands of caged sulfate, the integration was carried out between 1,692–1,685 cm^{-1} and 1,603–1,596 cm^{-1}.

8.4.2.2 Absorbance Changes upon Acidification

The photolytically induced absorbance changes of caged sulfate alone and together with $A\beta_{1-28}$ are presented in Fig. 8.9. The spectral region shown includes the absorption of the amide I′ normal modes of the peptide, to which the dominant contribution is from the C = O stretching vibration of the peptide groups. The prime indicates that the amide groups are deuterated in the D_2O solvent, which abolishes the contribution of the N–H bending vibration that is present in the amide I vibration of unexchanged [^{1}H] amide groups. The amide I′ vibrational frequency depends on the secondary structure of peptides and proteins and has been used extensively for secondary structure analysis of proteins [4, 75–79]. The spectral region shown also

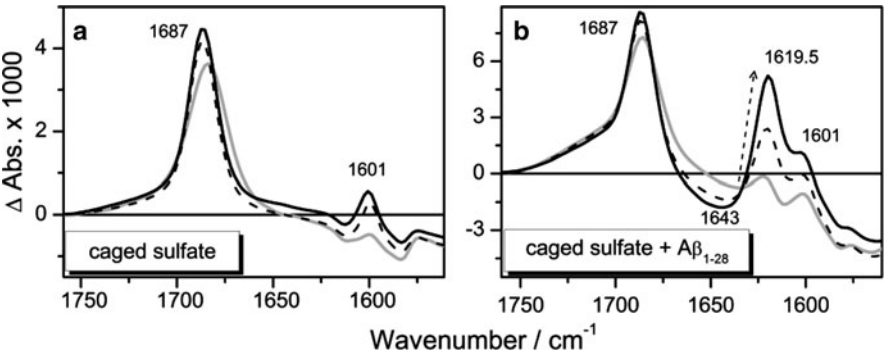

Fig. 8.9 Time-dependent infrared absorbance changes of caged sulfate and $A\beta_{1-28}$ peptide in the amide I′ region after a pD jump from ~8.5 to <6.0. Spectra were recorded in D_2O and the spectral changes triggered by photolysis of caged sulfate. They represent absorbance changes due to photolysis that were observed in the following averaged time intervals: 0.4–1 s (*grey solid line*); 9–20 s (*dashed line*); and 153–235 s (*black solid line*). (**a**) Time-dependent difference spectra of photolysis of 20 mM caged sulfate. (**b**) Time-dependent difference spectra in the additional presence of 0.83 mM $A\beta_{1-28}$. The *arrow* indicates the direction and shift of the 1,620 cm^{-1} band as a function of time

contains the absorption of other groups, for example of carbonyl groups of amino acid side chains and of the product of caged compound photolysis.

The photolysis of caged sulfate alone (Fig. 8.9a) gives rise to bands at 1,687 cm^{-1} and 1,601 cm^{-1} due to, respectively, the keto group and presumably a ring vibration of the photolysis by-product 2-nitrosoacetophenone **5**. The 1,687 cm^{-1} band, also present in Fig. 8.7, increases its intensity and shifts upon dimerization of the nitroso group [14, 27]. With the $A\beta_{1-28}$ peptide present, the release of protons upon photolysis of caged sulfate leads initially to the appearance of a positive band at 1,623.5 cm^{-1}, which shifts to 1,619.5 cm^{-1} and becomes much more intense in the later spectra. For simplicity, it is subsequently referred to as the 1,620 cm^{-1} band, i.e. including the shifted band in the early spectra. This band is not observed without $A\beta_{1-28}$ (see Fig. 8.9a) and can be assigned to β-sheet formation in the peptide. The increase of the broad negative band in the 1,660–1,630 cm^{-1} region is attributable to the disappearance of mainly random coil (1,643 cm^{-1}) structure. At higher frequencies, the 1,687 cm^{-1} band of the nitrosoacetophenone overlaps with a contribution from the peptide which is discussed in the next section.

8.4.2.3 Time Course of Peptide Aggregation

The time evolution of the area of the various infrared bands in the time-resolved spectra was studied by fitting one or two exponential functions to the kinetic data (see Fig. 8.10). The kinetics of NO dimerization of the photolysis by-product in the absence of peptide can be fitted to a single exponential with a time constant of ∼1.5 s for the bands at 1,601 and 1,687 cm^{-1}. The maximum amplitude is reached at ∼10 s (see Fig. 8.10a for the band at 1,687 cm^{-1}). The sigmoidal appearance of the single exponential curve in Fig. 8.10a is due to the logarithmic time scale. Figure 8.10a shows also the kinetics of the 1,687 cm^{-1} band in the presence of $A\beta_{1-28}$. The band exhibits at least a biexponential behaviour. The fast phase has a time constant of 1.8 s (with a relative amplitude of 0.60) and reaches its maximum at ∼10 s. This time constant agrees well with the value obtained for caged sulfate alone. The second exponential has a time constant of 67 s and a relative amplitude of 0.40. We attribute the fast phase predominantly to NO dimerization of the photolysis by-product **5** from caged sulfate [27] and the slow phase to a secondary structure transition of $A\beta_{1-28}$. The band position at 1,687 cm^{-1} is characteristic of antiparallel β-sheets. For peptides that form parallel β-sheets, such a band is not observed [80, 81], in line with theoretical predictions [82, 83]. The fast phase has also a contribution from a secondary structure change of the peptide as revealed by eliminating the effects of NO dimerization of the photolysis by-product from the spectra of the Aβ peptide. This was achieved by subtracting spectra of caged sulfate photolysis obtained in the absence of Aβ peptide [74]. The corrected spectra of Aβ aggregation showed a clear band at 1,687 cm^{-1} also for the fast phase, indicating the fast formation of antiparallel β-sheets. The formation of antiparallel β-sheets in both phases of $A\beta_{1-28}$ aggregation is in contrast to $A\beta_{1-42}$ aggregation, which does not produce a band at ∼1,690 cm^{-1}, because it forms parallel β-sheets [74]. The presence or absence

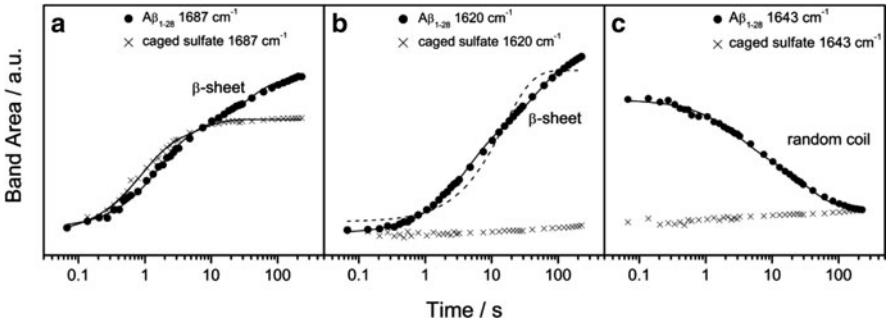

Fig. 8.10 Kinetics of Aβ aggregation induced by a pH decrease. (**a**) Evidence for the formation of antiparallel β-sheets. 1,687 cm^{-1} band after photolysis of caged sulfate in the absence or presence of $A\beta_{1–28}$. The *lines* show the exponential fits. The kinetics in the presence of $A\beta_{1–28}$ are due to rapid dimerization of the photolysis by-product and biphasic formation of antiparallel β-sheets by $A\beta_{1–28}$. (**b**) Kinetic evolution of the β-sheet band at 1,620 cm^{-1}. The *dashed line* corresponds to a single exponential fit; the continuous line shows a biexponential fit. (**c**) Kinetics of loss of random coil structure monitored at 1,643 cm^{-1}

of the hydrophobic segments 29–40 or 29–42 in the C-terminus of the Aβ full-length peptide seems therefore to determine whether parallel or antiparallel structure is adopted. Another explanation is the assumption of a stronger influence of the unprotected peptide ends in favour of antiparallel sheets for the shorter peptide.

Panel b of Fig. 8.10 shows the time evolution of a second β-sheet band at 1,620 cm^{-1} and panel c the time course of loss of random coil structure (1,643 cm^{-1}). Both time courses were fitted with two exponential functions. The time constants of fast and slow phases were ∼3 s and ∼40 s, respectively. The slower phase had a slightly larger relative amplitude. Controls with caged sulfate alone did not produce significant signals at these two wavenumbers, as shown in Figs. 8.10b, c.

8.4.2.4 Model of the Aggregation Process

As mentioned earlier, there was a shift from 1,623.5 to 1,619.5 cm^{-1} for the β-sheet band in the course of the experiment. Figure 8.11 illustrates the time dependence of the position of this band. As discussed in the original publication [74], there are several possible interpretations for this shift. The most natural one is a growth of β-sheet oligomers by docking of random coil peptides to existing sheets in the slow phase of aggregation. This will increase the number of strands per sheet, thus shifting the main β-sheet band near 1,620 cm^{-1} to lower wavenumbers and increasing the splitting between this band and the side band at 1,687 cm^{-1}. From the initial splitting of 63 cm^{-1} between high and low wavenumber β-sheet bands, the number of strands in the β-sheet of the early oligomers was estimated as 5–10. The further shift in the late phase indicates further strand addition and formation of larger aggregates with more than ten strands.

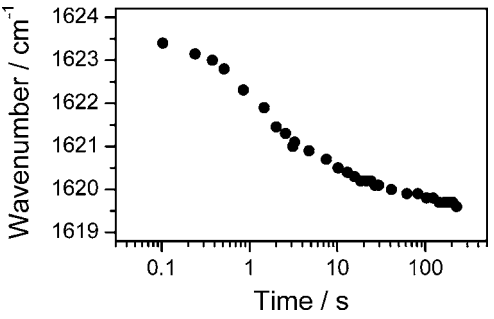

Fig. 8.11 Spectral shift of the main β-sheet band as a function of time

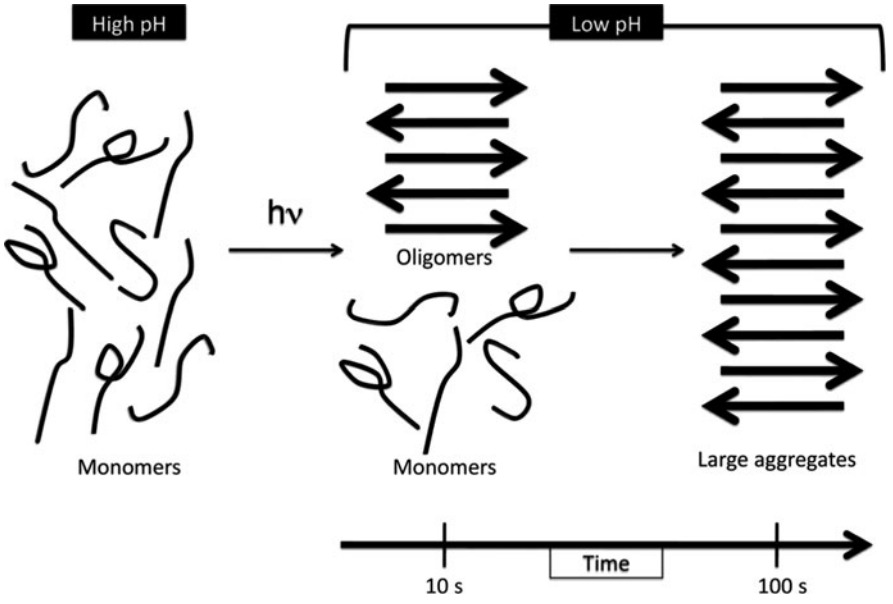

Fig. 8.12 A cartoon representation of pH-induced $A\beta_{1-28}$ aggregation

These data are in line with a nucleation-elongation mechanism that has been shown to operate in fibril formation in vitro [84]: a first rapid (time constant ∼3 s) formation of a β-sheet nucleus with limited number of strands is followed by a slower (time constant ∼40 s) formation of more extended β-sheets. The large aggregates are formed by adding unstructured peptides to the oligomeric β-sheet template. Figure 8.12 gives a cartoon illustration of these processes.

8.5 Outlook

Reaction-induced infrared difference spectroscopy is a powerful technique, as illustrated here, but it requires the ability to trigger the biological reaction of interest. Caged protons add a photoinduced pH drop to the toolbox of infrared spectroscopic

methods [45, 74]. Because of the ns timescale of the pH drop, the method has the advantage of being able to monitor early intermediates upon acidification, here the *early* oligomers of $A\beta_{1-28}$ aggregation. For the system discussed, the time resolution of the rapid scan spectrometer was appropriate (maximal 10 ms). If faster reactions are of interest, single wavelength techniques with sub-μs time resolution can be used. A pH jump may also be induced in the opposite direction, by use of a caged hydroxide [85], the photoexcitation of which transiently increases the pH for ∼100 ms. However, to date there are no published applications of the compound in biological infrared spectroscopy. The pH jump technique has a wide application range in the life sciences. Examples are structural transitions of other pH-sensitive peptides (e.g. poly-L-lysine), bacterial pH-sensitive toxins, activated in acidic cellular compartments such as endosomes and lysosomes, acid unfolding of proteins and structural changes induced by protonation of individual protein residues.

References

1. A. Barth, N. Bezlyepkina, J. Biol. Chem. **279**, 51888 (2004)
2. M.S. Braiman, K.J. Rothschild, Annu. Rev. Biophys. Biophys. Chem. **17**, 541 (1988)
3. C. Zscherp, A. Barth, Biochemistry **40**, 1875 (2001)
4. H. Fabian, W. Mäntele, in *Infrared Spectroscopy of Proteins*, ed. by J.M. Chalmers, P.R. Griffiths (Wiley, Chichester, 2002), p. 3399
5. C. Berthomieu, R. Hienerwadel, Photosynth. Res. **101**, 157 (2009)
6. F. Siebert, W. Mäntele, W. Kreutz, Biophys. Struct. Mech. **6**, 139 (1980)
7. K.J. Rothschild, M. Zagaeski, W.A. Cantore, Biochem. Biophys. Res. Commun. **103**, 483 (1981)
8. J.O. Alben, D. Beece, S.F. Bowne, L. Eisenstein, H. Frauenfelder, D. Good, M.C. Marden, P.P. Moh, L. Reinisch, A.H. Reynolds, K.T. Yue, Phys. Rev. Lett. **44**, 1157 (1980)
9. R. Vogel, F. Siebert, Curr. Opin. Chem. Biol. **4**, 518 (2000)
10. K. Gerwert, Curr. Opin. Struct. Biol. **3**, 769 (1993)
11. T. Noguchi, Photosynth. Res. **91**, 59 (2007)
12. A. Barth, W. Mäntele, W. Kreutz, FEBS Lett. **277**, 147 (1990)
13. V. Cepus, C. Ulbrich, C. Allin, A. Troullier, K. Gerwert, Methods Enzymol. **291**, 223 (1998)
14. A. Barth, in *Dynamic Studies in Biology*, ed. by M. Goeldner, R. Givens (Wiley-VCH, Weinheim, 2005), pp. 369
15. J.H. Kaplan, B. Forbush, J.F. Hoffman, Biochemistry **17**, 1929 (1978)
16. M. Goeldner, R. Givens (eds.) *Dynamic Studies in Biology* (Wiley-VCH, Weinheim, 2005)
17. J. White, K. Drabble, C.W. Wharton, Biochem. J. **306**, 843 (1995)
18. H. Fabian, D. Naumann, Methods **34**, 28 (2004)
19. E. Goormaghtigh, V. Raussens, J.-M. Ruysschaert, Biochim. Biophys. Acta **1422**, 105 (1999)
20. P.R. Rich, M. Iwaki, Mol. BioSyst. **3**, 398 (2007)
21. K. Fahmy, Biophys. J. **75**, 1306 (1998)
22. G. Panick, R. Malessa, R. Winter, G. Rapp, K.J. Frye, C.A. Royer, J. Mol. Biol. **275**, 389 (1998)
23. D. Moss, E. Nabedryk, J. Breton, W. Mäntele, Eur. J. Biochem. **187**, 565 (1990)
24. M. Lübben, K. Gerwert, FEBS Lett. **397**, 303 (1996)
25. M. Liu, E.-L. Karjalainen, A. Barth, Biophys. J. **88**, 3615 (2005)
26. A. Barth, K. Hauser, W. Mäntele, J.E.T. Corrie, D.R. Trentham, J. Am. Chem. Soc. **117**, 10311 (1995)

27. A. Barth, J.E.T. Corrie, M.J. Gradwell, Y. Maeda, W. Mäntele, T. Meier, D. R. Trentham, J. Am. Chem. Soc. **119**, 4149 (1997)
28. A. Barth, Spectroscopy **22**, 63 (2008)
29. M. Liu, A. Barth, J. Biol. Chem. **278**, 10112 (2003)
30. M. Liu, A. Barth, Biopolymers (Biospectroscopy) **67**, 267 (2002)
31. C. Toyoshima, T. Mizutani, Nature **430**, 529 (2004)
32. T.L.-M. Sørensen, J.V. Møller, P. Nissen, Science **304**, 1672 (2004)
33. M. Liu, A. Barth, J. Biol. Chem. **279**, 49902 (2004)
34. M. Liu, M. Krasteva, A. Barth, Biophys. J. **89**, 4352 (2005)
35. J. Andersson, A. Barth, Biopolymers **82**, 353 (2006)
36. M. Schwörer, J. Wirz, Helv. Chim. Acta **84**, 1441 (2001)
37. G. Wettermark, E.D. Black, L. Dogliotti, Photochem. Photobiol. **4**, 229 (1965)
38. Y.V. Il'ichev, J. Wirz, J. Phys. Chem. A **104**, 7856 (2000)
39. S. Khan, F. Castellano, J.L. Spudich, J.A. McCray, R.S. Goody, G.P. Reid, D.R. Trentham, Biophys. J. **65**, 2368 (1993)
40. G. Bonetti, A. Vecli, C. Viappiani, Chem. Phys. Lett. **269**, 268 (1997)
41. S. Abbruzzetti, C. Viappiani, J.R. Small, L.J. Libertini, E.W. Small, Biophys. J. **79**, 2714 (2000)
42. K. Janko, J. Reichert, Biochim. Biophys. Acta **905**, 409 (1987)
43. M. Gutman, D. Huppert, E. Pines, J. Am. Chem. Soc. **103**, 3709 (1981)
44. R.M.D. Nunes, M. Pineiro, L.G. Arnaut, J. Am. Chem. Soc. **131**, 9456 (2009)
45. T.P. Causgrove, R.B. Dyer, Chem. Phys. **323**, 2 (2006)
46. D.D. Perrin, *Ionisation Constants of Inorganic Acids and Bases in Aqueous Solution*, 2nd edn. (Pergamon, Oxford, 1982)
47. A. Barth, J.E.T. Corrie, Biophys. J. **83**, 2864 (2002)
48. S. Abbruzzetti, S. Sottini, C. Viappiani, J.E.T. Corrie, J. Am. Chem. Soc. **127**, 9865 (2005)
49. J.E.T. Corrie, V.R.N. Munasinghe, D.R. Trentham, A. Barth, Photochem. Photobiol. Sci. **7**, 84 (2008)
50. N.B. Colthup, L.H. Daly, S.E. Wiberley, *Introduction to Infrared and Raman Spectroscopy*, 2nd edn. (Academic Press, New York, 1975)
51. J.T. Sage, D. Morikis, P.M. Champion, Biochemistry **30**, 1227 (1991)
52. Z. Chi, S.A. Asher, Biochemistry **37**, 2865 (1998)
53. V. Palaniappan, D.F. Bocian, Biochemistry **33**, 14264 (1994)
54. S. Abbruzzetti, S. Sottini, C. Viappiani, J.E.T. Corrie, Photochem. Photobiol. Sci. **6**, 621 (2006)
55. J. Hardy, D.J. Selkoe, Science **297**, 353 (2002)
56. D.J. Selkoe, Nature **426**, 900 (2006)
57. F. Chiti, C.M. Dobson, Annu. Rev. Biochem. **75**, 333 (2006)
58. W.L. Klein, G.A. Krafft, C.E. Finch, Trends Neurosci. **24**, 219 (2001)
59. D.J. Selkoe, Annu. Rev. Neurosci. **17**, 489 (1994)
60. S.H. Pasternak, J.W. Callahan, D.J. Mahuran, J. Alzheimers Dis. **6**, 53 (2004)
61. S.J. Wood, B. Maleeff, T. Hart, R. Wetzel, J. Mol. Biol. **256**, 870 (1996)
62. D.A. Kirschner, H. Inouye, L.K. Duffy, A. Sinclair, M. Lind, D.J. Selkoe, Proc. Natl. Acad. Sci. U.S.A. **84**, 6953 (1987)
63. K. Ma, E.L. Clancy, Y.B. Zhang, D.G. Ray, K. Wollenberg, M.G. Zagorski, J. Am. Chem. Soc. **121**, 8698 (1999)
64. P.E. Fraser, D.R. McLachlan, W.K. Surewicz, C.A. Mizzen, A.D. Snow, J.T. Nguyen, D.A. Kirschner, J. Mol. Biol. **244**, 64 (1994)
65. J. Danielsson, R. Pierattelli, L. Banci, A. Gräslund, FEBS J. **274**, 46 (2007)
66. M. Cortijo-Arellano, J. Ponce, N. Durany, J. Cladera, Biochem. Biophys. Res. Commun. **368**, 238 (2008)
67. S.M. Decatur, Acc. Chem. Res. **39**, 169 (2006)
68. M.R. Nilsson, C.M. Dobson, Biochemistry **42**, 375 (2003)
69. A. Barth, Biochim. Biophys. Acta **1767**, 1073 (2007)
70. N. Benseny-Cases, M. Cocera, J. Cladera, Biochem. Biophys. Res. Commun. **361**, 916 (2007)
71. S.A. Petty, S.M. Decatur, J. Am. Chem. Soc. **127**, 13488 (2005)

72. M.R. Nilsson, Methods **34**, 151 (2004)
73. K. Matsuzaki, C. Horikiri, Biochemistry **38**, 4137 (1999)
74. A. Perálvarez-Marín, A. Barth, A. Gräslund, J. Mol. Biol. **379**, 589 (2008)
75. H. Susi, N. Timasheff, L. Stevens, J. Biol. Chem. **242**, 5460 (1967)
76. S.N. Timasheff, H. Susi, L. Stevens, J. Biol. Chem. **242**, 5467 (1967)
77. J.L.R. Arrondo, A. Muga, J. Castresana, F.M. Goñi, Prog. Biophys. Mol. Biol. **59**, 23 (1993)
78. M. Jackson, H.H. Mantsch, Crit. Rev. Biochem. Mol. Biol. **30**, 95 (1995)
79. E. Goormaghtigh, V. Cabiaux, J.-M. Ruysschaert, Subcell. Biochem. **23**, 405 (1994)
80. N. Yamada, K. Ariga, M. Naito, K. Matsubara, E. Koyama, J. Am. Chem. Soc. **120**, 12192 (1998)
81. P. Chitnumsub, W.R. Fiori, H.A. Lashuel, H. Diaz, J.W. Kelly, Bioorg. Med. Chem. **7**, 39 (1999)
82. Y.N. Chirgadze, N.A. Nevskaya, Biopolymers **15**, 627 (1976)
83. J. Kubelka, T.A. Keiderling, J. Am. Chem. Soc. **123**, 12048 (2001)
84. J.D. Harper, P. T. Lansbury, Annu. Rev. Biochem. **66**, 385 (1997)
85. S. Abbruzzetti, M. Carcelli, P. Pelagatti, D. Rogolino, C. Viappiani, Chem. Phys. Lett. **344**, 387 (2001)

Chapter 9
Examining Amyloid Structure and Kinetics with 1D and 2D Infrared Spectroscopy and Isotope Labeling

Lauren E. Buchanan, Emily B. Dunkelberger, and Martin T. Zanni

Abstract The combination of infrared spectroscopy and isotope labeling is a powerful toolset for providing residue-specific information about the structure and aggregation kinetics of amyloid peptides. In this chapter, we review a simple mathematical formalism to guide the interpretation of 1D- and 2DIR spectra of amyloid fibers. This formalism enables the design of isotope labeling schemes to extract precise structural features from IR spectra. We present 2DIR experiments on hIAPP and Aβ in which some of these strategies have been employed.

9.1 Introduction

More than 20 different diseases, including type II diabetes, Alzheimer's, Parkinson's, and Huntington's, are associated with the misfolding of proteins into amyloid fibrils. Although different proteins are associated with each disease and each protein has its own distinct sequence of amino acids and native state structure, the amyloid fibrils they form share many characteristics. In the fibers, the proteins adopt a cross-β configuration in which the individual strands of the β-sheet run perpendicular to the fibril axis. While high concentrations of amyloid deposits can disrupt organ function, there is evidence that the prefibril intermediates are actually the cytotoxic species responsible for type II diabetes, Alzheimer's, and other amyloid diseases [1, 2]. As a result, there is much interest in uncovering the aggregation mechanism and identifying the structures of transient intermediates.

Unfortunately, very little is known about these critical intermediates despite the variety of techniques – solid-state NMR, X-ray diffraction, electron spin resonance, circular dichroism, electron microscopy, and fluorescence spectroscopy – that are

M.T. Zanni (✉)
Department of Chemistry, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, USA
e-mail: zanni@chem.wisc.edu

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*, Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6_9,

often used to study protein folding. This dearth of information is mainly due to the fact that high structure resolution techniques cannot be applied to aggregates or do not have the time resolution needed to capture the short-lived intermediate. Optical techniques that do have the requisite time resolution cannot sufficiently resolve the structure. Moreover, even simple point mutations can alter the structures and aggregation of amyloids, making the bulky labels of ESR and fluorescence spectroscopy problematic.

Infrared (IR) spectroscopy has the potential to uncover many important details of the aggregation mechanism and the structure of the elusive intermediates. It has the necessary time resolution, which is as short as a few picoseconds when implemented with femtosecond laser pulses. It can be applied to protein aggregates as straightforwardly as dilute proteins. It also has bond-specific structural resolution when used in conjunction with isotope labeling. However, the structural resolution of IR spectroscopy is largely underutilized, because few studies have been performed to date on amyloids with isotope labels. Some of those performed thus far include ^{13}C labeling by Decatur [3–6] and $^{13}C{=}^{18}O$ labeling of the amino acid backbones by the Hochstrasser and Zanni research groups [7–11].

Most infrared studies of amyloids have employed linear IR spectroscopy [3–6, 12–16]. Of those that use isotope labels, Decatur and coworkers collected FTIR spectra of small isotope-labeled fragments of Aβ [3] and the prion protein PrP^c [4] as a function of time and looked for the frequency of the isotope-labeled feature to shift. They determined that both peptides first begin to form fibrils out of register and later rearrange to bring residues into register. Furthermore, they confirmed ssNMR results that the β-strands were oriented antiparallel within the β-sheet and they were able to pinpoint which residue was held perfectly in register in each β-sheet. Gai and coworkers [12] have discovered interesting facts about the hydration and surface conformations necessary for amyloid fiber formation, although isotope labeling has not yet been used.

There is also a subset of researchers applying two-dimensional infrared (2DIR) spectroscopy to amyloids. 2DIR spectroscopy has all the attributes of standard FTIR spectroscopy, with the additional advantages that it can be used to measure inhomogeneous and homogeneous linewidths, determine the secondary structures of individual residues using crosspeaks, provide angles between coupled carbonyl groups, and improve the resolution with one or more additional dimensions. There are currently two experimental research groups applying 2DIR spectroscopy to study amyloids. Hochstrasser and coworkers are studying the amyloid beta (Aβ) peptide that is involved in Alzheimer's disease [7, 8] and we are studying the human islet amyloid polypeptide (hIAPP) implicated in type 2 diabetes [9, 10]. In our studies, isotope labeling permits the kinetics of β-sheet formation, reputation, and alignment of the β-sheets to be monitored, as we describe below, all with bond-specific structural resolution. Theoretical work on the application of multidimensional spectroscopies to amyloids is being done by our collaborators Skinner, de Pablo, and coworkers, as well as Mukamel and coworkers.

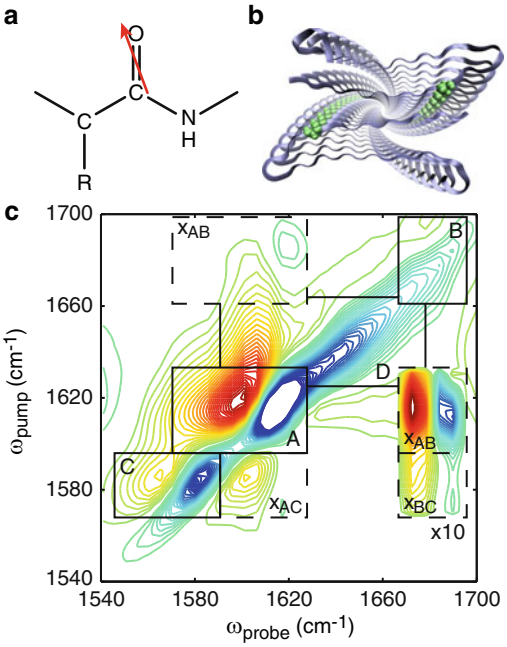

Fig. 9.1 (**a**) amide I mode. The dipole lies about 20° off the C=O bond, (**b**) hIAPP fibril, showing a single isotope-labeled residue per peptide, generated from an ssNMR structure [17], (**c**) example spectrum for an isotope-labeled sample of hIAPP. The diagonal modes, enclosed in *solid boxes*, are assigned amide I modes from the following secondary structures: (A) unlabeled amide I, (B) β-turn, (C) isotope-labeled amide I, and (D) random coil. The crosspeaks, enclosed in *dashed boxes*, arise when the corresponding regions of secondary structures are coupled

An illustrative example of a 2DIR spectrum for a fibril of singly labeled hIAPP is shown in Fig. 9.1. We typically excite the amide I stretching mode (Fig. 9.1a). By labeling one residue with $^{13}C{=}^{18}O$, that amide I mode is shifted down in frequency and isolated from the "bulk" unlabeled residues. The diagonal features are similar to those found in FTIR and indicate the secondary structure composition of the peptides. The peaks appear as doublets because the 2DIR pulse sequence probes both the fundamental ($0 \rightarrow 1$) and combination ($1 \rightarrow 2$) bands of each mode, colored blue and red in the spectrum, respectively. The spectrum is dominated by a pair of peaks (box A) that appear along the diagonal around 1,620 cm^{-1}, which are assigned to the asymmetric stretching mode along the parallel β-sheet axis for the unlabeled residues. A weaker doublet (box B) appears near 1,680 cm^{-1}, which is attributed to the residues in the turn region based on extensive empirical evidence. The doublet appearing along the diagonal around 1,580 cm^{-1} (box C) is the amide I band for the isotope-labeled residues. The last of the diagonal features is a broad set of peaks between 1,630 and 1,660 cm^{-1} created by segments of peptide with a random coil structural distribution (box D).

Other features appear off the main diagonal. These are crosspeaks that arise when different modes couple to each other. The most intense sets of crosspeaks result from coupling between the unlabeled β-sheet feature (box A) and the unlabeled turn feature (box B). These crosspeaks appear both above and below the diagonal (two boxes labeled x_{AB}). The other sets of crosspeaks result from the isotope-labeled mode coupling to either the β-sheet (box x_{AC}) or the turn (box x_{BC}).

The purpose of this review is to explain some of the ways in which 1D- and 2DIR spectroscopy can be applied to characterize amyloid fiber formation. The work of Decatur, Hochstrasser, and our group reveals specific structural information because single amino acids are isotope labeled to reveal the vibrational frequencies and lineshapes of individual residues. However, the observed frequencies cannot be interpreted in the same manner as the unlabeled residues, because their vibrational Hamiltonian is not the same. That is, one cannot assign α-helices or β-sheets according to the decades old empirical frequency assignments, even after subtracting the isotopic shift. The standard frequency assignments are not useful because the frequency shifts of the isotope labels are determined by a reduced Hamiltonian. Nonetheless, if one understands the details of the vibrational couplings involved, then isotope labeling experiments that extract much of the most important structural information with either 1D- or 2DIR spectra can be designed. Thus, the major aim of this review is to discuss these reduced Hamiltonians to aid in future labeling experiments. We start with a discussion of the vibrational modes in amyloid fibers, then discuss isotope labeling as it applies to amyloids, and end with a few examples of how couplings and labeling have been used to study amyloid fiber structure and formation.

9.2 Vibrational Modes of Amyloids

To design the best isotope labeling scheme, one needs to understand the vibrational modes of amyloid fibers. Amyloid fibers consist of large parallel β-sheets connected by disordered loops. The IR spectrum for a perfectly ordered, infinite parallel β-sheet would be exceedingly simple. Because of the 2D symmetry, only two vibrational modes would be observable, one around 1,620 cm^{-1} and a weaker mode at higher frequencies. To understand the origin of these peaks, we examine the vibrational Hamiltonian, shown in Fig. 9.2.

The vibrational Hamiltonian is written in the basis set of the individual carbonyl bonds (more precisely, the amide I vibrational modes), which is often called a local mode Hamiltonian. In this matrix, the local mode frequencies appear on the diagonal of the matrix. The frequency of each local mode is influenced by the extent to which it is hydrogen bonded and the electrostatics of its surrounding environment. Sophisticated methods that allow the local mode frequencies to be accurately calculated from molecular dynamics simulations [18–25] now exist. In a perfectly ordered and infinitely large β-sheet, all of the local mode frequencies will be identical and have a frequency of about 1,645 cm^{-1} in water. The off-diagonal

elements of the coupling Hamiltonian are set by the coupling strengths between the local modes, which depend on their distance, relative orientations, and mechanical anharmonicities between covalently bonded residues. These coupling strengths come from simple models, sophisticated ab initio calculations, and experiments [23], although the exact quantities do not alter our discussion here. The strongest coupling is between residues in adjacent strands, which is about $J_{01} = -8.8\,\mathrm{cm}^{-1}$. Several other couplings are also important (Fig. 9.3). When the Hamiltonian for this system is diagonalized, an eigenstate is generated for each oscillator in the system, each of which can contribute to the vibrational spectrum. Due to the symmetry of the sheet, only two of these modes are observed, which generates the characteristic

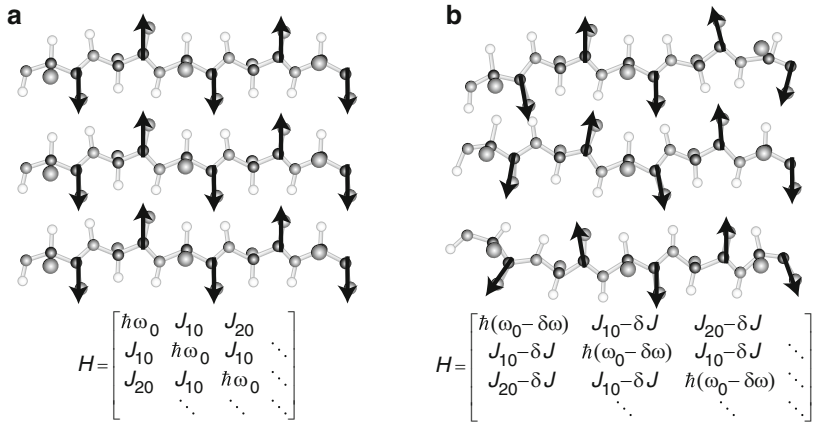

Fig. 9.2 Structures and vibrational Hamiltonians for (**a**) perfectly ordered, infinitely large parallel and (**b**) finite, structurally disordered β-sheets. The coupling constants, J_{mn}, are named such that m gives the separation between the coupled oscillators along the strand and n gives their interstrand separation

Coupling Constant	Strength (cm^{-1})
J_{10}	0.8
J_{20}	1.7
J_{01}	-8.8
J_{02}	-1.5
J_{11}	-1.1
$J_{1\text{-}1}$	1.5
$J_{\text{-}11}$	-1.0
$J_{\text{-}1\text{-}1}$	1.4

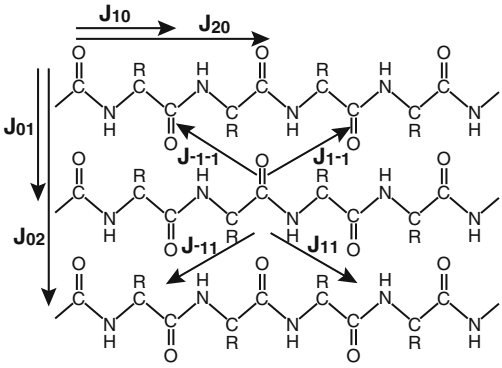

Fig. 9.3 The main coupling constants for a parallel β-sheet [23]. Again, the coupling constants, J_{mn}, are named such that m gives the separation between coupled oscillators along the strand and n gives the separation between strands

FTIR spectrum of a β-sheet (Fig. 9.4). The lower frequency mode appears at roughly 1,620 cm^{-1}, which we will refer to as the "asymmetric stretch" mode of the β-sheet since it creates chains of transition dipoles that lie perpendicular to the strands. The higher frequency mode, which we will call the "symmetric stretch," results from the oscillators vibrating in phase along the strand, but out of phase between strands.

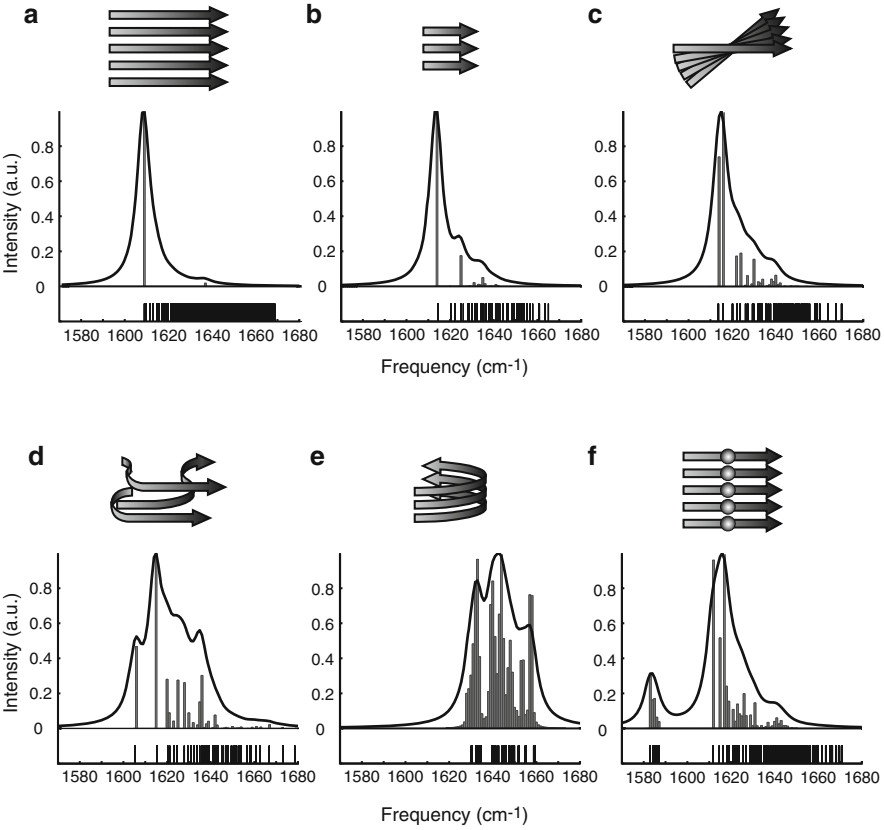

Fig. 9.4 Calculated schematic structures, FTIR spectra, and vibrational eigenstates for a variety of secondary structures of relevance to amyloid fibers. (**a**) Perfectly ordered, infinitely β-sheet. (**b**) Perfectly ordered, finite β-sheet. Reduced coupling at the edges causes some modes to appear at higher frequencies. (**c**) Twisted β-sheet. Twisting of the sheets around the fiber axis increases the distance between strands and thus decreases the coupling between transition dipoles. Twisting also breaks the symmetry of the two-dimensional β-sheet, which potentially creates more IR allowed transitions. Decreased coupling further broadens the β-sheet features and causes a slight shift to higher frequencies. (**d**) Structurally disordered β-sheet. Disorder changes coupling strengths or even eliminates coupling all together. Thus, many invisible modes appear, including a strong feature around 1,645 cm^{-1}, which results from uncoupled oscillators. (**e**) Turn or loop structure. Modes appear at both positive and negative coupling values. (**f**) Isotope-labeled β-sheet. The isotope-labeled peak modes appear around 1,585 cm^{-1}, well separated from the unlabeled features

However, there exist a large number of unseen vibrational modes spaced throughout the vibrational spectrum (see Fig. 9.4) whose existence is important for the reasons we give below.

While this ideal β-sheet provides a basic understanding of the spectroscopy/structure relationship, it is not very accurate and cannot be rigorously used to interpret IR spectra, because the "invisible" vibrational modes gain intensity in an imperfect fiber. The finite size, twist, turns, or loops, solvation, and structural disorder of the fibers cause the invisible modes to appear in the spectra. They also broaden and shift the β-sheet peaks and can generate new features. These effects are shown in Fig. 9.4a–e.

These effects have been discussed in the literature, and so we do not dwell on them here. However, to provide some quantitative insight into their origin, we discuss structural and environmental disorder, which is relevant to the infrared spectrum of all proteins. β-sheets do not start and stop abruptly, they fray. Thus, the edges of β-sheets have a broader distribution of structures, giving rise to variations in hydrogen bond strengths, solvation, and phi/psi angles that differ from one local mode to the next. As a result, the diagonal matrix elements in the vibrational Hamiltonian will differ, as can the off-diagonal elements (Fig. 9.2b). To estimate the effect of these structural distributions, consider that there is about a $10\,cm^{-1}$ difference in the frequency of hydrogen bonded versus non-hydrogen bonded amide-I mode. From this effect alone, the diagonal elements could differ from one another by much more than most of the off-diagonal coupling elements. Thus, to first order, one could just ignore the couplings, which would mean that those particular oscillators, even though they might still be defined as a β-sheet or a β-strand according to their phi/psi angles, would not contribute to the typical β-sheet features such as the mode at $1{,}620\,cm^{-1}$. Instead, they would contribute somewhere near $1{,}645\,cm^{-1}$, as if they were uncoupled. A similar effect occurs because of differences in the environment around the local vibrators, which can cause linewidth changes (and thus frequency differences) of around $5–10\,cm^{-1}$, which is once again comparable or larger than the coupling strengths. Each of these effects, as well as others that are not discussed here, will change the frequencies of the β-sheet modes and cause the invisible vibrational modes to gain intensity, thereby congesting the spectra. While we have not addressed each of these effects in detail, this brief overview illustrates the complexity in interpreting the infrared spectra of proteins and amyloid fibers.

Isotope labeling is the key to reducing the complexity of IR spectra and providing residue-specific structural resolution. When the backbone carbonyl of a residue is labeled with $^{13}C^{18}O$, the frequency of the amide I band for that residue is red shifted by approximately $54\,cm^{-1}$. The frequency shift does two things. First, it shifts the frequency enough that the labeled residues are resolved from the remaining unlabeled ones (Fig. 9.4f), so that the structures and environments corresponding to those particular residues can be studied. Second, the frequency shift allows their vibrational spectra to be interpreted largely independently from the unlabeled residues, so that the structural interpretation is simpler and more accurate. To understand the second point, consider the vibrational Hamiltonian once again.

In the matrix, we will need to subtract ~54 cm^{-1} from each of the diagonal elements that correspond to the isotope-labeled local modes. The off-diagonal elements will not change value, because isotope labeling does not change the structure of the fiber. Regardless, we may as well neglect the couplings between the labeled and unlabeled modes, because the diagonal frequency shifts are so much larger. But we cannot neglect the couplings between the isotope-labeled residues, because their frequency differences are small (depending on hydrogen bonding, solvation, and other differences discussed above). As a result, we get an effectively block-diagonal Hamiltonian, in which we can treat the isotope-labeled residues as independent of the others (Fig. 9.5b).

So, what we should expect to observe when we isotope label an amyloid peptide? Consider the vibrational Hamiltonian for parallel β-sheet, in which each strand is an amyloid-forming peptide. Since parallel β-sheets are in register, if a single residue in each peptide is isotope labeled (such as the 2nd residue, shown in Fig. 9.5b), then the labels will form a column that cuts through the β-sheet. As a result, we have a linear chain of oscillators that cuts across the strands and the isotope-labeled block of the Hamiltonian will have only J_{01} and J_{02} couplings. For an infinitely long fiber, the analytical solution to this classic physics problem is a series of vibrational eigenstates ranging from $E_1 + 2\beta$ to $E_2 - 2\beta$, where E_1 is the local mode frequency for the labeled oscillator and $\beta = J_{01} + J_{02}$. However, once again due to symmetry, only the mode at $E = E_1 + 2\beta$ appears in the spectrum. Thus, the spectrum does not look like a β-sheet, because it does not have the second observable mode which requires 2D symmetry. Moreover, the frequency shift is not 54 cm^{-1} from the unlabeled 1,620 cm^{-1} β-sheet mode unless by happenstance. There are many misstatements and mistaken conclusions in the literature stemming from a faulty understanding of these two facts.

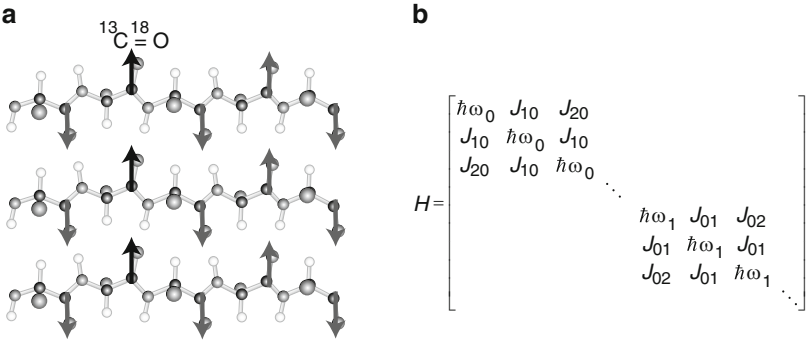

Fig. 9.5 (**a**) β-sheet with a single residue (*black*) isotope labeled in each strand, creating a linear chain of oscillators. (**b**) Block-diagonalized Hamiltonian for the labeled β-sheet, showing that the unlabeled and unlabeled Hamiltonians can be separated

9.3 Isotope Labeling Schemes

How do we use the isotope labels to obtain information on the fibril structure? There are a few ways. First, one obtains the coupling strength between the labeled residues. A large negative frequency shift is consistent with the coupling constants of parallel β-sheets. The most stable and inherently rigid regions of the β-sheets will have the largest frequency shifts, since diagonal disorder at these locations is minimized. Thus, the actual frequency is a measure of the β-sheet disorder. At the frayed ends of β-sheets where there is a large degree of disorder, small or negligible frequency shifts will be measured. In an actual fiber that contains a turn or a loop, the frequency shift can in fact be positive if the carbonyl groups are oriented perpendicular to the β-sheet axis.

Second, one can look at the effects that labeling has on the unlabeled infrared spectrum. The isotope-labeled residues leave a hole in the vibrational Hamiltonian of the remaining unlabeled amide I bonds, thus altering their spectral features depending upon where the disruption occurs. Simulations of β-sheets with 6, 8, or 10 strands and 10 residues per strand were performed, with the isotope label placed at different positions along the sheet [9]. In Fig. 9.6, the unlabeled β-sheet frequency versus label position is plotted. For β-sheet with 6 or 8 residues in each strand, we found that the asymmetric stretch frequency is perturbed the most when the column of isotope labels cuts the β-sheet in the middle into two smaller sheets, because smaller β-sheets have higher asymmetric stretch frequencies. For sheets with 10 or more residues, the largest frequency shift is caused by isotope labeling 3 or 4

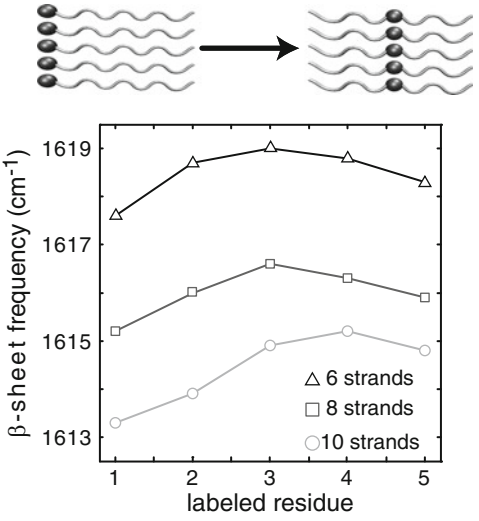

Fig. 9.6 Frequencies of the β-sheet feature plotted as a function of labeled residue position for 6, 8, and 10 strands of 10 residues

residues from the end, because sheets with only 3–4 residues per strand have a very high vibrational frequency.

The above two approaches are very powerful strategies for ascertaining structures in amyloid fibers and can be used with either 1D- or 2DIR spectroscopy. However, 2DIR spectroscopy has some additional benefits. In the above discussion, we argued that the coupling between the labeled and unlabeled residues can be ignored because its effects are small. However, the coupling still exists and gives rise to crosspeaks in the 2DIR spectra. Because these crosspeaks are indicative of the coupling strength between the unlabeled and labeled features, the intensity of the crosspeaks between labeled and unlabeled features can provide information on the position of the isotope label in the global fold. Shown in Fig. 9.7 are simulated 2DIR spectra for a β-sheet with either the first residue (Fig. 9.7a) or a central (Fig. 9.7b) residue labeled [9]. The β-sheet/label crosspeak is seven times larger when the isotope label occurs in the middle of the β-sheet rather than on the edge. The reason for this difference is that the strongest IR allowed mode for parallel β-sheets is the asymmetric stretch, in which residues in the middle of the strands have a larger contribution than those on the edge. Thus, if a residue on the edge of the β-sheet is labeled, the labeled eigenstate is going to be spatially separated by many residues from the transition dipole of the unlabeled eigenstate. However, if a more central residue is labeled then the full β-sheet will be broken apart into two smaller β-sheets, and the labeled and unlabeled transition dipoles will be closer together, increasing the coupling between the modes. Thus, the intensity of the β-sheet/label crosspeak can be used to gauge

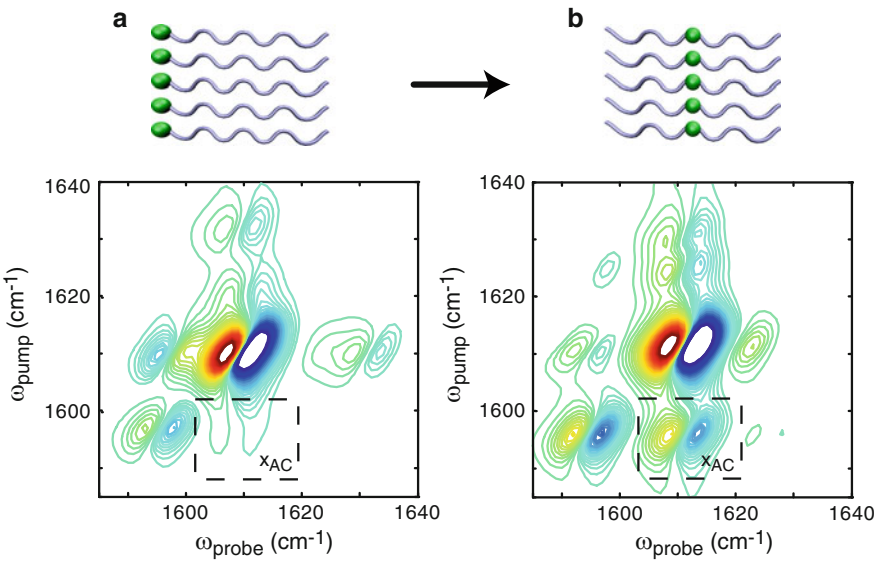

Fig. 9.7 Structures and simulated perpendicular 2DIR spectra for a seven-strand β-sheet labeled at either an (**a**) end residue or a (**b**) central residue. The *dashed box* highlights the crosspeak between the isotope-labeled mode and the unlabeled β-sheet mode

the position of the labeled residue within the β-strand. If the isotope label appears in a helix, loop, or turn, then crosspeaks to those secondary structures appear instead, as illustrated in Fig. 9.8 for the helical M2 transmembrane peptide.

A deletion strategy can also be applied to the crosspeaks. For example, the intensity of the x_{AB} crosspeak between the unlabeled β-sheet and the turn can be affected by isotope labeling, as shown in Fig. 9.9 [9]. If the labeled residue resides near the interface between the β-sheet and turn regions, the intensity of the

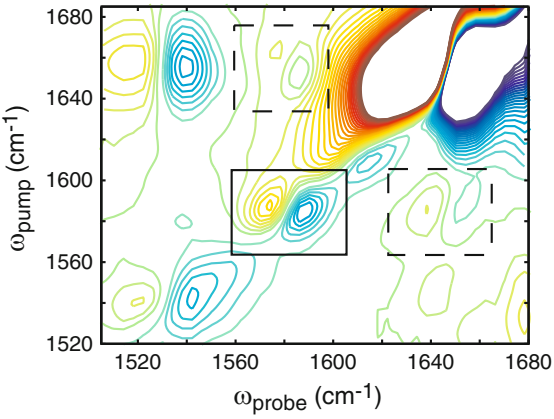

Fig. 9.8 2DIR spectrum of the transmembrane M2 peptide. The *solid box* highlights the isotope-labeled mode, while the *dashed boxes* show the crosspeaks between the isotope label and the alpha helix mode

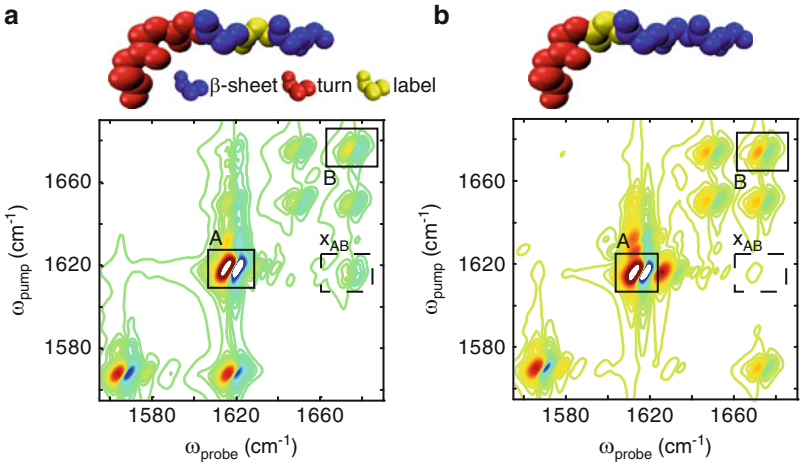

Fig. 9.9 Structures and simulated perpendicular 2DIR spectra for ideal seven-residue turn/β-sheet structures with the label either (**a**) buried in the β-sheet or (**b**) at in the interface between the turn and β-sheet regions

x_{AB} crosspeak is severely diminished. This effect is primarily due to the fact that the strongest coupling between the two structural regions is created by covalently bonded residues. When interfacial residues are labeled, a hole is created in the Hamiltonian between the β-sheet and turn modes, decreasing their coupling. The effect also occurs when residues near the interface are isotope labeled, because the label also alters the eigenstates of the β-sheet and turn region themselves. Therefore, residues near the interface of different secondary structures can be identified by the degree to which isotope labeling that residue suppresses the β-sheet/turn crosspeak in 2DIR spectra.

9.4 Vibrational Dynamics of Amyloids

In addition to coupling information contained in crosspeaks, 2DIR spectra also provide information on vibrational dynamics through the 2D lineshapes. Infrared lineshapes, either 1D or 2D, depend on the vibrational lifetime as well as the frequency distribution of the vibrational modes and the dynamics by which they interchange [26]. For instance, very fast dynamics create Lorentzian lines, whereas very slow or motionless dynamics cause Gaussian lineshapes. The fast limit gives rise to homogeneous dynamics, whereas the slow limit causes inhomogeneous broadening. Typically, IR lineshapes are neither Lorentzian nor Gaussian because the timescale for fluctuations of solvent around proteins is neither in the fast or slow dynamical range. Moreover, there is often more than one contribution to the lineshape. As a result, it is usually difficult if not impossible to accurately quantify the lineshape (and hence the dynamics) using 1D spectroscopies.

2DIR spectroscopy provides a means of quantifying vibrational dynamics. When a photon echo pulse sequence is used to collect a 2DIR spectrum, a symmetric 2D lineshape indicates that the system is homogeneously broadened, whereas elongation along the diagonal indicates that the vibrational mode has inhomogeneous character. A 2D lineshape analysis quantifies how much each contributes. To test whether the dynamics are in-between fast and slow, one collects a series of 2D IR spectra with nonzero time delays between the pump and probe pulses, and looks for changes in the 2D lineshape. A broadening of the anti-diagonal width or a change in the nodal slope are quantities that can be measured to determine the timescales of the dynamics.

An excellent example of how one can use 2DIR lineshapes to study protein structures is the ovispirin peptide bound to membrane bilayers [27]. Residues that lie near the surface of the membrane have very elongated lineshapes due to the backbone experiencing a disordered environment with strong electrostatics, giving rise to inhomogeneous broadening. In comparison, the 2D lineshapes for residues facing the membrane interior can be as much as $15\,cm^{-1}$ narrower, due to the much weaker electrostatics. By measuring a series of residues, one learns the secondary structure of ovispirin and a comparison to simulations gives its depth in the bilayer (Fig. 9.10).

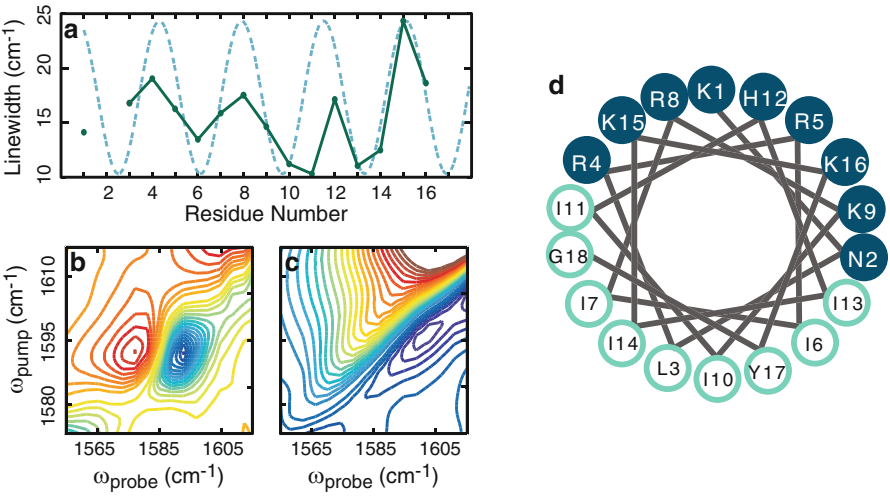

Fig. 9.10 (**a**) 2DIR diagonal linewidths for 15 isotope-labeled residues of ovispirin (*solid line*). The *dotted sinusoidal line* is plotted with a period of 3.6 amino acids. Example 2DIR spectra of the isotope-labeled mode for (**b**) I10 and (**c**) K15. K15 is significantly more inhomogeneously broadened than I10. (**d**) Helical diagram of ovispirin; hydrophilic residues are colored dark, while hydrophobic residues are light. The hydrophilic residues have larger measured linewidths than the hydrophobic ones, which agrees with predictions that the hydrophilic region of the helix should be more deeply buried in the lipid bilayer

9.5 Experimental Methods

We have previously published methods for synthesizing hIAPP [9, 28], as well as for synthesizing and purifying $^{13}C{=}^{18}O$ isotope-labeled amino acids for use in the Fmoc peptide synthesis [29]. Samples are denatured in d-HFIP and lyophilized. To initiate fibrillization, samples are dissolved in 20 mM deuterated potassium phosphate buffer at pH 7 to generate a peptide concentration between 500 μM and 1 mM. 5 μL of the solution is placed between CaF_2 windows with a 56 μm spacer. For kinetics scans, there is approximately 2 min of dead time between initiation of aggregation and the first collected 2DIR spectrum.

2DIR spectra are collected in a pump-probe beam geometry using a mid-IR pulse shaper to generate the required pulse train electronically rather than mechanically. We have published review articles on collecting 2DIR spectra with pulse shapers [11, 30–32]. Briefly, femtosecond pulses of 6.1 μm light are generated with a difference frequency generation-based optical parametric amplifier. Generally, around 3 μJ pulses are produced. Roughly 10% of each pulse is split off to serve as the probe (k_3), while the rest is sent into a mid-IR pulse shaper. The pulses are dispersed into the frequency domain by a 150 g/mm ruled grating and focused in the Fourier plane onto a Ge acousto-optic modulator. An arbitrary waveform generator is used to create a sinusoidal amplitude mask which amplifies and phase modulates the

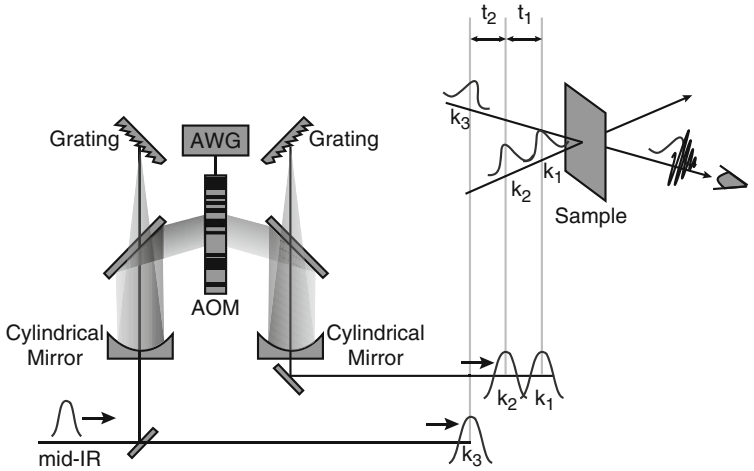

Fig. 9.11 Schematic 2DIR experimental setup

dispersed pulse before it is recombined with a second grating. By varying the period of the sine wave, we create two pump pulses in the time domain, k_1 and k_2, with a variable delay between them. The collinear pump pulses interact with the probe pulse at the sample, which emits a third-order electric field in the same direction as the probe beam. The probe beam also serves as the local oscillator (LO), which easily allows us to implement heterodyne detection. The combined third-order field and LO are dispersed through a monochromator and detected with an MCT array (Fig. 9.11).

Our spectra are generated by scanning over 2,560 fs of delay between the k_1 and k_2 pulses in 24 fs steps. Four π-phase shifted pump pulses are used to remove the transient absorption background and scatter to improve the signal-to-noise. Spectra can be taken with the polarization of the pump pulses (k_1 and k_2) either parallel or perpendicular to the probe pulse (k_3). Crosspeaks that arise from coupled transition dipoles that are nonparallel are better resolved in perpendicularly polarized spectra.

9.6 Experimental Data

We highlight two studies on amyloids using 2DIR spectroscopy that exploit the strategies and methodologies given above. The first study is our own, in which we have measured the aggregation kinetics of six individual residues of hIAPP (shown in Fig. 9.12) using isotope labeling [10]. In these experiments, we were able to show for the first time that fiber formation does not occur solely by a nucleation method as previously thought, but involves a competition with polypeptide folding. The second study is by Hochstrasser and coworkers, in which they have found a water pocket in fibers of Aβ through 2D lineshape analysis [8].

The most commonly used method for studying the kinetics of amyloid fiber formation is thioflavin-T ThT fluorescence [33–35]. ThT is a fluorescent dye that binds to β-sheets. As it binds, its maximum absorption and emission wavelengths both red shift and its fluorescence intensity increase [33]. Thus, by monitoring the ThT fluorescence spectrum, one measures the growth of amyloid fibers. It is a crude measure, however, because it has nonspecific binding to other structures than to β-sheets. Circular dichroism is also used, but it is most sensitive to α-helices (which are thought to play a role in membrane-catalyzed hIAPP aggregation), requires problematic deconvolution, and is not residue-specific either. We have followed the fibrillization of amylin with residue specificity using $^{13}C^{18}O$ isotope labeling and 2DIR spectroscopy [10]. Shown in Fig. 9.13 are a set of 2DIR spectra taken over time for a sample of hIAPP with an isotope label at Ala-25. The first spectrum was

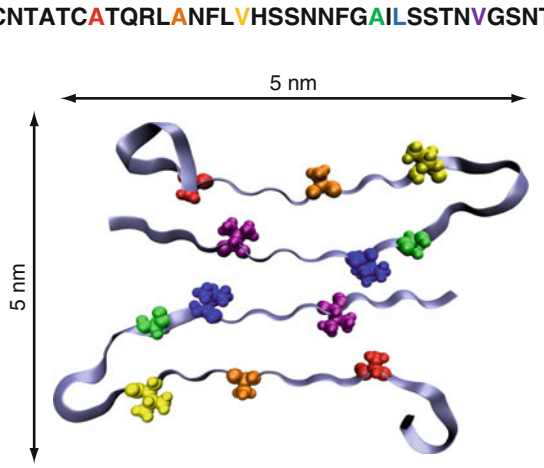

Fig. 9.12 The sequence of hIAPP with the six isotope-labeled residues highlighted. Cross section of an hIAPP fiber, taken from ssNMR data, showing the positions of the six labels relative to the regions of secondary structure. The labels are Ala-8 (*red*), Ala-13 (*orange*), Val-17 (*yellow*), Ala-25 (*green*), Leu-27 (*blue*), and Val-32 (*purple*)

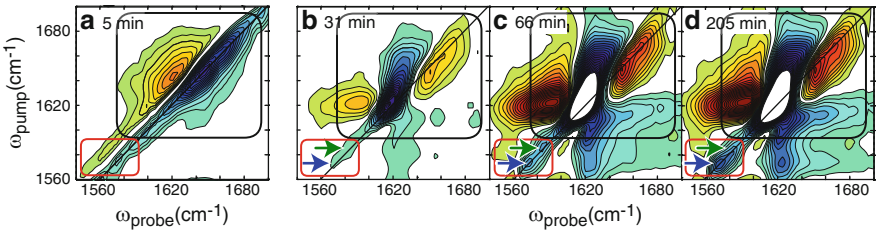

Fig. 9.13 Series of 2DIR spectra for hIAPP labeled at A25. (**a**) Spectrum taken 5 min after aggregation is initiated. (**b–d**) Difference 2DIR spectra at 31, 66, and 205 min after initiation. *Black boxes* enclose unlabeled features, while *red boxes* enclose labeled features. The *green* and *blue arrows* point out the two isotope-labeled peaks that appear initially

taken 5 min after aggregation was initiated. The subsequent spectra are plotted as the difference between the spectrum at that time and the first spectrum in order to better observe changes in the features.

At early times, the spectrum comprises a very broad doublet at 1,645 cm^{-1} (black box), resulting from a large distribution of random coil configurations, and an isotope-labeled feature at 1,580 cm^{-1} (red box). As time progresses, the random coil feature disappears while an unlabeled β-sheet doublet grows in at 1,617 cm^{-1}. The isotope-labeled feature initially appears as two doublets, until at much later times the lower frequency feature at 1,575 cm^{-1} dominates. Large crosspeaks grow in between the isotope-labeled peak and the unlabeled β-sheet, suggesting that Ala-25 is strongly coupled to the β-sheet. When the intensities of the β-sheet, label, and crosspeak are plotted as a function of time (Fig. 9.14), we can see that the kinetic curves are all sigmoidal, which is typical for amyloid kinetics with a well-established lag phase. Moreover, the crosspeaks and diagonal peaks have a nearly identical time to half-maximum (t_{50}), which indicates that Ala-25 is assembled directly into a β-strand structure as it is incorporated into the fibril and that Ala-25 has about "average" kinetics.

However, not all labels exhibit kinetics the same as that of the unlabeled residues. The kinetics of some isotope labels occur sooner and others later than the "average" as measured by the unlabeled features. This observation implies that different sites on the fiber are incorporated into their β-sheet structure at different points along the fiber formation mechanism. However, it is very difficult to directly compare

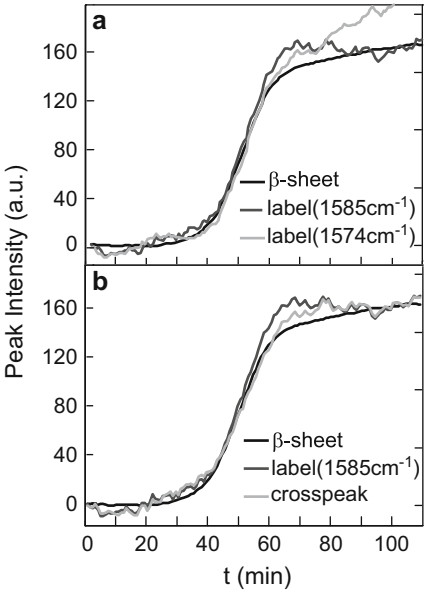

Fig. 9.14 Kinetics of the unlabeled β-sheet feature at 1,617 cm^{-1} compared with the (**a**) diagonal label features and (**b**) diagonal label feature and crosspeak

one experiment to another, because amyloid kinetics are not perfectly reproducible. From one experiment to the next, the t_{50} times for an identical experiment can vary considerably. Fortunately, the unlabeled β-sheet peak provides a convenient internal standard by which we can eliminate run-to-run variation. If each kinetics curve is scaled by the unlabeled β-sheet t_{50}, the scaled curves are identical to within signal-to-noise. Thus, to compare one kinetics experiment to another, we divide the x-axis by the t_{50} time of the unlabeled residues. The scaled kinetics curves will match if the residues fold via the same mechanism [36].

Figure 9.15 shows the scaled kinetic curves for the six isotope-labeled residues shown in Figure, as well as a plot of their scaled t_{50} times versus residues number. The plots reveal a dramatic difference in folding times, with some residues transitioning sooner than the unlabeled β-sheet feature ($t_{50} < 1$) and others folding later ($t_{50} > 1$). As the first residue to transition, Val-17 folds 10% faster than the unlabeled residues, while Val-32 is the last at 30% slower than the unlabeled ones. For a typical t_{50} folding time of 50 min, this lag corresponds to a 20 min separation between the folding of Val-17 and Val-32, with all the other residues falling in the middle.

How do we interpret the intensity change of the isotope-labeled features? We use the coupling Hamiltonian discussed above in Sect. 9.2. When a polypeptide has been incorporated into its final fiber structure, its isotope label lies in register, forming a linear chain of coupled oscillators. As a result, the frequency shifts by 2β, according to the given equation. Thus, the intensity change of the label is a measure of when that particular residue adopts its final fibrillar conformation.

These data allow us to propose a possible aggregation pathway for hIAPP (Fig. 9.16). As residues in the middle of the peptide are the first to transition, nucleation might occur near the loop of the fiber, followed by folding down the N- and C-terminal β-sheets. The N-terminal sheet appears to form in approximately half the time as the C-terminal, according to the two lines fit to the data in Fig. 9.15b. Since the crosspeaks develop at the same rate as the isotope-labeled peaks, we believe that the β-sheets form in register. However, we cannot say for certain that

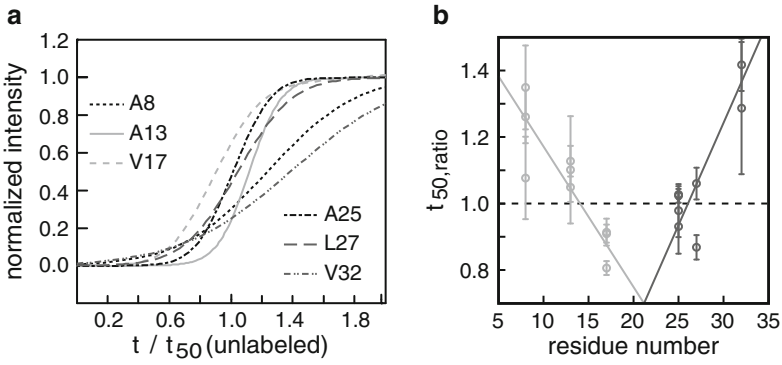

Fig. 9.15 (**a**) Sigmoidal fits of the kinetic traces for all six labeled peptides, scaled to the t_{50} of each scans unlabeled β-sheet feature. (**b**) t_{50} ratios measure for each experimental run

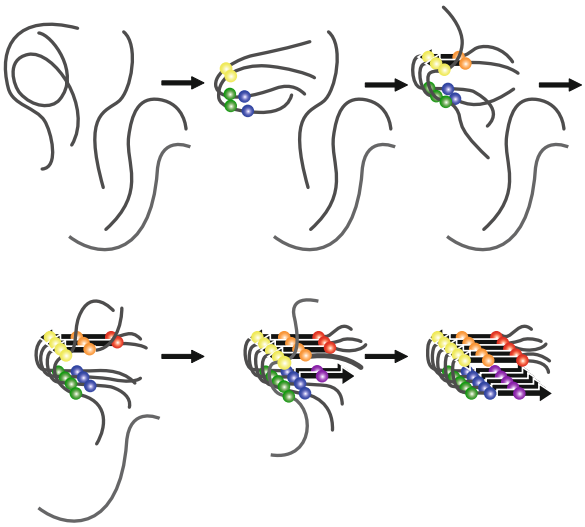

Fig. 9.16 Proposed aggregation mechanism for hIAPP. The turn region forms first, followed by propagation of the β-sheet down the N- and C-terminal sections, with the N-terminal β-sheet forming more quickly

there are not additional steps or that there are not competing pathways, but whatever the model is proposed, it must match our measurements which give the percentage of each residue in its formed β-sheet as a function of time. While this mechanism is still highly schematic, it provides some of the most detailed structural information currently known about peptide aggregation.

Our second experimental example relies on 2D lineshape analysis of $^{13}C^{18}O$-labeled Aβ fibrils to reveal the presence of individual water molecules trapped within the fibril core [8]. Any such water molecules could have a significant influence on the structure and formation kinetics of amyloid fibrils. Hochstrasser and coworkers measured the 2D lineshapes ^{18}O isotope-labeled samples of Aβ40 formed in D_2O (Fig. 9.17) to see whether any were inhomogeneously broadened. They found that residues L17, V18, L34, and V36 have larger inhomogeneous linewidths, than the other 14 residues that were measured. Moreover, the isotope-labeled mode for these residues appears at higher frequencies than the rest. Both these effects can be explained by the presence of water in the fibril core, perturbing the coupling by creating disorder in the β-sheets and hydrogen bonding to these residues.

Hochstrasser also found that the trapped water molecules are mobile, because they cause the vibrational frequencies of nearby amide groups to vary rapidly in time. The spectral diffusion, or waiting-time evolution of the 2DIR lineshape, was measured by varying the time delay between the pump and probe pulses (Fig. 9.18). It was found that the inverse slope, which is a measure of the inhomogeneous distribution, decayed over 1–2 ps for the residues with water nearby, whereas no

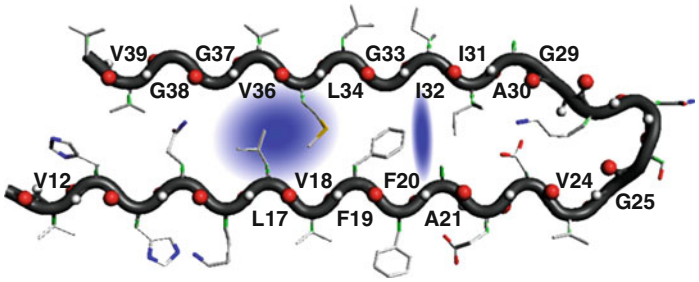

Fig. 9.17 Monomer of Aβ40 peptide with the $^{13}C{=}^{18}O$ labeled residues indicated. The *blue* areas are thought to contain D_2O molecules

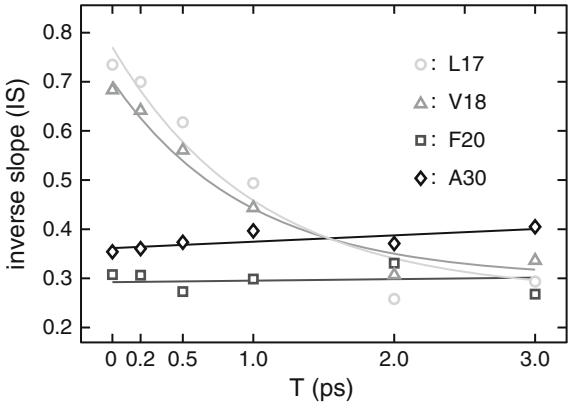

Fig. 9.18 Plot of inverse slope versus the time delay between pump and probe pulses for L17, V18, F20, and A30. L17 and V18 can be fit to exponential decay curves with time constants of around 1 ps

significant time evolution was observed for the other residues. These data support the idea that pockets of mobile water are trapped near those residues with larger inhomogeneous linewidths.

These effects are not due to the presence of bulk water in the fibril. Several of the residues have multiple peaks along the diagonal in the isotope-labeled region. Some of these even have obvious crosspeaks, such as in the case of I31. From the discussion of coupling for the isotope label in Sect. 9.2, we know that isotope-labeled modes can appear between the local mode frequency E_1 for a completely uncoupled oscillator and $E = E_1 + 2\beta$ for a perfectly ordered, infinite chain of coupled oscillators. Thus, the lowest energy peak in each spectrum is generated by a structure with a high degree of coupling between isotope labeled residues, while higher energy peaks results from excitons that are more localized on a single residue due to weakened coupling. The presence of crosspeaks confirms that the diagonal peaks must come from the same fibril. This suggests that within a fibril there are

finite clusters of disturbed structure, which means that only small pockets of water can be buried within the fibril.

From this evidence, Hochstrasser and coworkers concluded that small pockets of water molecules are trapped Aβ fibrils between residues 16–18 and 34–36. To a lesser extent, water may be trapped between residues 21 and 32 (Fig. 9.17). On average, they calculated that 1.2 water molecules were present inside fibrils for every monomer of Aβ40.

9.7 Summary

In this chapter, we have shown that infrared spectroscopies combined with isotope labeling can be a powerful tool for studying peptide structure and kinetics. Isotope labeling a single residue in a peptide isolates the vibrational mode of that residue from the unlabeled residues, resulting in significant changes to infrared spectra that can be interpreted with a simple vibrational Hamiltonians without having to account for the structure of every amino acid in the protein. The changes in the frequency of the labeled and unlabeled features can be measured with standard FTIR spectroscopy. 2DIR spectroscopy provides further information through crosspeaks and 2D lineshapes. Insightful experiments have already been performed using these approaches. With advances in peptides synthesis, it is becoming easier to synthesize larger peptides so that isotope labels as well as nonnative amino acids can be incorporated into proteins of 50 or 80 residues. Protein expression can be used as well. Thus, the range of systems for which infrared spectroscopy can be applied in a structurally specific manner is increasing rapidly.

Acknowledgements The authors would like to acknowledge Sang-hee Shim, David Strasfeld, and Yun Ling who contributed much of the data on hIAPP presented in this chapter and Yung Sam Kim and Robin Hochstrasser for providing the data on Aβ. We would also like to acknowledge our collaborators Juan de Pablo and Jim Skinner as well as funding from the NSF CHE-0832584 and NIH DK079895.

References

1. J.J. Meier, R. Kayed, C.-Y. Lin, T. Gurlo, L. Haataja, S. Jayasinghe, R. Langen, C.G. Glabe, P.C. Butler, Am. J. Physiol. Endocrinol. Metab. **291**, E1317 (2006)
2. A.E. Butler, J. Janson, W.C. Soeller, P.C. Butler, Diabetes **52**, 2304 (2003)
3. S.A. Petty, S.M. Decatur, J. Am. Chem. Soc. **127**, 13488 (2005)
4. R.A. Gangani, D. Silva, W. Barber-Armstrong, S.M. Decatur, J. Am. Chem. Soc. **125**, 13674 (2003)
5. S.A. Petty, S.M. Decatur, Proc. Natl. Acad. Sci. U.S.A. **102**, 14272 (2005)
6. S.M. Decatur, Acc. Chem. Res. **39**, 169 (2006)
7. Y.S. Kim, L. Liu, P.H. Axelsen, R.M. Hochstrasser, Proc. Natl. Acad. Sci. U.S.A. **105**, 7720 (2008)

8. Y.S. Kim, L. Liu, P.H. Axelsen, R.M. Hochstrasser, Proc. Natl. Acad. Sci. U.S.A. **106**, 17751 (2009)
9. D.B. Strasfeld, Y.L. Ling, R. Gupta, D.P. Raleigh, M.T. Zanni, J. Phys. Chem. B **113**, 15679 (2009)
10. S.-H. Shim, R. Gupta, Y.L. Ling, D.B. Strasfeld, D.P. Raleigh, M.T. Zanni, Proc. Natl. Acad. Sci. U.S.A. **106**, 6614 (2009)
11. C.T. Middleton, A.M. Woys, S. Mukherjee, M.T. Zanni, Methods **52**, 12 (2010)
12. S. Mukherjee, P. Chowdhury, F. Gai, J. Phys. Chem. B **111**, 4596 (2007)
13. S. Mukherjee, P. Chowdhury, F. Gai, J. Phys. Chem. B **113**, 531 (2009)
14. D.H.J. Lopes, A. Meister, A. Gohlke, A. Hauser, A. Blume, R. Winter, Biophys. J. **93**, 3132 (2007)
15. S. Jha, D. Sellin, R. Seidel, R. Winter, J. Mol. Biol. **389**, 907 (2009)
16. D.L. Sellin, L.-M. Yan, A. Kapurniotu, R. Winter, Biophys. Chem. **150**, 73 (2010)
17. S. Luca, W.-M. Yau, R. Leapman, R. Tycko, Biochemistry **46**, 13505 (2007)
18. A.S. Reddy, L. Wang, S. Singh, Y.L. Ling, L. Buchanan, M.T. Zanni, J.L. Skinner, J.J. de Pablo, Biophys. J. **99**, 2208 (2010)
19. A.S. Reddy, L. Wang, Y.-S. Lin, Y. Ling, M. Chopra, M.T. Zanni, J.L. Skinner, J.J. de Pablo, Biophys. J. **98**, 443 (2010)
20. W. Zhuang, D. Abramavicius, T. Hayashi, S. Mukamel, J. Phys. Chem. B **110**, 3362 (2006)
21. S. Bagchi, C. Falvo, S. Mukamel, R.M. Hochstrasser, J. Phys. Chem. B **113**, 11260 (2009)
22. N. Sengupta, H. Maekawa, W. Zhuang, C. Toniolo, S. Mukamel, D.J. Tobias, N.-H. Ge, J. Phys. Chem. B **113**, 12037 (2009)
23. J.-H. Choi, S. Ham, M. Cho, J. Phys. Chem. B **107**, 9132 (2003)
24. C. Lee, M. Cho, J. Phys. Chem. B **108**, 20397 (2004)
25. S. Hahn, S. Kim, C. Lee, M. Cho, J. Chem. Phys. **123**, 084905 (2005)
26. P. Hamm, M.T. Zanni, *Concepts and Methods of 2D Infrared Spectroscopy* (Cambridge University Press, Cambridge, 2011)
27. A.M. Woys, Y.S. Lin, A.S. Reddy, W. Xiong, J.J. de Pablo, J.L. Skinner, M.T. Zanni, J. Am. Chem. Soc. **132**, 2832 (2010)
28. Y.L. Ling, D.B. Strasfeld, S.-H. Shim, D.P. Raleigh, M.T. Zanni, J. Phys. Chem. B **113**, 2498 (2009)
29. P. Marek, A.M. Woys, K. Sutton, M.T. Zanni, D.P. Raleigh, Org. Lett. **12**, 4848 (2010)
30. S.-H. Shim, D.B. Strasfeld, Y.L. Ling, M.T. Zanni, Proc. Natl. Acad. Sci. U.S.A. **104**, 14197 (2007)
31. S.-H. Shim, M.T. Zanni, Phys. Chem. Chem. Phys. **11**, 748 (2009)
32. E.M. Grumstrup, S.-H. Shim, M.A. Montgomery, N.H. Damrauer, M.T. Zanni, Opt. Express **15**, 16681 (2007)
33. H. Naiki, K. Higuchi, M. Hosokawa, T. Takeda, Anal. Biochem. **177**, 244 (1989)
34. H. Naiki, K. Higuchi, K. Nakakuki, T. Takeda, Lab Invest. **65**, 104 (1991)
35. Y.C. Kudva, C. Mueske, P.C. Butler, N.L. Eberhardt, Biochem. J. **331**, 809 (1998)
36. S.B. Padrick, A.D. Miranker, Biochemistry **41**, 4694 (2002)

Index

Absorption coefficient, 154
Acetylene-type linker, 188
Aci-nitro intermediate, 204
Active site fragment, 186
Aggregates, 92, 101
Aggregation, 217, 230
Aggregation pathway, 233
Aggregation/Fibrillation, 135
Allosteric structural changes, 176
Alpha-helix conformation, 18
Alzheimer's disease, 207
Alzheimer's peptide, 207
Amide A, 53
Amide B, 53
Amide I, 53, 92, 148, 179, 182, 219, 220, 225
Amide I band, 208
Amide I mode, 223
Amide I normal mode, 17
Amide I′, 149
Amide II, 55, 179, 182
Amide III, 55
Amide protons, 57
Amino acid side-chain groups, 57
Amorphous aggregates, 54, 82
AMPB, 184
AMPP, 176, 184
Amyloid, 207
Amyloid beta peptides, 81
Amyloid deposits, 78
Amyloid fibers, 220, 231
Amyloid fibrils, 78, 85, 217, 234
Amyloid kinetics, 232
Amyloid plaques, 85
Amyloid protein Aβ40, 36
Amyloid structure, 9
Amyloid-peptide packing models, 36
Anomalous intensity distribution in the β sheet conformation, 34
Antiparallel sheet, 37
Anti-Stokes lines, 153
Anti-Stokes process, 153
APB, 178
Apo-form, 75
Arrhenius equation, 151
Arrhenius plot, 160
Aspartate, 59
Atomic resolution structures of amyloid fibrils, 10
ATP binding, 198
Azobenzene, 174, 176, 178
Azobenzene containing peptides, 175
Azobenzene peptides, 182, 184
Azopeptides, 184

Ball-type mixer, 97
Band fitting procedures, 43
Barstar, 59
Beta-sheet conformation, 18
Binding sites, 198
Bond lengths, 201
Bovine pancreatic RNase A, 41
Bulk water, 56
Buried helix, 105
Burst phase, 104

Ca^{2+}-ATPase, 198
Ca^{2+}-ions, 75
Caged compounds, 194
Caged protons, 202
Caged sulfate, 202

H. Fabian and D. Naumann (eds.), *Protein Folding and Misfolding*,
Biological and Medical Physics, Biomedical Engineering,
DOI 10.1007/978-3-642-22230-6, © Springer-Verlag Berlin Heidelberg 2012

Calmodulin, 59
Cassegrain-type infrared microscope, 98
Cavitation, 156
CD spectroscopy, 54, 180
Chain length, 92
Chain packing, 37
Chaperones, 12, 132
Charge-flux model, 25
Chemical denaturants, 60
Chemical stability, 173
Chemical structure, 196
Chromo-proteins, 171
Circular dichroism, 93, 180
Collagen, 187
Collagen model peptide, 18, 39
Collapse and search, 94
Component, 164
Concentration jump, 198
Conformation, 196
Conformational change, 199
Conformers, 173
Constructing the G and F matrices, 22
Continuous-flow rapid-mixing apparatus, 92
Cosolvents, 102, 133
Coupled oscillators, 21
Coupling, 149, 226, 228, 235
Coupling strengths, 221, 223, 225
Cross-strand coupling, 165
Cross-β architecture, 9
Crosspeaks, 218, 226, 228, 230, 232, 235
λ-Cro repressor, 59
Cryo-EM, 10
^{13}C-substituted groups, 23
Cyclic Azobenzene Peptides, 183
Cyclic system, 176
Cytochrome *c*, apomyoglobin, 93

Denaturants, 101
Density functional theory, 20
Desolvation, 106
Determination of the protein orientation at an interface, 45
Deuterium oxide, 56
Diagonal elements of the F matrix, 23
Dialysis-related amyloidosis, 78
Diamond anvil cell, 122
Dielectric effect, 27
Difference spectra, 71
Difference spectrum, 194
Diffusion micromixer, 64
Diffusive mixing device, 97
Dihedral angles, 91
Dimethylsulfoxide, 101
Dipolar coupling constant, 110
Dipole derivative, 27
2D-IR, 51
2DIR spectroscopy, 218
Diseases of protein misfolding, 2
Disordered environment, 228
Disordered structure, 105
Disulfide bridge, 72
2D-NMR, 180
Driven structural changes, 176
Dynamic light scattering, 72

Early stages of thermal denaturation, 39
Eigenvalues of the GF matrix, 31
Eigenvector matrix, 31
Elastin, 142
Electric dipole in a peptide group, 26
Electric field, 196
Electron micrographs, 81
Electron paramagnetic resonance, 8, 81
Electronic coupling, 175
Energy landscape, 2, 3, 92
Engineered variants, 72
Enthalpic or entropic barriers, 191
Enzymatic activity, 173
Enzymatic reactions, 138
EPR, 10
Equilibrium constant, 149
Euler rotation matrix, 47
Experimental dead time, 69

F matrix, 20
Femtosecond, 173
Femtosecond laser, 181
Fibril, 212
Filamentous species, 79
Flexibility, 58
Flexible construct, 175
Flow-through, 63
Folding energy landscape, 2
Fourier self-deconvolution, 43, 102
Frequency correlations in globular proteins, 41
Frustration, 2
FTIR spectroscopy, 54
Fulgides, 174

G matrix, 20
G matrix elements, 22
Gaussian fitting, 102
Global fitting analysis, 103
Glutamate, 59

Guanidinium, 101
Guanidinium chloride, 60

H/D exchange, 188
β-hairpin, 176
Hairpin, 185
Hairpin formation, 161
Hamiltonian, 221, 224, 228, 233
Hamster prion protein, 85
Heat transfer, 155
Heat-treated assemblies, 81
Heavy water, 62
α-helices, 53
3_{10} helix, 75
Helix dynamics, 161
Helix-to-coil transition, 158
Heme-binding maquettes, 186
Hemithioindigo, 174
Hemoglobin, 59
High pressure, 117
Holo-form, 75
Homogeneous dynamics, 228
Homogeneously broadened, 228
Horse heart myoglobin, 42
HTI, 174, 175
Hydration, 92
Hydration shells, 56
Hydrodynamic Stokes, 79
Hydrogen bonding, 92, 196
Hydrogen bonding network, 188
Hydrogen bonds, 185, 187
Hydrogen exchange, 7, 10
Hydrogen–deuterium exchange, 104
Hydrogen/deuterium (H/D) exchange, 55

Indolylfulgides, 174
Infrared absorption spectroscopy, 92
Infrared diode laser, 158
Infrared (IR) spectroscopy, 218
Infrared spectroscopy, 17, 179
Inhomogeneous, 228
Inhomogeneous and homogeneous linewidths, 218
Inhomogeneous broadening, 228
Inhomogeneous linewidths, 234
Insulin, 81
Intensity of a normal mode, 32
π−π interaction, 24
Interactions between the peptide groups, 18
Intermediate states, 72
Intermediates, 6, 92
Interstrand coupling, 163
Intrinsic frequency, 110
Intrinsically disordered, 5
Inverse G matrix, 22
Ion mobility mass spectrometry, 9
IR bands, 53
IR spectroscopy, 53
IRRAS simulations, 45
Irregular structures, 54
Isoelectric point, 102
Isomerization, 173, 180
Isosbestic points, 108
Isotope editing of peptide structures, 18
Isotope labeled, 220, 232, 234
Isotope labeling, 32, 33, 162, 200, 218, 223, 230
Isotope labels, 224–227, 232
Isotope-editing, 162
Isotopic substitution, 101

Kinetics, 150, 210, 234

α-lactalbumin, 74
β-lactoglobulin, 103
Lactose synthase, 74
Laser flash experiments, 181
Laser-induced temperature jump, 95
Light switchable peptides, 172
Light-Triggered Peptide, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191
Lineshapes, 228, 234
Lineshape functions, 32
Linewidth, 223
Linking groups, 172, 175
Local mode, 220
β-loop, 176
Loop, 105
LPS monolayer, 47
Lysozyme, 81

Magic angle spinning, 81
Main chain collapse, 110
Main chain dehydration, 110
Major histocompatibility complex class I, 78
Mass-weighted displacement coordinates, 22
Mercaptoproline, 188
Mercury cadmium telluride (MCT) detector, 98
Methods, 4
Methylene spacer, 178
Metmyoglobin, 206
Michelson interference spectrometer, 98

Microfluidic devices, 98
β_2-microglobulin, 55
Mid infrared, 182
Mixing dead time, 93
Modified GF matrix method, 21
Molecular dynamics simulations, 6
Molecular mass, 79
Molten globule, 93
Molten-globule state, 75
Monolayers at the air–water interface, 45
Multiple labels, 166
Multistate folder, 160
Multistranded sheet structures, 24
Musculoskeletal system, 78
Myoglobin, 206

Native state, 92
Native-like structure, 66
Natively unfolded, 5
NMR, 180
Normal coordinate calculations, 19
Normal modes of vibration, 19
Nucleotide, 199

Octameric aggregate structures, 86
Off-diagonal elements of the F matrix, 25
Oligomers, 8, 211
Optical filter, 98
Optical theory of Kuzmin, 45
Oscillating electric dipole in the amide I mode, 26

Parallel organization of the β-strands, 81
Parallel sheet, 37
Peptide bond, 91
Peptide moieties, 172
Peptide moiety, 177
Peptides, 53
Pericyclic ring-opening, 173
pH jump, 202
pH-induced unfolded state, 101
Phosphoenzyme, 200
Photo-acoustic effects, 156
Photochromic switching unit, 172
Photochromism, 174
Photoisomerization, 173
Photolysis, 205
Photoresponsive, 175
Picosecond, 173
Pneumatic drive, 62
Point dipole approximation for the interaction energy, 26
Poly(*N*-isopropylacrylamide), 139
Polyglutamic acid, 103
Post-translational modifications, 12
Potential energy of a collection of static dipoles, 26
Primary sequence, 91
Prionoses, 85
Prion diseases, 85
Prion protein, 65
Protein backbone, 53
Protein-bound water, 56
Protein folding, 91, 124, 127, 207
Protein folding and misfolding in vivo, 11
Protein only hypothesis, 85
Protein subunits, 72
Proteins, 53
Proteostasis, 13
Protofibrillar structures, 87
Protofibrils, 8
Protofibrils, oligomers, 8
Protofilaments, 79
Pseudo-amino acid, 175
Pulse shaper, 229
Pump-probe beam geometry, 229
Pump-probe experiment, 181

Quantum efficiency, 173

Radius of gyration, 93, 104
Raman conversion, 154
Random coil, 211
Rapid mixing, 61, 92
Rapid-scan FTIR spectroscopy, 61
Rate constants, 150
Rate-determining step, 107
Ratio of absorptions of s- and p-polarized, 45
Reaction-induced infrared difference spectroscopy, 194
Reflectance–absorption of a thin film, 45
Relaxation dispersion, 7
Relaxation times, 159, 164
Relaxational processes, 184
Residual structures, 6
Resolution enhancement, 102
Resolvation, 185
Ribonuclease A, 43
Ribonuclease T1, 59
Ribosome, 11
Rigid coupling, 175

Ring opening, 174
RNase T1 variant, 71

β-sandwich, 78
Scattering techniques, 93
Second-derivative calculation, 102
Second-derivative spectrum, 109
Secondary structures, 54, 93, 147, 208
Secondary structure content, 93
Secondary structure formation, 110
Secondary-structure elements, 148
β-sheets, 53, 107, 210, 220, 222–226, 228, 231, 232
SH3 domain, 81
Signal-to-noise ratio, 61, 99
Simple harmonic oscillator, 21
Simulating the amide I contour, 31
Simulation of isotopic substitution, 34
Single labels, 166
Single molecule fluorescence, 5
Single molecule FRET, 7
Single-chain monellin, 93
Site-specific dynamics, 162
Site-specific frequency shifts, 162, 164
Small angle X-ray scattering, 93
Solid-state NMR, 11
Solubility, 177
Solution mixing, 95
Solution NMR, 7
Solution NMR spectroscopy, 5
Solvated helix, 105
Solvent, 177
Solvent heating, 151
Spectral diffusion, 234
Spectral probes, 166
Spectrometer, 157
Stability diagram, 119
Stability variations, 165
Static mixers, 100
Stationary spectroscopy, 177
Stimulated Raman effect, 152
Stokes line, 153
Stokes process, 153
Stopped-flow, 63, 93
β-strands, 54
Structural and environmental disorder, 223
Structural relaxation, 176
Structure, 58
Subdiffraction microscopy, 173
Substitution of ^{13}C, 33
Switching chromophore, 173
Switching cycles, 173
Synchronization, 181
α-synuclein, 81

T-jump kinetics, 159
T-shaped mixing channel, 97
Temperature jumps, 61, 151, 154
Temperature stability, 155
Tertiary contacts, 93
Tertiary structures, 62, 186
The double-jump technique, 102
Theoretical modeling, 180
Thermal denaturation, 18
Thermal lensing, 156
Thermal reconversion, 178
Thermoresponsive polymers, 139
Through hydrogen bond interaction, 29
Through hydrogen bond interaction force constant, 24
Through hydrogen bond $\pi-\pi$ interaction, 30, 31
Through-space transition–dipole interaction, 25, 38
Through valence bond interaction, 30
Time delay, 181
Time resolution, 194
Time-resolved, 93
Time-resolved infrared spectroscopy, 171
Time to half-maximum (t_{50}), 232
Toxicity, 8
Trans–cis isomerization, 66
Transition dipole coupling, 54, 68, 103, 149
Transition dipole moment, 25
Transition dipole–dipole interaction, 20
Transition dipole interaction, 29
Transition dipoles, 20, 222, 226, 230
Transition states, 6
Transmissible neurodegenerative disorders, 85
Transthyretin, 81
Trifluoroethanol, 103
Trigger, 194
Triple helix, 187
Triple-helical structure, 39
Triplet–triplet energy transfer, 180
Tryptophan fluorescence, 110
Turbulent mixing devices, 96
Turns, 53, 105
Two-dimensional infrared (2D-IR) spectroscopy, 183
Two-state folder, 160
Two-state folding process, 150
Two-state model, 158
Tyrosine, 59

Ultrafast absorption spectroscopy, 183
Ultrafast processes, 172
Ultrafast structural changes, 191
Ultrafast structural dynamics, 180
Unfolded ensemble, 4
Unfolded state, 92
Urea, 60, 101
Urea-unfolded state, 104

Valence bond interaction force constant between peptide groups, 30
Φ-value, 92
Φ-value analyses, 6
Vibrational dynamics, 228
Vibrational energy, 184
Vibrational excess energy, 174, 185
Vibrational Hamiltonian, 220, 223, 224
Vibrational modes, 220
Visible and ultraviolet spectroscopy, 179

Weakly coupled oscillator model, 19
Wilson GF matrix method, 19

X-ray crystal structures, 66

Zipper-like folding mechanism, 187