Solution NMR with large macromolecular assemblies – the GroE chaperonin system

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of

Doctor of Natural Sciences

presented by:

Jocelyne Fiaux
Chimiste diplomée (Université de Lausanne, Switzerland)
born on July 14th, 1975
citizen of Hermenches, VD

accepted on the recommendation of:

Prof. K. Wüthrich, examiner
Prof. R. Glockshuber, co-examiner
2002
Acknowledgements

I would like to thank Prof. Kurt Wüthrich for supervising my Ph.D. thesis, and for his continuous support and guidance during my doctoral work. He offered me the chance to work on an ambitious project and provided the best environment for successful research.

My thanks also go to Prof. Rudi Glockshuber for accepting the task of co-examiner for this work.

I would like to thank Dr. Roland Riek for teaching me the basics of solution NMR with large molecules and for providing support at all stages of this work. I thank him for his generosity, his openness, and for the many enjoyable discussions. He also developed the experiments and theoretical descriptions reported in chapter 3, and this work will be published separately (R. Riek, J. Fiaux, E. B. Bertelsen, A. L. Horwich and K. Wüthrich, to be published).

I would like to thank Prof. Arthur L. Horwich for the excellent collaboration. I thank Dr. Eric B. Bertelsen for his remarkable contribution to this project. He prepared the SR1 NMR samples, and initiated many experiments reported in chapter 6. His advice, his patience and kindness were also invaluable for our teamwork.

My thanks also go to Prof. Thomas Szyperski for teaching me methods for the investigations of cellular metabolism in the first year of my Ph.D. thesis. I benefitted a lot from the good collaboration and from the help and support he provided.

I would like to thank my colleagues for their precious contribution to
this work: the preliminary 3D structure calculations were done in collaboration with PD Dr. P. Güntert. For their careful reading and valuable comments on my thesis, I would also like to thank Sophie Bonjour, Sebastian Hiller, and Dominik Lysek. I also thank Rudolf Baumann for the technical assistance with the many electronic problems I could create.

Finally, I thank all colleagues in the group for the good working atmosphere, for their help and support, for their friendliness. I also thank Daniel and my friends for forgiving me for the long days spent in the lab instead of with them. I thank my family for their continuous support through good and bad times.
# Table of Contents

## ACKNOWLEDGEMENTS

## TABLE OF CONTENTS

## SUMMARY

## RÉSUMÉ

## ABBREVIATIONS

## 1. INTRODUCTION

1.1 NMR with large macromolecular structures

1.2 The GroE chaperonins

   *Molecular chaperones*

   *The chaperonins*

   *The GroEL-GroES system*

   *The GroE cycle*

   *Prospects and open questions*

1.3 Aims and scope of this thesis project

## 2. PREPARATION OF GROE PROTEINS FOR NMR MEASUREMENTS

2.1 Requirements for solution NMR studies of large molecules

2.2 Stable isotope labeling of the GroE chaperonins

   *Uniform $^{15}$N,$^{2}$H- and $^{13}$C,$^{15}$N,$^{2}$H-labeling*

   *$[^{15}$N,$^{2}$H]-Leu [$^{2}$H]-chaperonins*

2.3 Purification and biophysical characterization of the chaperonins

   *GroEL purification*
3. Solution NMR Techniques for Large Supramolecular Structures ............................................ 53

3.1 Sample conditions ............................................. 54
3.2 NMR equipment .................................................... 54
3.3 Water suppression in NMR experiments with large structures .................. 55
3.4 One-dimensional spectra ........................................... 57
   One-dimensional $^1$H NMR experiments .................................. 57
   Pulse length calibration ............................................... 57
   Adjustment of the recycle delay ........................................ 57
3.5 Two-dimensional correlation experiments ........................................ 61
   2D $[^{15}\text{N},^1\text{H}]-\text{HSQC}$ and $[^{15}\text{N},^1\text{H}]-\text{TROSY}$ .................... 61
   2D $[^{15}\text{N},^1\text{H}]-\text{CRIPT-TROSY}$ .................................. 61
   2D $[^{15}\text{N},^1\text{H}]-\text{CRINEPT-TROSY}$ ................................. 69
   2D $[^{15}\text{N},^1\text{H}]-\text{CRINEPT-HMQC-[}^1\text{H}]-\text{TROSY}$ ....................... 69
   A comparison of the 2D $[^{15}\text{N},^1\text{H}]-\text{CRINEPT-TROSY}, 2D [^{15}\text{N},^1\text{H}]-\text{CRINEPT-HMQC-[}^1\text{H}]-\text{TROSY}$ and 2D $[^{15}\text{N},^1\text{H}]-\text{CRIPT-TROSY}$ experiments ........................................... 70
3.6 $^{15}\text{N}$ and $^1\text{H}$ linewidths in TROSY-type experiments ...................... 76
3.7 Influence of the deuteration level of the proteins ................................... 79
3.8 Discussion .......................................................... 82

4. Backbone Assignment of the GroES Co-chaperonin .............................................. 85

4.1 NMR spectroscopy .................................................... 86
4.2 Assignment procedure ................................................. 87
4.3 Secondary structure of GroES ........................................ 94
4.4 Calculation of the 3D-fold of GroES ....................................... 94
4.5 Chemical shift list ..................................................... 100
5. **GroES–GroEL Interactions Mapped by Solution NMR** 105

5.1 $^{15}$N,$^2$H-GroES–chaperonin complex formation followed by NMR 106

5.2 Mapping of the GroES–GroEL interaction surface 112

5.3 Discussion 116

6. **NMR Studies of GroEL and SR1** 119

6.1 2D correlation spectra of $^{15}$N,$^2$H-labeled GroEL and SR1 119

6.2 Leucine-specific $^{15}$N-labeling of GroEL and SR1 126

6.3 Nucleotide and GroES binding to GroEL and SR1 128

7. **Discussion and Perspectives** 133

**Literature** 139

**Appendix I. Metabolic Flux Ratio Analysis in Micro-organisms Using Fractional $^{13}$C-Labeling and 2D NMR** 163

**Appendix II. Backbone Resonance Assignment of GroES** 171

**Publication List** 193

**Curriculum Vitae** 195
Summary

Nuclear magnetic resonance has been a method of choice for structural characterization of proteins and nucleic acids up to ~25 kDa and the introduction of transverse relaxation-optimized spectroscopy (TROSY) already has extended the size range of structures amenable to solution NMR investigations to about 150 kDa. The large size of many protein complexes however still represents a challenge for NMR studies, and the development of NMR approaches for investigations of larger molecular sizes is of considerable interest. This thesis describes studies of very large supramolecular structures in the range of 200-900 kDa with solution NMR, starting from the previously established TROSY and CRINEPT techniques.

Labeling of these large molecular assemblies with stable isotopes is essential for NMR investigations. Several labeling protocols were established for uniform and residue-specific labeling on a deuterated background. These experiments required knowledge about the amino acid biosynthetic pathways and built on experience obtained from a first research project on investigations of the central carbon metabolism in micro-organisms using $^{13}$C-labeling of amino acids and 2D NMR. This work is summarized in the Appendix 1 in the form of a short introduction of the method applied, and a collection of the published work.

Refined implementations of experiments based on the TROSY and CRINEPT principles were established and applied to the *E. coli* chaperonin proteins GroE labeled with $^{15}$N and $^2$H. In this molecular chaperone system, GroEL, a double ring-shaped homo-tetradecamer, interacts with the dome-shaped heptameric co-chaperonin GroES in the
presence of ATP to form a chamber that is the site of productive protein folding. Spectra were recorded for GroES (72 kDa), GroEL (800 kDa), and the single-ring variant of GroEL, SR1 (400 kDa). Detailed analyses of the NMR signals provided a guideline for adequate selection of the experimental schemes, the water suppression technique, and parameters such as the polarization transfer delay and the recycle delay.

Investigations of complexes of isotope-labeled GroES with SR1 or GroEL using these techniques revealed that nearly complete $^{15}\text{N},^{1}\text{H}$-correlation spectra can be obtained at 472 and 872 kDa. Furthermore, observation of chemical shift changes in GroES upon complex formation with the chaperonins enabled a mapping of GroES surface areas involved in intermolecular contacts. The data showed that only the mobile loop of GroES undergoes a major conformational change upon binding to GroEL and becomes structured. Interestingly, residues in this region exhibit variable dynamic properties and raise interesting possibilities for the binding mode of these proteins. Applications of the same NMR techniques to isotope-labeled GroEL and SR1 yielded preliminary results on nucleotide and GroES binding.

Quite generally, the collection of informative solution NMR spectra of structures with molecular weights up to 900 kDa and the manifestation in these spectra of conformational changes upon complex formation opens entirely new possibilities for studying functional interactions and structural dynamics in large macromolecular assemblies.
Résumé

La spectroscopie de résonance magnétique nucléaire (RMN) est devenue une technique puissante pour étudier la dynamique et la structure des protéines et acides nucléiques de poids moléculaire inférieur à 25 kDa. Déjà, l’introduction de la technique TROSY ("transverse relaxation-optimized spectroscopy") a permis d’étendre la classe de macromolécules accessibles à l’étude par RMN en solution à une taille d’environ 150 kDa. La dimension importante de nombreux complexes protéiques représente toutefois un défi considérable pour l’application de la RMN; c’est pourquoi il existe un intérêt croissant pour le développement de nouvelles approches permettant l’étude de grosses macromolécules à l’aide de la RMN. La présente thèse décrit l’étude de structures supramoléculaires de l’ordre de 200-900 kDa par RMN en solution, à partir des principes de TROSY et CRINEPT ("cross-correlated relaxation-induced polarization transfer") récemment établis.

Le marquage isotopique des protéines à étudier est indispensable à l’étude RMN. Différents protocoles ont été développés pour le marquage uniforme et spécifique, ainsi que la deutération des macromolécules. Ces méthodes présupposent, entre autres, la connaissance des voies métaboliques de synthèse des acides aminés. Ainsi, ces expériences s’appuient sur la compréhension acquise au cours d’un premier projet de recherche impliquant l’étude du métabolisme central de micro-organismes à l’aide de marquage $^{13}$C et de la RMN à deux dimensions. Les résultats de ce travail sont présentés en annexe sous forme d’une brève introduction à la méthode appliquée et d’une collection de publications.

Des corrélations hétéronucléaires basées sur les techniques TROSY et CRINEPT ont été optimalisées et appliquées à l’étude des protéines
chaperonines GroE de *E. coli*. Le chaperon moléculaire GroEL, composé de 2 anneaux de 7 sous-unités identiques, a pour fonction d’aider au repliement de substrats protéiques en association avec son cofacteur heptamérique GroES. Des spectres ont été mesurés pour GroES (72 kDa), GroEL (800 kDa) et SR1 (400 kDa), un variant de GroEL constitué d’un seul anneau heptamérique. L’analyse détaillée des résonances observées a fourni des indications précises pour la sélection des séquences appropriées, de la technique de suppression du solvent et de paramètres tels que le temps de transfert d’aimantation ou le délai entre les scans.

L’étude, à l’aide de ces techniques, de complexes entre GroES marquée et SR1 ou GroEL a démontré que les spectres de corrélation obtenus pour les tailles de 472 et 872 kDa, respectivement, contiennent un ensemble de résonances pratiquement complet. De plus, l’observation de changement du déplacement chimique pour certaines résonances appartenant à GroES lors de sa liaison à GroEL a permis de délimiter la surface d’interaction des chaperonines. Les données mettent en évidence la transition de la boucle mobile de GroES vers une conformation structurée dans le complexe avec GroEL. Le fait notable que les résidus de cette partie de la molécule possèdent des propriétés dynamiques variables soulève des questions intéressantes quant au mode de liaison de ces protéines. L’application des mêmes techniques aux protéines marquées GroEL et SR1 ont parallèlement fourni des résultats préliminaires au sujet des modes d’interactions des chaperonines avec des nucléotides et avec la co-chaperonine GroES.

De manière générale, la disponibilité de spectres porteurs d’informations pour des structures moléculaires de taille importante et l’observation, dans ces spectres, de transitions conformationnelles ouvrent de nouvelles voies pour l’étude des interactions fonctionnelles et de la dynamique pour de tels complexes.
 Abbreviations

1D, 2D, 3D  1-dimensional, 2-dimensional, 3-dimensional
ADP       adenosine 5’-diphosphate
AMP-PNP   adenyllyl- imidodiphosphate
ANS       4,4’-bis(1-anilino-8-naphtalenesulfonic acid)
ATP       adenosine 5’-triphosphate
ATP-γS    adenosine-5’-O- (3-thiotriphosphate)
CD        circular dichroism
CRINEPT   cross-correlated relaxation-enhanced polarization transfer
CRIPT     cross-correlated relaxation-induced polarization transfer
CSA       chemical shift anisotropy
cT        constant-time
DHNA      7,8-dihydroneopterin aldolase from *Staphylococcus aureus*
DSS       2,2-dimethyl-2-silapentane-5-sulfonic acid
DTT       1,4-dithiothreitol
EDTA      ethylenediaminetetraacetic acid
EM        electron microscopy
HMQC      heteronuclear multiple-quantum correlation
HSQC      heteronuclear single-quantum correlation
INEPT     insensitive nuclei enhanced by polarization transfer
METAFoR   metabolic flux ratio
NMR       nuclear magnetic resonance
NOE       nuclear Overhauser effect
NOESY     nuclear Overhauser effect spectroscopy
PFG       pulsed field gradient
RMSD      root mean square deviation
SDS-PAGE  sodium dodecyl sulfate polyacrylamide gel electrophoresis
SR1 single-ring variant 1 of GroEL
TROSY transverse relaxation-optimized spectroscopy
Introduction

1.1 NMR with large macromolecular structures

Nuclear magnetic resonance spectroscopy (NMR), one of the principal techniques in structural biology, has developed into the method of choice to determine three-dimensional atomic resolution structures of proteins and nucleic acids in solution and to investigate dynamic features and intermolecular interactions of macromolecules (Ferentz and Wagner, 2001; Kay, 1998; Wüthrich, 1986; Wüthrich, 1995; Zuiderweg, 2001). Current technologies, including isotope labeling (Goto and Kay, 2000; LeMaster, 1994; Lian and Middleton, 2001), triple resonance experiments (Bax and Grzesiek, 1993; Cavanagh et al., 1996; Clore and Gronenborn, 1991; Montelione and Wagner, 1989) and the more recent transverse relaxation-optimized spectroscopy (TROSY) technique (Pervushin et al., 1997; Pervushin, 2000), have enabled the characterization of three-dimensional structures up to 40 kDa (Clore and Gronenborn, 1998; Ferentz and Wagner, 2000; Kay, 2001; Kay and Gardner, 1997; Zuiderweg, 2002) and studies of proteins in
supramolecular structures of 100 kDa or larger (Riek et al., 2000; Wider and Wüthrich, 1999; Wüthrich, 1998). The large size of many protein complexes however still represents a challenge for NMR studies and the development of NMR techniques for investigation of larger molecular sizes is of considerable interest, for example for characterization of large proteins that cannot be crystallized, or of dynamics and molecular interactions in large complexes. Therefore, recent technical advances aim at enabling acquisition of solution NMR spectra of systems of several hundreds of kDa (Riek et al., 1999; Riek et al., 2000; Salzmann et al., 1998; Salzmann et al., 1999a–d; Salzmann et al., 2000; Wider and Wüthrich, 1999; Wüthrich, 1998).

With increasing molecular weight, a number of problems arise that limit the resolution and sensitivity of the NMR spectra: (i) faster transverse spin relaxation, (ii) sensitivity when only limited concentrations can be achieved, and (iii) signal overlap due to the higher number of resonances (Riek et al., 2000; Wider and Wüthrich, 1999).

The rate of transverse relaxation increases with molecular size and leads to line broadening and signal losses during the experiment. It thus limits the size of structures amenable to study by solution NMR. A direct answer to that problem is the use of the TROSY technique (Pervushin et al., 1997; Pervushin, 2000) that exploits constructive interference between the dipole–dipole coupling and the chemical shift anisotropy. It allows to partially suppress transverse relaxation during the evolution and acquisition periods of a NMR experiment, improving the resolution and/or the sensitivity (Figure 1.1, C and D). For molecular sizes above 200 kDa, for which transverse relaxation during polarization transfers becomes a limiting factor, application of CRINEPT (cross-correlated relaxation-enhanced polarization transfer) (Riek et al., 1999) yields a significant gain in sensitivity (Figure 1.1 B).
CRINEPT combines the conventional INEPT transfer (Morris and Freeman, 1979) with cross-correlated relaxation-induced polarization transfer (CRIPT), since cross-correlated relaxation becomes an efficient polarization transfer mechanism for very large molecules. Furthermore, in the CRINEPT scheme, TROSY is also active during the polarization transfer steps, yielding a fully transverse relaxation-optimized experiment (see Riek et al., 2000, Box 1 for a more complete description of these aspects) (Figure 1.1). Deuteration of the protein sample can also be used to reduce relaxation by suppressing additional relaxation sources such as dipole-dipole coupling to remote protons (Gardner and Kay, 1998 and references therein; Riek et al., 2000).

Sensitivity limitations occur when working with large molecular structures that are not soluble to typical concentration used for NMR studies (~1mM). The recent development of NMR cryo-probeheads already improved the sensitivity of NMR measurements with protein solution by a factor 2-3 (as measured in our laboratory. R. Horst, personal communication), so that it should nowadays be possible to work with lower concentrations.

As far as the spectral overlap is concerned, several isotope labeling schemes have been proposed that result in simplification of the spectra (Goto and Kay, 2000). Fractional deuteration (Gardner and Kay, 1998 and references therein; Lian and Middleton, 2001), amino acid selective labeling (Kainosho, 1997; Lian and Middleton, 2001; Muchmore et al., 1989), and the recent approach of segmental labeling (Cowburn and Muir, 2001; Otomo et al., 1999a, b; Xu et al., 1999; Yamazaki et al., 1998; Yu, 1999; see also single chain labeling in oligomers, Simplaceanu et al., 2000) provide strategies to reduce the spectral complexity.

In the present study, we chose to work with homo-oligomeric
1.1 NMR WITH LARGE MACROMOLECULAR STRUCTURES

A

Magnetization Transfer (b)  Evolution (c)  Magnetization Transfer  Acquisition (d)

\[ ^{1}\text{H} \rightarrow ^{15}\text{N} \]

\[ ^{15}\text{N} \rightarrow ^{1}\text{H} \]

B

\[ ^{1}\text{H} \rightarrow ^{15}\text{N} \]

CRINEPT

5 ms

C

\[ ^{1}\text{H} \]

decoupling

\[ ^{15}\text{N} \]

conv.

\[ ^{1}\text{H} \]

TROSY

0 – 100 ms

D

\[ ^{1}\text{H} \]

conv.

\[ ^{15}\text{N} \]

decoupling

\[ ^{15}\text{N} \]

TROSY

0 – 100 ms
proteins. In these proteins, one subunit defines the set of resonances observed in NMR and because of the equivalence of all subunits in the symmetrical assembly, one obtains NMR spectra of manageable complexity in spite of the large overall size of the structure studied. In addition, samples can be prepared at high molar concentration in subunit, so that the sensitivity is not limiting. Therefore, work with homo-oligomeric proteins allows us to specifically address the first limitation discussed above, the transverse spin relaxation. Starting with the established TROSY schemes and the CRINEPT principle that have been so far tested up to a size of 200 kDa, we aim at obtaining high-resolution two-dimensional (2D) solution NMR spectra of macromolecular structures in the size range of 200-900 kDa. These 2D spectra should provide the basis for studying intermolecular contacts, conformational changes and dynamics in large protein assemblies.

Figure 1.1 Schematic representation of a two-dimensional (2D) \([^{15}\text{N},^{1}\text{H}]\)-correlation experiment using conventional NMR, TROSY or CRINEPT-TROSY (from Riek et al., 2000).
A. Magnetization pathway in a 2D \([^{15}\text{N},^{1}\text{H}]\)-correlation experiment. The magnetization transfers are indicated with a dotted line and the decay of the signal due to relaxation during the frequency labeling period is drawn schematically.
B. First magnetization transfer element of the correlation experiment of A, as implemented in conventional NMR and in CRINEPT-TROSY. In conventional NMR (upper part), the magnetization is transferred via scalar spin-spin coupling in the INEPT element. In CRINEPT, cross-correlated relaxation-induced polarization transfer is used in addition to the INEPT transfer (see also chapter 3.5). Moreover, there is no decoupling pulse in the CRINEPT magnetization transfer period and TROSY is active, so that the resulting decay of the signal is slower.
C. and D. Frequency labeling periods on \(^{15}\text{N}\) and \(^{1}\text{H}\), respectively, as present in conventional NMR and in TROSY. Decoupling during the evolution and acquisition periods is applied in conventional NMR and causes collapse of the four multiplet components resulting from \(^{15}\text{N}-^{1}\text{H}\) scalar coupling into a single line. In contrast, no decoupling is applied during the TROSY evolution and acquisition periods. Of the four components that show different transverse relaxation rates, only the component with the most favorable relaxation properties is selected (not shown). The resulting signal decay is slower than in conventional NMR.
The system chosen for this work consists of the *Escherichia coli* chaperonins GroEL (14-mer, 800 kDa), GroES (7-mer, 72 kDa), and the single-ring variant of GroEL, SR1 (7-mer, 400 kDa) (see section 1.2 and references therein). Current knowledge on their structures, functions and molecular mechanisms is briefly reviewed in the next section.

1.2 The GroE chaperonins

*Molecular chaperones*

The successful folding of newly translated polypeptide chains into their correct three-dimensional conformation is essential to the viability of the cell. From Anfinsen’s pioneering work (Anfinsen, 1973), we know that polypeptide chains *in vitro* can spontaneously acquire their properly folded state from the sole information of their primary structure. *In vivo*, however, molecular crowding in the cell results in very unfavorable conditions for protein folding, that favour aggregation. The processes of folding, oligomeric assembly as well as transport and degradation of proteins thus require participation of helper proteins termed molecular chaperones.

Molecular chaperones are defined as “a class of unrelated families of proteins that have in common the ability to assist the non-covalent assembly of other protein-containing structures *in vivo*, but which are not permanent components of these structures when they are performing their normal biological functions” (Ellis, 1996) or, as an alternative phrasing, a chaperone is a protein “that binds to and stabilizes an otherwise unstable conformer of another protein, and by controlled binding and release of the substrate protein facilitates its correct fate in vivo, be it folding, oligomeric assembly, transport to another subcellular compartment or controlled switching between
active and inactive conformations” (Hendrick and Hartl, 1993; Ellis, 1996). Their common features are the interaction with non-native polypeptides, the stabilization of intermediate forms and the prevention of aggregation. These ubiquitous proteins have been shown to play essential roles in catalysis of correct protein folding and of assembly of oligomeric proteins, recovery from cellular stress, translocation of protein across membranes and in protein degradation (Fink and Goto, 1997; Horwich et al., 1999; Saibil, 2000; Tomoyasu et al., 2001).

**The chaperonins**

The chaperonins are a family of sequence-related molecular chaperones involved in ATP-dependent promotion of protein folding in all types of cells (Ellis, 1996; Hemmingsen et al., 1988). They are essential for the cell viability at all temperatures. There are two distinct classes of chaperonins: the Group I or GroE chaperonins found in eubacteria, mitochondria and chloroplasts, and the Group II or TCP1 chaperonins found in archaeabacteria and in the eukaryotic cytosol. They have in common an oligomeric assembly in two rings of seven to nine subunits of about 60 kDa defining a central cavity, within which the protein substrate binds, and the catalysis of ATP hydrolysis to ADP (Steinbacher and Ditzel, 2001).

These proteins serve at least three functions (Ellis, 1996; Grallert and Buchner, 2001): (i) They prevent aggregation of non-native polypeptides by binding of the partially folded intermediates and reducing the concentration of these aggregation-prone species. (ii) They release bound substrates in a cavity, allowing folding in a protected environment without intermolecular interactions, and (iii) they bind and possibly unfold polypeptides trapped in a non-productive conformation, the binding and release process giving them a new chance to fold correctly. *In vivo*, the chaperonins protect proteins from
aggregation after denaturation by cellular stress and indeed many of them are induced by stresses such as high temperature.

**The GroEL-GroES system**

The *Escherichia coli* GroEL-GroES chaperonin system is one of the best studied molecular chaperone system. GroEL (cpn60) is a homooligomer of fourteen 57 kDa subunits (Hemmingsen *et al.*, 1988). The crystal structure of unliganded GroEL at 2.8 Å resolution (Braig *et al.*, 1994, 1995) as well as many electron microscopy studies (Chen *et al.*, 1994; Roseman *et al.*, 1996) revealed that the subunits are arranged in two rings stacked back-to-back with nearly perfect seven-fold symmetry (Figure 1.2). The cylindrical assembly is 135 Å in diameter and 145 Å in height. Its central channel forms two cavities, one in each ring, that are separated by the confluence of the C-terminal disordered tail of each GroEL subunit.

The GroEL subunit (547 amino acids) folds into three distinct domains (Figure 1.3): The well-defined equatorial domain (residues 1-133 and 409-547; ”E” in Figure 1.3) forms most of the inter- and intra-ring contacts and contains the ATPase active site. The more mobile apical domain (residues 191-376; ”A” in Figure 1.3) is located at the ends of the cylinder and is responsible for binding of the protein substrate and of the co-chaperonin GroES. Finally, an intermediate domain (residues 134-190 and 377-408; ”I” in Figure 1.3) connects the equatorial and the apical domains.

GroES is the co-chaperonin that interacts with GroEL in the process of assisting protein folding (Chandrashekar *et al.*, 1986; Tilly *et al.*, 1981). It is a dome-shaped heptameric ring of 10 kDa subunits (97 amino acids), about 75 Å in diameter and 30 Å high (Hunt *et al.*, 1996). The GroES subunit core structure is an irregular β-barrel from which two loops are
protruding. A β-hairpin arches upwards and inwards from the ring and forms the top of the dome, while the other loop (Glu 16-Ala 32) extends from the bottom outer rim of GroES (Figure 1.3). Earlier proton NMR studies indicated that this second loop is mobile and unstructured in free GroES and becomes immobilized upon binding to GroEL (Landry et al., 1993). GroES binds at one end of the GroEL cylinder in the presence of nucleotides and forms a cap over the GroEL cavity (Ishii et al., FEBS Letters, 1992; Langer et al., 1992; Mayhew et al., 1996; Roseman et al., 1996; Saibil et al., 1991; Schmidt et al., 1994; Weissman et al., 1996).

These detailed structural analyses together with numerous remarkable biochemical experiments have provided a gradual better understanding of the function and mechanism of the GroEL-GroES system in vitro (for reviews, see Fenton and Horwich, 1997; Grallert and Buchner, 2001; Ranson et al., 1998; Sigler et al., 1998; Thirumalai and Lorimer, 2001; Xu and Sigler, 1998). Polypeptides are first bound to hydrophobic sites on the large oligomeric structure of GroEL, then trapped inside a chamber as the chaperonin changes conformation and binds ATP and GroES, and finally allowed to fold in the protected environment of the cavity before being released (Figure 1.4). This GroE functional cycle in vitro is described in more details below. In vivo, there is evidence for assistance of GroEL in the folding of newly synthesized polypeptide chains as well as in the maintenance of preexisting proteins both under normal and heat-stress conditions (Ewalt et al., 1997; Houry et al., 1999; Houry, 2001).

The GroE cycle

I. Polypeptide binding and conformational states of bound substrates

GroEL interacts with a wide variety of unfolded and partially folded proteins. About 50% of the *E. coli* soluble proteins form complexes with
Figure 1.2 3D reconstruction of GroEL-nucleotide and GroEL-GroES-nucleotide complexes at 30 Å resolution from electron cryo-microscopy. (Chen et al., 1994; Roseman et al., 1996, Ranson et al., 2001; Roseman et al., 2001. Picture from the Birbeck College Chaperone Group homepage http://www.cryst.bbk.ac.uk/~ubcg16z/chaperone.html). A. Crystal structure of GroEL (Braig et al., 1995) converted to electron density and filtered to 30 Å resolution. B. From left to right: Electron cryo-microscopy reconstructions of unliganded GroEL, GroEL-ADP, GroEL-AMP-PNP, and GroEL-ATP complexes. C. From left to right: Electron cryo-microscopy reconstructions of GroEL-GroES-ADP, GroEL-GroES-AMP-PNP, GroEL-GroES-ATP asymmetrical (bullet-like) complexes and GroES-GroEL-GroES-AMP-PNP symmetrical (football-like) complexes. Binding of nucleotides results in expansion of the GroEL structure and induces twisting of the apical domains to different extents. The bullet-shaped complexes also exhibit different apical domain conformations. Detailed analysis of these images suggested a variety of ligand-dependent apical domain twist, hinge rotation and interring contacts changes.
GroEL in vitro when denatured (Viitanen et al., 1992). On the other hand, in vivo substrates of GroEL have been suggested to represent 5-30% of the cytoplasmic E. coli proteins under normal growth conditions, preferentially of the size range of 20-60 kDa (Ellis and Hartl, 1996; Ewalt et al., 1997; Horwich et al., 1993; Houry et al., 1999; Lorimer, 1996). Remarkably, these proteins known to interact with GroEL show no homology, which excludes a sequence-based recognition motif (Houry et al., 1999). Moreover, GroEL has only low affinity for native proteins. The chaperonin is thus thought to recognize exposed hydrophobic surfaces of a broad range of non-native protein conformations and bind them in its central cavity. Indeed, hydrophobic interactions constitute the major contribution to GroEL-substrate binding (Chen and Sigler, 1999; Fenton et al., 1994; Lin et al., 1995; Richarme and Kohiyama, 1994; Wang et al., 1999a), although electrostatic interactions may also contribute (Perrett et al., 1997).

Electron microscopy (Braig et al., 1993; Chen et al., 1994; Falke et al., 2001; Ishii et al., 1992; Langer et al., 1992), site-directed mutagenesis (Fenton et al., 1994) and X-ray crystallography (Chatellier et al., 1999; Chen and Sigler, 1999) have located the substrate binding site at the inner top rim of the apical domain. This region shows inherent flexibility in the absence of substrate or GroES, as indicated by small differences in the various structures determined for the GroEL apical domain and by the high temperature factors (Boisvert et al., 1996; Braig et al., 1994; Braig et al., 1995; Chen and Sigler, 1999; Sigler et al., 1998; Zahn et al., 1996a). This accounts for the necessary plasticity of the binding site to recognize a broad spectrum of different proteins. So far, GroEL-substrate complexes have been studied by electron microscopy (Chen et al., 1994; Falke et al., 2001), but more detailed atomic resolution information has only revealed the interactions of GroEL with peptides (Buckle et al., 1997; Chen and Sigler, 1999; Feltham and Gierasch, 2000;
Kobayashi et al., 1999; Landry et al., 1991, 1992; Tanaka and Fersht, 1999; Wang et al., 1999). These peptides bind in the groove between two parallel α-helices in the site that is to become occupied by the loops of GroES. The GroEL complexes with larger proteins have not been crystallized to date, possibly because the bound substrates occur as ensembles of largely unstructured non-native states (Feltham and Giersasch, 2000). Recent work from the Horwich laboratory demonstrated that some substrates bind to a minimum of three to four adjacent apical domain sites (Farr et al., 2000).

The conformations assumed by bound substrates however remain to be elucidated. A number of studies have aimed at characterizing the conformations recognized by GroEL and the conformations of the polypeptides while stably bound to GroEL (for reviews, see Bukau and Horwich, 1998; Fenton and Horwich, 1997). Unfolding or refolding of a substrate in the presence of GroEL enabled identification of the structures or conditions for interaction with GroEL (Gervasoni et al., 1996, 1998; Katsumata et al., 1996; Martin et al., 1991; Zahn and Plückthun, 1994). Binding of preformed metastable intermediates of the substrate proteins to GroEL was also investigated (Hayer-Hartl et al., 1994; Lillie and Buchner, 1995). Finally, conformations of the polypeptides bound to GroEL have been characterized directly using methods such as tryptophan fluorescence, ANS binding, protease resistance or H/D exchange experiments (Chen et al., 2001; Gervasoni et al., 1996; Goldberg et al., 1997; Hayer-Hartl et al., 1994; Martin et al., 1991; Mendoza et al., 1992; Robinson et al., 1994; Weissman et al., 1994; Zahn et al., 1994). These studies showed that the substrate polypeptide in the complexes are susceptible to proteolytic digestion and that tryptophan side-chains lie in a environment of intermediate polarity between that of the unfolded state and native state. The polypeptide-GroEL complexes also possess exposed hydrophobic surfaces. It follows that
the bound substrates are found in a collapsed but loosely packed conformation. The extent of residual secondary and tertiary structure in the complexes was probed in H/D exchange experiments in several studies. The degree of protection that was observed ranges from very low or inexistent, suggesting destabilized secondary structure or transient unfolding (Robinson et al., 1996; Zahn et al., 1994; Zahn et al., 1996c) to higher degrees indicating molten globule-like structure, a core of partially protected secondary structure, or even native-like topology (Chen et al., 2001; Gervasoni et al., 1996; Goldberg et al., 1997; Gross et al., 1996). Different proteins thus appear to form transient and stable complexes with GroEL in different states. It was suggested that the determining factor for the interaction is the exposure of hydrophobic patches and that preservation of the native structure in the bound substrate is possible as long as sufficient hydrophobic surfaces are present (Gervasoni et al., 1996). Note that in these experiments, protection arising from residual structure cannot be distinguished from protection due to direct interactions with GroEL and spatial resolution of the exchange data is required for identification of protected core structure. This is also a problem when investigating whether the substrate is further unfolded upon release in the GroEL cavity (Shtilerman et al., 1999).

II. Encapsulation (nucleotides and GroES binding)

ATP binding to one GroEL ring is diffusion controlled and highly cooperative (Gray and Fersht, 1991; Jackson et al., 1993). It induces a concerted allosteric transition in this ring from a T state, with high substrate affinity and low ATP affinity, to a R state with a reduced affinity for substrate proteins and a greater affinity for ATP (Horovitz et al., 2001; Inbar and Horovitz, 1997; Ma et al., 2000; Makio et al., 2001). Conversely, there is negative cooperativity for ATP binding between the rings, so that the transition of the double-ring structure from TR to RR
is unfavorable and requires high ATP concentrations (Burston et al., 1995; Horovitz et al., 2001; Yifrach and Horovitz, 1995, 1996). Cryo-electron microscopy studies showed significant conformational changes in the GroEL-ATP\textsubscript{7} complexes when compared to unliganded GroEL: the apical domains of one ring twist with respect to the equatorial domain and move upwards (Figure 1.2) (Chen et al., 1994; Roseman et al., 1996; Roseman et al., 2001; White et al., 1997). Surprisingly, these movements were not observed in the crystal structure of GroEL-ATP\textsubscript{γS}\textsubscript{14} (Boisvert et al., 1996), which resembles that of unliganded GroEL, possibly due to prevailing of the lattice forces over movements of the flexible apical domains or to imperfect mimicry of ATP by ATP\textsubscript{γS} (Roseman et al., 2001; Sigler et al., 1998).

In the GroE cycle, the conformational change triggered by ATP binding to one ring is followed by the fast association of GroES (Burston et al., 1995). GroES binding induces a more extensive movement of the apical domains in the \textit{cis} ring (Xu et al., 1997; Figure 1.3) and leads to encapsulation of a bound substrate. The resulting GroEL-GroES-ATP\textsubscript{7} complex is the folding active species (Weissman et al., 1995, 1996).

The asymmetrical GroEL-GroES-ADP\textsubscript{7} complex structure (Xu et al., 1997)

The crystal structure of the GroEL-GroES-ADP\textsubscript{7} ternary complex provided atomic resolution information on the allostERIC movements of GroEL and large conformational changes occurring upon GroES binding. As depicted in Figure 1.3, GroES caps one end of GroEL and locks its \textit{cis} ring in a new conformation where the apical domains have moved upwards by 60° and completed a 90° twist with respect to the equatorial domain (Note: \textit{cis} refers to the location of the ligands or rings with respect to GroES). The hydrophobic binding sites of the \textit{cis} ring are rotated into subunit interfaces and contacts with GroES, and the cavity is now lined with a hydrophilic surface. These rearrangements result in
the doubling of the volume of the GroEL cavity and would cause release of a bound substrate into the now hydrophilic environment of the cavity, giving it a chance to fold. The second considerable change in the structure is the rotation and downward movement of the intermediate domain to close the nucleotide binding site and bring the active aspartate residue in proper position for catalysis of ATP hydrolysis. GroES on the other hand retains its dome-like structure upon binding and its mobile loop extends downwards to form contacts with hydrophobic residues of GroEL’s apical domain. The loop that was unstructured and highly flexible in free GroES adopts a β-hairpin conformation in contact with GroEL (Landry et al., 1993; Shewmaker et al., 2001; Xu et al., 1997). Interestingly, as for substrate binding, not all seven binding sites on GroEL are necessary for GroES and as few as two interaction sites are enough to form a relatively stable complex (Farr et al., 2000).

III. ATP hydrolysis and polypeptide folding

Because of the concerted movements that occurred in the cis ring upon GroES binding, seven ATP molecules are locked in the active site of the complex and committed to hydrolysis. Therefore, ATP hydrolysis is “quantized” in the presence of GroES (Burston et al., 1995; Kad et al., 1998; Todd et al., 1994). Completion of one ATPase cycle takes approximately 15 s at room temperature and ATP hydrolysis is thought to act as a timer controlling the lifetime of the encapsulated substrate state (Burston et al., 1995; Hayer-Hartl et al., 1995; Rye et al., 1999; Todd et al., 1994). ATP hydrolysis is K⁺-dependent and results in the formation of the asymmetrical GroEL-GroES-ADP₇ complex in which negative cooperativity between rings for ATP and substrate binding to the trans ring is reduced (Rye et al., 1999).
**IV. GroES, ADP and polypeptide release**

The release of all three ligands GroES, ADP and the substrate protein is triggered by binding of ATP to the *trans* ring and accelerated by the
presence of another substrate molecule on this ring (Rye et al., 1997; Rye et al., 1999; Weissman et al., 1994). GroES and ADP are released, the substrate protein set free, folded or not (Burston et al., 1996), and the hydrophobic surface of the central GroEL cavity is restored. So the double ring structure serves to transmit signals for the alternating binding and release of substrates between the rings. Single-ring GroEL mutants are indeed arrested in the GroEL-GroES-ADP$_7$ state since they lack the trans ring (Horwich et al., 1998; Weissman et al., 1995, 1996).

It was demonstrated that the substrate protein is released with each cycle of ATP hydrolysis (Brinker et al., 2001; Sigler et al., 1998; Todd et al., 1994; Weissman et al., 1994) and that it can undergo several rounds of binding and release to reach its native state (Chaudhuri et al., 2001; Ranson et al., 1997; Rye et al., 1997; Weissman et al., 1994).

Prospects and open questions

A wealth of genetical, biochemical and structural studies have advanced our understanding of the structures and molecular details involved in the GroE chaperone cycle. A number of issues however remain to be solved (see also discussion in Grantcharova et al., 2001; Sigler et al., 1998; Weissman, 2001): First, much more is known on the conformations and cycle events for the chaperonin system itself than on the fate of protein substrates during cycles of assisted folding. The conformations of bound substrates and their interactions with GroEL have not yet been elucidated in details. There is evidence for at least partial unfolding of the polypeptides associated with binding to GroEL, that could occur either through catalyzed unfolding (Zahn et al., 1996b,c) or through thermodynamic partitioning if GroEL preferentially binds less folded states (Grantcharova et al., 2001, Walter et al., 1996; Zahn and Plückthun, 1994), but different substrates show different degrees of unfolding (see before; see also Fenton and Horwich, 1997;
1.2 The GroE chaperonins

Similarly, the nature of the events associated with ATP and GroES binding during the first second of encapsulation is still a major question, that a recent study however started to address (Shtilterman et al., 1999).

Finally, the principle of the chaperonin action on its substrates is also a matter of debate. Two competing models have been proposed (Brinker et al., 2001; Wang and Weissman, 1999): The “Anfinsen cage” model explains GroEL-mediated folding assistance by prevention of off-pathway aggregation through binding and encapsulation of the proteins in the cage-like structure of the GroEL-GroES complex (Brinker

---

**Figure 1.4** Schematic representation of the GroEL-GroES reaction cycle (reproduce with permission from Grantcharova et al., 2001).

The GroEL rings are shown in purple, and GroES in light blue; the nucleotides are depicted with green boxes and the substrate protein with a dark blue line. Non-native polypeptide binds in the open trans ring of a pre-existing GroEL-GroES-ADP complex. In the presence of ATP (T), GroES binding to the same ring induces large conformational changes in the GroEL cis ring and triggers release of the polypeptide in the cavity, where it can fold. After ~10 s, ATP hydrolysis (T→D) primes the complex for ADP and polypeptide release. It is then ATP binding to the trans ring that triggers the discharge of the cis GroES and substrate, whether folded or not. This release is accelerated by the binding of non-native polypeptide. Renewed GroES binding to the other ring forms the next folding-active cis complex and starts a new cycle.

Grantcharova et al., 2001; Horovitz, 1998 and references therein).
et al., 2001; Ellis, 1994). On the other hand, the “iterative annealing” model suggests that kinetically trapped folding intermediates can be unfolded by GroEL and are given a chance to refold upon release from the chaperonin, in solution or into the cavity (Betancourt et al., 1999; Thirumalai and Lorimer, 2001; Wang and Weissman, 1999). Recently, Brinker et al. demonstrated that folding assistance by GroEL and folding rate enhancement as observed for some substrates require confinement of the unfolded protein in the GroEL-GroES cavity, suggesting that enclosure in the cage could alter the folding energy landscape (Brinker et al., 2001). In another recent study, Chaudhuri et al. describe GroEL-mediated folding of a substrate too large to be encapsulated that proceeds through a trans ternary complex in a new type of cycle (Chaudhuri et al., 2001). These results evoke that the interactions and mechanisms involved in GroEL-mediated folding could vary with different types of substrates. Clearly, other experiments along these lines will be needed to further improve the understanding of the fate of substrates in the GroE cycle.

Among other aspects of the GroE cycle that are not yet clearly understood, the magnitude of the domain movements induced by nucleotide binding has recently been addressed (Chen et al., 1994; Roseman et al., 1996; Roseman et al., 2001; White et al., 1997). Understanding the different effects of ATP and ADP (Rye et al., 1997; Todd et al., 1995) and the details of the mechanisms that trigger opening and closing of the GroEL cavity requires direct visualization of the ATP complex (Roseman et al., 2001) and probably atomic resolution studies of these assemblies. Another question to answer is what role the symmetric GroES-GroEL-GroES complexes observed by electron microscopy (Azem et al., 1994; Schmidt et al., 1994) have in the GroE cycle and their possible involvement in an associative mechanism for switching from one folding active complex to another remains to be
investigated (Azem et al., 1995; Diamant et al., 1995; Hayer-Hartl et al., 1995; Horowitz et al., 1999; Llorca et al., 1996; Török et al., 1996). Finally, the question of the mechanism of the GroE chaperone cycle under in vivo conditions should be addressed and set against the in vitro results.
1.3 Aims and scope of this thesis project

Based on the developments of solution NMR and aspects described in chapters 1.1 and 1.2, the goals of this work are as follows. Starting from the established TROSY technique (Pervushin et al., 1997; Pervushin, 2000) and applications with the 110 kDa DHNA (Salzmann et al., 1998; Salzmann et al., 1999a,b; Salzmann et al., 2000), and using the principle of CRINEPT (Riek et al. 1999), we aim at collecting two-dimensional correlation spectra of molecular sizes between 200 and 900 kDa and establishing refined implementations of the experimental schemes using homo-oligomeric proteins. The proteins chosen for these experiments are the *E. coli* chaperonin GroEL (800kDa), its co-chaperonin GroES (72 kDa) and a single-ring variant of GroEL, SR1 (400 kDa) which properties have been described in details in section 1.2. Once the techniques established and characterized, they should be applied to study specific aspects of the chaperonin system, such as formation of complexes of GroEL with GroES, nucleotides or polypeptide substrates, in a collaboration with Dr. E. B. Bertelsen and Prof. Dr. A. L. Horwich at Yale University.

This thesis is thus organized in several sections. As a first step, protocols must be established for production of the chaperonins in various isotope-labeling combinations. Chapter 2 thus describes the protocols used for the expression, labeling, purification and characterization of the different samples of GroEL and GroES. Then, the NMR experimental schemes implementing TROSY and CRINEPT will be described and compared in chapter 3. Chapter 3 provides a guideline for 2D measurements with macromolecular structures in the size range of 150-900 kDa including optimal selection of the relaxation delay, the polarization transfer, and the frequency labeling periods. In chapter 4, the backbone assignment of the 72 kDa co-chaperonin GroES is described, that will serve as a tool to test the potential of TROSY/
CRINEPT-based NMR experiments with structures of sizes up to 900 kDa (chapters 3 and 5) and to study GroES in the GroEL-GroES complex (chapter 5). Further analysis of the NMR data provided the secondary structure and overall fold of GroES. Chapter 5 reports results of the study of the conformations and dynamics of GroEL-bound GroES, using the techniques and tools described in the preceding chapters. It is thus the central part of this work. In chapter 6, these experiments are then applied to study GroEL and its single-ring variant SR1, and their complexes with GroES and nucleotides. Finally, chapter 7 summarizes the results achieved in this work, and discusses the problems encountered, open questions and directions for future work on NMR studies of large supramolecular systems.

In Appendix 1, the results of a separate research project involving metabolic studies using $^{13}$C-labeling and two-dimensional NMR are exposed (Szyperski, 1997; Szyperski, 1998; Szyperski et al., 1999), in form of a short description of the method applied, and a collection of the published work.
Chapter 2

Preparation of GroE proteins for NMR measurements

2.1 Requirements for solution NMR studies of large molecules

In order to perform studies of large macromolecular systems by solution NMR, a number of requirements must first be met at the level of sample preparation.

First, acquisition of high quality spectra with good sensitivity and spectral resolution rely on the availability of a high-concentration homogeneous sample. High purity and homogeneity of the sample are needed to ensure clean spectra with narrow lineshapes. Typically, solution NMR is applied at a minimum concentration of 1 mM in the compound to study. With the development of the cryoprobe techniques, NMR spectra can nowadays be obtained within reasonable measurement time for solution concentration as low as 300 μM. This still
represents 10-15 mg of a 100 kDa protein for one NMR sample. It follows that availability of a high-yield expression system capable of producing tens of mg of the proteins or compounds of interest is a prerequisite for such studies. Furthermore, this material must be soluble enough to reach the high concentration required. For the example of a 100 kDa protein, concentrations of the order of 100 mg/ml must be achieved for work with a normal NMR probehead, or 30 mg/ml when a cryoprobe is available. Note also that the increased viscosity in these solutions can make manipulation of the sample more difficult and possibly increase the molecular tumbling time which affects NMR properties of the sample.

The second critical aspect is the production of stable-isotope labeled samples to allow use of heteronuclear experimental schemes. Thus, the organism bearing the expression system must show reasonable growth and expression on isotope-labeled (mostly with $^{15}$N and $^{13}$C) minimal media or on synthetic isotope-labeled rich media (Martek Biosciences Corp., Columbia, MD USA; Cambridge Isotope Lab, Andover, MA USA; Silantes, München, Germany). For molecules above 30 kDa, uniform $^2$H-labeling (deuteration) is further necessary to reduce relaxation by dipole-dipole coupling of the amide protons to remote protons (i.e. all other protons) that leads to line broadening and signal loss. Deuteration implies most often growth of the expression system host on D$_2$O-based\footnote{D$_2$O designates $^2$H$_2$O.} media, which leads to a reduction of the growth rate and lower biomass yields. Expression on the particular media required for isotope labeling is not always possible, in particular for complex eukaryotic systems, and several methods have been described to overcome expression difficulties in D$_2$O and achieve at least partial isotope enrichment with a good yield (Cai \textit{et al.}, 1998; Lian and
Finally, NMR experiments of several days length will be recorded with the compound to study. The sample must therefore be stable towards aggregation and degradation over the time course of such experiments.

2.2 Stable isotope labeling of the GroE chaperonins

Isotope labeling of macromolecules for NMR studies aims at increasing the sensitivity and resolution of the experiments and simplifying spectra as well as narrowing linewidths. $^{13}$C and $^{15}$N incorporation into nucleic acids and recombinant proteins enabled the routine use of heteronuclear multi-dimensional NMR experiments for assignments of the resonances. As larger systems are investigated using solution NMR, amino acid selective labeling is more and more often applied to reduce the spectral complexity and help the assignment procedure. Deuteration of the samples also becomes necessary at large molecular weights, and most often requires growth of the host strain for expression in D$_2$O. This section describes the various isotope labeling schemes applied for NMR study of the large-sized chaperonins.

**Uniform $^{15}$N,$^{2}$H- and $^{13}$C,$^{15}$N,$^{2}$H-labeling**

Expression of the uniformly-labeled chaperonins was achieved in *E. coli* BL21(DE3) bacteria growing on D$_2$O-based minimal medium supplemented with isotopically-labeled nutrients (Box 2.1). We found that stepwise adaption of the cells to increasing levels of D$_2$O was not necessary to achieve efficient growth and expression in D$_2$O in this system. However, care was taken that only one variable was changed at
### Box 2.1 Composition of the media for isotope labeling

#### A. M9 minimal medium

- 880 ml water (H₂O or D₂O)
- 100 ml M9 salt solution 10X
- Autoclave in a flask
- 2 ml 1M MgSO₄, autoclaved separately
- 1 g NH₄Cl, dissolved in 5 ml and filter sterilized
- 10 ml Vitamin mix
- 20 ml 20% Glucose or Acetate, autoclaved
  - *add in a sterile way*
  - *check pH*
  - add antibiotics (Ex: 100 µg/ml Ampicillin)

**Note:** complementation of 1 l of M9 medium with 2 ml of solution Q can increase growth and protein expression efficiency.

#### B. Celtone-spiked medium

- 870 ml water (H₂O or D₂O)
- 100 ml M9 salt solution 10X
- Autoclave in a flask
- 2 ml 1M MgSO₄, autoclaved separately
- 1 g NH₄Cl, dissolved in 5 ml and filter sterilized
- 10 ml Vitamin mix
- 10 ml 20% Glucose or Acetate, autoclaved
- 1-2g Celtone powder (Martek) dissolved in 10 ml and filter sterilized
  - *add in a sterile way, check pH*
  - add antibiotics for specific amino acid labeling, add the labeled amino acid at induction (Ex: 75 mg¹⁵N-Leu /g celtone)

---

#### M9 salts solution 10X

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na₂HPO₄·2H₂O</td>
<td>85.1 g/l</td>
</tr>
<tr>
<td>KH₂PO₄</td>
<td>30 g/l</td>
</tr>
<tr>
<td>NaCl</td>
<td>5 g/l</td>
</tr>
</tbody>
</table>

#### Solution Q

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCl 5M</td>
<td>8 ml/l</td>
</tr>
<tr>
<td>FeCl₂·4H₂O</td>
<td>5 g/l</td>
</tr>
<tr>
<td>CaCl₂·2H₂O</td>
<td>184 mg/l</td>
</tr>
<tr>
<td>H₃BO₃</td>
<td>64 mg/l</td>
</tr>
<tr>
<td>CoCl₂·6H₂O</td>
<td>18 mg/l</td>
</tr>
<tr>
<td>CuCl₂·2H₂O</td>
<td>4 mg/l</td>
</tr>
<tr>
<td>ZnCl₂</td>
<td>340 mg/l</td>
</tr>
<tr>
<td>Na₂MoO₄·2H₂O</td>
<td>605 mg/l</td>
</tr>
<tr>
<td>MnCl₂·4H₂O</td>
<td>40 mg/l</td>
</tr>
</tbody>
</table>

*Autoclave to sterilise*

---

#### Vitamin mix

- Thiamine: 50 mg
- d-biotin (stored at 0-4°C): 10 mg
- Choline chloride: 10 mg
- Folic acid: 10 mg
- Niacinamide: 10 mg
- D-pantothenic acid: 10 mg (stored at 0-4°C)
- Pyridoxal (stored at -20°C): 10 mg
- Riboflavin: 1 mg

*Dissolve in 100 ml H₂O or D₂O*

*Filter sterilise*

---

¹ Based on R. Zahn, personal communication and Sambrook et al., 1989.
each of the subsequent growth steps, such as H₂O to D₂O or glucose to acetate, to ensure steady growth (Gardner and Kay, 1998). Prewarming of the medium before inoculation and good aeration of the cultures are also essential to avoid long lag phases and achieve efficient growth (Box 2.2).

The carbon source used in the D₂O-based medium was chosen according to the protein deuteration level required. Selecting protonated glucose as the sole source of carbon in >98% D₂O minimal medium yielded a deuteration level of approximately 85% (Table 2.1, line 1). The advantages of this medium composition are the growth efficiency and low cost. In counterpart, residual protonation occurs in the side chains due to direct incorporation of intact glucose fragments into the amino acids along their biochemical synthesis pathways (Neidhardt et al., 1996; Voet and Voet, 1995; See also Szyperski, 1995 and Appendix 1). Growth on protonated acetate as an alternative should result in higher and random deuterium incorporation, but provides lower final cell densities and yield.

For complete deuteration of all non-exchangeable sites, perdeuterated carbon sources are necessary, such as [²H₇]-D-glucose, [²H₈]-glycerol or [²H₄]-acetate. We used [²H₄]-acetate as the sole carbon source in a D₂O-based minimal medium supplemented with ¹⁵NH₄Cl to obtain ¹⁵N-labeled chaperonins proteins with a deuteration level as high as 97% (Table 2.1, line 2; Box 2.1 A; Box 2.2). The low cost of deuterated acetate compared to deuterated glucose (6 $/g vs. 370 $/g, respectively, Martek Biosciences Corp., Columbia, MD, USA.) compensates for the lower yield of the cultivations. Note however that not all host strain can grow and express protein on acetate and that use of deuterated glucose or of commercially available rich media based on labeled algal hydrolysate may be necessary (Table 2.1, line 3). In
2.2 Stable Isotope Labeling of the GroE Chaperonins

Box 2.2 Protocol for protein expression in D₂O

Cultivation in minimal medium / acetate / D₂O:

- Inoculate a LB preculture from one single colony picked from a Petri dish of freshly transformed E. coli cells. Incubate the preculture at 37°C for 4-5 h until the cells reach their exponential growth phase.

- Inoculate 20 ml of prewarmed minimal medium / glucose / H₂O 1:50 from the LB preculture and incubate it for approx. 8-10 h at 37°C.

- When the cultivation reaches the exponential growth phase with an OD above 1, use it to inoculate a similar quantity of prewarmed minimal medium / acetate / H₂O by 1:50 dilution and incubate again.

- Use the acetate medium preculture in turn for inoculation of 50 ml of prewarmed minimal medium / deuterated acetate / D₂O supplemented with $^{15}$NH₄Cl.

- After sufficient bacterial growth in this culture (OD >0.3, after ~24 h), take a 5 ml aliquot and induce it with IPTG. Check expression of the desired protein after 2-3 h by SDS-PAGE gel electrophoresis.

- If the expression is normal, use the rest of the deuterated preculture to inoculate 1:30 to 1:50 the prewarmed labeled main cultures.

- Induce protein expression in the exponential growth phase (OD 0.4-0.6 in the case of GroEL or GroES expression in BL21(DE3) cells).

- Add fresh antibiotics to the deuterated cultures after 24 h.

Typically, the final deuterated cultivation take 36-48 h and the whole procedure ~4 days. At all steps, the volume of the medium does not exceed 1/10 to 1/8 of the volume of the flask to ensure proper aeration through agitation.
Cultivation in minimal medium / glucose / D$_2$O:

- Inoculate a LB preculture from one single colony picked from a Petri dish of freshly transformed E. coli cells. Incubate the preculture at 37°C for 4-5 h until the cells reach their exponential growth phase.

- Inoculate a 20 ml preculture in prewarmed minimal medium with glucose as carbon source 1:50 from the LB preculture and incubate it for approx. 8-10 h at 37°C.

- When the cultivation reaches the exponential growth phase with an OD above 1, use it to inoculate 50 ml of prewarmed D$_2$O-based minimal medium supplemented with $^{15}$NH$_4$Cl and incubate for 5-8 hours until the cells reach the exponential phase. Growth in D$_2$O is slowed down by a factor ~2 compared to H$_2$O-based medium.

- Take a 5 ml aliquot and induce it with IPTG. Check expression of the desired protein after 2-3 h by SDS-PAGE gel electrophoresis.

- If the expression is normal, use the rest of the deuterated preculture to inoculate 1:30 to 1:50 the prewarmed labeled main cultures.

- Induce protein expression in the exponential growth phase (OD 0.5-0.7 in the case of GroEL or GroES expression in BL21(DE3) cells).

---

Auto clave and dry all flasks as well as possible. To reduce the residual H$_2$O amount for perdeuteration, the flasks can be rinsed with sterile D$_2$O before use and the salts desiccated.

Prepare all solutions in D$_2$O. Do not autoclave them, but filter sterilize to avoid reintroducing water.

Adaptation to D$_2$O in steps of increasing D$_2$O content was reported to yield slightly higher final cell densities (Hochuli, 1999), but was here found to increase the risk of expression loss in the many steps. Use of only one adaptation step as described here results in efficient cell growth and protein expression for the system studied.

Growth in D$_2$O is slow and affects respiratory activity of the cells (Hochuli, 1999; Hochuli et al., 2000), so make sure that the cultivations are at least not O$_2$-limited (fill the flasks with only one 10th of their total volume). When inoculating a medium on acetate, the lag phase can be rather long (up to 1-2 days in D$_2$O and deuterated acetate). Prewarming the medium and ensuring a good oxygenation of the cultivations is critical to avoid long lag phases. Slow growth and long lag phase can cause problem of loss of expression. Make sure that you check expression before starting the main cultivation.

<table>
<thead>
<tr>
<th>Medium</th>
<th>Growth rate (h$^{-1}$)</th>
<th>Generation time (h)</th>
<th>Max OD$_{600}$ nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>M9 4g/l Glu, H$_2$O</td>
<td>0.6</td>
<td>1.2</td>
<td>4</td>
</tr>
<tr>
<td>M9 2g/l Glu, H$_2$O</td>
<td>0.6</td>
<td>1.2</td>
<td>2</td>
</tr>
<tr>
<td>M9 4g/l Acetate, H$_2$O</td>
<td>0.4</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td>M9 4g/l Glu, D$_2$O</td>
<td>0.35</td>
<td>1.9</td>
<td>2-2.5</td>
</tr>
<tr>
<td>M9 4g/l $[^2$H$_4$]-Acetate, D$_2$O</td>
<td>0.1</td>
<td>6.5</td>
<td>1.2</td>
</tr>
</tbody>
</table>
contrast, when preparing $^{13}$C,$^{15}$N,$^2$H-labeled proteins, glucose is the most advantageous $^{13}$C-labeled carbon source, whether deuterated or not.

One drawback of recombinant protein expression on D$_2$O-based media is the retention of deuterons at a number of protected backbone amide sites in the core of the proteins that do not exchange to protons even though the purification is carried out in H$_2$O-based buffers. Indeed, re-measurement of [${}^{15}$N,${}^1$H]-correlation spectra for [U-98% $^{15}$N; U-$\sim$85% $^2$H]- SR1 after a period of 7 months at 4˚C revealed about 15 new peaks, probably originating from amide groups characterized by slow exchange rates (data not shown). Introduction of protons at all exchangeable sites requires chemical denaturation and renaturation of the proteins, or incubation at higher temperature to catalyse the NH exchange. Denaturation and renaturation cannot be applied easily to complex multimeric proteins without significant loss of labeled material. A protocol was thus developed for intermediate-level deuteration in H$_2$O-based cultivations, that was here more specifically used for specific amino acid labeling (Table 2.1, line 4, see below).

Table 2.1 provides a summary of the various labeling strategies described here for obtaining uniformly labeled and deuterated, as well as residue-specific $^{15}$N-labeled and deuterated chaperonins. The $^{15}$N label and deuteron incorporation levels are listed for each procedure, together with its respective yield, advantages and disadvantages. Table 2.1 should help choosing a labeling strategy depending for example on the deuteration level requirements. One important alternative to the procedures presented here for uniform labeling that has not been tested here is the use of LB rich medium to achieve high cell density and resuspension of the cells in isotope-supplemented and deuterated
Table 2.1 Labeling strategies for uniform and specific labeling of proteins

<table>
<thead>
<tr>
<th>MEDIUM DESCRIPTION</th>
<th>DEUTERATION LEVEL</th>
<th>( ^{15})N LABELING OCCUPANCY</th>
<th>YIELD/L CULTURE</th>
<th>COST/50 MG PROTEIN</th>
<th>ADVANTAGES/ DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deuterated, uniformly ( ^{15})N-labeled protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 1 Minimal medium on glucose, \( D_2O, ^{15}\)NH_4Cl       | ~85% of aliphatics, lower in aromatic side-chains | uniform, 100%       | GroES: 65 mg     | 300-400 $          | + low cost, easily performed  
- residual protonation in side chains |
|                                                          |                   |                                  | SR1: 50 mg\(^a\) |                    |                          |
| 2 Minimal medium on \([^{2}H_4]\)-acetate, \( D_2O, ^{15}\)NH_4Cl | ~98% | uniform, 100%       | GroEL:24 mg      | 1000-1400 $        | + high level of deut.  
- slow growth, low yield    |
|                                                          |                   |                                  | GroES:20 mg      |                    |                          |
| 3 Martek-9dN on \( D_2O \)                          | ~98% | uniform, 100%       | SR1: 20 mg\(^a\) | 2500 $             | -low yield with BL21, expensive |
| 4 Minimal medium on glucose spiked with Celtone-dN in \( H_2O \) | (~30-50% in pulse labeling)\(^b\) | uniform, 100%       | GroES: 38 mg\(^b\) | 300-400 $          | + no N-H/N-D exchange problems  
- expensive |
|                                                          |                   |                                  | GroEL:50 mg\(^b\) |                    |                          |
|                                                          |                   |                                  | SR1: 35 mg\(^b\) |                    |                          |
| Deuterated and specifically \( ^{15}\)N-Leu labeled protein |                   |                                  |                  |                    |                          |
| 5 Martek-9d in \( D_2O + ^{15}N, ^{2}H\)-Leu          | ~97% | specific, ~70-80% expected | SR1: 15 mg\(^a\) | 3150 $             | - light scrambling, use DL39(DE3) to prevent it. |
| 6 Protonated synthetic rich medium in \( H_2O + ^{15}N, ^{2}H\)-Leu | specific, only Leu (10%) | specific,100% | GroEL:126 mg    | 175$              | + no N-H/N-D exchange problems, high yield, low cost |
| 7 Minimal medium on glucose spiked with Celtone-d in \( H_2O + ^{15}N, ^{2}H\)-Leu | ~60% | specific, ~70-80% expected | GroEL:20 mg     | 890$              | + no N-H/N-D exchange problems  
- light scrambling, use DL39(DE3) to prevent it. |
|                                                          |                   |                                  | SR1: 5 mg\(^a\) (IDL39 cells!) |                    |                          |

\(^a\) Contribution of E. B. Bertelsen  
\(^b\) Pulse labeling: labeled celtone added at induction
medium shortly before induction (Marley et al., 2001). This procedure should achieve high levels of $^{15}$N and $^{13}$C labeling as well as partial deuteration of the protein with a good protein expression yield.

$[^{15}N,2H]$-Leu $[^{U-2H}]$-chaperonins

When $^{15}$N-labeling specific types of amino acids, the main concerns are isotopic dilution and scrambling of the label as a result of transamination. It follows that knowledge of the amino acid biosynthetic pathways is necessary when planning such an experiment (See for example Appendix 1). Specific labeling is usually achieved on rich medium containing all amino acids of which one or several are $^{15}$N-labeled. Previous studies have demonstrated that supplementing the growth medium with excess of unlabeled amino acids can reduce cross-labeling and that this procedure works best for leucine or phenylalanine, which terminate a metabolic pathway (Muchmore et al., 1989; Senn et al., 1987; Wüthrich, 1991). We decided to prepare $[^{15}N]$-Leu labeled chaperonins to reduce the spectral complexity and enable a more accurate count of the observed $[^{15}N,1H]$-cross peaks. However, for such large proteins, deuteration is necessary for observation of resonances. We found that deuteration of the $^{15}$N-labeled amino acid alone was not sufficient: $[^{15}N,2H]$-Leu GroEL was prepared in a synthetic rich medium (Table 2.1, line 6; Muchmore et al., 1989) containing all natural isotope abundance amino acids except for $[^{15}N,2H]$-Leu. In the $[^{15}N,1H]$-correlation spectra for this protein (Figure 2.1 A, B), at most 2 peaks could be observed, corresponding to the two leucine residues of the C-terminal flexible tail of the GroEL subunit. No other resonance could be detected due to too fast transverse relaxation in the mostly protonated protein. Besides, the $[^{15}N,1H]$-HSQC spectrum for $[^{15}N,2H]$-Leu GroEL denatured in 6M GuHCl shown in Figure 2.2 A provided a check of the $^{15}$N-labeling of the sample and yielded the
expected 41 leucine resonances. Achieving specific $^{15}\text{N}$-labeling and uniform deuteration necessitates a deuterated rich medium. Since preparation of synthetic rich medium from individual deuterated amino acid is not feasible financially, we worked with commercially available deuterated algal hydrolysates (Martek 9 and Celtone media for example) to which a slight excess $^{15}\text{N},^2\text{H}$-labeled Leu was added. Note that the composition of these media corresponds to a minimal medium supplemented with different amounts of an amino acid mix. Typically, 125 mg of labeled leucine were added to 1 L of deuterated Martek 9-d medium in D$_2$O just before induction (Table 2.1, line 5). For the fully deuterated $[^{15}\text{N},^2\text{H}]-\text{Leu}$ [U$^2\text{H}$]-SR1 prepared using this procedure (contribution of Dr. E. B. Bertelsen), 10-15 resonances could be observed in the $[^{15}\text{N},^1\text{H}]$-correlation spectra (Figure 2.1 C, D; see also section 6.2).

To avoid problems of incomplete N-H/N-D exchange mentioned above, most of the $[^{15}\text{N},^2\text{H}]-\text{Leu}$ labeling was also performed in a H$_2$O-based medium of approximately the same composition (Box 2.1, B) prepared from deuterated algal hydrolysate powder and protonated glucose. With this protocol, a deuteration level of 60% can be achieved (Table 2.1, line 7). Use of deuterated glucose only increased the deuteration to 67%. This approach enables preparation of partially deuterated proteins for systems that do not express well in D$_2$O-based media and for cases where re-introduction of protons at all exchangeable sites after expression on D$_2$O is problematic. The amount of algal hydrolysate powder and glucose used can be adjusted for optimal yield-to-cost ratio. As an alternative, the powder can be used to spike D$_2$O-based minimal medium for better growth of the bacteria (mostly equivalent to use of commercial rich media). The $[^{15}\text{N},^1\text{H}]$-correlation spectra for $[^{15}\text{N},^2\text{H}]-\text{Leu}$ [U$~60%$ $^2\text{H}$]-GroEL obtained from this procedure are shown in Figure 2.1 E, F. The deuteration level proved
2.2 Stable Isotope Labeling of the GroEL Chaperonins

Figure 2.1 $[^{15}\text{N},^{1}\text{H}]$-TROSY and $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY spectra of various $[^{15}\text{N},^{2}\text{H}]$-Leu specifically-labeled GroEL produced with different deuteration levels and strategies to suppress label scrambling.

**A.** and **B.** $[^{15}\text{N},^{2}\text{H}]$-Leu labeled GroEL prepared by addition of $^{15}\text{N},^{2}\text{H}$-Leu to an unlabeled rich synthetic medium in H$_2$O. For this mostly protonated protein, only 1-2 leucine resonances can be detected due to too fast transverse relaxation. In spite of the abundance of all amino acids in the medium, 2-10% residual scrambling of the label, mainly to Ala and Gly residues, can be observed (indicated by arrows). This leads here to strong peaks that could interfere with observation of the leucine resonances because the residues are located in a mobile region of GroEL, namely its C-terminal flexible tail. The protein was expressed in BL21(DE3) cells carrying a pET plasmid. The sample was 110 µM in GroEL monomers in 25 mM potassium phosphate, pH 6.15 and 20 mM KCl.
C and D. [\(^{15}\text{N},^{2}\text{H}\)] Leu \([U->97\%^{2}\text{H}]\)-labeled SR1 prepared by addition at induction of \(^{15}\text{N},^{2}\text{H}\)-Leu to a deuterated minimal medium supplemented with an amino acid mix (Martek 9-d). High deuteration enables observation of \(~10\) resonances in the \([^{15}\text{N},^{1}\text{H}]\)-CRIPT-TROSY spectrum. In the medium used, reduced amounts of amino acids result in extensive scrambling of the label, as observed from the \([^{15}\text{N},^{1}\text{H}]\)-TROSY spectrum (indicated by arrows. Compare to A and E.). The protein was expressed in BL21(DE3) cells carrying a pET plasmid. The sample was 70 \(\mu\)M in GroEL monomers in 25 mM potassium phosphate, pH 6.15 and 20 mM KCl. (Note that the low concentration of this sample might limit the number of cross peaks observed compared to E and F.)

E and F. [\(^{15}\text{N},^{2}\text{H}\)] Leu \([U\approx60\%^{2}\text{H}]\)-labeled GroEL prepared by addition of \(^{15}\text{N},^{2}\text{H}\)-Leu to an unlabeled minimal medium in H\(_2\)O supplemented with a deuterated amino acid mix, as expressed from transaminase-deficient DL39 cells. The deuteration level obtained is sufficient to allow observation of up to 15 resonances in the \([^{15}\text{N},^{1}\text{H}]\)-CRIPT-TROSY spectrum. As observed from the \([^{15}\text{N},^{1}\text{H}]\)-TROSY spectrum, this strategy suppresses scrambling of the label efficiently. The sample was 110 \(\mu\)M in GroEL monomers in 25 mM potassium phosphate, pH 6.15 and 20 mM KCl.

The second concern in the preparation of \(^{15}\text{N},^{2}\text{H}\)-Leu \(^2\text{H}\)-labeled proteins is the minimization of label scrambling. To illustrate this aspect, Figure 2.2 B shows the \([^{15}\text{N},^{1}\text{H}]\)-HSQC spectrum of denatured leucine-labeled GroEL prepared by addition of \(^{15}\text{N},^{2}\text{H}\)-Leu to a deuterated minimal medium supplemented with an amino acid mix (Martek 9-d), using BL21(DE3) cells (same sample as in Figure 2.1 C,D, but in denatured form). Beside the 41 main cross peaks corresponding to the leucine resonances, numerous weaker peaks are observed that complicate the analysis of the NMR data. This is also observable in the \([^{15}\text{N},^{1}\text{H}]\)-TROSY spectrum of the native protein and pointed out with arrows in Figure 2.1, C. For comparison, \(^{15}\text{N},^{2}\text{H}\)-Leu-labeled GroEL prepared in a synthetic rich medium containing large quantities of
natural isotope abundance amino acids to which $^{15}$N,$^2$H-Leu was added revealed much less scrambling of the label. The $[^{15}$N,$^1$H]-correlation spectra for this protein in denatured (Figure 2.2 A) and native (Figure 2.1 A, B) forms revealed a few percent of scrambling of the label to residues like alanine (7-10%) and glycine (1-2%). However, although the scrambling looks minimal in denatured form (compare Figure 2.2, A and B), these few labeled residues can give rise to strong peaks in the spectra of the native protein when located in flexible regions of the protein (indicated with arrows in Figure 2.1A)).

Finally, use of the transaminase-deficient strain DL39 was necessary to completely avoid this cross-labeling (Figure 2.1 E, F and Figure 2.2 C). Although the use of transaminase-deficient or auxotrophic strains often causes yield reduction, this approach for selective labeling was found here and by others (Muchmore et al., 1989; Lian and Middleton, 2001) to be more reliable and less growth conditions-dependent.

Table 2.2 Suppliers for isotope-labeled compounds and host cells used for expression of chaperonins.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Supplier</th>
<th>Compound</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>D$_2$O &gt;97%</td>
<td>Martek Biosciences Corp.,</td>
<td>Celtone-d</td>
<td>Martek Biosciences Corp.,</td>
</tr>
<tr>
<td></td>
<td>Columbia, Md, USA</td>
<td>or dN powder</td>
<td>Columbia, Md, USA</td>
</tr>
<tr>
<td>D$_2$O &gt;99%</td>
<td>Cambridge Isotope Laboratories,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Andover, MA, USA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$[^{15}$N]-NH$_4$Cl</td>
<td>Cambridge Isotope Laboratories,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Andover, MA, USA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$[^{13}$C$_6$]-glucose</td>
<td>Martek Biosciences Corp.,</td>
<td>Martek 9-d</td>
<td>Martek Biosciences Corp.,</td>
</tr>
<tr>
<td></td>
<td>Columbia, Md, USA</td>
<td>or 9-dN</td>
<td>Columbia, Md, USA</td>
</tr>
<tr>
<td>$[^2$H$_4$]-acetate</td>
<td>Martek Biosciences Corp.,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Columbia, Md, USA</td>
<td>E. coli</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BL21(DE3)</td>
<td></td>
</tr>
<tr>
<td>L-Leu-d$_{10}$$^{15}$N</td>
<td>Cambridge Isotope Laboratories,</td>
<td>$E. coli$</td>
<td>Gift of E. B. Bertelsen$^a$</td>
</tr>
<tr>
<td></td>
<td>Andover, MA, USA</td>
<td>DL39(DE3)</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ The original strain can be obtained from the E. coli Genetic Stock Center, Yale University, New Haven, CT, USA
2. Preparation of GroEL proteins for NMR measurements

Figure 2.2 $[^{15}\text{N},^{1}\text{H}]$-HSQC of the various $[^{15}\text{N},^{2}\text{H}]$-Leu specifically-labeled GroEL denatured in 6M GuHCl, as a check for scrambling of the $^{15}\text{N}$-label

A. $[^{15}\text{N},^{2}\text{H}]$-Leu labeled GroEL prepared by addition of $^{15}\text{N},^{2}\text{H}$-Leu to a rich synthetic medium in $\text{H}_2\text{O}$, as expressed from BL21 cells (same preparation as Figure 2.1 A and B). In spite of the abundance of all amino acids in the medium, slight residual scrambling of the label can be observed, for example in the region indicated by a circle (see also Figure 2.1 A). The inset shows a zoom on two well resolved peaks of the correlation spectrum. In the right panel, observation of a doublet shifted downfield of the main peak in the $^{15}\text{N}$ dimension indicates exchange of 10-20% of the deuteron at position $\alpha$ of Leu for a proton. In an identical preparation performed in $\text{D}_2\text{O}$ shown in the left panel, this doublet is indeed not present. The sample was 0.5 mM in GroEL monomers in 6M guanidinium chloride, 25 mM potassium phosphate, pH 6.15. The HSQC spectrum was recorded for 15 h, acquired data size 300*1600 complex points, $t_{1,\text{max}} = 250$ ms, $t_{2,\text{max}} = 267$ ms.

B. Denatured $[^{15}\text{N},^{2}\text{H}]$-Leu [U->97% $^{2}\text{H}$]-labeled SR1 prepared by addition at induction of $^{15}\text{N},^{2}\text{H}$-Leu to a deuterated minimal medium supplemented with an amino acid mix (same preparation as Figure 2.1 C and D). Beside the 41 main cross peaks corresponding to leucine resonances extensive scrambling of the $^{15}\text{N}$-label can be observed (compare A and B). The sample was ~100 $\mu$M in GroEL monomers in 6M guanidinium chloride, 25 mM potassium phosphate, pH 6.15. The HSQC spectrum was recorded on a Varian Inova 400 instrument for 30 h, acquired data size 300*1600 complex points, $t_{1,\text{max}} = 250$ ms, $t_{2,\text{max}} = 267$ ms. The INEPT delay was 5.4 ms and the recycle delay 1 s in spectra of A-C.
2.3 Purification and biophysical characterization of the chaperonins

GroEL purification

GroEL was purified as described in Box 2.3 (adapted from Zahn, 1994). Each fraction of the subsequent steps was analyzed by SDS-PAGE gel electrophoresis using 12% polyacrylamide gels (Figure 2.3 A, B). Although the purified sample was homogeneous as seen from the gel electrophoresis, it could still contain a small amount of peptides bound to GroEL. If present, these peptides could not be observed by NMR (Figure 3.5) and GroEL was not further purified (Procedures involving gel filtration in presence of ATP or 20% methanol (Weissman et al., 1995), disassembly to monomers (Lissin et al., 1990; Ybarra and Horowitz, 1995a,b), affigel batch chromatography (E. B. Bertelsen, personal communication; Martin, 2000; Weber, 2000 and references therein), etc. are nevertheless available for such purification. Pulse labeling can also be applied for NMR studies (E. Bertelsen, personal communication)). This procedure yielded >100 mg of $[^{15}\text{N},^{2}\text{H}]-\text{GroEL}$ from 1 L of minimal medium with glucose as sole carbon source and 24 mg of purified $[^{15}\text{N}]$-GroEL from 1 L of minimal medium in D$_{2}$O with $[^{2}\text{H}_{4}]-\text{acetate}$ as sole carbon source (Table 2.1). In some
cases, purification of GroEL from cells grown on deuterated acetate in D$_2$O showed signs of degradation at the C-terminus. Reduced expression time, use of protease inhibitors during purification, or addition of a chromatography “polishing” step should help avoiding this problem.

SR1 was purified as described previously (Horwich et al., 1998). All samples were contributed by Dr. E. Bertelsen. Natural isotope abundance SR1 was purified from cells grown in LB medium, with a yield of 100 mg from 1 L of culture. About 10 mg of purified \( [U\rightarrow$98\%$^{15}$N; \( U\rightarrow$97\%$^{2}$H$]\)-SR1 were obtained from 1 L of culture in deuterated medium.

GroEL and SR1 should not be frozen for storage, but sterile filtered and kept at 4°C. Although one freezing and thawing step does not significantly reduce the ATPase activity (E. Bertelsen, personal communication), it damages part of the protein, which gives rise to sharp signals of denatured protein in the NMR spectrum (data not shown).

**GroES purification**

GroES was purified according to the procedure described in Box 2.4 (developed with the precious advice of B. Gsell, Roche, Basel). Fractions of each step were analyzed by SDS-PAGE gel electrophoresis on 18% polyacrylamide gels (Figure 2.3, C–E). GroES is very sensitive to protease in its accessible mobile loop. Cleavage was observed in particular between L27 and T28. Purification at 4°C and in presence of potent protease inhibitors is necessary to obtain intact GroES. This procedure yielded 20 mg of \( [U\rightarrow$98\%$^{15}$N; \( U\rightarrow$97\%$^{2}$H$]\)-GroES and 65 mg of purified \( [U\rightarrow$97\%$^{13}$C; \( U\rightarrow$98\%$^{15}$N; \( U\sim$85\%$^{2}$H$]\)-GroES from 1 L of minimal medium in D$_2$O with \( ^2$H$_4$]-acetate and \( ^{13}$C$_6$]-D-glucose,
Box 2.3 Purification of GroEL (adapted from Zahn, 1994)

Buffers:

<table>
<thead>
<tr>
<th>Buffer 1</th>
<th>Buffer 2</th>
<th>Buffer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(extraction buffer)</td>
<td>(chromatography buffer)</td>
<td>(NMR buffer)</td>
</tr>
<tr>
<td>100 mM TrisCl at pH 8.1</td>
<td>50 mM TrisCl at pH 7.2</td>
<td>25 mM K phosphate</td>
</tr>
<tr>
<td>10 mM DTT</td>
<td>2 mM DTT</td>
<td>pH 6.15</td>
</tr>
<tr>
<td>0.1 mM EDTA</td>
<td>0.1 mM EDTA</td>
<td>20 mM KCl</td>
</tr>
<tr>
<td>0.1 mg/ml DNase I</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Purification procedure:

2-3g of cells (wet weight) expressing GroEL were redissolved in 30-50 ml extraction buffer 1 and lysed by sonication on ice. This step was repeated with the remaining pellets until no cell pellet could be seen after centrifugation (10 min., 9000 x g).

The lysate was made 30% in ammonium sulfate by slow addition of salt at 4°C and the solution was stirred for 1h. The lysate was centrifuged for 30 minutes at 35,000 g. The soluble fraction was then made 80% ammonium sulfate, stirred for 1h at 4°C and centrifuged. The pellets containing GroEL can be frozen at this stage in liquid nitrogen and stored at -80°C.

The pellets were resuspended in the buffer 2 and dialysed overnight against the same buffer.

The extract was loaded onto a DEAE-Sepharose fast flow (Pharmacia) column (200 ml) equilibrated in buffer 2 and washed with 3 column volumes. GroEL was eluted with a 0–500 mM NaCl gradient in buffer 2 over 10 column volumes. GroEL elutes between 360 and 420 mM NaCl.

The fractions containing GroEL were concentrated (Biomax-100 Ultrafree-15, Millipore) and applied to a Superdex 200 column (Pharmacia) equilibrated in buffer 3 in aliquots of 300-400ul.

Notes:

GroEL should NOT be frozen, but sterile filtered and kept at 4°C. Freezing and thawing damages part of the protein, which gives rise to sharp signals of denatured protein in the NMR spectrum.

An alternative for further purification of GroEL is application to a MonoQ column in buffer 2 and elution with a 200-500mM NaCl gradient. This polishing step is efficient for removal of the damaged protein fraction after a freeze and thaw procedure.

This purification scheme yielded ~150 mg of purified unlabeled GroEL for 1 L of culture in minimal medium, or 24 mg of purified [U-\(^{15}\)N; U-\(^{2}\)H]-GroEL for 1 L of culture in deuterated minimal medium with acetate as carbon source.
Box 2.4 Purification of GroES

Buffers:

<table>
<thead>
<tr>
<th>Buffer 1</th>
<th>Buffer 2</th>
<th>Buffer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(extraction buffer)</td>
<td>(chromatography buffer)</td>
<td>(NMR buffer)</td>
</tr>
<tr>
<td>10 ml B-PER bacterial protein extraction reagent (Pierce) pH 8.0</td>
<td>50 mM BisTrisCl pH 6.2</td>
<td>25 mM K phosphate pH 6.15</td>
</tr>
<tr>
<td>10% sucrose</td>
<td>2 mM DTT</td>
<td>20 mM KCl.</td>
</tr>
<tr>
<td>1 mg/ml DNase I, 10 mM MgCl₂</td>
<td>0.1 mM EDTA</td>
<td>protease inhibitors</td>
</tr>
<tr>
<td>5 mM isopropylfluorophosphate</td>
<td>protease inhibitors</td>
<td></td>
</tr>
<tr>
<td>a cocktail of protease inhibitors (Complete, Roche)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Purification procedure:

1g of cells (wet weight) was resuspended in 10 ml extraction buffer 1 and stirred at 4°C for 45 min. The lysate was cleared by centrifugation and the extraction of the pellets was repeated 2 times at 4°C and 4 times at room temperature.

The extraction fractions 4–7 were pooled and applied to a 30 ml Source Q (Pharmacia) column equilibrated in buffer 2 and washed for 3 column volumes. GroES was eluted with a 0–200 mM NaCl gradient in buffer 2 over 15 column volumes. GroES elutes between 120 and 160 mM NaCl.

The pooled fractions were concentrated (Biomax-30 Ultrafree-15, Millipore) and purified on Superdex 200 (Pharmacia) in buffer 3 in 300-400ul aliquots.

Notes:

GroES is very sensitive to protease and should be kept on ice and purified rapidly.

The yield was 20 mg of purified \([U^{15}N, U^{97%2H}]\)-GroES, or 65 mg of purified \([U^{13}C, U^{15}N; ~85%2H]\)-GroES respectively, from 1 L of culture in deuterated minimal medium.
respectively, as sole carbon source.

**Chaperonin characterization**

GroES and GroEL were characterized using the following biophysical methods to assess their native state, concentration and activity. Concentrations of the purified samples were determined using both UV absorption at 280 nm with absorption coefficient from Landry *et al.*, 1996 and Zahn *et al.*, 1996b, and the Bradford protein determination (Bradford, 1976) calibrated with a bovine serum albumin solution. Non-denaturing gel electrophoresis at 4% polyacrylamide showed that both proteins were of the correct overall size and able to form complexes in the presence of nucleotides (data not shown). Size exclusion chromatography on a Superdex-200 column (Pharmacia) was used to check the multimeric assembly of the proteins and remove possible aggregates (data not shown). ATPase activity measurement on several protein preparations according to the Lanzetta phosphate determination (Lanzetta *et al.*, 1979) showed that GroEL hydrolyzes ATP and that presence of GroES reduces the ATPase activity (data not shown). Finally, circular dichroism spectra and temperature denaturation curves were collected to check folding and correct assembly of the subunits (Figure 2.4). Purity and proper folding can also be checked in the 1D $^1$H-spectra recorded with the different samples prepared, shown in Figure 2.5. Further NMR characterization of these proteins was one aim of this work and is described in chapter 3.
Figure 2.4  Circular dichroism analysis of the GroES and GroEL labeled preparations.
A. Far-UV CD spectrum of 68% $^2$H-labeled GroEL. The solution was 1.7 mg/ml in 25 mM potassium phosphate buffer pH 6.15, 20 mM KCl. The spectrum is dominated by the signal from the $\alpha$-helical parts of the proteins, consistent with previous observations (Khandekar et al., 1993).
B. Far-UV CD spectrum of $>98\%^{15}$N, $>97\%^{2}$H-labeled GroES. The spectrum corresponds to that reported previously, with a low negative molar ellipticity due to the predominance of $\beta$-sheets (Boudker et al., 1997).
C. Temperature denaturation curve of $^2$H-labeled GroEL measured following the circular dichroism signal at 222 nm. The transition mid-point indicates an approximate denaturation temperature of 63˚C, in agreement with previous observations (Zahn, 1994).
D. Temperature denaturation curve of $^{15}$N,$^2$H-labeled GroES measured following the circular dichroism signal at 222 nm (Boudker et al., 1997). The solution was 1.0 mg/ml in 25 mM potassium phosphate buffer pH 6.15, 20 mM KCl. The transition mid-point indicates an approximate denaturation temperature of 78˚C, in agreement with previous report (Boudker et al., 1997).
2.3 Purification and Biophysical Characterization of the Chaperonins

Figure 2.3 SDS-PAGE gel electrophoresis of different fractions of the GroEL and GroES purification.

A. Fractions along the GroEL purification. A Coomassie-stained SDS-12% (w/v) polyacrylamide gel is shown. Lane 1: Molecular weight markers. Lane 2 and 3: Whole cell extract of DL39(DE3) carrying the pET-EL plasmid, uninduced. Lane 4 and 5: Whole cell extract of DL39(DE3) carrying the pET-EL plasmid, after induction with IPTG. The cells were grown in a minimal medium in H2O supplemented with a deuterated amino acid mix and with [15N,2H]-Leu. Lane 6: Pooled fractions of GroEL after ion exchange chromatography. Lane 7: Purified [15N,2H]-Leu GroEL.

B. Purified unlabeled GroEL. A Coomassie-stained SDS-12% (w/v) polyacrylamide gel is shown. Lane 1: Purified unlabeled GroEL. Lane 2: Molecular weight markers.

C. Whole cell extract of BL21(DE3) cells carrying the pET-ES plasmid. A Coomassie-stained SDS-18% (w/v) polyacrylamide gel is shown. The cells were grown in 2H,15N,13C-labeled minimal medium and induced with IPTG (Lane 1 and 2). Lane 3: Molecular weight markers.

D. Subsequent 13C,15N,2H-labeled GroES extraction fractions using B-PER buffer. A Coomassie-stained SDS-18% (w/v) polyacrylamide gel is shown. Lane 1: 1st extract. Lane 2: 2nd extract. Lane 3: 3rd extract. Lane 4: 4th extract. Lane 5: Molecular weight markers. Lane 6: 5th extract. Lane 7: 6th extract. Lane 8: 7th extract. Lane 9: 8th extract. Lane 10: Pellets after 8 steps of extraction. Lane 11: pooled extracts 4-8. Lane 12: Molecular weight markers.

E. Purified 2H,15N-labeled GroES fractions. A Coomassie-stained SDS-18% (w/v) polyacrylamide gel is shown. Lane 1-9: Fractions of purified 15N,2H-labeled GroES as eluting from a Superdex-200 size-exclusion chromatography column. Lane 10: Molecular weight markers.
Figure 2.5 1D $^1$H-spectra of the various GroEL and GroES samples prepared.
A. [U-98% $^{15}$N; U-$>$97% $^2$H]-GroEL, prepared according to the strategy of line 2 in Table 2.1.
B. [$^{15}$N, 2H]-Leu [U-60% $^2$H]-GroEL, prepared according to the strategy of line 7 in Table 2.1.
C. [$^{15}$N, $^2$H]-Leu GroEL, prepared according to the strategy of line 6 in Table 2.1.
D. [U-98% $^{15}$N; U-$>$97% $^2$H]-GroES, prepared according to the strategy of line 2 in Table 2.1.
E. [U-97% $^{13}$C; U-98% $^{15}$N; U-$\sim$85% $^2$H]-GroES, prepared according to the strategy of line 1 in Table 2.1.
F. [U-98% $^{15}$N; U-$\sim$50% $^2$H]-GroES, prepared according to the strategy of line 4 in Table 2.1 and pulse labeling.
The asterisks designate peaks that do not belong to the GroE proteins and originate from protease inhibitors, or indicate distortion of the spectrum around 4.7 ppm due to the large water resonance. The spectra have been recorded with the pulse sequence of Figure 3.1 with recycling delays of 1 s.
This chapter describes refined implementations of the recently introduced NMR techniques TROSY (transverse relaxation-optimized spectroscopy) (Pervushin et al., 1997) and CRINEPT (cross-correlated relaxation-enhanced polarization transfer) (Riek et al., 1999) that enable solution NMR studies with larger structures. Applications are described using $^{15}$N,$^2$H-labeled homo-oligomeric proteins with molecular weights of 110 to 800 kDa.
3.1 Sample conditions

For the work described in this thesis, more than 20 different samples of chaperonins and their complexes were prepared. All SR1 and DHFR samples were prepared by Dr. E. B. Bertelsen, as well as a few GroEL and the unlabeled GroES samples, whereas J. Fiaux contributed the labeled samples of GroES and GroEL. All samples were studied in H₂O solution of 25 mM potassium phosphate at pH 6.15, 20 mM KCl, containing as well 5% D₂O, 0.02% NaN₃, a buffer composition referred to as NMR buffer. To the samples containing GroES, a small amount of a protease inhibitor cocktail (Complete, Roche, Basel) was added for prevention of proteolytic degradation. The sample temperature was either 25°C, or 35°C for some experiments with the GroEL and GroEL complexes.

3.2 NMR equipment

Since high field strength B₀ benefits TROSY as well as CRINEPT-type polarization transfers (Pervushin et al., 1997; Pervushin, 2001; Riek et al., 1999; Wüthrich, 1998), the NMR experiments were carried out on Bruker DRX 750 and DRX 800 spectrometers at a ¹H resonance frequency of 750 and 800 MHz, respectively, using a triple-resonance probe with an actively shielded z-gradient coil. Details of the acquisition parameters and data handling after acquisition are given in the figure captions for the individual experiments. Specific features of the water handling techniques, and of individual one- and two-dimensional NMR experiments for use with very large structures are discussed in the following. The spectra were processed with the program PROSA (Güntert et al., 1992) and analyzed with the program XEASY (Bartels et al., 1995).
3.3 Water suppression in NMR experiments with large structures

Compared to corresponding experiments commonly used for studies with structures of molecular weights in the range 10–30 kDa, the following several points of sections 3.3 and 3.4 are novel traits of the 2D $^{15}$N,$^2$H-correlation NMR experiments designed for studies of large structures, that will be described in section 3.5.

In NMR experiments with large structures, the water magnetization has to be along the z-axis during the entire duration of the measurement. Improper water handling results in a dramatic decrease in sensitivity over the whole spectral width, as demonstrated here with one-dimensional (1D) $^1$H NMR spectra of uniformly $^{15}$N,$^2$H-labeled GroEL (Figure 3.1). The 1D $^1$H NMR spectrum of GroEL recorded with the pulse sequence of Figure 3.1 A contains a broad envelope with sizeable intensity from about 6 to 10 ppm (Figure 3.1B). The corresponding $^1$H NMR spectrum recorded with solvent suppression by presaturation during the recycle delay shows similar features, except that the signal intensity is about ten-fold reduced. Similar sensitivity losses were observed in multidimensional experiments with less than optimal water suppression (data not shown). All our NMR experiments for large structures therefore used water flip-back techniques with individual adjustment of each individual water-selective pulse, as described below. The dramatic influence of the water handling on the intensities of the protein signals implicates strong interactions between the water magnetization and the protein magnetization, which affect the effective longitudinal $^1$H relaxation rate (Wang et al., 1999; M. Hohwy, D. Braun, G. Wider and K. Wüthrich, to be published).

In each NMR experiment each water-selective pulse was adjusted individually for optimal water flip-back quality. For example, since
radiation damping tends to oppose driving the water magnetization from the +z-axis to the transverse plane (Chen et al., 1999), such as by the first soft pulse in the scheme of Figure 3.3 A, higher power for the selective water pulse will be required than when returning the water to the +z axis, such as by the second soft pulse in the scheme of Figure 3.3 A, where radiation damping supports the soft-pulse.
3.4 One-dimensional spectra

One-dimensional $^1$H NMR experiments

The scheme of Figure 3.1 A for 1D $^1$H NMR experiments with large structures contains a water-selective 90° pulse followed by a 90° hard pulse, so that the water magnetization is along the +z-axis during the signal acquisition period $t$. A pulsed magnetic field gradient is applied during the interscan delay to destroy possible residual in-phase magnetization.

Pulse length calibration

Because of the strong water-protein interactions, the 90° ($^1$H) pulse-length cannot conveniently be determined with conventional approaches (Cavanagh et al., 1996). Therefore, it was calibrated with a pulse sequence of a single hard pulse phase-shifted by 45° relative to the phase of the acquisition, during which the real and imaginary parts of the free induction decay are recorded simultaneously. The pulse is exactly 180° when the first points of both the real and imaginary parts of the free induction decay are zero (M. Salzmann, personal communication).

Adjustment of the recycle delay

The adjustment of the relaxation delay, $\Delta$, for obtaining maximal sensitivity of an NMR experiment within a given total measuring time is primarily based on considerations of the effective longitudinal relaxation times and the desirable pulse rotation angle (Ernst et al., 1987). To obtain a basis for estimating a favorable recycle time for the set-up of the experiments in Figure 2, we measured the proton $T_1$-values.
of a group of structures with molecular weights from 40 to 800 kDa (Figure 3.2) in H₂O solution. (Note that these experiments were carried out on a Varian Inova 400 instrument in contrast to all other measurements performed at high field.) For GroEL the value of the effective longitudinal relaxation time in H₂O, \( T_{1H^2O}(\^H) \), was thus found to be 0.3 s, for SR1 it was 0.4 s and for DHNA at 20˚ C it was 0.8 s. Overall, for \( \tau_c \) values longer than about 100 ns, a plot of the experimental \( T_{1H^2O}(\^H) \) values versus the overall rotational correlation time \( \tau_c \) levels off and even shows a trend towards shorter \( T_{1H^2O}(\^H) \) values for larger sizes. The experimental values for \( T_{1H^2O}(\^H) \) in Figure 6 are also much shorter than what one would predict from the formalism of established relaxation theory (Abragam, 1961; Ernst et al., 1987). This apparent discrepancy appears to be the result of intramolecular motions and exchange-mediated interactions with the solvent water that are not properly accounted for in conventional relaxation theory (Wang et al., 1999; M. Hohwy, D. Braun, G. Wider and K. Wüthrich, to be published). The presence of strong coupling between the water \(^1H\) magnetization and the \(^1H\) magnetization of the macromolecular solute is supported by the observation that the effective longitudinal relaxation time for SR1 in D₂O solution, \( T_{1D^2O}(\^H) = 6 \) s, is much longer than the value measured in H₂O solution (Figure 3.2 B). Furthermore, the longitudinal relaxation times of amide protons measured in H₂O solution increased up five-fold when the water resonance was saturated during the recovery delay Δ (data not shown, see Figure 3.1).

An approximate qualitative fit of the experimental data in H₂O and D₂O solutions (Figure 3.2) was obtained with the model calculations described elsewhere (Riek R., Fiaux J., Bertelsen E. B., Horwich A. L. and Wüthrich K., to be published), which account for the overall rotational tumbling of a solid sphere representing the structure considered and for intramolecular mobility of the \(^{15}N–^1H\) groups.
Figure 3.2 Measurement of the longitudinal proton relaxation times $T_1(1H)$ for a selection of proteins in the size range 40–800 kDa.

A. Pulse sequence used to measure $T_1(1H)$. The experiment starts with a 180° pulse, which is followed by the recovery delay $\Delta$ and a 90° observation pulse. The water resonance is held along $+z$ during the entire experiment by applying three water-selective 90° pulses, which are indicated by the curved shapes before and after the 180° pulse, and before the 90° hard pulse. The pulsed magnetic field gradients during the interscan delay and the recovery delay had a duration of 1 ms and amplitudes of 10 and 20 G/cm, respectively.

B. The black squares represent experimental $T_1(1H)$ values measured at 400 MHz on a Varian INOVA spectrometer for the following $^{15}$N,$^2$H-labeled proteins: Gyrase (43 kDa, estimated $\tau_c = 17$ ns at 20° C), DHNA at 4° and 20° C, SR1 at 20° C, and GroEL at 20° C (for details on these proteins see caption to Figure 3.4). For each structure, 30 measurements were recorded with the pulse sequence of A, using recovery delays $\Delta$ between 10 ms and 3 s. The experimental relaxation time $T_1(1H)$ at a proton frequency of 400 MHz was fitted as described elsewhere (Riek R., Fiaux J., Bertelsen E. B., Horwich A. L., Wüthrich K., to be published) with the following model assumptions: We consider a $^{15}$N–$^1$H group located in a β-sheet, with the remote protons $^1$HN($i$–1), $^1$HN($i$+1) and $^1$HN($j$) at distances of 4.3, 4.3 and 3.3 Å, respectively (Wüthrich, 1986), fast internal motion with $\tau_e = 100$ ps, order parameter $S^2 = 0.8$ (Lipari and Szabo, 1982). For the fit of the $T_1(D_2O)(1H)$ data, an additional remote proton was introduced, which is assumed to be in fast exchange with the bulk water at a distance of 4.0 Å, so that it is in Boltzmann equilibrium.
within the structure. For a satisfactory fit of the relaxation times measured in a H$_2$O solution, the overall relaxation rate was expressed by equation (1),

$$\frac{1}{T_1^{H_2O}(^{1}H)} = \frac{1}{T_1^{D_2O}(^{1}H)} + \rho_{^{1}H^{1}O}$$

(1)

$\rho_{^{1}H^{1}O}$ accounts for dipole–dipole coupling to additional remote hydrogen nuclei $^{1}H^{1}O$, which are assumed to be in fast exchange with the bulk water so that they would not be seen in D$_2$O solution. Examples would be the hydroxyl protons of Thr, Tyr, and Ser. In (1) it is assumed that the exchange of $^{1}H^{1}O$ with the bulk water is sufficiently fast so that cross relaxation between $^{1}H^{1}N$ and $^{1}H^{1}O$ is negligibly small. The smooth curve in Figure 3.2 was obtained with the assumption that a single $^{1}H^{1}O$ spin in fast exchange with the bulk water is present at a distance of 4.0 Å from the observed $^{1}H^{1}N$ spin.

$T_1^{D_2O}(^{1}H)$ and $T_1^{H_2O}(^{1}H)$ are similar for $\tau_c$ values shorter than about 5 ns, whereas the interaction with remote protons in fast exchange with H$_2$O becomes dominant for long $\tau_c$ values. In practice, the smooth curves in Figure 3.2 provide a starting point for estimating the relaxation delay $\Delta$ for new structures with known molecular weight.
3.5 Two-dimensional correlation experiments

2D $[^{15}\text{N}, ^1\text{H}]-\text{HSQC}$ and $[^{15}\text{N}, ^1\text{H}]-\text{TROSY}$

Experiments commonly used for studies with structures of molecular weights up to 50 kDa rely on the 2D $[^{15}\text{N}, ^1\text{H}]-\text{HSQC}$ and $[^{15}\text{N}, ^1\text{H}]-\text{TROSY}$ schemes for detection of the fingerprint of $^{15}\text{N}-^1\text{H}$ moieties in a protein (Mori et al., 1995; Pervushin et al., 1997). Applications of these experiments to larger structure sizes however quickly reach a limit: Whereas the $[^{15}\text{N}, ^1\text{H}]-\text{TROSY}$ scheme yields complete fingerprints for proteins up to ~150 kDa (see for example Salzmann, 1999 and Figure 4.2), it is inefficient in detecting cross peaks for the 400 and 800 kDa GroEL single-ring variant SR1 and GroEL wild-type proteins. In Figure 3.5 showing the $[^{15}\text{N}, ^1\text{H}]-\text{TROSY}$ spectrum of $^{15}\text{N}, ^2\text{H}$-labeled GroEL, only resonances from a C-terminal flexible tail of GroEL can be observed (see section 6.1). Other pulse schemes have then been implemented for measurement of correlation spectra of large structures which are described in the following paragraphs.

2D $[^{15}\text{N}, ^1\text{H}]-\text{CRIPT-TROSY}$

The 2D $[^{15}\text{N}, ^1\text{H}]-\text{CRIPT-TROSY}$ experiment (Figure 3.3 A) is transverse relaxation-optimized for its entire duration (Table 3.1). It uses TROSY (Pervushin et al., 1997; Pervushin, 2001) during the evolution and acquisition periods, and CRIPT (Riek et al., 1999) for the polarization transfer. 2D $[^{15}\text{N}, ^1\text{H}]-\text{CRIPT-TROSY}$ starts with the $^1\text{H}$ magnetization generated by the first $90^\circ(\text{H})$ pulse, which is then transferred to heteronuclear antiphase magnetization, using cross-correlated relaxation between $^{15}\text{N}-^1\text{H}^\text{N}$ dipole–dipole coupling and $^1\text{H}^\text{N}$ CSA during the time period T (Brüschweiler and Ernst, 1991; Riek...
Figure 3.3 2D $^{15}\text{N},^1\text{H}$-correlation experiments applied in the work with large macromolecules.

A. 2D $^{15}\text{N},^1\text{H}$-CRIPT-TROSY (Riek et al., 1999)
B. 2D $^{15}\text{N},^1\text{H}$-CRINEPT-TROSY (Riek et al., 1999)
C. 2D $^{15}\text{N},^1\text{H}$-CRINEPT-HMQC-$^1\text{H}$-TROSY (Riek et al., 1999)
D. 2D $^{15}\text{N},^1\text{H}$-TROSY (Pervushin et al., 1997)
E. 2D $^{15}\text{N},^1\text{H}$-HSQC (Mori et al., 1995)

In all experiments the radio-frequency pulses on $^1\text{H}$ and $^{15}\text{N}$ were applied at 4.8 ppm and 119 ppm, respectively. The narrow and wide black bars indicate non-selective 90° and 180° pulses. The curved shapes on the $^1\text{H}$ line represent Gaussian-shaped selective 90° pulses on the water resonance, which have a length of 0.7–1.0 ms. The power and the phase of each water-selective pulse were adjusted by a trial-and-error approach to ensure optimal alignment of the water magnetization along the z-axis during the entire experiment (see text). On the line marked PFG, rectangles indicate duration and amplitude of pulsed magnetic field gradients applied along the z-axis: G₀, 500 µs, 11 G/cm; G₁, 300 µs, 19 G/cm; G₂, 300 µs, 39 G/cm; G₃, 300 us, 19 G/cm; G₄, 300 us, 24 G/cm, G₅, 500 us, 13 G/cm; G₆, 300 us, 9 G/cm; G₇, 500 us, 18 G/cm. Quadrature detection in the $t_1(^{15}\text{N})$ dimension is achieved by the States-TPPI method applied to the phase $\psi_1$. 
The length of the period $T$ was adjusted in the range 1.0 to 5.4 ms, depending on the size of the structure studied (see text). All radio-frequency pulses are applied with phase $x$, unless a different phase is indicated in the figure: A. $\phi_1 = \{x, x, -x, -x\}$, $\psi_1 = \{-x, x, -x, -x\}$; $\psi_2 = \{x, -x\}$, $\phi_1 = \{-x, x\}$, $\psi_1 = \{x, -x\}$; two free induction decays are recorded for each $t_1$ increment, with $\psi_2 = \{x, x\}$ and $\psi_2 = \{-x, -x\}$, respectively, and are added with a 90° phase shift in both dimensions. B. $\psi_1 = \{x, -x\}$, $\psi_1 = \{x, -x\}$. B. $\psi_1 = \{y, y, -x, x\}$, $\phi_1 = \{x, x, x, -x, -x, -x, -x\}$. Two free induction decays are recorded per $t_1$ delay, with $\psi_1$ incremented by 90° in between and stored as the real and imaginary parts of the interferogram in $t_1$. E. $\psi_1 = \{x, -x\}$, $\phi_1 = \{x, -x\}$.

---

_etch et al., 1999; Riek R., Fiaux J., Bertelsen E. B., Horwich A. L. and Wüthrich K., to be published_

$$\langle 2H_y N_z \rangle (T) = \sinh (R_{\text{cor}} T) \exp (-R_H T) \langle H_y \rangle (0),$$

(2)

with

$$R_H = \left[ \frac{2}{9} (\gamma_H B_0 \Delta \sigma_H)^2 + \frac{1}{2} (\hbar \gamma_H \gamma_N / r_{HN})^2 \right] \frac{2}{5} \tau_c + \frac{1}{2T_1 (N)} + \frac{1}{T_{2r} (H)}$$

(3)

and

$$R_{\text{cor}} = \frac{2}{3} (\gamma_H B_0 \Delta \sigma_H) (\hbar \gamma_H \gamma_N / r_{HN}) P_2 (\theta) \frac{2}{5} \tau_c$$

(4)

$H$ and $N$ stand for $^1H_N$ and $^{15}N$, $R_H$ is the transverse relaxation rate of $^1H_N$, and $R_{\text{cor}}$ the relaxation rate due to cross correlation between $^1H_N$ CSA (first term of (4)) and $^{15}N-^1H$ dipole–dipole coupling (second term of (4)). The Legendre polynome of second order, $P_2 (\theta)$, accounts for deviations of the angle $\theta$ between the dipole–dipole vector and the principal axis of the CSA tensor from 0°. $1/T_1 (N)$ is the longitudinal relaxation rate of $^{15}N$, and $1/T_{2r} (H)$ is the contribution to the transverse
relaxation rate of $^1\text{H}^\text{N}$ due to dipole–dipole coupling with remote protons, $H_i$, at distances $r_{HHi}$. During the CRIPT transfer time, $T$, the proton chemical shift evolution is refocused by a 180° pulse, which also decouples the protons from $^{15}\text{N}$. At the end of the period $T$, 90° pulses on $^{15}\text{N}$ and $^1\text{H}$ generate heteronuclear antiphase coherence $2H_{Z}^{\text{N}}N_{y}$, which is then frequency-labeled during $t_1$. The hard 90° pulses on $^{15}\text{N}$ and $^1\text{H}$ applied at the end of $t_1$ transfer the coherences back to antiphase coherence $2H_{Z}^{\text{N}}N_{y}$, which is acquired during $t_2$. To minimize the overall length of the pulse sequence, no multiplet component selection is applied. In principle, the 2D $[^{15}\text{N},^1\text{H}]$-CRIP-TROSY experiment thus retains all four multiplet components of the $^{15}\text{N}$–$^1\text{H}$-moieties, with two negative components upfield-shifted and two positive components downfield-shifted along the $^{1}\text{H}$ chemical shift axis (Table 3.1, Figure 5.2).

In the practice of experiments with molecular sizes of about 200 kDa and above, however, transverse relaxation largely suppresses the unwanted three multiplet components. The water magnetization is kept along the z-axis during the whole duration of the experiment, using the three water-selective pulses indicated by curved shapes in Figure 3.3. The phase cycling scheme proposed in an earlier version of the 2D $[^{15}\text{N},^1\text{H}]$-CRIP-TROSY experiment (Figure 2a of Riek et al., 1999) was extended to achieve improved suppression of artifacts.

The CRIPT delay, $T$, needs to be adjusted for the molecular size studied, since cross-correlated relaxation between dipole–dipole coupling and CSA increases with increasing rotational correlation time $\tau_c$ (Riek et al., 1999). On the basis of (2), $T$ is given by

$$
T = \frac{1}{R_{\text{cor}}} \text{ArcCosh} \left( \frac{R_H}{\sqrt{2R_H - R_{\text{cor}}^2}} \right)
$$

Assuming for large structures that $1/T_{2r}(H) = \sum_i \left( \frac{\hbar \gamma_H^2}{2r_{HHi}^3} \right)^2 \tau_c$
3.5 TWO-DIMENSIONAL CORRELATION EXPERIMENTS

>> $1/T_1(N)$, one then has from (5) that $T$ is to a good approximation inversely proportional to the effective rotational correlation time $\tau_c$ (see the additional data given in Figure 3.4 below). As these considerations imply, the 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRIPT-TROSY}$ experiment (Figure 3.3 A) yields good results for very large structures, with molecular weights beyond 200 kDa.

The spectra of $^{15}\text{N},^{2}\text{H}$-labeled GroEL in Figure 3.5 show that transverse relaxation-optimization during the polarization transfers is indispensable in NMR experiments with large structures. In the experimental schemes of Figure 3.3, A, B and C, the standard INEPT transfer (Morris and Freeman, 1979) has therefore either been supplemented or substituted by a transfer scheme based on cross-correlated relaxation. Whereas the efficiency of INEPT transfers decreases with increasing size because of rapid transverse relaxation, the efficiency of cross correlated relaxation-induced polarization transfers increases for longer values of the rotational correlation time $\tau_c$ ((2)–(5)) (Riek et al., 1999). Figure 3.4, A and B, illustrates the experimental determination of the optimal CRIPT transfer time for SR1 and GroEL. The black squares in Figure 3.4 C represent the thus measured optimal transfer times for five structures with estimated $\tau_c$-values between 50 and 350 ns. The experimental data have been approximated with a model calculation described in the figure caption, and the resulting smooth curve in Figure 3.4 provides a starting point for estimating the delay to be used in the experiments of Figure 3.3 for a given system. In our experience, however, it pays to measure new CRIPT build-up curves (Figure 3.4, A and B) for optimizing the transfer delays in each new system studied.

Measurement of the optimal CRIPT transfer period $T$ can also be used, via the curve in Figure 3.4, to estimate the overall rotational
Figure 3.4 Experimental measurement of the CRIPT transfer delays at a proton frequency of 750 MHz.

A. Plot of the relative NMR peak heights, $I_{rel}$, as a function of the duration of the CRIPT transfer for the 400 kDa protein SR1. The buildup curve was obtained from a series of measurements with the pulse sequence of Figure 3.3 A. The transfer delay, $T$, was incremented stepwise from 0.8 ms to 9.0 ms between the individual experiments.

B. Same as A for the 800 kDa protein GroEL. $T$ was incremented between 0.4 ms and 6.8 ms. The arrows indicate the $T$ values with optimal transfer efficiency that were derived from this data.

C. CRIPT transfer delays, $T$, providing optimal transfer efficiency at a proton frequency of 750 MHz. The black squares represent experimental optimal $T$-values measured with the experiments of A and B for the following $^{15}$N,$^2$H-labeled proteins: DHNA (110 kDa, estimated value for $\tau_c = 45$ ns at 20°C, 75 ns at 4°C), SR1 (400 kDa, $\tau_c = 155$ ns at 25°C), GroES in a complex with SR1 (472 kDa, $\tau_c = 185$ ns at 25°C), GroEL (800 kDa, $\tau_c = 310$ ns at 35°C). The $\tau_c$ values were estimated from the molecular weight and the temperature used, assuming a rigid spherical shape of the structures studied, and the previously reported experimental $\tau_c$ values for DHNA (Riek et al., 1999) were used as an additional reference. The experimental data were fitted using (5) (see also (2)–(4)) with the following parameters: $\tau_{HN} = 1.04 \text{ Å}, \Delta \sigma_{H} = 15$ ppm, $\vartheta = 10^\circ$, dipole–dipole coupling with the remote protons expected in an $\alpha$-helix, i.e. $^1H^N(i-2), ^1H^N(i-1), ^1H^N(i+1)$ and $^1H^N(i+2)$ at distances of 4.2, 2.8, 2.8, and 4.2 Å, respectively (Salzmann et al., 1998; Wüthrich, 1986).
correlation time $\tau_c$ of a new structure. This is of practical interest, since the use of the standard method of calculating $\tau_c$ from the ratio of $T_1(^{15}N)$ and $T_2(^{15}N)$ (Cavanagh et al., 1986) is usually impractical for large structures, due to the low sensitivity of the experiments used to measure the relaxation times.

**Figure 3.5** Heteronuclear 2D $[^{15}N, ^{1}H]$-correlation spectra of the uniformly $^{15}N, ^{2}H$-labeled 800 kDa tetradecamer protein GroEL from *E. coli*.  
**A.** 2D $[^{15}N, ^{1}H]$-TROSY spectrum, measuring time = 20 hours, acquired data size 256*1024 complex points, $t_{1,\text{max}} = 32$ ms, $t_{2,\text{max}} = 100$ ms.  
**B.** 2D $[^{15}N, ^{1}H]$-CRIPT-TROSY spectrum, measuring time = 20 hours, acquired data size 100*1024 complex points, $t_{1,\text{max}} = 10$ ms, $t_{2,\text{max}} = 100$ ms. Prior to Fourier transformation the data were multiplied along the $t_1$-dimension with a sine function shifted by 10˚ (DeMarco and Wüthrich, 1976), and in the $t_2$-dimension with an empirically optimized exponential function.
2D $[^{15}\text{N,}^{1}\text{H}]$-CRINEPT-TROSY

The 2D $[^{15}\text{N,}^{1}\text{H}]$-CRINEPT-TROSY experiment (Figure 3.3 B) (Riek et al., 1999) provides an alternative as well as a complementation of 2D $[^{15}\text{N,}^{1}\text{H}]$-CRIPT-TROSY for studies of structures with molecular weights larger than 200 kDa. During CRINEPT-based polarization transfers, transverse relaxation-optimization is active and polarization is transferred via a combination of cross-correlated relaxation and scalar coupling. Similar to 2D $[^{15}\text{N,}^{1}\text{H}]$-CRIPT-TROSY, 2D $[^{15}\text{N,}^{1}\text{H}]$-CRINEPT-TROSY includes transverse relaxation-optimization during its entire duration (Table 3.1, Figure 3.3 B). 2D $[^{15}\text{N,}^{1}\text{H}]$-CRINEPT-TROSY contains two CRINEPT transfer periods, and a third delay of length $T$ which is used to select for those two multiplet components of the $^{15}\text{N}$–$^{1}\text{H}$ four-line fine structure that have the slowest and fastest transverse relaxation rates (Table 3.1). Since no pulses are applied during the CRINEPT elements (Figure 3.3 B) there is free proton chemical shift evolution, which leads to a sensitivity loss by a factor 2 when the coherence from the first CRINEPT element is refocused during the second CRINEPT element. A detailed description of the 2D $[^{15}\text{N,}^{1}\text{H}]$-CRINEPT-TROSY experiment is presented in Riek et al., 1999.

2D $[^{15}\text{N,}^{1}\text{H}]$-CRINEPT-HMQC-$[^{1}\text{H}]$-TROSY

The 2D $[^{15}\text{N,}^{1}\text{H}]$-CRINEPT-HMQC-$[^{1}\text{H}]$-TROSY experiment (Figure 3.3 C) (Riek et al., 1999) is included here as an alternative to 2D $[^{15}\text{N,}^{1}\text{H}]$-CRINEPT-TROSY that has higher intrinsic sensitivity, since there is no loss of magnetization attributable to the proton chemical shift evolution during the two CRINEPT elements. Part of this advantage is offset because there is no TROSY compensation during $t_1$, which results in broad resonances along the $^{15}\text{N}$ dimension. Since decoupling is applied during $t_1$ but not during acquisition (Figure 3.3 C), there are two
multiplet peaks per $^{15}\text{N}-^{1}\text{H}$-moiety, which are shifted about 45 Hz (= 0.5 $^{1}J_{\text{HN}}$) upfield along the $^{15}\text{N}$ chemical shift axis when compared to the TROSY component (Table 3.1), i.e., their position along $\omega_{1}(^{15}\text{N})$ corresponds to that of the single peak in conventional, broad-band-decoupled COSY spectra.

A comparison of the 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRINEPT-TROSY}$, 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRINEPT-HMQC-[^{1}\text{H}]-TROSY}$ and 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRIPT-TROSY}$ experiments

In the 2D $[^{15}\text{N},^{1}\text{H}]-\text{TROSY}$ spectrum of the $^{15}\text{N},^{2}\text{H}$-labeled heptameric 78 kDa protein GroES, 89 out of 94 expected $^{15}\text{N}-^{1}\text{H}$ signals have been identified (chapter 5; Fiaux et al., 2002). For $^{15}\text{N},^{2}\text{H}$-labeled GroES in a complex with $^{2}\text{H}$-labeled SR1, which has a molecular weight of 472 kDa, 78 of the expected 94 cross-peaks are detected in the 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRINEPT-TROSY}$, 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRINEPT-HMQC-[^{1}\text{H}]-TROSY}$ and 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRIPT-TROSY}$ spectra (Figure 3.6), showing that TROSY-optimization throughout the entire pulse sequence (Table 3.1) enables the recording of virtually complete $^{15}\text{N}-^{1}\text{H}$ fingerprints for a structure size of approximately 500 kDa. The following is a more detailed evaluation of the relative merits of the three experiments of Figure 3.3 A–C for studies of very large structures.

The 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRINEPT-TROSY}$ experiment uses CRINEPT transfers to ensure that cross-peaks of $^{15}\text{N}-^{1}\text{H}$-moieties from flexible as well as structured regions of the protein are detected. Use of TROSY in both dimensions (Table 3.1) results in narrow line shapes. This is readily seen by comparison of the resolution in the spectral region highlighted by the rectangles in Figure 3.6. The 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRINEPT-TROSY}$ experiment is somewhat less sensitive when compared to the other two experiments of Figure 3.3, due to the longer pulse sequence used for
Table 3.1 *Survey of the salient features of the presently described solution NMR experiments for studies of large structures*

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Polarization Transfer</th>
<th>Evolution/Acquisition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nr&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Mechanism&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>[15N,1H]-TROSY</td>
<td>3</td>
<td>$^3J_{HN}$</td>
</tr>
<tr>
<td>[15N,1H]-CRIPT-TROSY</td>
<td>1</td>
<td>R&lt;sub&gt;cor&lt;/sub&gt;</td>
</tr>
<tr>
<td>[15N,1H]-CRINEPT-TROSY</td>
<td>2&lt;sup&gt;e&lt;/sup&gt;</td>
<td>$^3J_{HN}$ + R&lt;sub&gt;cor&lt;/sub&gt;</td>
</tr>
<tr>
<td>[15N,1H]-CRINEPT-HMQC-[1H]-TROSY</td>
<td>2</td>
<td>$^3J_{HN}$ + R&lt;sub&gt;cor&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Number of polarization transfers.

<sup>b</sup> $^3J_{HN}$ indicates polarization transfer based on scalar couplings (INEPT)<sup>31</sup>, R<sub>cor</sub> indicates polarization transfer based on cross-correlated relaxation (CRIPT) (see text).

<sup>c</sup> ‘+’ indicates that transverse relaxation-optimization is active, ‘-’ indicates that there is no transverse relaxation-optimization.

<sup>d</sup> The square indicates the positions of the four multiplet components of a $^{15}$N--$^1$H two-spin system in a non-decoupled [15N,1H]-correlation experiment. The $\omega_2(^1H)$ frequency is along the horizontal axis and the $\omega_1(^{15}N)$ frequency along the vertical axis. The multiplet components selected by the individual experiments are indicated by ‘+’ for positive intensity, and ‘-’ for negative intensity, and the size of the circles indicates the anticipated relative peak height.

<sup>e</sup> There are 3 periods T in the pulse sequence (Figure 3.3 B), of which the second one is used for multiplet component selection (see text).

<sup>f</sup> There is a 50% loss of magnetization due to the free $^1$H chemical shift evolution during the first CRINEPT element.
multplet component selection (Table 3.1). For structure sizes above about 200 kDa, transverse relaxation ensures that only the TROSY-component of the remaining two-component multiplet (Table 3.1) is seen.

The 2D \[^{15}N,^1H\]-CRINEPT-HMQC-[\(^1H\)]-TROSY experiment uses CRINEPT to detect cross-peaks of \(^{15}N–^1H\)-moieties from both flexible and more rigidly structured protein regions. Along the \(\omega_1(^{15}N)\) dimension the peaks have broad line-widths because of the absence of TROSY-compensation in the indirect dimension (Table 3.1). This is clearly seen when the 2D \(^{15}N,^1H\)-CRINEPT-HMQC-[\(^1H\)]-TROSY spectrum of Figure 3.6 D is processed identically to the data sets of A and C. With the digital filtering of Figure 3.6 B, a similar spectrum could be generated as for the fully TROSY-based spectra of Figure 3.6 A and C, but this processing also introduced artifactual ‘wiggles’ of the most intense cross-peaks (Figure 3.6 E), which correspond to mobile \(^{15}N–^1H\)-groups in a loop of the protein. For flexible \(^{15}N–^1H\)-moieties one also observes two multiplet components (Table 3.1), whereas for cross-peaks from structured regions of the protein the rapid transverse relaxation suppresses the lower-field fine structure component (Figure 3.6 B). Overall, the 2D \(^{15}N,^1H\)-CRINEPT-HMQC-[\(^1H\)]-TROSY experiment has been used primarily for initial screening of the feasibility of NMR observation with new structures, where one can exploit the intrinsic high sensitivity of this scheme.

The 2D \(^{15}N,^1H\)-CRIPT-TROSY experiment with its single polarization transfer (Table 3.1) provides high sensitivity for the observation of peaks with fast transverse relaxation. These strong peaks then typically arise from \(^{15}N–^1H\)-moieties located in well-structured regions of the protein core, whereas peaks originating from flexible loops or chain ends are suppressed. Similarly, peaks that might arise
from admixtures of smaller protein impurities would be at least partially filtered away. The filtering effect can be tuned with the experimental set-up. With a short transfer delay, CRIPT selects for $^{15}\text{N}^{-1}\text{H}$-moieties with a long effective rotational correlation time $\tau_c$. This is illustrated in Figure 3.6 with the circled cross-peak, which is only observable in the CRINEPT-based spectra of A and B, but is suppressed in the 2D $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY experiment of C. Comparison with 2D $[^{15}\text{N},^{1}\text{H}]$-TROSY spectra, where the peaks from amide protons with short transverse relaxation times are suppressed, confirmed that the missing peaks in Figure 3.6 originate from $^{15}\text{N}^{-1}\text{H}$ groups with slow transverse relaxation. In this particular sample, the considered peak originates from 5-10% residual free GroES at 72 kDa.

A limitation of the 2D $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY scheme arises because $^{15}\text{N}^{-1}\text{H}$ moieties with small $^{1}\text{H}$ CSA or a large angle $\vartheta$ may not be detected due to absence or near-absence of cross-correlated relaxation, $R_{\text{cor}}$ (4). These peaks will normally be present in CRINEPT-based experiments and can therefore not \textit{a priori} be distinguished from peaks representing either flexible regions of a large structure or admixtures of smaller proteins. At a structure size of 500 kDa, an antiphase multiplet pattern (Table 3.1) is visible only for a small number of peaks (Figure 3.6 E), which all appear to originate from flexible loops. This demonstrates nicely that the multiplet component selection by rapid transverse relaxation works well for structures in the size range of interest for the use of the 2D $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY experiment.
3.5 Two-Dimensional Correlation Experiments

\[ \omega_2(1^H) \] (ppm) \[ \omega_1(1^N) \] (ppm)

\[ \omega_2(1^H) \] (ppm) \[ \omega_1(1^N) \] (ppm)

A

B

C

D

E

\[ \omega_2(1^H) \] (ppm) \[ \omega_1(1^N) \] (ppm)

\[ \omega_2(1^H) \] (ppm) \[ \omega_1(1^N) \] (ppm)
3. SOLUTION NMR TECHNIQUES FOR LARGE SUPRAMOLECULAR STRUCTURES

Figure 3.6 2D $^{15}$N-$^1$H-correlation spectra recorded using the experimental schemes of Figure 3.3 with uniformly $^{15}$N,$^2$H-labeled GroES bound to SR1 in a complex with molecular weight 500 kDa.

A. 2D $^{15}$N-$^1$H-CRINEPT-TROSY.
B. 2D $^{15}$N-$^1$H-CRINEPT-HMQC-$^1$H-TROSY.
C. 2D $^{15}$N-$^1$H-CRIPT-TROSY. A cross-peak corresponding to a mobile $^{15}$N–$^1$H-group is circled. The rectangle indicates a spectral region that is discussed in the text. The arrows indicate the positions of the cross sections shown in E. The acquired data size is 100*1024 complex points, with $t_{1,\text{max}} = 10$ ms and $t_{2,\text{max}} = 100$ ms. Prior to Fourier transformation a sine bell was applied in the $t_1$-dimension, with phase shifts of $10^\circ$ for A and $30^\circ$ for B and C (DeMarco and Wüthrich, 1976), and an empirically optimized exponential function was applied in the $t_2$-dimension.

D. Illustration of the impact of different digital filtering on NMR spectra of large structures obtained with the experiments of A, B and C. Here, the data of the 2D $^{15}$N-$^1$H-CRINEPT-HMQC-$^1$H-TROSY experiment of B have been treated identically to those of the 2D $^{15}$N-$^1$H-CRINEPT-TROSY and 2D $^{15}$N-$^1$H-CRIPT-TROSY experiments of A and C, i.e., a sine bell shifted by $30^\circ$ was applied along the $t_1$-dimension (see text).

E. Cross sections through the 2D $^{15}$N–$^1$H correlation spectra of uniformly $^{15}$N,$^2$H-labeled GroES bound to SR1. The positions of the cross sections are indicated by arrows and the appropriate lettering a–c. The asterisks identify transformation artefacts due to the $10^\circ$-shifted sine bell function used to enhance the resolution of the 2D $^{15}$N-$^1$H-CRINEPT-HMQC-$^1$H-TROSY spectrum. The + and – signs below the trace c2 indicate the antiphase character of multiplets observed in the 2D $^{15}$N-$^1$H-CRIPT-TROSY spectrum.
3.6 $^{15}$N and $^1$H linewidths in TROSY-type experiments

The $^{15}$N and $^1$H linewidths were measured in 2D $[^{15}$N,$^1$H]-CRIP-TROSY spectra of $^{15}$N,$^2$H-labeled DHNA at 4°C, $^{15}$N,$^2$H-labeled SR1 at 25°C, and $^{15}$N,$^2$H-labeled GroEL at 25°C (Figure 3.7). For DHNA, the average linewidths is 50 Hz along the $^{15}$N dimension and 30 Hz along the $^1$H dimension. For SR1, average values of 60 Hz were observed for both the $^{15}$N and $^1$H linewidths, and for GroEL these quantities are 75 Hz and 110 Hz, respectively. In the two larger proteins, SR1 and GroEL, a much broader linewidth distribution was observed than in DHNA.

The straight lines in the two displays of Figure 3.7 suggest a tentative qualitative rationalization of the experimental linewidths based on the following model considerations: We assume that the linewidths in the two dimensions of a 2D $[^{15}$N,$^1$H]-correlation experiment are proportional to the effective transverse relaxation rates during the frequency labeling period, and the acquisition period, respectively (Cavanagh et al., 1986). Thus, in TROSY-type experiments the linewidths are assumed to be proportional to the difference between the relaxation rates $R_H$ and $R_{cor}$ (for $^1$H these expressions are given by (3) and (4); the corresponding expressions for $^{15}$N, $R_N$ and $R_{cor}$ are obtained by exchanging the indices H and N in (3) and (4)) (Goldman, 1984; Pervushin et al., 1997)

\[
R_H - R_{cor} = \frac{1}{T_{2}^{\text{eff}} (^1\text{H})} = \pi \delta v_2
\]

(6)

\[
R_N - R_{cor} = \frac{1}{T_{2}^{\text{eff}} (^{15}\text{N})} = \pi \delta v_1
\]

(7)
δν₂ is the full width at half-height of a peak with Lorentzian lineshape along the ω₂(¹H) dimension, and T²ₑffective(¹H) is the effective transverse relaxation rate of ¹H during acquisition. Similarly, δν₁ is the full width at half-height of a peak with Lorentzian lineshape along the ω₁(¹⁵N) dimension, and T²ₑffective(¹⁵N) is the effective transverse relaxation rate of ¹⁵N during evolution. In addition to the τc dependence, variations of the ¹H and ¹⁵N linewidths are expected to arise from the different microenvironments of the individual amide protons in the protein structure. For example, for a given τc value the linewidths of a ¹⁵N–¹H-moiety located in a β-sheet of a uniformly ¹⁵N,²H-labeled protein are expected to be much smaller than the linewidths in an α-helix of the same structure. The linewidth predictions based on this model do not give a rationale for the experimental observations in DHNA. For SR1 and GroEL, the linewidth predictions for a ²H-labeled protein fall within the range covered by the experimental values, although the experimental data extend over a much wider range. The wide distribution of linewidths in SR1 and GroEL could possibly be due to the presence of additional remote protons not considered in the model calculations, and to local conformational exchange. Note however that these measurements were obtained from observable cross peaks in the [¹⁵N,¹H]-CRIPT-TROSY spectra, which only account for 20% of the total number of expected resonances. As discussed in chapter 6.1, the reason for the observation of only a small proportion of the resonances is not known and it is thus possible that the global picture considering all expected resonances would be very different. For a more quantitative interpretation, this analysis should be repeated for the GroEL- or SR1-bound GroES, which yield nearly complete fingerprints in the [¹⁵N,¹H]-CRIPT-TROSY spectra.
3.6 15N AND 1H LINewidths in TROSY-type Experiments

Figure 3.7 NMR linewidths measured in 2D [15N,1H]-CRIPT-TROSY spectra of H2O solutions of the proteins DHNA, SR1 and GroEL at a 1H resonance frequency of 750 MHz.

In A. and B., respectively, the distributions of the 1H and 15N linewidths in these proteins are displayed by black bars which represent the number of peaks with a given linewidth along the vertical axis. For SR1 and GroEL only about 20% of the expected 15N–1H cross peaks were observed (see text). The straight lines represent linewidth predictions based on (6) and (7) for an amide proton located either in an α-helix of a uniformly 15N,2H-labeled protein, α(2H), in an α-helix of a uniformly 15N-labeled protein, α(1H), in a β-sheet of a uniformly 15N,2H-labeled protein, β(2H), or in a β-sheet of a 15N-labeled protein, β(1H). The following model parameters were used: \( r_{HN} = 1.04 \text{ Å} \), \( \Delta \sigma_H = 15 \text{ ppm} \), \( \vartheta = 10^\circ \), and dipole–dipole coupling with remote protons as follows\(^2\): α(2H): \( ^1\text{HN}(i-2), ^1\text{HN}(i-1), ^1\text{HN}(i+1) \) and \( ^1\text{HN}(i+2) \) at distances of 4.2, 2.8, 2.8, and 4.2 Å. α(1H): \( ^1\text{H}^a(i), ^1\text{H}^a(i-1), ^1\text{H}^a(i-2), ^1\text{H}^a(i-3), ^1\text{H}^a(i-4) \) and one \( ^1\text{H}^\beta(i) \) at distances of 2.6, 3.5, 4.4, 3.4, 4.2, 2.5 Å. β(2H): \( ^1\text{HN}(i-1), ^1\text{HN}(i+1) \) and \( ^1\text{HN}(j) \) at distances of 4.3, 4.3 and 3.3 Å. β(1H): \( ^1\text{H}^a(i), ^1\text{H}^a(i-1), ^1\text{H}^a(j), one ^1\text{H}^\beta(i) \) and one \( ^1\text{H}^\beta(i-1) \) at distances of 2.8, 2.2, 3.2, 2.5 and 3.2 Å.
3.7 Influence of the deuteration level of the proteins

The $[^{15}\text{N},^{1}\text{H}]$-correlation experiments developed, and mainly the $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY scheme, have been applied for measurements of proteins of the *E. coli* chaperonin system prepared according to different labeling strategies, achieving different levels of deuteration. Satisfactory results have so far been obtained with proteins that had been deuterated to the extent of at least ~80%, as obtained from D$_2$O-based minimal medium with protonated glucose. For example, correlation spectra of $^{15}$N-labeled and ~80% deuterated or perdeuterated SR1 were indistinguishable within the error range due to possible small protein concentration differences. In another example, comparison of spectra for the 500 kDa structure of SR1-bound GroES in ~80% deuterated and perdeuterated form show that the pattern of peaks is the same in both cases, with the exception of 5-6 missing peaks among the well-dispersed resonances (Figure 3.8). The resonances in the ~80% deuterated case seem generally weaker, with intensities corresponding to ~80% of the perdeuterated protein cross peak intensities for similar protein concentration. It is legitimate to ask whether this intensity difference is due to the lower deuteration level or to uncertainties in the protein concentration measurement and protein losses over the time course of the experiment. Production of proteins deuterated to the extent of ~80% generally offers a better yield than for perdeuterated proteins and thus a potential intensity loss could be compensated through higher protein quantities in the sample. However, it cannot be excluded that specific resonances in the protein experience faster relaxation due to their particular proton neighborhood and could be missing in the correlation spectra. Along this direction, the $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY scheme applied to $^{15}$N-labeled and ~50% deuterated GroEL as obtained from a pulse labeling experiment using $^{15}$N-labeled and deuterated celtone powder in H$_2$O (see section 2.2)
failed to detect most of the cross peaks observed in the perdeuterated protein (data not shown). Our experience shows that the extent of deuteration is critical for observation of resonances in 200-800 kDa structures and should be of at least ~80%.

To investigate the influence of deuteration of the protein partner on the ease of observation of the labeled component in a protein-protein complex, we prepared complexes of $^{15}$N-labeled and perdeuterated GroES bound to perdeuterated or natural isotope abundance SR1 or GroEL in the presence of ADP. Strikingly, there was no significant effect of chaperonin deuteration on the GroES [$^{15}$N,$^1$H]-correlation spectra and more than 85% of the expected resonances could be observed in both cases (data not shown). In the particular case of GroES-GroEL or GroES-SR1 interactions, the direct contacts are mediated by the side-chains of only 3 residues and the whole structure might be relatively dynamic. We thus cannot exclude that we do not see any effect of the deuteration of the chaperonin because (i) the peaks for residues that make direct contact to GroEL or SR1 are already absent from the spectra of the perdeuterated complexes, or (ii) this particular molecular assembly retains mobility at the protein-protein interface (see also section 5.2 and 5.3) and the residues mediating the contacts do not behave like a 472 or 972 kDa particle.

The discovery that deuterium labeling is needed only for the NMR-observed macromolecular component in GroES-GroEL complexes may greatly facilitate further applications for study of this system. In addition to a dramatic reduction of the labor and expense for the NMR sample preparation, the natural isotope abundance ensures that signals from unlabeled components are suppressed in the presently used experiments. Nevertheless, whether this observation is still valid for other protein-protein complexes remains to be tested.
Figure 3.8 Effect of deuteration on the detection of correlation spectra for a 472 kDa assembly. Comparison of the $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY spectra of $[U-^{13}\text{C}; \ U^{15}\text{N}; \ U-80\% \ ^{2}\text{H}]$-GroES (A) and $[U-^{15}\text{N}; \ U-97\% \ ^{2}\text{H}]$-GroES (B) both in complex with unlabeled SR1 and ATP. The main differences between the spectra are observed for 5-6 peaks that are missing or much weaker in the region delineated by the red rectangles. C. Cross-sections in the $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY spectra of A and B in the $^{1}\text{H}$ dimension. The cross-sections position are indicated by arrows with corresponding numbers. D. Cross-sections in the $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY spectra of A and B in the $^{15}\text{N}$ dimension.
3.8 Discussion

The results described in this chapter validate that the combined use of TROSY during the frequency labeling and acquisition periods and CRINEPT or CRIPT for polarization transfer provides a basis for NMR observations of structures with molecular weights of 500 kDa and beyond. Among the four 2D correlation experiments considered, 2D $^{15}$N,$^1$H-TROSY enables observation of resonance lines from amide groups in flexible chain ends or surface loops, whereas 2D $^{15}$N,$^1$H-CRIPT-TROSY suppresses these resonances and emphasizes the signals originating from well-structured parts of the proteins, which are subject to rotational diffusion on a time scale that is largely determined by the overall size of the structure. These two experiments thus provide a survey of all the resonance lines that may originate from a given structure. The two additional experiments, 2D $^{15}$N,$^1$H-CRINEPT-TROSY and 2D $^{15}$N,$^1$H-CRINEPT-HMQC-$^1$H-TROSY, which both use CRINEPT for the polarization transfers, provide a complete spectrum in a single experiment, but suffer from drawbacks on sensitivity and resolution, respectively. If the combination of the entire group of experiments is helpful for the spectral analysis, the conjunction of the $^{15}$N,$^1$H-TROSY and $^{15}$N,$^1$H-CRIPT-TROSY spectra is optimal to provide the complete information in good-appearance spectra. The filtering effects of the 2D $^{15}$N,$^1$H-TROSY and 2D $^{15}$N,$^1$H-CRIPT-TROSY experiments, as well as the multiplet patterns of the latter, provide direct information on dynamic properties of the molecular regions observed by these experiments.

In Figure 3.2 and Figure 3.4, and in the discussion on water suppression, we provide general guidelines for the set-up of the experiments of Figure 3.3 for studies of structures in the molecular weight range from 200 kDa to 800 kDa. We use these guidelines in our own work, but based on the experience gained with GroES, SR1 and
GroEL we would like to re-emphasize that it is worthwhile to optimize the set-up of these experiments anew for each new system studied, which may include a new calibration of the water flipback pulses, a new determination of the optimal polarization transfer time, and measurements of the longitudinal proton relaxation times.

Limitations in applications of the above-described experiments tend to arise from the fact that a percentage of the amide groups in large structures may not readily be observable in $^{[15\text{N},1\text{H}]}$-correlation experiments due to incomplete exchange of the deuterons from the D$_2$O-based expression medium against protons, even after long standing in H$_2$O solution. The implication is that among the large structures studied in this work, GroEL, SR1 and GroES bound to SR1 or GroEL, only the SR1-bound GroES had completely protonated amide groups, as evidenced by NMR spectra recorded with free GroES in H$_2$O solution (Figure 4.2). The experiments with the GroES/SR1-complex are therefore of particular importance for the validation of the presently described experimental procedures (see chapter 5).

Potential practical applications of the procedures described in this chapter include investigations of variable mobility in different parts of large macromolecular structures, chemical shift mapping of intermolecular contact areas in supramolecular structures, and the use of the 2D $^{[15\text{N},1\text{H}]}$-fingerprints as a basis for screening of potential drug molecules for binding to receptor structures in the size range of interest here. Variations of the chemical shifts in the $^{[15\text{N},1\text{H}]}$-fingerprints may further provide indications on conformational changes of a given macromolecular structure upon interactions with other molecules. In many of these potential applications the apparent limitation arising from lack of $^2\text{H} \rightarrow ^1\text{H}$ exchange from the core of large proteins might even be used in support of particular projects, by providing a
simplification of the spectra as well as initial indications for the assignment of the observed resonance lines to surface areas of the structure considered.

Further work along the same line should attempt implementation of the same principles used for these two-dimensional spectra in three-dimensional experiments. In particular, development of triple resonance experiments and NOESY schemes for very large molecules is of considerable interest for potential assignment of such macromolecular structures. This task remains a challenge since additional transfer steps inevitably imply sensitivity losses.
Backbone assignment of the GroES co-chaperonin

In this section, the sequence-specific backbone assignment of GroES is presented. Assignment of the backbone resonances has two objectives: First, identification of the resonances is a platform for the detailed study of GroES interactions with GroEL, which is the subject of chapter 5. Second, this assignment and further data analysis enable the determination of the secondary structure and of the overall fold of GroES, and may provide a basis for future work toward further structural characterization of GroES in solution.
4.1 NMR spectroscopy

Assignment of the $^1$H$^N$, $^{15}$N and $^{13}$C$^\alpha$ resonances of GroES was achieved with the triple resonance experiments $[^{15}$N,$^1$H]-TROSY-HNCA, $[^{13}$C]-ct-$[^{15}$N,$^1$H]-TROSY-HNCA and $[^{15}$N,$^1$H]-TROSY-HNCACB, and a 3D $[^1$H,$^1$H]-NOESY-$[^{15}$N,$^1$H]-TROSY spectrum. Table 4.1 lists the experiments with their main recording parameters. All triple resonance experiments were recorded using a 3.5 mM sample of $[^U$-$^{13}$C; $[^U$-$^{15}$N; $[^U$-$\sim$85% $^2$H]-GroES in NMR buffer (see section 3.1). The spectra were recorded on a Bruker DRX-800 spectrometer at 25°C. The 3D $[^1$H,$^1$H]-NOESY-$[^{15}$N,$^1$H]-TROSY spectrum was recorded with a 3.5

<table>
<thead>
<tr>
<th>Experiment</th>
<th>HNCA</th>
<th>$[^{13}$C]-ct-HNCA</th>
<th>HNCACB</th>
<th>NOESY-TROSY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb. data points</td>
<td>$t_1$</td>
<td>42 ($^{15}$N)</td>
<td>42 ($^{15}$N)</td>
<td>42 ($^{15}$N)</td>
</tr>
<tr>
<td></td>
<td>$t_2$</td>
<td>40 ($^{13}$C)</td>
<td>130 ($^{13}$C)</td>
<td>95 ($^{13}$C)</td>
</tr>
<tr>
<td></td>
<td>$t_3$</td>
<td>1024 ($^1$H)</td>
<td>1024 ($^1$H)</td>
<td>1024 ($^1$H)</td>
</tr>
<tr>
<td>Evolution time</td>
<td>$t_{1\text{max}}$</td>
<td>18 ms</td>
<td>18 ms</td>
<td>18 ms</td>
</tr>
<tr>
<td></td>
<td>$t_{2\text{max}}$</td>
<td>8 ms</td>
<td>27 ms</td>
<td>8 ms</td>
</tr>
<tr>
<td></td>
<td>$t_{3\text{max}}$</td>
<td>71 ms</td>
<td>71 ms</td>
<td>71 ms</td>
</tr>
<tr>
<td>Spectral width</td>
<td>$\omega_1$</td>
<td>30 ppm</td>
<td>30 ppm</td>
<td>30 ppm</td>
</tr>
<tr>
<td></td>
<td>$\omega_2$</td>
<td>25 ppm</td>
<td>25 ppm</td>
<td>59 ppm</td>
</tr>
<tr>
<td></td>
<td>$\omega_3$</td>
<td>18 ppm</td>
<td>18 ppm</td>
<td>18 ppm</td>
</tr>
<tr>
<td>Recording time</td>
<td></td>
<td>30 h</td>
<td>102 h</td>
<td>77 h</td>
</tr>
<tr>
<td>Reference</td>
<td></td>
<td>Salzmann et al., 1998</td>
<td>Salzmann et al., 1999d</td>
<td>Salzmann et al., 1999b</td>
</tr>
<tr>
<td>Figure</td>
<td></td>
<td>4.3</td>
<td>4.4</td>
<td>4.5</td>
</tr>
</tbody>
</table>
mM sample of \([U-^{15}N; U\rightarrow97\%\, ^2H]-\text{GroES}\) in NMR buffer at pH 6.15 (see section 3.1) on a Bruker DRX-750 instrument at 25°C. The details of the parameter settings are given in Table 4.1 and in the figure captions of Figures 4.3-4.6.

4.2 Assignment procedure

GroES is a homo-heptamer of 97-residue subunits (sequence in Figure 4.1). Its three-dimensional structure has been determined by X-ray crystallography (Hunt et al., 1996) and shows a seven-fold symmetry. Our experiments confirmed that the symmetrically arranged subunits generate only one set of resonances in the NMR spectra, which correspond to a single subunit of 97 residues.

In the 2D \([^{15}N, \, ^1H]\)-TROSY spectrum of \([U-^{15}N; U-^2H]-\text{GroES}\), 89 of the 94 expected resonances were identified (Figure 4.2). Observation of virtually complete \(^{15}N-^1H\) correlation spectra indicates that the N-D/N-H exchange is complete in GroES after solubilisation of the protein in H\(_2\)O buffers. This contrasts to problems reported for other homo-oligomeric proteins (Salzmann, 1999; chapter 6.1) where peaks were missing in the spectra due to incomplete N-D/N-H exchange after producing proteins in perdeuterated medium.

Sequential assignments for the GroES backbone resonances were primarily obtained from the HNCA experiment. Figure 4.3 shows \([\omega_2(13C), \omega_3(1H)]\) strips of this 3D \([^{15}N, \, ^1H]\)-TROSY-HNCA for residues 74 to 85, in which the sequential connectivities established are indicated with straight lines. In cases where the HNCA experiment did not permit unambiguous assignment of the sequential connectivities due to \(^{13}C\)\(^\alpha\) chemical shift degeneracy, the resonances could be distinguished and connectivities were established in the \([^{13}C]\)-ct-\([^{15}N, \, ^1H]\)-TROSY-HNCA
4.2 Assignment Procedure

Figure 4.1 Amino acid sequence of GroES, indicated in the one-letter code (Hemmingsen et al., 1988)

![Amino acid sequence of GroES](image)

**Figure 4.2** $[^{15}\text{N},^{1}\text{H}]-\text{TROSY}$ spectrum of $[^{15}\text{N},^{2}\text{H}}\text{-GroES}$. 89 of the 94 expected resonances are observed. The spectrum was recorded as described in Figure 5.1 A.
spectrum that benefits from a higher resolution in the \( \omega_2^{13C} \) dimension (Figure 4.4). Alternatively, the \( ^{13}C^\beta \) chemical shift dispersion could be used in the \([^{15}N,^1H]\)-TROSY-HNCACB spectrum to confirm the \( ^1H^N,^{13}C^\alpha \) connectivities with corresponding \( ^1H^N,^{13}C^\beta \) connectivities (Figure 4.5). Finally, the search for corresponding NOE cross peaks in the 3D \([^1H,^1H]\)-NOESY-[^{15}N,^1H]-TROSY experiment helped confirming the unambiguous assignment of sequential connectivities (Figure 4.6). Supplementary information containing the details of the strips that lead to resonance assignment in these spectra is presented in Appendix 2.

The sequence-specific assignment of GroES was supported by the use of the program MAPPER (Güntert et al., 2000). This program makes use of the \( ^{13}C^\alpha \) and/or \( ^{13}C^\beta \) chemical shift of manually connected residues in a stretch to map the amino acid segment to matching positions along the protein sequence. Sequential connectivities for the mobile loop segment of residues 17-32 in GroES was easily obtained since these residues give rise to very intense resonances in all triple resonances experiments. The resulting 21-residue stretch could only be mapped to a unique position in the GroES amino acid sequence and provided a starting point for the sequence-specific assignment. Eventually, input from 13 amino acid segments of 3 to 21 residues into MAPPER resulted in one self-consistent sequence-specific assignment. Residues 88-94 for which connectivities were not identified in the triple resonance experiments were assigned using the NOE pattern in the \([^1H,^1H]\)-NOESY-[^{15}N,^1H]-TROSY spectrum.

The sequence-specific assignment of GroES \( ^1H^N,^{15}N \) and \( ^{13}C^\alpha \) backbone resonances was completed to 96%. The remaining 4 residues could not be identified due to missing strips for Ile 11, Ile 45, and Ile 66. In addition, 50% of the \( ^{13}C^\beta \) resonances could also be assigned.
Further work will focus on collecting more resonance assignments for GroES. First, detection of the $^{13}$CO resonances in a $^{15}$N, $^1$H-TROSY-HNCO experiment might help identifying the missing resonances mentioned above, thanks to the inherent sensitivity of this scheme. Besides, in order to assign more of the $^{13}$C$\beta$ resonances than obtained so far, a HNCACB experiment should be recorded with parameter settings.

**Figure 4.3** $[\omega_2(^{13}$C), $\omega_3(^1$H)] strips from the 3D $[^{15}$N, $^1$H]-TROSY-HNCA spectrum of $^{13}$C, $^{15}$N, $^2$H-labeled GroES. The sequential connectivities are shown with lines and the sequence-specific assignment obtained are indicated as the one-letter code for the amino acid and the residue number in the amino acid sequence. In the scheme of the upper right corner, the nuclei correlated in the triple resonance experiment are indicated with filled circles. The experiment was recorded with a 3.5 mM solution of GroES (see text) according to Salzmann et al., 1998. The $[^{15}$N,$^1$H]-TROSY-HNCA spectrum was recorded for 30 h with a data size of 42 x 40 x 1024 complex points, $t_{1\text{max}} = 17.5$ ms, $t_{2\text{max}} = 8.0$ ms, $t_{3\text{max}} = 70.7$ ms. The INEPT transfer time was 3.5 ms and the relaxation delay 1s.
such as to transfer as much magnetization as possible to the $^{13}\text{C}\beta$ atoms. Knowledge of more $^{13}\text{C}\beta$ chemical shift values will help refining the analysis of the secondary structure elements with the parameter $\Delta\delta^{13}\text{C}\alpha - \Delta\delta^{13}\text{C}\beta$ (see section 4.3). In a next step, side-chain assignment could be attempted using newly-developed experiments for assignment of deuterated proteins (Lin and Wagner, 1999; Nietlispach et al., 1996)

**Figure 4.4** $[\omega_2^{(13}\text{C}), \omega_3^{(1}\text{H})]$ strips from the 3D $[^{13}\text{C}]$-constant time $[^{15}\text{N}, \: ^{1}\text{H}]$-TROSY-HNCA spectrum of $^{13}\text{C}, \: ^{15}\text{N}, \: ^{2}\text{H}$-labeled GroES. The improvement of resolution due to the constant time evolution on $^{13}\text{C}$ is visible upon comparison with Figure 4.3. The sequential connectivities are shown with lines and the sequence-specific assignment obtained are indicated as the one-letter code for the amino acid and the residue number in the amino acid sequence. The experiment was recorded with a 3.5 mM solution of GroES (see text) according to Salzmann et al., 1999d. The spectrum was recorded for 102 h with a data size of 42 x 130 x 1024 complex points, $t_{1\text{max}} = 17.5$ ms, $t_{2\text{max}} = 27.0$ ms, $t_{3\text{max}} = 70.7$ ms. The INEPT transfer time was 3.5 ms and the relaxation delay 1s.
Figure 4.5 $[\omega_2(13C), \omega_3(1H)]$ strips from the 3D $[^{15}N, ^1H]$-TROSY-HNCACB spectrum of $^{13}C$, $^{15}N$, $^2H$-labeled GroES. The sequential connectivities are shown with lines, both for the $C_\alpha$ cross peaks and the $C_\beta$ cross peaks whenever possible (see text). The sequence-specific assignment obtained are indicated as the one-letter code for the amino acid and the residue number in the amino acid sequence. The experiment was recorded with a 3.5 mM solution of GroES (see text) according to Salzmann et al., 1999b. The $[^{15}N,^1H]$-TROSY-HNCACB spectrum was recorded for 77 h with a data size of 42 x 95 x 1024 complex points, $t_{1\max} = 17.5$ ms, $t_{2\max} = 8.0$ ms, $t_{3\max} = 70.7$ ms. The INEPT transfer time was 3.5 ms and the relaxation delay 1 s.

and partial deuteration strategies.
Figure 4.6  $[\omega_1(1\text{H}), \omega_3(1\text{H})]$ strips from the 3D $[^1\text{H},^1\text{H}]-\text{NOESY}-[^{15}\text{N},^1\text{H}]-\text{TROSY}$ spectrum of $^2\text{H},^{15}\text{N}$-labeled GroES. The NOE connectivities are shown with dashed lines and the sequence-specific assignments are indicated as the one-letter code for the amino acid and the residue number in the amino acid sequence. The network of connectivities of residues 87-90 arise from the presence of a short helix. In the other strips, long-range NOE’s across the β-strand can be observed and are labeled with red lettering. NOE’s to V10, V12 and R14 are intersubunit NOE’s. The experiment was recorded for 57 hours with a 3.5 mM solution of GroES (see text) according to Pervushin et al., 2000, acquired data size of 180 x 30 x 1024 complex points, $t_{\text{1max}} = 18.0$ ms, $t_{\text{2max}} = 13.0$ ms, $t_{\text{3max}} = 98.0$ ms, mixing time 160.0 ms. The INEPT transfer time was 3.4 ms and the relaxation delay 0.9 s.
4.3 Secondary structure of GroES

Backbone sequential assignment can provide information on the secondary structure of the protein based on chemical shift analysis. Observed deviations of the $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ chemical shift from their residue-specific average value in random coil conformation is used as basis for the identification of secondary structure elements along the protein sequence (Spera and Bax, 1991; Wishart and Sykes, 1994). In Figure 4.7, the parameter $\Delta \delta^{13}\text{C}^\alpha = \delta^{13}\text{C}^\alpha(\text{observed}) - \delta^{13}\text{C}^\alpha(\text{random coil})$ is plotted versus the amino acid sequence. Several regions are observed for which consecutive positions show negative $\Delta \delta^{13}\text{C}^\alpha$ values, indicating a $\beta$-sheet conformation of these positions in the protein. The identified $\beta$-sheet strands correlate well with the secondary structure elements observed in the crystal structure (Hunt et al., 1996) shown schematically at the top of Figure 4.7. The short segment adopting $3\times10$ helical conformation in GroES is also reflected in the chemical shift data as a short stretch with positive $\Delta \delta^{13}\text{C}^\alpha$. In addition, the present analysis suggests that the region comprising residues 17 to 32 is unstructured. This is consistent with the previously reported dynamically disordered nature of this segment in GroES (Hunt et al., 1996; Landry et al., 1993; Shewmaker et al., 2001) and with the overall much higher intensity observed in NMR for these residues (Figure 4.2). This region is generally referred to as the mobile loop of GroES.

4.4 Calculation of the 3D-fold of GroES

Assignment of the $^\text{H}_\text{N}$-$^\text{H}_\text{N}$ NOE cross peaks in the 3D $[^1\text{H},^1\text{H}]$-NOESY-$[^{15}\text{N},^1\text{H}]$-TROSY was completed combining manual assignments with those resulting from a calculation using the program CANDID (Herrmann et al., 2002). The respective intensities of these cross peaks were then used as distance constraints in a structure
calculation to test the self-consistency of the data set and attempt delineation of the fold of GroES.

The structure calculation was performed using the program DYANA (Güntert et al., 1997) for one single subunit of GroES. It included distance constraints calculated from the H^N-H^N NOE peak intensities as well as constraints for the backbone torsion angles Φ and Ψ derived from the ^13C chemical shifts using the program TALOS (Cornilescu et al., 1999). Intersubunit NOE’s were eliminated. The final input consisted of

**Figure 4.7** Identification of the secondary structure elements from analysis of the ^13Cα chemical shift (Spera and Bax, 1991; Wishart and Sykes, 1994). ΔδCα was obtained as the difference between experimental chemical shifts in GroES and the corresponding random coil values and plotted vs. the amino acid sequence (Schwarzinger et al., 2000; Wishart and Sykes, 1994). The parameter ΔδCα at each position represent an average over three consecutive residues. Negative and positive values of ΔδCα indicate that a residue adopts β-sheet-like and helix-like conformation, respectively. The secondary structure elements present in the crystal structure of GroES (Hunt et al., 1996) are drawn schematically above the plot. Note that β-strand 1 and 9 interact with a neighboring subunit.
157 upper-limit distance constraints and 88 angle constraints (Table 4.2, Calculation A). The experimental constraints only concern the backbone atoms since the protein was perdeuterated and their number is limited. One expects however that they are sufficient to define the overall fold of the GroES subunit. For comparison, the same calculation was performed including additional distance constraints for H-bonds obtained from the crystal structure of GroES (calculation B). The results of these two calculations are shown in Figure 4.8. The resulting characteristic parameters are listed in Table 4.2. RMSD values apply to residues 6 to 15 and 34 to 91, excluding the mobile loop and the undefined N- and C- termini (see below).

Table 4.2 Structure calculation for a GroES subunit with the program DYANA

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Calculation A</th>
<th>Calculation B</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE upper distance constraints</td>
<td>157</td>
<td>157</td>
</tr>
<tr>
<td>Dihedral angle constraints</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>Constrained H-bonds</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td>Residual target function, Å²</td>
<td>0.24 +/- 0.01</td>
<td>0.94 +/- 0.31</td>
</tr>
<tr>
<td>Residual NOE violations, number &gt;0.1 Å</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Maximum violation, Å</td>
<td>0</td>
<td>0.53</td>
</tr>
<tr>
<td>Residual angle violations &gt; 2.0°</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Average RMSD for residues 6-15, 34-91, Å</td>
<td>3.99 +/- 1.59</td>
<td>1.64 +/- 0.36</td>
</tr>
</tbody>
</table>

Figure 4.8 shows a bundle of the 20 best DYANA conformers obtained for one subunit of GroES. The absence of constraint violations in the best 20 conformers confirms the self-consistency of the assignment data. Although the bundle of Figure 4.8 A shows low precision, the overall backbone fold of GroES can be recognized. The irregular β-barrel is formed in these conformers (central part of the figure) and the mobile
loop (lower left corner) samples many varying conformations. Figure 4.8 provides a comparison of the NMR conformers with the smallest residual target function value, in red (calculation A) and blue (calculation B), to the backbone trace of the GroES crystal structure, in green (Hunt et al., 1996). The RMSD for backbone atoms in the structured part of GroES is 4.6 Å between the mean NMR structure and the crystal structure.

Residues 2-5 and 92-96 at the N- and C- terminii of the polypeptide chains are part of a β-sheet formed with a strand of the neighboring subunit in the GroES assembly. Since the structure calculation did not account for neighboring subunits, these regions lack the restraints defining their conformations and appear as undefined. Eliminating this artefact of calculation requires the involvement of at least 3 subunits in the calculation and imply development of the calculation program that are beyond the scope of this preliminary test. As expected, inclusion of additional constraints from H-bonds significantly improves the definition of the fold. The secondary structure information obtained from the NMR data, the long-range NOE’s and the fold calculation A should enable to derive the topology of the GroES subunit. From this topology, one could in turn derive constraints for H-bonds. It should thus be possible to achieve a similar result without information from the crystal structure.

A reasonably well defined fold from H\(^N\)-H\(^N\) distance and backbone angle constraints has been reported for the β-barrel structure of OmpX (Fernandez et al., 2001). Comparison of these results with those obtained for GroES shows that the fold obtained for GroES seems of slightly lower precision. This probably originates from the less regular β-sheet structure and the missing intersubunits constraints. To improve the precision of the calculated fold in future work, the particularly attractive
approach involving selective protonation of methyl groups on a deuterated background could also be applied (Gardner and Kay, 1997, 1998; Goto et al., 1999; Rosen et al., 1996): the observed methyl resonances can be assigned using experiments that correlate the methyl group with the sequential amide resonances (Hilty, C., Fernandez, C., Wider, G., Wüthrich, K., manuscript in preparation; Lin and Wagner, 1999). Subsequent detection of CH$_3$-CH$_3$ and CH$_3$-H$^N$ NOE’s in $^{13}$C- and $^{15}$N-edited NOESY experiments can be used as a source of additional distance constraints for calculation of a refined fold. Alternatively, measurement of residual dipolar couplings could provide additional constraints to further improve the precision of the obtained structure.

Although a crystal structure of GroES is available (Hunt et al., 1996), it remains of interest to study GroES in solution and address questions relating to its dynamics and its interactions with GroEL. Moreover, further efforts to assign the GroES resonances and investigate its structure using solution NMR are valuable for the development and test of NMR methods applicable to large and deuterated proteins.
4. BACKBONE ASSIGNMENT OF THE GROES CO-CHAPERONIN

Figure 4.8 3D fold obtained from the DYANA calculation for a GroES subunit.
A. Superposition of the 20 best DYANA conformers (see Table 4.2). The conformer with the smallest residual target function value is shown with a thicker line in the bundle. The conformers were aligned according to best global fit of the backbone atoms for residues 6-15 and 34-91.
B. Superposition of the 20 best DYANA conformers obtained from a calculation including additional H-bond constraints (see Table 4.2).
C. Superposition of the conformers with the smallest residual target function value from calculations without (red) and with (blue) additional H-bond constraints with the backbone trace from the GroES crystal structure (green, Hunt et al., 1996).

The NMR structures were calculated using the torsion angle dynamics protocol of DYANA and were started each from 200 random conformers. The 20 conformers with the smallest residual target function value were used to define the final fold. The figures were drawn using Molmol (Koradi et al., 1996).
4.5 Chemical shift list

Table 4.3 lists the chemical shift value for $^1$H,$^15$N,$^{13}$C$^\alpha$, $^{13}$C$^\beta$ (when available), and side-chains $^1$H and $^{15}$N (when available) in GroES. All chemical shifts are referenced to DSS.

<table>
<thead>
<tr>
<th>Residue</th>
<th>$\delta(^1$H$\uparrow$) (ppm)</th>
<th>$\delta(^{15}$N$\uparrow$) (ppm)</th>
<th>$\delta(^{13}$C$^\alpha$) (ppm)</th>
<th>$\delta(^{13}$C$^\beta$) (ppm)</th>
<th>others (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met 1</td>
<td>54.9</td>
<td></td>
<td></td>
<td>32.2</td>
<td></td>
</tr>
<tr>
<td>Asn 2</td>
<td>8.90</td>
<td>123.0</td>
<td>52.6</td>
<td>39.3</td>
<td>$\delta(^{15}$N$\delta^2)$ 114.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\delta(^1$H$\delta^2)$ 6.82, 7.67</td>
</tr>
<tr>
<td>Ile 3</td>
<td>7.82</td>
<td>119.8</td>
<td>59.3</td>
<td>38.8</td>
<td></td>
</tr>
<tr>
<td>Arg 4</td>
<td>9.18</td>
<td>127.9</td>
<td>51.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 6</td>
<td>8.14</td>
<td>114.3</td>
<td>53.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>His 7</td>
<td>9.20</td>
<td>117.9</td>
<td>56.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp 8</td>
<td>8.98</td>
<td>120.4</td>
<td>52.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 9</td>
<td>7.68</td>
<td>119.6</td>
<td>55.2</td>
<td></td>
<td>$\delta(^{15}$N$\epsilon)$ 86.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\delta(^1$H$\epsilon)$ 7.20</td>
<td></td>
</tr>
<tr>
<td>Val 10</td>
<td>8.79</td>
<td>115.0</td>
<td>59.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 12</td>
<td>9.40</td>
<td>129.0</td>
<td>59.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys 13</td>
<td>8.91</td>
<td>125.4</td>
<td>53.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 14</td>
<td>9.29</td>
<td>129.5</td>
<td>57.8</td>
<td>30.3</td>
<td></td>
</tr>
<tr>
<td>Lys 15</td>
<td>8.73</td>
<td>126.8</td>
<td>54.5</td>
<td>32.5</td>
<td></td>
</tr>
<tr>
<td>Glu 16</td>
<td>8.63</td>
<td>124.2</td>
<td>55.6</td>
<td>29.3</td>
<td></td>
</tr>
<tr>
<td>Val 17</td>
<td>8.43</td>
<td>121.9</td>
<td>61.8</td>
<td>32.0</td>
<td></td>
</tr>
<tr>
<td>Glu 18</td>
<td>8.52</td>
<td>125.6</td>
<td>56.1</td>
<td>29.5</td>
<td></td>
</tr>
<tr>
<td>Thr 19</td>
<td>8.30</td>
<td>117.5</td>
<td>61.5</td>
<td>69.4</td>
<td></td>
</tr>
<tr>
<td>Lys 20</td>
<td>8.40</td>
<td>124.4</td>
<td>55.8</td>
<td>32.3</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.3 Chemical shift list of $^{13}$C,$^{15}$N,$^2$H-GroES from *E. coli*

<table>
<thead>
<tr>
<th>Residue</th>
<th>$\delta^{(1\text{HN})}$ (ppm)</th>
<th>$\delta^{(15\text{N})}$ (ppm)</th>
<th>$\delta^{(13\text{C}^\alpha)}$ (ppm)</th>
<th>$\delta^{(13\text{C}^\beta)}$ (ppm)</th>
<th>others (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser 21</td>
<td>8.38</td>
<td>117.5</td>
<td>57.7</td>
<td>63.8</td>
<td></td>
</tr>
<tr>
<td>Ala 22</td>
<td>8.44</td>
<td>126.3</td>
<td>52.5</td>
<td>18.1</td>
<td></td>
</tr>
<tr>
<td>Gly 23</td>
<td>8.33</td>
<td>108.3</td>
<td>45.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly 24</td>
<td>8.14</td>
<td>108.9</td>
<td>44.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 25</td>
<td>7.78</td>
<td>120.7</td>
<td>60.4</td>
<td>37.7</td>
<td></td>
</tr>
<tr>
<td>Val 26</td>
<td>8.27</td>
<td>126.1</td>
<td>62.0</td>
<td>31.7</td>
<td></td>
</tr>
<tr>
<td>Leu 27</td>
<td>8.44</td>
<td>127.7</td>
<td>54.5</td>
<td>41.2</td>
<td></td>
</tr>
<tr>
<td>Thr 28</td>
<td>8.08</td>
<td>114.9</td>
<td>61.3</td>
<td>69.7</td>
<td></td>
</tr>
<tr>
<td>Gly 29</td>
<td>8.46</td>
<td>111.4</td>
<td>45.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 30</td>
<td>8.27</td>
<td>117.0</td>
<td>58.4</td>
<td>63.4</td>
<td></td>
</tr>
<tr>
<td>Ala 31</td>
<td>8.27</td>
<td>125.7</td>
<td>52.3</td>
<td>18.2</td>
<td></td>
</tr>
<tr>
<td>Ala 32</td>
<td>7.95</td>
<td>122.4</td>
<td>51.8</td>
<td>18.2</td>
<td></td>
</tr>
<tr>
<td>Ala 33</td>
<td>7.98</td>
<td>123.3</td>
<td>51.9</td>
<td>18.4</td>
<td></td>
</tr>
<tr>
<td>Lys 34</td>
<td>8.16</td>
<td>120.5</td>
<td>55.3</td>
<td>32.7</td>
<td></td>
</tr>
<tr>
<td>Ser 35</td>
<td>8.67</td>
<td>117.9</td>
<td>57.8</td>
<td>63.9</td>
<td></td>
</tr>
<tr>
<td>Thr 36</td>
<td>8.26</td>
<td>117.9</td>
<td>60.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 37</td>
<td>7.90</td>
<td>119.9</td>
<td>53.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly 38</td>
<td>8.95</td>
<td>107.1</td>
<td>45.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 39</td>
<td>8.89</td>
<td>121.6</td>
<td>53.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 40</td>
<td>9.19</td>
<td>126.8</td>
<td>63.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 41</td>
<td>9.51</td>
<td>130.6</td>
<td>55.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala 42</td>
<td>7.54</td>
<td>118.1</td>
<td>51.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 43</td>
<td>8.81</td>
<td>115.4</td>
<td>59.6</td>
<td>35.1</td>
<td></td>
</tr>
<tr>
<td>Gly 44</td>
<td>8.10</td>
<td>111.3</td>
<td>44.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn 45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>54.0</td>
</tr>
<tr>
<td>Gly 46</td>
<td>7.88</td>
<td>105.6</td>
<td>43.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg 47</td>
<td>8.01</td>
<td>119.8</td>
<td>54.9</td>
<td>31.3</td>
<td></td>
</tr>
<tr>
<td>Ile 48</td>
<td>8.38</td>
<td>124.4</td>
<td>60.3</td>
<td>36.7</td>
<td></td>
</tr>
</tbody>
</table>
## 4.5 Chemical Shift List

**Table 4.3 Chemical shift list of $^{13}$C, $^{15}$N, $^2$H-GroES from E. coli**

<table>
<thead>
<tr>
<th>Residue</th>
<th>$\delta^{(\mathrm{HN})}$ (ppm)</th>
<th>$\delta^{(\mathrm{N})}$ (ppm)</th>
<th>$\delta^{(\mathrm{C}^\alpha)}$ (ppm)</th>
<th>$\delta^{(\mathrm{C}^\beta)}$ (ppm)</th>
<th>others (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu 49</td>
<td>8.89</td>
<td>129.5</td>
<td>53.6</td>
<td>40.8</td>
<td></td>
</tr>
<tr>
<td>Glu 50</td>
<td>8.73</td>
<td>120.8</td>
<td>58.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn 51</td>
<td>7.65</td>
<td>114.0</td>
<td>52.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly 52</td>
<td>8.19</td>
<td>108.9</td>
<td>44.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 53</td>
<td>7.70</td>
<td>120.8</td>
<td>55.8</td>
<td>29.3</td>
<td></td>
</tr>
<tr>
<td>Val 54</td>
<td>8.22</td>
<td>120.4</td>
<td>60.4</td>
<td>32.1</td>
<td></td>
</tr>
<tr>
<td>Lys 55</td>
<td>9.09</td>
<td>130.8</td>
<td>51.3</td>
<td>32.7</td>
<td></td>
</tr>
<tr>
<td>Pro 56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 57</td>
<td>8.52</td>
<td>119.2</td>
<td>54.3</td>
<td>40.2</td>
<td></td>
</tr>
<tr>
<td>Asp 58</td>
<td>9.54</td>
<td>122.0</td>
<td>54.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 59</td>
<td>7.05</td>
<td>113.4</td>
<td>60.0</td>
<td>33.0</td>
<td></td>
</tr>
<tr>
<td>Lys 60</td>
<td>8.83</td>
<td>122.4</td>
<td>53.5</td>
<td>35.0</td>
<td></td>
</tr>
<tr>
<td>Val 61</td>
<td>8.23</td>
<td>120.0</td>
<td>65.4</td>
<td>30.3</td>
<td></td>
</tr>
<tr>
<td>Gly 62</td>
<td>8.92</td>
<td>117.2</td>
<td>44.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp 63</td>
<td>8.01</td>
<td>122.1</td>
<td>55.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 64</td>
<td>8.87</td>
<td>123.1</td>
<td>57.8</td>
<td>34.8</td>
<td></td>
</tr>
<tr>
<td>Val 65</td>
<td>9.13</td>
<td>117.4</td>
<td>57.4</td>
<td>34.2</td>
<td></td>
</tr>
<tr>
<td>Ile 66</td>
<td></td>
<td></td>
<td>59.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe 67</td>
<td>8.32</td>
<td>123.9</td>
<td>54.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn 68</td>
<td>8.89</td>
<td>118.3</td>
<td>51.7</td>
<td>39.2</td>
<td>$\delta^{(15N\delta^2)}$ 114.3, $\delta^{(1H\delta^2)}$ 6.93, 7.80</td>
</tr>
<tr>
<td>Asp 69</td>
<td>8.41</td>
<td>125.9</td>
<td>52.9</td>
<td>40.0</td>
<td></td>
</tr>
<tr>
<td>Gly 70</td>
<td>7.73</td>
<td>109.4</td>
<td>44.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr 71</td>
<td>8.31</td>
<td>120.6</td>
<td>59.4</td>
<td>37.2</td>
<td></td>
</tr>
<tr>
<td>Gly 72</td>
<td>8.60</td>
<td>109.5</td>
<td>44.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 73</td>
<td>7.52</td>
<td>122.5</td>
<td>62.3</td>
<td>31.0</td>
<td></td>
</tr>
<tr>
<td>Lys 74</td>
<td>8.59</td>
<td>128.7</td>
<td>53.6</td>
<td>33.1</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.3 Chemical shift list of $^{13}\text{C}$,$^{15}\text{N}$,$^2\text{H}$-GroES from E. coli

<table>
<thead>
<tr>
<th>Residue</th>
<th>$\delta(1\text{HN})$ (ppm)</th>
<th>$\delta(15\text{N})$ (ppm)</th>
<th>$\delta(13\text{C}^\alpha)$ (ppm)</th>
<th>$\delta(13\text{C}^\beta)$ (ppm)</th>
<th>others (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser 75</td>
<td>8.56</td>
<td>120.0</td>
<td>57.1</td>
<td>63.5</td>
<td></td>
</tr>
<tr>
<td>Glu 76</td>
<td>8.96</td>
<td>127.5</td>
<td>54.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys 77</td>
<td>8.63</td>
<td>122.6</td>
<td>54.9</td>
<td>32.9</td>
<td></td>
</tr>
<tr>
<td>Ile 78</td>
<td>8.58</td>
<td>125.0</td>
<td>59.4</td>
<td>38.4</td>
<td></td>
</tr>
<tr>
<td>Asp 79</td>
<td>9.06</td>
<td>128.8</td>
<td>55.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn 80</td>
<td>8.55</td>
<td>113.3</td>
<td>54.0</td>
<td>$\delta(15\text{N}^{82})$ 114.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\delta(1\text{H}^{82})$ 6.86,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.55</td>
<td></td>
</tr>
<tr>
<td>Glu 81</td>
<td>7.93</td>
<td>120.5</td>
<td>54.7</td>
<td>31.2</td>
<td></td>
</tr>
<tr>
<td>Glu 82</td>
<td>8.58</td>
<td>123.9</td>
<td>56.1</td>
<td>29.6</td>
<td></td>
</tr>
<tr>
<td>Val 83</td>
<td>8.95</td>
<td>121.5</td>
<td>59.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 84</td>
<td>9.03</td>
<td>120.0</td>
<td>52.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 85</td>
<td>9.09</td>
<td>120.9</td>
<td>60.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met 86</td>
<td>8.68</td>
<td>124.2</td>
<td>53.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 87</td>
<td>9.39</td>
<td>116.3</td>
<td>57.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu 88</td>
<td>8.77</td>
<td>126.7</td>
<td>59.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser 89</td>
<td>8.28</td>
<td>111.4</td>
<td>60.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp 90</td>
<td>7.76</td>
<td>120.1</td>
<td>55.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 91</td>
<td>8.29</td>
<td>119.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu 92</td>
<td>8.86</td>
<td>125.3</td>
<td>55.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala 93</td>
<td>7.45</td>
<td>115.1</td>
<td>50.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile 94</td>
<td>8.65</td>
<td>121.1</td>
<td>59.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val 95</td>
<td>9.00</td>
<td>127.4</td>
<td>61.3</td>
<td>31.0</td>
<td></td>
</tr>
<tr>
<td>Glu 96</td>
<td>8.54</td>
<td>128.4</td>
<td>55.3</td>
<td>30.1</td>
<td></td>
</tr>
<tr>
<td>Ala 97</td>
<td>8.03</td>
<td>130.4</td>
<td>53.7</td>
<td>19.2</td>
<td></td>
</tr>
</tbody>
</table>
GroES–GroEL interactions mapped by solution NMR

In the previous chapters, the necessary protocols for labeled chaperonins preparation have been described and assessed, and appropriate NMR schemes developed for observation of these large structures. The present chapter establishes the utility of these techniques for studies of macromolecular complexes with sizes up to 900 kDa. Together with the knowledge of the backbone assignment of GroES as described in chapter 4, these experiments could be applied to study the interactions of GroES with GroEL. The coming sections describe these investigations and their results.
5.1 $^{15}$N,$^{2}$H-GroES–chaperonin complex formation followed by NMR

GroES constitutes a system of choice for assessing the performances of the NMR techniques described in chapter 3. Its size of 72 kDa is within the range of observable proteins with the established [$^{15}$N,$^{1}$H]-TROSY scheme and provides a control for detectability of its $^{15}$N-$^{1}$H moieties with all chosen experiments. GroES can then be incorporated into a much larger structure upon binding to GroEL or to its single-ring variant SR1, and thus studied at three different structure sizes, namely 72, 472 and 872 kDa. Moreover, in contrast to GroEL and SR1, perdeuterated and $^{15}$N-labeled GroES yields a nearly complete fingerprint in [$^{15}$N,$^{1}$H]-correlation spectra. Studies of GroES in complexes with GroEL or SR1 thus provide an opportunity for unambiguous tests of the potential of TROSY/CRINEPT-based NMR experiments with structures in the size range up to 900 kDa.

For free GroES, 89 out of 94 expected backbone amide $^{15}$N-$^{1}$H cross peaks were detected in 2D [$^{15}$N,$^{1}$H]-TROSY and 2D [$^{15}$N,$^{1}$H]-CRIP-TROSY experiments (Figure 5.1, A and B), and sequence-specific backbone resonance assignments were obtained using TROSY-triple resonance experiments (see chapter 4). Complexes were obtained in four experiments, by addition to $^{15}$N,$^{2}$H-labeled GroES of either perdeuterated and $^{15}$N-depleted GroEL or SR1, or to natural isotope-abundance GroEL or SR1 in the presence of ADP. The comparison of $^{15}$N-depleted vs. natural isotope-abundance chaperonin was used to obtain experimental proof that all the resonance lines observed in the complexes originated from GroES. Perdeuteration of chaperonin was carried out to assess whether this would improve the observation of GroES in the complexes. No significant effect of chaperonin deuteration on the GroES [$^{15}$N,$^{1}$H]-correlation spectra was detected (see also section
3.7), and therefore most of the spectra shown in this chapter were recorded using GroES complexes with unlabeled chaperonins.

Complex formation during stepwise addition of the deuterated chaperonins was evident from decrease of the GroES TROSY cross peak intensities. Examination of spectra measured at different stages of the titration showed how certain cross peaks gradually disappear while others appear in other regions of the spectra upon stepwise addition of GroEL or SR1 to GroES (data not shown). After addition of one equivalent of either SR1 or GroEL, nearly all of the TROSY resonances observed with free GroES had disappeared (compare Figure 5.1, A and C, and Figure 5.3, A and D). The peaks are almost certainly missing because transverse relaxation in the complexes is too fast to allow observation of the NMR signals with the TROSY experimental scheme (see chapter 3). Yet the TROSY spectra of the complexes contain a small number of peaks, attributed to regions of GroES that remain mobile (Figure 5.1 C and Figure 5.3 D, see discussion below). One of these resonances was identified as that of the C-terminal residue Ala 97 of GroES (Figure 5.3D, peak 97).

The incorporation of GroES into a much larger structure is also readily apparent in $^{[15N,1H]}$-CRIPT-TROSY spectra. In contrast to the TROSY experiment, a majority of the $89$ $^{[15N,1H]}$ cross peaks seen in free GroES (Figure 5.1 B) are observed in the $^{[15N,1H]}$-CRIPT-TROSY spectra of the complexes (Figure 5.1 D and Figure 5.3 D and E). Moreover, for nearly all cross peaks only the most slowly relaxing one of the four components of the multiplet pattern (Figure 5.2 E) is observable after addition of one equivalent of either SR1 or GroEL (Figure 5.1 D, Figure 5.2 B, and Figure 5.3 B, E and F). The other three components are broadened beyond detection in the large complexes (see section 3.5).
5.1 15N,2H-GROES–CHAPERONIN COMPLEX FORMATION FOLLOWED BY NMR

- Panel A: 
  - 1H NMR spectrum with peaks at ω2(1H) 9.0 and 8.0 ppm.

- Panel B: 
  - 1H NMR spectrum with peaks at ω2(1H) 9.0 and 8.0 ppm.

- Panel C: 
  - 1H NMR spectrum with peaks at ω2(1H) 9.0 and 8.0 ppm.

- Panel D: 
  - 1H NMR spectrum with peaks at ω2(1H) 9.0 and 8.0 ppm.

- Panel E: 
  - 1H NMR spectrum with peaks at ω2(1H) 9.0 and 8.0 ppm.

- Molecular structures:
  - A-B: 72 kDa
  - C-E: 472 kDa

- Chemical shifts:
  - ω1(15N) for 15N
  - ω2(1H) for 1H
Figure 5.1  Two-dimensional (2D) $[^{15}N,^1H]$-correlation spectra at 25°C of the uniformly $^{15}N,^2H$-labeled co-chaperonin GroES free in solution, and in a complex with the unlabeled GroEL variant SR1. GroES is a 72 kDa oligomeric protein with 7 identical subunits (Hunt et al., 1996; Chandrasekhar et al., 1986), of which one is colored in gold in the otherwise violet structure at the top of the upper right panel. The second drawing in the upper right shows a molecular model of the 472 kDa complex of GroES and SR1 based on the structure of the GroEL-GroES-ADP complex by Xu et al., 1997, with SR1 colored in blue. In A-E, SR1 is not isotope-labeled and therefore not visible with the presently used NMR experiments.

A. and B.: free $[^{15}N,>97%^2H]$-GroES recorded with A. 2D $[^{15}N,^1H]$-TROSY, and B. 2D $[^{15}N,^1H]$-CRIPT-TROSY.

C.-E.: $[^{15}N,>97%^2H]$-GroES bound to natural isotope abundance SR1 recorded with C. 2D $[^{15}N,^1H]$-TROSY, D. 2D $[^{15}N,^1H]$-CRIPT-TROSY, and E. 2D $[^{15}N,^1H]$-CRINEPT-TROSY.

The concentration of free GroES was 0.3 mM in heptamer. The $[^{15}N,^1H]$-TROSY spectrum (A) was recorded for 3 h with a data size of 150 x 1024 complex points, $t_{1\text{max}} = 37.5$ ms, $t_{2\text{max}} = 98$ ms, INEPT transfer time = 3.4 ms. The $[^{15}N,^1H]$-CRIPT-TROSY experiment (B) was recorded for 10.5 h, with a data size of 100 x 1024 complex points, $t_{1\text{max}} = 25$ ms, $t_{2\text{max}} = 98$ ms, CRIPT magnetization transfer time = 2.8 ms. The recycle delay was 1s. Prior to Fourier transformation a sine bell function shifted by 10° (DeMarco and Wüthrich, 1976) was applied to the $[^{15}N,^1H]$-CRIPT-TROSY data in the indirect dimension, and an empirically optimized exponential function was used in the direct dimension.

The GroES–SR1 complex was studied in a solution containing 0.15 mM $[^{15}N,>97%^2H]$-GroES and 0.24 mM natural isotope abundance SR1. A total of 34 mg of SR1 was added to 4.2 mg of GroES in the presence of 10 mM ADP. The $[^{15}N,^1H]$-TROSY and $[^{15}N,^1H]$-CRIPT-TROSY spectra were recorded and processed with the same parameters as in A and B, respectively. The $[^{15}N,^1H]$-CRINEPT-TROSY experiment was recorded for 10.5 h, with a data size of 100 x 1024 complex points, $t_{1\text{max}} = 25$ ms, $t_{2\text{max}} = 98$ ms, CRINEPT transfer time = 2.8 ms, same digital filtering as in B. A recycle delay of 0.6 s was used for these three experiments.
The few cross peaks that are also present in the $[^{15}\text{N},^1\text{H}]$-TROSY spectra of the complexes (Figure 5.1 C and Figure 5.3 D) show one or two of the other multiplet components in the $[^{15}\text{N},^1\text{H}]$-CRIPT-TROSY spectra, consistent with the idea that these residues are mobile in the complex. Finally, the $[^{15}\text{N},^1\text{H}]$-CRINEPT-TROSY spectrum (Figure 5.1 E) contains the cross peaks from both the structured and the more mobile regions of the chaperonin-bound GroES, but this experiment has lower sensitivity than the $[^{15}\text{N},^1\text{H}]$-CRIPT-TROSY scheme (see section 3.5).

**Figure 5.2** Multiplet patterns in $[^{15}\text{N},^1\text{H}]$-CRIPT-TROSY spectra of structures with molecular weights 72 and 472 kDa.

A. Experimental spectrum of $[^{15}\text{N},>97\%^2\text{H}]$-GroES (same data set as Figure 5.1 B).

B. Corresponding spectral region of $[^{15}\text{N},>97\%^2\text{H}]$-GroES bound to SR1 (same data set as Figure 5.1D).

C. and D. Cross-sections from A and B at the frequencies indicated by the arrows.

E. Theoretical multiplet pattern for $[^{15}\text{N},^1\text{H}]$-CRIPT-TROSY cross peaks. The four components due to the $^1\text{H}$-$^{15}\text{N}$ scalar coupling have different linewidths, which reflect their different transverse relaxation rates. Since the $^1\text{H}$ magnetization is detected in antiphase, the two components along the horizontal $^1\text{H}$-dimension have opposite sign, whereas the two components along the $^{15}\text{N}$-dimension have equal sign. For very large structures all but the “sharp/sharp” fine structure component are broadened beyond detection due to fast relaxation. In all spectra, positive contour lines are black and negative contour lines are blue.
5.2 Mapping of the GroES–GroEL interaction surface

Comparison of the $^{15}$N,$^1$H-CRIPT-TROSY spectra of GroES in the free form or when bound to either SR1 or GroEL revealed that the complex formation caused significant chemical shift changes only for a discrete set of resonance lines, and that there were several cross peaks in bound GroES that could not be correlated with any of the peaks in free GroES (identified by red circles in Figure 5.3). Since nearly identical numbers of $^{15}$N–$^1$H correlation peaks were observed for free and bound GroES, and the majority of the cross peaks are in closely similar positions in free and chaperonin-bound GroES, we chose the following approach for a systematic comparison of free and bound GroES, which has been previously validated with a different protein–protein complex (Pellecchia et al., 1999): Each sequence-specific resonance assignment obtained for free GroES (see chapter 4) was transferred to the nearest peak in the spectra of bound GroES, provided that the two peak positions were within a distance of 1.0 ppm along the $^{15}$N chemical shift axis, and within 0.2 ppm along the $^1$H chemical shift axis. If within this tolerance range no peak was observed in the spectrum of the bound GroES, no assignment was transferred from free to bound GroES. In Figure 5.3, B, C, D and E, the peaks thus assigned in bound GroES are identified with black numbers indicating the sequence positions. Based on these tentative assignments, the chemical shift variations were evaluated as $\Delta \delta_{av} = \{0.5[\Delta \delta(^1H)^2 + (0.2 \Delta \delta(^{15}N)^2)]\}^{1/2}$ (Pellecchia et al., 1999). Resonances for which no assignments could be transferred were assumed to undergo large chemical changes and $\Delta \delta_{av}$ was set to a default value of 0.3 ppm. The results of this procedure are shown in Figure 5.4, in which regions with $\Delta \delta_{av} \geq 0.1$ ppm are depicted in red.

There are a number of striking observations to be made about this result: (i) Most of the cross peaks assigned to residues 17–32 in free GroES could not be assigned in the complexes by the aforementioned
procedure, while for a majority of the other residues the chemical shifts in free and bound GroES could be correlated. The “new” cross peaks in the complexes (red circles in Figure 5.3) are thus likely to correspond to the residues 17–32 of the mobile loop in GroES (Hunt et al., 1996; Landry et al., 1993). (ii) In free GroES, the peaks of the mobile loop have random coil chemical shift values, reflecting a disordered conformation. By contrast, the “new” peaks appearing upon complex formation show a large chemical shift dispersion, indicating that the mobile loop of GroES has become ordered upon binding to GroEL. The other parts of GroES do not appear to undergo a major conformational change. (iii) Most of the $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY cross peaks in the spectra of the complexes consist of only the most slowly relaxing fine structure component (Figure 5.2 E), showing that the corresponding residues of GroES are immobilized relative to SR1 or GroEL, respectively, and their Brownian motions are restricted to those of the complex. However, a few of the backbone $^{15}\text{N}$–$^{1}\text{H}$ cross peaks of bound GroES that are located in new positions when compared to free GroES show two or three fine structure components in the $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY spectra. These peaks also appear in the $[^{15}\text{N},^{1}\text{H}]$-TROSY spectra. This indicates that at least some regions of GroES that undergo conformational changes upon binding to the chaperonins nevertheless retain significant mobility in the complex. (iv) Virtually identical results were obtained for GroES in complexes with either SR1 or GroEL (Figure 5.1 and Figure 5.3).
5.2 MAPPING OF THE GROES–GROEL INTERACTION SURFACE

A. TROSY of free $[^{15}\text{N},^{2}\text{H}]$-GroES

B. CRIP of SR1 bound $[^{15}\text{N},^{2}\text{H}]$-GroES

C. TROSY of GroEL-bound $[^{15}\text{N},^{2}\text{H}]$-GroES

D. CRIP of GroEL-bound $[^{15}\text{N},^{2}\text{H}]$-GroES

E. CRIP of GroEL-bound $[^{15}\text{N},^{2}\text{H}]$-GroES at 35°C
Figure 5.3 Chemical shift changes in $^{15}\text{N},^{2}\text{H}$-labeled GroES upon binding to natural isotope abundance SR1 or GroEL.

A. $[^{15}\text{N},^{1}\text{H}]$-TROSY spectrum of free GroES at 25°C, which provides the GroES reference chemical shifts (same data set as Figure 5.1A).

B. $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY spectrum of GroES bound to SR1 in the presence of ADP and at 25°C (same data set as Figure 5.1 D).

C. $[^{15}\text{N},^{1}\text{H}]$-TROSY spectrum of GroES bound to GroEL at 25°C.

D. $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY spectrum of GroES bound to GroEL at 25°C.

E. Same as E. at 35°C.

In C–E, the sample contained 0.13 mM [U-$^{15}$N; $U>$97% $^{2}$H]-GroES and 0.15 mM natural isotope abundance GroEL, and was prepared by addition of a total of 55 mg GroEL to 4.4 mg of GroES in the presence of 10 mM ADP.

The $[^{15}\text{N},^{1}\text{H}]$-TROSY spectrum (C) was recorded for 14 h, with a data size of 120 x 1024 complex points, $t_{1\text{max}} = 30$ ms, $t_{2\text{max}} = 98$ ms. The $[^{15}\text{N},^{1}\text{H}]$-CRIPT-TROSY spectra of (D) and (E) were recorded for 15 h each, with a data size of 64 x 1024 complex points, $t_{1\text{max}} = 16$ ms, $t_{2\text{max}} = 98$ ms, recycle delay = 0.3 s, same digital filtering as in Figure 5.1B. In C the INEPT transfer time was 3.4 ms, in D and E the CRIPT transfer time was 1.4 ms.

In all spectra, positive contour lines are black and negative contour lines are blue. The sequence-specific assignments for free GroES indicated by black numbers in A. were obtained from TROSY-triple resonance experiments with a $^{2}$H,$^{13}$C,$^{15}$N-labeled protein sample (chapter 4). Tentative assignments for the two complexes indicated in B, D and E have been obtained by transferring assignments of free GroES to nearby peaks of bound GroES (see text). Peaks of bound GroES that appear in new chemical shift regions when compared to free GroES are circled in red and have not been individually assigned (see text).
5.3 Discussion

In conclusion, the results presented here demonstrate that TROSY/CRINEPT-based $^{15}$N-$^1$H correlation experiments (Pervushin et al., 1997; Riek et al., 1999; see chapter 3) can be applied to collect solution NMR spectra and perform chemical shift mapping of intermolecular interactions in complexes of sizes up to 900 kDa. For the *E. coli* chaperonin system, our data show that the mobile loop of GroES is involved in a major conformational change upon binding to GroEL, while the rest of the molecule is largely unaffected by the interaction. These observations are compatible with previous spectroscopic and crystallographic studies of GroES interactions with GroEL (Landry et al., 1993, 1996; Shewmaker et al., 2001; Xu et al., 1997). Moreover, the individual $^{15}$N-$^1$H cross peaks appearing in new chemical shift positions upon complex formation (red circles in Figure 5.3), which are likely to arise from the polypeptide segment of bound GroES that corresponds to the disordered loop in free GroES, have variable dynamic properties, consistent with a binding mode in which some of the mobile loop residues are immobilized in the complex while others retain a relatively high degree of mobility.

Full interpretation of the results now require assignment of these resonances: The chemical shift mapping procedure applied here has only a statistical value and is meaningful when the shifting resonances cluster in one region of the protein 3D structure, which is the case for GroEL-bound GroES. Due to the slow exchange regime, assignments for GroES in the complex should be determined independently for final results. This seems feasible for the few resonances of GroEL- or SR1-bound GroES that are detected in the TROSY experiments. Owing to
their apparent mobility, these resonances could be detected using the TROSY scheme. Although they are weak in this experiment, it is probable that they could be detected in optimized 3D TROSY-HNCA or TROSY-HNCO, as well as NOESY experiments, and possibly assigned, which we are currently attempting (R. Horst, J. Fiaux). The situation is however different for all other resonances that are characterized by

---

**Figure 5.4** Molecular model of one subunit of GroES in the crystal structure of the GroEL-GroES-ADP$_7$ complex (Xu et al., 1997) colored according to the chemical shift changes observed upon complex formation with GroEL or SR1. The chemical shift variations were evaluated as $\Delta\delta_{av} = \{0.5[\Delta\delta(^1H)^2 + (0.2 \Delta\delta(^15N))^2]\}^{1/2}$ in the spectra of Figure 5.3 (Pellecchia et al., 1999). Red: $\Delta\delta_{av} \geq 0.1$ ppm; this includes the peaks for which no corresponding peak was assigned in bound GroES (see text). All the residues in this class are identified by red lettering indicating the amino acid type and the sequence position. Dark blue: $\Delta\delta_{av} < 0.1$ ppm. Light blue: no information due either to spectral overlap or to absence of the resonances in free GroES.
much faster transverse relaxation. Development of fully transverse-relaxation optimized experiments for assignments in structures above 200 kDa or possibly new assignment strategies must be achieved before addressing this question.

Further applications of the presently used approach may be greatly facilitated by the discovery that deuterium labeling is needed only for the NMR-observed macromolecular component in GroES-GroEL complexes, as evidenced by the fact that exactly the same cross peaks were observed in the $^{15}\text{N}^{-1}\text{H}$ correlation spectra of complexes with deuterated and $^{15}\text{N}$-depleted SR1 or GroEL, or with the natural isotope abundance chaperonins (data not shown). In addition to a dramatic reduction of the labour and expense for the NMR sample preparation, the natural isotope abundance ensures that signals from unlabeled components are suppressed in the presently used experiments.
In this chapter, we apply the methods described in the previous sections to isotope-labeled GroEL and SR1. The long-term goal is to address questions relative to the molecular mechanisms in the GroE system, involving for instance nucleotide binding, GroES binding and polypeptide substrate binding.

6.1 2D correlation spectra of $^{15}\text{N},^{2}\text{H}$-labeled GroEL and SR1

Using the techniques described in details in chapter 3 and samples prepared according to the protocols of chapter 2, we have recorded 2D correlation spectra of $^{15}\text{N},^{2}\text{H}$-labeled GroEL and $^{15}\text{N},^{2}\text{H}$-labeled GroEL single ring variant, SR1. Figure 6.1 and Figure 6.2 show the set of five correlation spectra used for observation of resonances in the 800 and 400 kDa proteins, respectively.

In GroEL the C-terminal 24-residue polypeptide segment is disordered (Braig et al., 1994, 1995) and probably highly flexible. Consistently, this segment is sensitive to proteolytic digestion (Langer et al., 1992; Weissman et al., 1995). This expectation from biochemical, EM
and X-ray studies is supported by observation of a small number of strong resonances with random coil $^1$H chemical shifts (Wüthrich, 1986) in the 2D [$^{15}$N,$^1$H]-TROSY and [$^{15}$N,$^1$H]-HSQC spectra of GroEL and SR1, and by the quasi-absence of these peaks in the 2D [$^{15}$N,$^1$H]-CRIPT-TROSY spectrum recorded with a CRIPT delay of $T = 1.4$ ms and 2.8 ms, respectively (Figure 6.1 and 6.2, A-C). Assignment of these peaks to the C-terminal flexible tail of GroEL was confirmed by the absence of these resonances in similar measurements with a SR1 construct in which this segment had been truncated (data not shown). On the other hand, the 2D [$^{15}$N,$^1$H]-CRIPT-TROSY spectrum contains a number of cross-peaks showing large chemical shift dispersion, which belong to $^{15}$N–$^1$H-moieties from structured parts of the molecules. This illustrates again that the 2D [$^{15}$N,$^1$H]-CRIPT-TROSY experiment provides also a selection scheme for $^{15}$N–$^1$H spin systems with long $\tau_c$ values.

In the 2D [$^{15}$N,$^1$H]-CRINEPT-HMQC–[$^1$H]-TROSY experiment that uses CRINEPT to detect cross-peaks of $^{15}$N–$^1$H-moieties from both flexible and more rigidly structured protein regions, the peaks have broad line-widths along the $\omega_1$(15N) dimension because of the absence of TROSY-compensation in the indirect dimension (Riek et al., 1999; see also chapter 3). Presence of both intense sharp resonances from the C-terminal tail of GroEL and very broad peaks from structured regions leads to baseline correction problems during the transformation of the spectra. The 2D [$^{15}$N,$^1$H]-CRINEPT-HMQC–[$^1$H]-TROSY experiment is then not a favorable scheme for study of the chaperonins and is mostly applied in initial measurements with new structures for its intrinsic high sensitivity. The 2D [$^{15}$N,$^1$H]-CRINEPT-TROSY experiment on the other hand yields a better resolved spectrum, but it somewhat less sensitive than the [$^{15}$N,$^1$H]-CRIPT-TROSY. Again, the presence of intense resonances from the C-terminal flexible tail of GroEL hinders the observation of the resonances from structured regions. In conclusion, a
combination of $[^{15}\text{N},^{1}\text{H}]-\text{TROSY}$ and $[^{15}\text{N},^{1}\text{H}]-\text{CRIPT-TROSY}$ experiments is the preferred approach for NMR signal observation in large proteins.

The peaks in the 2D $[^{15}\text{N},^{1}\text{H}]-\text{CRIPT-TROSY}$ spectrum of both $^{15}\text{N},^{2}\text{H}$-labeled GroEL and SR1 recorded in H$_2$O (Figure 6.1 and 6.2, C) account for only about 20% of the expected $^{15}\text{N}--^{1}\text{H}$ correlation peaks. This observation will be further investigated. At the present state of the investigation we cannot exclude any of the following possible explanations for the apparent scarcity of observable signals: (i) Since GroEL was expressed on a D$_2$O-based nutrient and has not been unfolded during the purification, some of the missing amide groups are probably not NMR-observable because they remain deuterated due to slow $^{2}\text{H}^{\text{N}}/^{1}\text{H}^{\text{N}}$ exchange. This explanation is supported by the observation that a few new peaks appeared after prolonged standing of the SR1 in H$_2$O solution, and that perdeuterated GroES with complete protonation of the amide groups (chapter 5) in a complex with unlabeled GroEL yielded spectra containing nearly all expected cross-peaks (Figure 5.1 C-E and Figure 5.3 D-F). However, this explanation does not account for all the missing peaks (see section 6.2). (ii) Alternatively, the broad distribution of the linewidths in GroEL (Figure 3.7) suggests that some missing peaks might be broadened beyond detection. (iii) Individual peaks may represent two or several overlapping resonances, and (iv) there may also be line broadening due to internal conformational exchange. An important observation is that the almost exact same subset of peaks are observed in SR1 and in GroEL, although their molecular weights are very different. This suggests that the number of missing peaks might be due to an intrinsic property of this particular molecular assembly, such as conformational exchange. Assignment of the observed peaks would be crucial to understand what
6.1 2D correlation spectra of $^{15}\text{N},^{2}\text{H}$-labeled GroEL and SR1

\begin{itemize}
  \item A. $[^{15}\text{N},^{1}\text{H}]$-HSQC
  \item B. $[^{15}\text{N},^{1}\text{H}]$-TROSY
  \item C. $[^{15}\text{N},^{1}\text{H}]$-CRIP-TROSY
  \item D. $[^{15}\text{N},^{1}\text{H}]$-CRINEPT-TROSY
  \item E. $[^{15}\text{N},^{1}\text{H}]$-CRINEPT-HMQC
\end{itemize}

$^{15}\text{N},^{2}\text{H}$-Gr\text{O}EL:
**Figure 6.1** Heteronuclear 2D correlation spectra of the uniformly $^{15}$N,$^2$H-labeled 800 kDa tetradecamer protein GroEL from *E. coli*, measured at 35°C.

**A.** $[^{15}$N,$^1$H]-HSQC spectrum, measuring time = 11 hours, acquired data size 300*1024 complex points, $t_{1,\text{max}} = 75$ ms, $t_{2,\text{max}} = 100$ ms. The INEPT transfer delay was 5.4 ms and the recycle delay 0.3 s.

**B.** $[^{15}$N,$^1$H]-TROSY spectrum, measuring time = 10 hours, acquired data size 200*1024 complex points, $t_{1,\text{max}} = 50$ ms, $t_{2,\text{max}} = 100$ ms. The INEPT transfer delay was 5.4 ms and the recycle delay 0.3 s. Shortening the polarization transfer delay to 3.4 ms in A or B did not affect the peak pattern observed.

**C.** $[^{15}$N,$^1$H]-CRIPT-TROSY spectrum, measuring time = 10 hours, acquired data size 100*1024 complex points, $t_{1,\text{max}} = 25$ ms, $t_{2,\text{max}} = 100$ ms. The CRIPT transfer delay was 1.4 ms and the recycle delay 0.3 s. Prior to Fourier transformation the data were multiplied along the $t_1$-dimension with a sine function shifted by 10° (DeMarco and Wüthrich, 1976), and in the $t_2$-dimension with an empirically optimized exponential function.

**D.** $[^{15}$N,$^1$H]-CRINEPT-TROSY spectrum, measuring time = 10 hours, acquired data size 100*1024 complex points, $t_{1,\text{max}} = 25$ ms, $t_{2,\text{max}} = 100$ ms. The CRINEPT transfer delay was 2.0 ms and the recycle delay 0.3 s. The spectra was transformed as indicated for C.

**E.** $[^{15}$N,$^1$H]-CRINEPT-HMQC spectrum, measuring time = 10 hours, acquired data size 100*1024 complex points, $t_{1,\text{max}} = 50$ ms, $t_{2,\text{max}} = 100$ ms. The delays and transformation procedure were as for D.
6.1 2D correlation spectra of $^{15}$N,$^2$H-labeled GroEL and SR1

A. $^{[15N, 1H]}$-HSQC
B. $^{[15N, 1H]}$-TROSY
C. $^{[15N, 1H]}$-CRIP-TROSY
D. $^{[15N, 1H]}$-CRINEPT-TROSY
E. $^{[15N, 1H]}$-CRINEPT-HMQC

[$^{15N, 2H}$]-SR1:

A. $^{[15N, 1H]}$-HSQC
B. $^{[15N, 1H]}$-TROSY
C. $^{[15N, 1H]}$-CRIP-TROSY
D. $^{[15N, 1H]}$-CRINEPT-TROSY
E. $^{[15N, 1H]}$-CRINEPT-HMQC
are the factors limiting detection of the cross peaks, but standard techniques are not applicable anymore to such large sizes and high number of residues.

The isolated apical domain of a GroEL subunit (residues 193-335) has been studied by Fersht and coworkers using solution NMR among other techniques (Kobayashi et al., 1999). In a comparison of the $^{15}$N-$^1$H correlation spectra of the 18 kDa ‘minichaperone’, for which sequence-specific assignments have been determined, with that of the entire GroEL, a few similar peak patterns can be recognized. Transfer for those peaks of the assignment of the minichaperone to the $[^{15}$N,$^1$H]-CRIP-TROSY spectrum of GroEL yields tentative assignment of
approximately 12 isolated resonances (data not shown). However, since only 20% of the peaks in GroEL are detected, it cannot be excluded that these peaks belong to residues of the other two domains of GroEL that would show a resonance in the same spectral region. Moreover, there may be chemical shift differences for resonances in the isolated domain vs. the whole GroEL molecule. Therefore, these are only speculative preliminary assignments.

When measuring 2D correlation experiments with isotope-labeled SR1 samples, we also found that this mutant shows increased tendency towards dissociation. The resulting monomers are unstable and contiguously unfold. This eventually leads to accumulating non-native species over time which generate strong resonances in the $^{[15N,1H]}$-TROSY and $^{[15N,1H]}$-HSQC spectra, but do not affect the $^{[15N,1H]}$-CRIPT-TROSY spectra. Measurements with SR1 were thus performed with freshly purified samples.

### 6.2 Leucine-specific $^{15N}$-labeling of GroEL and SR1

To address the questions relative to the observation of only 20% of the expected resonances in uniformly labeled GroEL and SR1, we performed residue-specific labeling experiments that should reduce the spectral complexity. As described in section 2.2, leucine is well suited for such labeling and $^{[15N,2H]}$-Leu labeled and partially deuterated chaperonins were prepared. Proper labeling of the proteins was checked in $^{[15N,1H]}$-correlation spectra of the denatured forms in 6 M guanidinium chloride and the 41 expected leucine resonances could be observed (Figure 2.2). The 2D $^{[15N,1H]}$-TROSY spectra of native $^{[15N,2H]}$-Leu $^{[U\sim 60\%2H]}$-GroEL only shows the resonance of Leu 531 in the C-terminal flexible tail of GroEL, as expected from the results obtained with uniformly labeled proteins, and possibly that of Leu 523
overlapping with the more intense Leu 531 peak (Figure 6.3). In the 
\([^{15}\text{N}(strtolower{,}^{1}\text{H})]\)-CRIPT-TROSY spectra, approximately 15 resonances can be counted for both GroEL and SR1 and correspond to 37% only of the expected resonances. Thus, the proportion of resonances that can be detected in these proteins is also low, as for the uniformly-labeled chaperonins. We then can exclude that a significant number of peaks was not identified because of spectral overlap. Moreover, the leucine-labeled GroEL and SR1 have been prepared from H\text{2}O-based cultivation media (see section 2.2) and do not suffer from incomplete $^{2}\text{H}^\text{N}/^{1}\text{H}^\text{N}$ exchange. Therefore, another reason needs to be invoked to explain the small proportion of observed resonances in these proteins, and it follows that slow $^{2}\text{H}^\text{N}/^{1}\text{H}^\text{N}$ exchange in the perdeuterated and uniformly-labeled proteins is probably only responsible for part of the

**Figure 6.3** 2D $\left[^{15}\text{N}(strtolower{,}^{1}\text{H})\right]$-TROSY (A) and $\left[^{15}\text{N}(strtolower{,}^{1}\text{H})\right]$-CRIPT-TROSY (B) correlation spectra of $\left[^{15}\text{N},^{2}\text{H}\right]$-Leu $\left[U\sim60\%\;^{2}\text{H}\right]$-GroEL. The parameter settings are as described for Figure 6.1 B and C. The sample was prepared as described in section 2.2 and was 110 $\mu$M in GroEL tetradecamer in NMR buffer (see section 4.1).
missing peaks in the spectra of section 6.1. In this discussion, it should also be noted however that due to their preparation in H₂O-based media, the specifically-labeled proteins are only deuterated to ~60% (Table 2.1). Residual protonation is thus likely to limit the detection of leucine cross peaks (see section 3.7). In conclusion, examination of a perdeuterated leucine-labeled sample of GroEL or SR1 after unfolding and refolding would be necessary to finally solve this issue. The refolding of GroEL has been reported as troublesome (Mizobata and Kawata, 1994; Price et al., 1993; Ybarra and Horowitz, 1995b), but should be attempted anyway in future work.

6.3 Nucleotide and GroES binding to GroEL and SR1

Nucleotide and GroES binding to GroEL trigger a series of conformational changes and modulate the interplay of the different allosteric states which play a major role in the GroEL cycle. In an attempt to study these effects using solution NMR, we measured correlation spectra of SR1 and GroEL in the presence of nucleotides and in complexes with nucleotides and GroES.

Upon addition of ADP or of ATP that rapidly hydrolyzes to ADP, a few cross-peaks in the [¹⁵N,¹H]-CRIPT-TROSY spectra of GroEL and SR1 showed chemical shift changes. Examples of such cross-peaks are indicated with arrows in Figure 6.4. This confirms the occurrence of some conformational changes upon nucleotide binding. Since no assignment of the [¹⁵N,¹H]-CRIPT-TROSY resonances are available as yet, we unfortunately cannot interpret these observations. Intriguingly, some of the cross-peaks in new spectral regions were also observed in a SR1 sample that was kept at 4°C for 7 months, in which newly exchanged Hᴺ are detected. It also remains unclear whether the changes
observed concern residues in direct contact with the nucleotide binding region, or whether nucleotide binding causes cross peak shifts indirectly by transmitting a conformational change to remote parts of the molecules, such as the apical domain. Cryo-electron microscopy studies of such GroEL-nucleotide complexes revealed changes in the interferring contacts as well as in the orientation of the apical domain and different conformations with different nucleotides (Roseman et al., 1996; Roseman et al., 2001). If some changes are indeed manifested in the NMR spectra, they are few and it is likely that their observation is limited by the total number of resonances detected. Clearly, solving the question of missing peaks is a prerequisite for a more complete observation of the effects of nucleotide binding on the correlation spectra of the chaperonins. Moreover, new assignment strategies need to be developed to identify at least part of the resonances and enable an interpretation of the observed chemical shift variations. Important questions to address with these experiments would then concern the differential effect of ATP vs. ADP or AMP-PNP on the GroEL conformation and what molecular details makes the ATP complex the only folding-active species whereas ADP or AMP-PNP are not sufficient for productive folding (Rye et al., 1997; Weissman et al., 1996).

Complexes of SR1 with GroES in the presence of nucleotides on the other hand showed a completely different $^{15}\text{N}-^{1}\text{H}$ fingerprint in the $[^{15}\text{N},^{1}\text{H}]-\text{CRIPT-TROSY}$ experiment compared to free SR1 (compare Figure 6.4 A and C). This is consistent with the large domain movements reported in the case of the GroEL-GroES-ADP crystal structure (Xu et al., 1997). The clear manifestation of conformational changes in these spectra opens possibilities for studying the chaperonin complexes, but once again, assignment of at least some resonances is required for interpretation of the spectra. When repeating the experiment with wild-type GroEL however, very different correlation spectra were obtained.
Figure 6.4 Chemical shift induced in the $[^{15}\text{N},^1\text{H}]-\text{CRIPT-TROSY}$ spectrum of $[^{15}\text{N},^{2}\text{H}]-\text{SR1}$ by binding of nucleotides and GroES.

A. $[^{15}\text{N},^1\text{H}]-\text{CRIPT-TROSY}$ spectrum of $[^{15}\text{N},^{2}\text{H}]-\text{SR1}$ that provide the reference chemical shifts. The arrows point to examples of cross peaks that shift or disappear upon binding of ADP to SR1 (compare A and B).

B. $[^{15}\text{N},^1\text{H}]-\text{CRIPT-TROSY}$ spectrum of $[^{15}\text{N},^{2}\text{H}]-\text{SR1}$ in the presence of ADP. The arrows indicate examples of cross peaks that show chemical shifts variations when compared to SR1 in absence of nucleotides.

C. $[^{15}\text{N},^1\text{H}]-\text{CRIPT-TROSY}$ spectrum of $[^{15}\text{N},^{2}\text{H}]-\text{SR1}$ bound to GroES in the presence of ADP. The peak pattern is different from that of SR1 alone, as shown for example by the few cross peaks marked with arrows that appear in new locations.

All spectra were recorded and transformed as indicated in Figure 6.2 C. The sample was ~$150\ \mu\text{M}$ in SR1 heptamer in NMR buffer. B was obtained after addition of 5 mM ADP and C after addition of unlabeled GroES in a 1.5 to 1 ratio to SR1 in presence of 5 mM ADP.
The most striking effect of GroES binding to GroEL was a line-broadening effect that severely degrades the appearance of the spectrum (compare Figure 6.5 A and B). Moreover, no peaks were observable in regions of dispersed chemical shifts between 9.5 and 10.5 ppm in the $^1$H dimension, where new peaks were clearly observed for GroES-bound SR1 (compare Figure 6.4 C and Figure 6.5 B). Analysis of the samples by size exclusion chromatography proved that the complexes were of the expected size and showed no sign of aggregation or inhomogeneity. Varying the GroES to GroEL ratio also did not change the appearance of the spectrum. Some possible explanations for the

---

**Figure 6.5** Effect of GroES binding on the 2D $[^{15}N,^{1}H]$-CRIPT-TROSY spectrum of GroEL.

**A.** $[^{15}N,^{1}H]$-CRIPT-TROSY spectrum of $[^{15}N; U-85\%{^2}H]$-GroEL (data measured as described in Figure 6.1 C).

**B.** $[^{15}N,^{1}H]$-CRIPT-TROSY spectrum of $[^{15}N; U-85\%{^2}H]$-GroEL bound to unlabeled GroES in the presence of ADP. The difference in overall appearance of the spectrum compared to A is due to line broadening (see text). The sample was 70 $\mu$M in $[^{15}N,^{2}H]$-GroEL tetradecamer, 70 $\mu$M in unlabeled GroES and 5 mM in ADP. This sample is a contribution of Dr. E. B. Bertelsen. The spectrum was recorded as described in Figure 6.1 C.
observed effects are: (i) Broadening due to chemical exchange. There is a possibility that at the high protein concentration of our experiment GroES might be involved in a dynamic exchange process at intermediate rates on the NMR time scale. This exchange would not take place in the SR1-GroES complexes because the single-ring mutant lacks the trans ring responsible for triggering GroES release. (ii) Broadening due to the presence of multiple conformation in the sample, which could be free GroEL, GroEL in a 1:1 complex with GroES (“bullet-like” complex) and/or GroEL in a 1:2 complex with GroES (“football-like” complex). In the SR1-GroES complex, only one arrangement is possible which facilitates its observation in the NMR spectra. More experiments are here needed to clarify these results and further exploration of the GroES-bound conformation of chaperonins 60 using solution NMR is probably best done in the model system of SR1.
Discussion and perspectives

The results presented in this work establish that TROSY/CRINEPT-based \([^{15}\text{N},^{1}\text{H}]-\)correlation experiments can be used to collect informative solution NMR spectra of structures with molecular weights up to 900 kDa. In particular, detection of fingerprints for the isotope-labeled subunit in large macromolecular assembly can be of direct use for example for ligand search procedure such as SAR by NMR (Structure-activity relationship by NMR, Shuker et al., 1996). The NMR schemes refined here also yield direct information on the dynamics of the studied structure: with the comparison of the 2D TROSY and CRIPPT-TROSY spectra, the observation of the CRIPPT cross peak multiplet pattern, or measurement of the optimal CRIPPT transfer delays, the relative mobility of different regions can be easily assessed. Finally, it is demonstrated that chemical shift mapping can be employed to characterize protein-protein interactions in complexes of such large sizes and identifying conformational changes. These three aspects directly serve the purpose of expanding our understanding of macromolecular machines.
For the *E. coli* chaperonin system, these NMR experiments together with appropriate isotope labeling schemes enabled a mapping of the GroES–GroEL interaction surface, observation of a conformational change in a loop region, and characterization of the relative dynamics of different parts of the complex (Fiaux *et al.*, 2002). Full interpretation of the available NMR data now requires independent assignment of the resonances in the large complex (see below). Besides, further applications of the presently used approach, with *E. coli* chaperonins as well as with other supramolecular systems, may be greatly facilitated by the discovery that deuterium labeling is needed only for the NMR-observed macromolecular component in GroES-GroEL complexes.

We thus now possess tools for NMR observation of the large molecular system of chaperonin, opening new possibilities for addressing specific biological questions such as the conformational changes in the GroE cycle and the fate of bound substrates. However, when applied to isotope-labeled GroEL itself or to its variant SR1, these methods only provided preliminary results and did not yet as such enable conclusions on the observed phenomena. The main issue now hindering further analyses is that of understanding why only 20% of the expected resonances can be detected in GroEL and SR1, whereas nearly complete fingerprints could be measured for GroES complexes of the same sizes. Explanations relating to incomplete $^2$H/$^1$H exchange and degree of deuteration of the molecules should be possible to assess with comparison experiments of a new set of protein preparations. Besides, there is a possibility that only parts of the molecules with sufficient mobility yield detectable cross-peaks. In such a case, we would mainly observe resonances from the apical domain, which has been reported to have intrinsic flexibility (Braig *et al.*, 1995; Sigler *et al.*, 1998). The various experiments performed and comparison with NMR spectra of the individual apical domain (Kobayashi *et al.*, 1999) still did not show what
7. DISCUSSION AND PERSPECTIVES

proportion of the total detected resonances arises from the apical domain.

In addition, in all experiments attempted here to study GroEL and its complexes with nucleotides and GroES (chapters 5 and 6), as well as to solve the questions relative to missing peaks in the GroEL correlation spectra (chapter 6), developing new assignment strategies appear as crucial since current procedures are not anymore applicable in these large complexes. This remains a difficult task in view of the sizes of the molecules. Since the pulse programs must be very short for detection of signals at large sizes, it could be that a solution comes from combination of homo-nuclear techniques and spectral simplification means. Clearly, intensive work will be needed until detailed structural and mechanical information can be gained from these experiments.

With respect to the question of upper size limits for the structures that can be investigated with presently available techniques, several sources of limitation have been encountered in this work. First, the linewidths in the 900 kDa assemblies are such that they already limit detection and resolution of the cross peaks in some cases. Besides, very different results were obtained for the proportion of observed resonances in the two different supramolecular structures investigated at this large size (GroEL and GroES in complex with GroEL), although the observed cross peaks have similar linewidths. Whether missing peaks in GroEL is a problem due to molecule-specific properties and to which extent the nearly complete correlation spectra for GroEL-bound GroES is related to its particular dynamics remain unresolved questions. General applicability of the here-described methods should then be checked by investigations of other systems of similar size. Another problem is the inevitable compromise between achieving high deuteration and ensuring complete N-D/N-H exchange of the amide protons when
preparing NMR samples of large supramolecular structures for which refolding or re-assembly is difficult.

NMR investigations of molecular chaperones comprise to date determination of the three-dimensional structure of smaller proteins or individual domains (Berjanskii et al., 2000; Ellgard et al., 2001a,b; Martinez-Yamout et al., 2000; Morshauser et al., 1995; Morshauser et al., 1999; Pellecchia et al., 1996, 1998, 2000; Qian et al., 1996; Szyperski et al., 1994; Wang et al., 1998), identification of very flexible regions (Carver et al., 1995, 1996; Carver, 1999; Huang et al., 1999a; Lindner et al., 2000; van de Klundert et al., 1998; Volkert et al., 1999) or other dynamics analysis (Bertelsen et al., 1999; Huang et al., 1999b), characterization of folding transition or conformational changes (Bann et al., 2002; Swain et al., 2001), and study of binding interactions of partner proteins (Briknarova et al., 2001; Frickel et al., 2002; Greene et al., 1998; Hua et al., 2001; Karlsson et al., 1998; Landry and Gierasch, 1991; Landry et al., 1992; Landry et al., 1993; Nieba-Axmann et al., 1997; Pellecchia et al., 1999; Pellecchia et al., 2000; Soti et al., 1998; Walse et al., 1997; Wang et al., 1999; Zahn et al., 1994; Zahn et al., 1996b,c). Apart from the investigation of the rather large FimC-FimH complex using TROSY (Pellecchia et al., 1998; Pellecchia et al., 1999), the large majority of these applications have been involving single domains or fragments of the larger molecular chaperones in view of the size limitation when using solution NMR spectroscopy. Clearly, the methods described in this work can be of considerable interest in these cases: with the notable exception of three-dimensional structure determination, they should allow to address the same questions in the complete molecular assembly.

In contrast to TROSY, the CRINEPT principle has not yet been applied to current biomolecular NMR research projects. Further developments such as potential three-dimensional and NOESY schemes as well as
labeling strategies for spectral simplification are certainly required to complement presently available techniques and make them more generally applicable. With this work, we nevertheless hope to have demonstrated their performance and utility for studies of large macromolecular structures.
Literature


Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A.,


Acad Sci U S A 98(6): 3133-3138.


Natl Acad Sci U S A 92(18): 8100-8104.


Qin, J., Vinogradova, O. and Gronenborn, A. M. (2001). “Protein-protein interactions...


direct alternation of folding-active rings.” Cell 97(3): 325-338.


Metabolic flux ratio analysis in microorganisms using fractional $^{13}\text{C}$-labeling and 2D NMR

A.1 From fractional $^{13}\text{C}$-labeling of amino acids to metabolic flux ratios

Knowledge of cellular metabolism is necessary to understand the biochemistry and physiology of cellular growth and a prerequisite for rational metabolic engineering in biotechnological applications. Detailed investigations of the cellular metabolism may address different levels such as the topology of reaction networks, the thermodynamics and kinetics of the enzymatic reactions, their molecular mechanism or the in vivo regulation of the metabolic pathways. Although large bodies of information have accumulated on various model organisms or pathways, the regulation of metabolic pathways and the response of an
organism to varying growth conditions remains difficult to predict. Therefore, direct experimental measurements of the regulatory response of an organism are of considerable value. In this section, we present a method based on fractional $^{13}$C-labeling and 2D NMR that enables characterization of a bioreaction network topology and determination of metabolic flux ratios in micro-organisms (Szyperski, 1995). The brief description of the principles underlying this method (see also reprints at the end of this section for a more complete description) will be followed by a presentation of several applications in metabolic studies of microorganisms.

Biosynthetically directed fractional $^{13}$C-labeling of amino acids is achieved by growing cells in minimal medium containing a mixture of 10% $[^{13}$C$_6]$-glucose and 90% natural isotope abundance glucose. Since the resulting isotope enrichment is uniform at all carbon positions, the main observable is the $^{13}$C-$^{13}$C scalar coupling fine structure. $^{13}$C-$^{13}$C connectivities can be traced in the metabolic intermediates, and thus this method correspond to a “labeling” of carbon-carbon bonds. Routinely, the $^{13}$C-$^{13}$C scalar coupling fine structures are detected for each carbon position in the amino acids obtained from complete biomass hydrolysis using 2D $[^{13}$C, $^1$H]-correlation NMR spectroscopy. In the $[^{13}$C, $^1$H]-HSQC experiment (Ruben and Bodenhausen, 1980), all relevant peaks are resolved and separation of the amino acids is not necessary prior to NMR measurement. Figure A.1 shows an example of a region in the 2D $[^{13}$C, $^1$H]-HSQC spectrum measured for an amino acid hydrolysate obtained from Saccharomyces cerevisiae cells. The $^{13}$C-$^{13}$C scalar fine structure is observed along the $\omega_1(^{13}$C) dimension and consists of the superposition of several multiplets arising from the different $^{13}$C isotopomers (Figure A.2 A). Quantitative analysis of this fine structure can unravel to which extent the observed carbon atom has neighboring carbons originating from the same source molecule of
Figure A.1 Region of the $^{13}$C,$^{1}$H-HSQC spectrum containing $^{13}$C-$^{1}$H cross peaks of $\beta$-, $\gamma$- and $\varepsilon$-position of amino acids in a hydrolysate of cellular protein from S. cerevisiae. The resonance assignments are given using the one letter code of the amino acids. The complete spectrum was recorded on a DRX-500 Bruker instrument for 8 hours, acquired data size 1500*256 complex points, $t_{1\text{max}}=350$ ms, $t_{2\text{max}}=100$ ms, according to Ruben and Bodenhausen, 1980. The recycle delay was 1.8 s. The carrier was set to 42.5 ppm in the $^{13}$C dimension and 2.9 ppm in the $^{1}$H dimension, and the spectral width to 34.0 and 5.0 ppm, respectively. A similar spectrum with the carrier positioned at 126.0 ppm in the $^{13}$C dimension and 7.0 ppm in the $^{1}$H dimension was recorded for detection of the aromatic resonances.
glucose: The $^{13}$C fine structure is decomposed into its multiplet components, which are integrated to yield their relative intensities and thus the relative abundance of the corresponding isotopomers (Figure A.2 A). Using probabilistic equations that account for the contribution of natural $^{13}$C abundance to these isotopomers, the relative abundances of intact carbon fragments arising from a single source molecule of glucose are calculated (Szyperski, 1995) (Figure A.2 B). When tracing for example the intact carbon fragment $C^\alpha-C^\beta$ in Asp (Figure A.2), the same information can be obtained from the data for positions Asp-$\alpha$ and Asp-$\beta$. In a third step, the data obtained from such redundant carbon position are correlated to yield the final relative abundance of carbon-carbon connectivities originating from a single glucose source molecule (Figure A.2 C).

Once the observed fine structures have been translated into fragment abundances (Figure A.2 C), the breakdown of the carbon skeleton of the source glucose molecule can be interpreted. The patterns of intact carbon fragments in the amino acids reflect those present in their direct metabolic precursors, thus yielding information on eight principal intermediates in the central carbon metabolism, i.e. ribose-5-phosphate, erythrose-4-phosphate, 3-phosphoglycerate, phosphoenolpyruvate, pyruvate, acetyl-CoA, 2-oxoglutarate and oxaloacetate. The metabolic origin of these intermediates can be inferred from the presence or absence of certain intact carbon fragments that are diagnostic for a given pathway. This allows the identification of the metabolic pathways that are activated under the conditions of the experiment as well as the derivation of ratios of metabolic fluxes leading to the intermediates.

The here-presented method, coined METAFoR (metabolic flux ratio) analysis, provides information on the flux distribution in several key
Figure A.2  Determination of relative abundances of carbon-carbon connectivities from the $^{13}C-^{13}C$ scalar coupling fine structure (figure from Szyperski et al., 1999).

A. Integration of the fine structure yields the relative intensities of the multiplet components, and thus the relative abundance of the corresponding isotopomers (shown on the right). $I_s$, $I_d$, $I_{d^*}$, and $I_{dd}$ indicate the relative intensities of the singlet, a doublet split by a small one-bond scalar coupling, a doublet* split by a large one-bond scalar coupling, and a corresponding doublet of doublets. As an example, the fine structure detected for the $\alpha$-carbon of Asp is shown.

B. The relative abundances of intact carbon fragments arising from a single source molecule of glucose are calculated using a system of probabilistic equations (Szyperski, 1995). Connectivities arising from a single source molecule are depicted in bold. In analogy to the notation for the multiplets, the values $f^{(1)}$, $f^{(2)}$, $f^{(2*)}$, and $f^{(3)}$ denote the fraction of Asp molecules that contain, respectively, an $\alpha$-carbon without neighbor from the same source molecule, a $C^{\alpha}-C^{\beta}$ fragment, a $C^{\alpha}-C=O$ fragment, and a $C^{\beta}-C^{\alpha}-CO$ fragment originating from a single source molecule.
pathways, such as glycolysis, the pentose phosphate pathway, the tricarboxylic acid cycle and the C1-metabolism (Szyperski, 1995). The pattern of activated pathways represents a “fingerprint” of the regulatory state of the cell. It thus well-suited for analysis of metabolic response of organisms to changes in the growth conditions or mutations for example in “model microbes” for biotechnology (Emmerling et al., 2002; Fiaux et al., 1999; Frey et al., 2001; Hochuli et al., 2000; Maaheimo et al., 2001; Sauer et al., 1997; Sauer et al., 1999), as well as for exploration of metabolic network and amino acid biosynthesis pathways in less well-characterized organisms (Hochuli et al., 1999).

A.2 Applications to the study of *Escherichia coli* metabolism

In the following pages, several applications of METAFoR analysis to study the metabolic responses of *Escherichia coli* are presented in the form of published work. In a first part, two studies address the problem of oxygen limitation in bioprocesses. Such microaerobic conditions are encountered frequently in nature as well as in high cell density reactors of biotechnological processes (Konz et al., 1998; Spiro and Guest, 1991). They affect the growth and metabolite production rates of microbes in particular in large-scale biotechnological processes. In the first paper (Fiaux et al., 1999), we show that *E. coli* cells cultivated under oxygen-limitation show a fully fermentative metabolism in spite of the presence of oxygen, a result that is at variance with normally assumed aerobic
configurations. In an attempt to relieve the bacteria from this limitation, mutants cells expressing variants of the *Vitreoscilla* haemoglobin protein were found to have improved growth properties and consistently exhibited a more respirative metabolism (Frey *et al.*, 2001).

Another study addressed the adaptation potential of *E. coli* metabolism to mutations or moderate overexpression of enzymes at important nodes of the central metabolism, such as phosphofructokinase, pyruvate kinase, pyruvate decarboxylase and alcohol dehydrogenase (Sauer *et al.*, 1999; Emmerling *et al.*, 2002). *E. coli* central carbon metabolism was found be remarkably robust since the distribution of the main carbon fluxes was not altered in presence of genetic variations. More significant flux changes were seen in response to varying environmental conditions such as ammonia or carbon limitation.

### A.3 Extension of METAFOReR analysis to simple eukaryotic cells

As noticed from the examples cited in the above section, METAFOReR analysis has been applied to many prokaryotes. To further develop the method, we aimed at enabling such analysis with simple eukaryotes such as yeast cells, which are pivotal eukaryotic model organisms for many areas of medicine and biotechnology. In these organisms, the key challenge is the compartmentation of the cell, which leads to dissection of central carbon metabolism in sub-networks located in either the cytosol or in organelles, e.g. mitochondria or peroxisomes.

The extension of the biosynthetic fractional labeling protocol for investigating eukaryotic systems is presented in two publications: The first one describes the development of the analysis model for
Saccharomyces cerevisiae (Maaheimo et al., 2001) and a second one will report an application to study metabolic differences between S. cerevisiae and the much less well-characterized yeast Pichia stipitis (Fiaux et al., manuscript in preparation). Key differences observed between the two yeast strains include a remarkably low use of the non-oxidative pentose phosphate pathway for glucose catabolism in S. cerevisiae compared to P. stipitis batch cultures, and a fully respirative metabolism of P. stipitis batch cultures, which contrasts with the predominantly respiro-fermentative metabolic state of S. cerevisiae that experiences catabolite repression.
Backbone resonance assignment of GroES

In this section, the details of GroES backbone resonance assignment is shown: the strips from the 3D spectra should provide direct evidence for the connectivities established and for the resulting assignments.

Part A contains the complete set of strips from the $[^{15}\text{N},^1\text{H}]-\text{TROSY-HNCA}$ spectrum. The sequential connectivities are shown with lines and the sequence-specific assignments obtained are indicated as the one-letter code for the amino acid and the residue number in the amino acid sequence. These connectivities could be confirmed in the $[^{13}\text{C}]-\text{ct-}[^{15}\text{N},^1\text{H}]-\text{TROSY-HNCA}$ spectrum for resonances of sufficient intensities. Part B shows the instances where the high resolution of this spectrum was necessary to unambiguously establish connectivities. Further confirmation of the established connectivities can be obtained from the $[^{15}\text{N},^1\text{H}]-\text{TROSY-HNCACB}$ spectrum, and part C displays the corresponding regions. Finally, the 3D $[^1\text{H},^1\text{H}]-\text{NOESY-[}^{15}\text{N},^1\text{H}]-\text{TROSY}$ spectrum was used to confirm the sequence-specific assignment obtained, and to complete assignment in regions of missing strips in the triple resonance experiments, as shown in Part D. Note that generally combined used of all four spectra was applied to obtain unambiguous assignment. A final check of the backbone assignment is obtained by the absence of violations in a preliminary 3D fold calculation using DYANA, which demonstrates self-consistency of the data set.
Part A: $^{[15N,1H]}$-TROSY-HNCA

HNCA

| N2   | 123.0 | I3   | 119.6 | P5   | 127.9 | L6   | 117.9 | H7   | 120.4 | D8   | 119.6 | R9   | 115.0 | V10  | 129.0 | I11  | 125.4 | K13  | 129.5 | R14  | 126.8 | K15  | 124.2 | E16  | 124.2 |
|------|-------|------|-------|------|-------|------|-------|------|-------|------|-------|------|-------|------|-------|------|-------|------|-------|------|-------|------|-------|------|-------|------|

$\omega_1^{(13C)}$ (ppm)
N. B.: strips for 17-28 are plotted with a higher cut-off for contour levels than the rest of the spectrum because the resonances for these residues are very intense (mobile loop).
HNCA

\( \omega_1 (^{13}C) \) (ppm)

\( \omega_3 (^1H) \) (ppm)

<table>
<thead>
<tr>
<th>L41</th>
<th>A42</th>
<th>V43</th>
<th>G44</th>
<th>G45</th>
<th>R47</th>
<th>L49</th>
<th>E50</th>
<th>N51</th>
<th>G52</th>
<th>E53</th>
</tr>
</thead>
<tbody>
<tr>
<td>130.6</td>
<td>118.1</td>
<td>115.4</td>
<td>111.3</td>
<td>105.6</td>
<td>119.8</td>
<td>129.5</td>
<td>120.8</td>
<td>114.0</td>
<td>108.9</td>
<td>120.8</td>
</tr>
</tbody>
</table>

HNCA
<table>
<thead>
<tr>
<th>13C (ppm)</th>
<th>1H (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNCA</td>
<td></td>
</tr>
<tr>
<td>V54 120.4</td>
<td>8.22</td>
</tr>
<tr>
<td>K55 130.8</td>
<td>9.10</td>
</tr>
<tr>
<td>L57 119.2</td>
<td>8.53</td>
</tr>
<tr>
<td>D58 122.0</td>
<td>9.55</td>
</tr>
<tr>
<td>V59 113.4</td>
<td>8.84</td>
</tr>
<tr>
<td>K60 122.4</td>
<td>8.24</td>
</tr>
<tr>
<td>V61 120.0</td>
<td>8.93</td>
</tr>
<tr>
<td>G62 117.2</td>
<td>8.01</td>
</tr>
<tr>
<td>D63 122.1</td>
<td>8.87</td>
</tr>
<tr>
<td>I64 123.1</td>
<td>9.13</td>
</tr>
<tr>
<td>V65 117.4</td>
<td>8.33</td>
</tr>
<tr>
<td>F67 123.9</td>
<td>8.33</td>
</tr>
<tr>
<td></td>
<td>N68</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
</tr>
<tr>
<td>Value</td>
<td>118.3</td>
</tr>
</tbody>
</table>

HNCA

\[ \omega_1^{(13C)} \]

\[ \omega_3^{(1H)} \] (ppm)
See also NOESY spectrum for assignment of the region 87-93.
(see also NOESY spectrum for assignment of the region 87-93)
Part B: $^{13}$C-$^{15}$N, $^1$H-TROSY-HNCA
ct-HNCA

G29  S30  A31  A32  A33  K34  S35
111.4 117.0 125.7 122.4 123.3 120.5 117.9

ω₁(^13C) (ppm)

ω₃(^1H) (ppm)
Besides confirmation of the complete connectivities set shown in Part B, use of the $^{[13\text{C}]}$-ct-$^{[15\text{N},1\text{H}]}$-TROSY-HNCA spectrum could resolve the segment G23 G24 I25 and provide unambiguous sequence-specific assignment since the motif GG is unique in the GroES sequence (see page 178). Similarly, this spectrum enabled resolution of resonances for the segment A31 A32 A33 K34, as shown on page 179. The right panel provides a zoom on these peaks.

For amino acid 44-47, G44 and G46 could be distinguished and the assignment confirmed in spite of the missing strip for N45 (see also NOESY strips for this segment). The connectivities in the segment of residues 69-75 could also be confirmed unambiguously. These data are shown on page 180.
The data from the $^{15}{\text{N}}-^{1}{\text{H}}$-TROSY-HNCACB on pages 182-185 show the confirmation of the $\mathrm{C}^\alpha-\mathrm{C}^\alpha$ connectivities established using the $^{15}{\text{N}}-^{1}{\text{H}}$-TROSY-HNCA and $^{13}{\text{C}}$-ct-$^{15}{\text{N}}-^{1}{\text{H}}$-TROSY-HNCA spectra by the $\mathrm{C}^\beta-\mathrm{C}^\beta$ connectivities. These connectivities are shown with blue lines. These data also help obtaining a sequence-specific assignment since the $^{13}{\text{C}}^\beta$ chemical shift can be used as information for the mapping of a segment onto the GroES sequence.
<table>
<thead>
<tr>
<th></th>
<th>118.3</th>
<th>125.9</th>
<th>109.4</th>
<th>120.6</th>
<th>109.5</th>
<th>122.5</th>
<th>128.7</th>
<th>120.0</th>
<th>127.5</th>
<th>122.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>N68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D69</td>
<td>118.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G70</td>
<td>109.4</td>
<td>120.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y71</td>
<td>125.9</td>
<td>109.5</td>
<td>122.5</td>
<td>128.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G72</td>
<td>109.5</td>
<td>120.6</td>
<td>122.5</td>
<td>128.7</td>
<td>120.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V73</td>
<td>122.5</td>
<td>128.7</td>
<td>120.0</td>
<td>127.5</td>
<td>122.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNCACB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω₁⁻¹³C</td>
<td>8.90</td>
<td>8.41</td>
<td>7.73</td>
<td>8.32</td>
<td>8.60</td>
<td>7.53</td>
<td>8.59</td>
<td>8.56</td>
<td>8.96</td>
<td>8.64</td>
</tr>
<tr>
<td>ω₃¹H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table lists the chemical shifts for various positions in the HNCACB spectrum, with columns for N68, D69, G70, Y71, G72, V73, K74, S75, E76, and K77, along with the corresponding chemical shifts in ppm. The spectrum includes peaks at various ω₁⁻¹³C and ω₃¹H positions.
Part D: [$^1\text{H},^1\text{H}$]-NOESY-[$^{15}\text{N},^1\text{H}$]-TROSY

The strips of the [$^1\text{H},^1\text{H}$]-NOESY-[$^{15}\text{N},^1\text{H}$]-TROSY spectrum shown here belong to residues in two segments forming an antiparallel β-sheet: Residues 94-97 of one GroES subunit define a β-strand making contact to the neighboring strand of residues 1-5 of the next subunit. The NOE connectivities show as red lines thus confirm the assignment, and delineate the topology of the β-sheet and of the intersubunit contacts. Identification of an asparagine in position 2 also provide additional confirmation of the sequence-specific assignment. Diagonal peaks are designated with a red dot.
These strips show the NOE connectivities for the residue stretch following that of p. 18 in the GroES sequence. Since position 5 is a proline and no strip could be found for position 11 in the spectra, this segment could not be directly connected to its neighbors. Sequence-specific assignment was however helped by the identification of an arginine in position 9.
The NOE connectivities shown here confirm the assignment around the missing resonance for position 45.
These strips document the assignment of the side-chain resonances for N$^{\delta2}$, H$^{\delta21}$ and H$^{\delta22}$ of Asn 68.
The NOE pattern observed for residues 87-90 provide evidence for the presence of a short helical segment. Resonances for residue 91 are missing in all triple resonance experiments. These resonances could be identified in the \[^{1}H,^{1}H\]-NOESY-[\(^{15}N,^{1}H\)]-TROSY through NOE’s to its neighboring positions. Note that the diagonal peak for Ile 91 probably disappears due to overlap with wiggles originating from the strong resonance of Thr 19 in a \(^{15}N\)-plane close by.
Publication list


NMR techniques for large molecular and supramolecular structures.” *Journal of the American Chemical Society*, in press.

Name: Jocelyne Fiaux

Date of birth: July 14th 1975

Place of birth: Nyon (VD)

Citizenship: Hermenches (VD), Switzerland

Education:

1998-2002 Ph.D. student at the Institute for Molecular Biology and Biophysics, ETH Switzerland.

Ph.D. thesis under the supervision of Prof. Dr. K. Wüthrich.

1996-1998 4th year of studies at Iowa State University of Science and Technology, Ames IA, USA.

Diploma thesis in biochemistry under the supervision of Prof. Dr. R. W. Thornburg.

Howard Hughes Research Assistantship under the supervision of Prof. Dr. N. M. Kostic.

1993-1996 Undergraduate studies in chemistry at the University of Lausanne.

1990-1993 Maturité fédérale, CESSNOV, Yverdon.