

Diss. ETH Nr. 12684

**Expression, Stabilities, and NMR Studies of
Hydration of Recombinant Variants of the DNA
Binding Domain of Phage 434 Repressor**

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH

for the degree of
Doctor of Natural Sciences

presented by

Hideo Iwai

M.Sc. University of Tokyo, Japan

born on May 8, 1968

citizen of Japan

accepted on the recommendation of
Prof. Dr. Kurt Wüthrich, Referent
Prof. Dr. Rudolf Glockshuber, Korreferent
Dr. Gerhard Wider, Korreferent

1998

Acknowledgment

I wish to acknowledge my thesis advisor Prof. Kurt Wüthrich for giving me this great opportunity to work in his outstanding group and for his guidance. It was a marvellous experience.

I thank Prof. Rudolf Glockshuber and Dr. Gerhard Wider for accepting to be co-examiners of this thesis.

I would like to specially thank PD Dr. Walter Sidler for his support.

I thank the numerous co-workers in the Wüthrich group for their helps, for collaboration, and for many stimulating discussions. In particular I would like to thank: Dr. Walfrido Antuch, Ruedi Baumann, Prof. Martin Billiter, PD Dr. René Brunisholz, Dr. Fred Damberger, César Fernández, Dr. Gerhard Frank, Dr. Peter Güntert, Dr. Chi-yen King, Dr. Reto Koradi, Dr. Francisco López García, Dr. Peter Luginbühl, Dr. Anders Öhman, Dr. Maurizio Pellecchia, Dr. Gregg Siegal, Dr. Alexander Sobol, Dr. Thomas Szyperski, Dr. Ralf Zahn. I also thank many co-workers in the institute for their help and simulating discussions.

I would like to thank specially to Prof. Yoji Arata and Prof. Ichio Shimada for inviting me to the field of biomolecular NMR and for their enormous and continuous supports and their encouragements.

I would like to thank the Japan Society for the promotion of the science for the financial support of part of my study.

I would like to thank my friends for their supports, and finally I thank my parents for their supports and understanding.

This thesis is dedicated to the memory of my two grandmothers and PD Dr. Walter Sidler, who died during this dissertation.

1. Summary	1
2. Introduction	5
3. Hydration study of dh434(0–63) by NMR spectroscopy .	16
3.1. Introduction	16
3.2. Detection of surface hydration water by NMR spectroscopy	16
3.3. Protein design for the hydration study	17
3.4. Discussion	31
3.5. Conclusions and outlook	42
4. Preparation of variants of the DNA binding domain of Phage 434 repressor, 434(1–63)	46
4.1. Introduction	46
4.2. Description of the expression system	46
4.3. Methods for mutagenesis	47
4.4. Preparation of a chimeric DNA binding domain of 434 repressor and P22c2 repressor	52
4.5. Preparation of des-hydroxyl 434(0–63)	55
4.6. Construction, expression, and purification of the DNA binding domain of P22c2 repressor(5–68)	64
4.7. Overexpression and purification of an unstable 434 variant, [E35Q]-434(1–63)	66
4.8. Materials	70
5. Stabilities of variants of the DNA binding domain of phage 434 repressor	71
5.1. Thermal denaturation	71
5.2. Chemical denaturation	76
5.3. Discussion	78
6. Structure determination and Characterization of dh434(0–63) by NMR	82
6.1. ¹ H resonance assignments of dh434(0–63) and secondary structure determination	82
6.2. Solution structures of dh434(0–63)	85
6.3. Comparison between dh434(0–63), 434(1–63), and P22c2(1–76)	91

6.4. ^{15}N relaxation measurements of dh434(0–63)	104
6.5. H/D exchange rates of dh434(0–63) and 434(1–63)	107
6.6. ^{13}C and ^{15}N resonance assignments of dh434(0–63)	109
6.7. Stereospecific assignments of methyl groups by fractionally 10% ^{13}C labelling	115
6.8. Empirical correlation between $^{13}\text{C}_\gamma$ chemical shifts and conformations in valines.	117
6.9. Summary of Chapter 6	122
7. NMR techniques	123
7.1. Pulse scheme for 'artifact-free' water-protein NOEs detection ..	123
7.2. The 2D NMR Experiments H(C)CO ₂ and HCCO ₂	129
8. References	133
9. APPENDICES	145
Curriculum vitae	153

Abbreviations

1D,2D,3D	<i>one, two, three-dimensional</i>
2QF	<i>two-quantum filtered</i>
434(1–63)	<i>N-terminal 63-residue DNA-binding domain of the 434 repressor</i>
434(1–69)	<i>N-terminal 69-residue DNA-binding domain of the 434 repressor</i>
BPTI	<i>bovine pancreatic trypsin inhibitor</i>
bp	<i>base pair</i>
COSY	<i>2D correlation spectroscopy</i>
CPMG	<i>Carr-Purcell-Meiboom-Gill</i>
CW	<i>continuous wave</i>
CD	<i>circular dichroism</i>
ct	<i>constant time</i>
dh434(0–63)	<i>des-hydroxyl 434(0–63)</i>
DSS	<i>2,2-dimethyl-2-silapentane-5-sulfonate</i>
E.COSY	<i>2D exclusive correlation spectroscopy</i>
<i>E.coli</i>	<i>Escherichia Coli</i>
EDTA	<i>Ethylenediaminetetraacetic acid</i>
HSQC	<i>heteronuclear single-quantum correlation</i>
GdmCl	<i>Guanidium Chloride</i>
IPTG	<i>Isopropyl-β-D-thiogalactopyranoside</i>
NMR	<i>nuclear magnetic resonance</i>
NOE	<i>nuclear Overhauser enhancement</i>
NOESY	<i>2D NOE spectroscopy</i>
OD	<i>optical density</i>
P22c2(1–76)	<i>N-terminal 76-residue DNA-binding domain of the P22c2 repressor</i>
ppm	<i>parts per million</i>
PAGE	<i>polyacrylamide gel electrophoresis</i>
PCR	<i>polymerase chain reaction</i>
REDAC	<i>redundant dihedral angle constraints</i>
rf	<i>radio-frequency</i>
RMSD	<i>root mean square deviation</i>
ROE	<i>nuclear Overhauser enhancement in the rotating frame</i>
SDS	<i>sodium dodecyl sulphate</i>
TMS	<i>tetramethylsilane</i>
TOCSY	<i>2D total correlation spectroscopy</i>
TPPI	<i>time-proportional phase incrementation</i>

1. Summary

Water is ubiquitous in biological processes and plays essential roles in the hydrophobic effect, in protein stability, protein folding, molecular association and protein functions. NMR spectroscopy and diffraction methods with single crystals by x-rays and neutrons are the only methods which can provide the locations of water molecules at atomic resolution. This dissertation describes the characterization of the hydration shell of a protein by combination of NMR spectroscopy and protein engineering.

Inherent difficulties in the study of the surface hydration of a protein by NMR spectroscopy with detection of NOEs originates from the fact that NMR spectroscopy can not experimentally separate NOEs due to long-lived water molecules from NOEs via chemical exchange with water of rapidly exchanging labile protons such as hydroxyl protons. Moreover, NOEs from long-lived waters or chemical exchange may conceal weak signals from surface hydration water molecules. In order to overcome such inherent difficulties for the unambiguous identification of water molecules, protein engineering was applied to remove the majority of interfering residues (Chapter 4). The DNA binding domain of the 434 repressor, 434(1–63), was chosen as a model protein because no internal water molecules were found either in the crystal or the NMR structure, and it is soluble up to 20 mM in aqueous solution and can be efficiently expressed and purified from an *E.coli* expression system. A new variant of 434(1–63) was designed for the hydration studies, des-hydroxyl 434(0–63), or short dh434(0–63), which contains no hydroxyl groups.

In the course of the construction of dh434(0–63), several other variants of 434(1–63) and the homologous DNA binding domain of phage P22c2, P22c2(1–76), were constructed. In Chapter 5 the stabilities of these variants are characterized. Interestingly a chimeric protein from parts of 434(1–63) and parts of P22c2(1–76) was found to be significantly more stable than either 434(1–63) or P22c2(5–68). dh434(0–63) was derived from this hyper-stable chimeric protein.

The characterization by NMR spectroscopy revealed that the newly designed dh434(0–63) has a nearly identical backbone conformation to that of 434(1–63) (Chapter 6). For the hydration studies of this protein, it was necessary to optimize the pulse schemes for detection of water–protein interactions at high magnetic fields in order to obtain artifact-free spectra (Chapter 7). Following optimization of both the protein and the NMR experiments, nearly the complete surface of the protein could be characterized.

Protons showing either a positive NOE and a positive ROE, or no NOE and a positive ROE, which is indicative of short-lived water molecules with residence time of less than ~300 psec, were found to cover 88% of the surface, i.e., nearly the entire surface of the protein is covered with highly mobile water.

There remains some ambiguity in the characterization of longer-lived water molecules, but there are indications that the surface structure of the protein is a more important factor than the amino acid type in the determination of the residence times of water molecules, because no clear-cut difference could be detected between the hydration of polar and non-polar amino acids, whereas the presence of longer-lived water molecules near surface cavities was indicated.

Zusammenfassung

Wasser ist allgegenwärtig in biologischen Vorgängen und spielt eine entscheidende Rolle für den hydrophoben Effekt, die Stabilität, Faltung und Funktion von Proteinen sowie für die molekulare Erkennung. NMR-Spektroskopie und die Streuung von Röntgen- oder Neutronenstrahlen an Einkristallen sind die einzigen Methoden, die die Position von Wassermolekülen mit atomarer Auflösung zu liefern vermögen. Diese Dissertation präsentiert die Charakterisierung der Hydratationschale eines Proteins mit Hilfe von NMR Spektroskopie und Protein Engineering.

Die Untersuchung der Hydratisierung der Oberfläche eines Proteins mit Hilfe von NMR-Spektroskopie wird dadurch erschwert, daß NOEs zu langlebigen Wassermolekülen experimentell nicht von NOEs unterschieden werden können, die durch chemischen Austausch von labilen Wasserstoffatomen, zum Beispiel Hydroxylprotonen, mit Wasser zustande kommen. Außerdem können NOEs, die von langlebigem Wasser oder chemischem Austausch herrühren, schwache Signale von Oberflächenhydratisierungswassermolekülen verdecken. Um diese grundlegenden Schwierigkeiten bei der eindeutigen Identifikation von Wassermolekülen zu umgehen, wurden die störenden Aminosäurereste mit Methoden des Protein Engineering entfernt (Kapitel 4). Die DNA-bindende Domäne des 434-Repressors, 434(1–63), wurde als Modellsystem gewählt, weil einerseits weder in der Kristall- noch in der NMR-Struktur interne Wassermoleküle gefunden wurden und andererseits das Protein bis zu einer Konzentration von 20 mM in Wasser löslich ist und mit Hilfe eines E.coli Expressionssystems effizient hergestellt und gereinigt werden kann. Für die Hydratisierungsstudien wurde eine neue Variante von 434(1–63) hergestellt, des-hydroxyl 434(0–63), kurz dh434(0–63), die keine Hydroxylgruppen enthält.

In Verlauf des Herstellung von dh434(0–63) wurden verschiedene andere Varianten von 434(1–63) und der homologen DNA-bindenden Domäne des Phagen P22c2, P22c2(1–76), konstruiert. In Kapitel 5 werden die Stabilitäten dieser Varianten charakterisiert. Interessanterweise erwies sich ein chimäres Protein, kombiniert aus Teilen von 434(1–63) und P22c2(1–76) als signifikant stabiler als sowohl 434(1–63) als auch P22c2(5–68). dh434(0–63) wurde von diesem hyperstabilen chi-

mären Protein abgeleitet.

Die Untersuchung mit Hilfe der NMR-Spektroskopie zeigte für dh434(0–63) eine fast identische Rückgratskonformation wie für 434(1–63) (Kapitel 6). Für die Hydratisierungsstudien an diesem Protein mußten die Pulssequenzen zur Detektion von Wasser-Protein Wechselwirkungen bei hohen Magnetfeldern optimiert werden, um artefaktfreie Spektren zu erhalten (Kapitel 7). Nach der Optimierung des Proteins und der NMR Experimente konnte nahezu die gesamte Oberfläche des Proteins charakterisiert werden.

Auf 88% der Proteinoberfläche wurden Protonen gefunden, die entweder einen positiven NOE und einen positiven ROE, oder keinen NOE und einen positiven ROE zeigten, was kurzlebige Wassermoleküle mit einer Kontaktlebensdauer von weniger als ~300 psec anzeigt, d.h. fast die ganze Proteinoberfläche ist von hochmobilen Wassermolekülen bedeckt.

Eine gewisse Unsicherheit bezüglich der Charakterisierung langlebiger Wassermoleküle bleibt bestehen. Aber es gibt Anzeichen, daß die Oberflächenstruktur des Proteins die Kontaktlebensdauer von Wassermolekülen stärker beeinflusst als der Aminosäuretyp, weil kein klarer Unterschied in der Hydratisierung von polaren und unpolaren Aminosäuren festgestellt werden konnte, wogegen langlebige Wassermoleküle bei Oberflächenvertiefungen zu erkennen waren.

2. Introduction

Water is the ubiquitous solvent for biological systems. Despite its simple chemical structure, water is a particularly complex solvent and an excellent solvent for a wide variety of solute molecules and ions, and for suspending colloidal particles and biological molecules such as proteins, DNA, viruses, and cells. The importance of water has long been appreciated in a wide range of fields and extensive studies of water have been performed by experimentalists and theorists (Eisenberg & Kauzmann, 1969; Ben-Naim, 1974; Franks, 1983; Ruplay & Careri, 1991; Weshhof, 1993; Gregory, 1995). However, water is still a mysterious liquid and in spite of much effort very little is known about the detailed structure and dynamics of water molecules in particular at the macromolecular interfaces. Because of the lack of knowledge about the detailed structure of water, water structure has conveniently been used for explaining unexpected experimental results, for example, “polywater” or “memory of water” (Rousseau *et al.*, 1971; Davenas *et al.*, 1988; Benveniste *et al.*, 1988, 1991, 1994).

The hydrophobic interaction

The hydrophobic interaction is one of the forces caused by solvation effects. The hydrophobic interaction is generally the entropic driving force that tends to bring together two non-polar groups in aqueous solution (Tanford, 1980). It is significantly temperature sensitive and can be modulated by changing the properties of the solvent, for example by adding ions. The hydrophobic effect is generally accepted as an important driving force in many process such as protein folding, protein-protein interaction, protein stability, and enzyme-ligand interaction. However, the molecular origin of the hydrophobic interaction itself is still unclear and for example protein denaturation with increasing temperature appears to be contradictory because the entropic force, *i.e.* hydrophobic interaction should increase with an increase in temperature. The entropic consideration of the hydrophobic effect leads to the concept

of the “iceberg”, where water molecules around hydrophobic sites are ordered, structured, or “frozen” to decrease the entropy, as suggested by Frank & Evans (1945). The “iceberg” concept is also applied to proteins (Klotz, 1958). The concept of “iceberg” is often extended to the “hydration shell” on the protein surface, which is an old idea of water structure, envisaging a layer of water molecules surrounding molecules and particles, which provides a protective barrier against the approach of two hydrophilic surfaces or groups. This idea has long been used for the explanation of the molecular mechanism of the effect of ions and cosolute on protein stability (Collins & Washabaugh, 1985; Timasheff & Arakawa, 1989). X-ray diffraction has supported these concepts of “iceberg”, ordered water molecules, or “frozen water molecules” around non-polar groups by the observation of ordered water molecules in small compounds and in proteins (Saenger, 1987; Baker *et al.*, 1987; Teeter, 1984; Teeter *et al.*, 1993; Kumaraswamy *et al.*, 1996). However, there are only a few cases in which ordered water molecules at a hydrophobic surface have been observed despite thousands of crystal structures of proteins, which have been solved up to now (Protein Data Bank, Bernstein *et al.*, 1977). The picture of a hydration shell thus provided by diffraction methods and the concept of ordered water have suppressed a crucial aspect of the hydration in proteins, namely, its dynamics. It is worth noting that so far there is no clear experimental evidence that water molecules around non-polar side chains are more ordered than around polar residues for proteins in solution. Difference of water structures around polar and non-polar side chains at atomic resolution is of special interest.

Protein stability

The stability of the native structure of a protein is maintained by a delicate energy balance. This is manifested by the fact that only small perturbations in a structure such as single point mutations may result in large changes in stability (for example, see Chapter 5). It has also long been known that some cosolutes such as urea and guanidium chloride break the balance denaturing the native structure and that different salts and cosolutes effect the stability and solubility of proteins (Kauzmann, 1959). The different effects of ions on proteins were systematically summarized more than one hundred years ago in the Hofmeister series, providing

an important basis for common practices of protein crystallization and preservation of enzymes activities (Hofmeister 1888; Collins & Washabaugh, 1985; Cacace *et al.*, 1997). A molecular model of water structures at interfaces has been used to interpret the molecular mechanism of the effects of the Hofmeister series (Collins & Washabaugh, 1985). Water molecules appear to be both the cement that fills crevices between protein building elements such as helices and sheets, and the lubricant that allows motion of these building elements. The broader role of water in stabilizing protein structure is still a subject of controversy (Das *et al.*, 1989; Berndt *et al.*, 1993; Matthews, 1996; Buckle *et al.*, 1996; Takano *et al.*, 1997).

Protein (un)foldi

The early events in protein folding are fast processes. A recent study has suggested that the initial collapse to a compact structure can happen within micro second time scale (Hagen *et al.*, 1996). During this fast process the water molecules on the surface of proteins have to be excluded in order to allow formation of the core of the protein. There have been speculations of the role of water molecules in protein unfolding and folding. For example it has been proposed that In “molten globule” states, or “compact denatured” states, water molecules penetrate more readily into the core of proteins than native structures (Ptitsyn, 1992). Some evidence such as proton exchange rates in “molten globule” structure somehow support this idea (Chyan *et al.*, 1993). It has also been proposed that molten globule states may be intermediate states in the folding process. Statistical analysis of the location of water molecules in crystal structures revealed that the N and C terminus of helices are often found to be highly hydrated (Baker & Hubbard, 1984; Blundell *et al.*, 1983; Sundaralingam & Sekhurudu, 1989; Thanki *et al.*, 1991). The insertion of water molecules in the termini of helices are proposed to be an intermediate step in the unfolding and folding of α helices (Sundaralingam & Sekharudu, 1989). Water seems to be a crucial player in the unfolding process. Detailed knowledge of water-protein interactions will be beneficial for protein engineering in order to modify the folding process and to prevent unfolding of proteins.

Water and proton exchange

Proteins and nucleic acids contain many hydrogens in polar groups that are in continual exchange with hydrogens of the solvent. It has long been accepted that protein hydrogen exchange depends on some aspects of structural mobility (see review by Englander & Kallenbach, 1984). Hydrogen isotope exchange has long been recognized and used as a powerful method to characterize structural and conformational dynamics of proteins in folded states as well as during folding (Wagner *et al.*, 1984; Roder *et al.*, 1986; Radford *et al.*, 1992).

As mechanisms of proton exchange in proteins, local structural unfolding models, or solvent penetration models have been proposed. It is apparent that the solvent molecules play a central role to cause proton exchange in proteins. The detailed interaction between water and protein will shed light on the mechanism of hydrogen exchange such as relevance of residence times of hydration water molecules to the hydrogen exchange rates.

Protein functions

Experimental observations indicate that hydration of protein is indispensable for enzyme activities. X-ray crystallography often reveals bound water molecules in the active sites of enzymes. The importance of such water molecules in catalytic reaction is well recognized (for example, Warshel *et al.*, 1989). Water molecules have been found to be involved in inter molecular interactions such as, protein-protein, ligand-protein, and DNA-protein complexes (Janin & Chothia, 1990; Harrison & Aggarwal, 1990). Water molecules also appear to be an important element for molecular recognition (Janin & Chothia, 1990; Billeter *et al.*, 1996).

Location of water molecules in proteins

The positions of water molecules are mostly determined or identified by x-ray crystallography (Saengar, 1987; Jeffery & Saengar, 1991; Baker, 1995). Some water molecules are also observed at the same position by NMR spectroscopy as found in x-ray crystallography (see review, Gerotheranassis, 1994). Water molecules are often found in the core of protein as integral parts of protein structures, which are referred to “interior” water molecules. In addition, water molecules are also found on

the surface of proteins. Statistical analysis of the locations of water molecules in crystal structures reveals several patterns. The preference of water molecules to oxygen atoms rather than nitrogen atoms has been noted (Baker, 1980; Blake *et al.*, 1983; James & Sielecki 1983; Baker & Hubbard 1984). A predominance of interactions with main chain groups rather than side chains is also found (Thanki *et al.*, 1988; Baker, 1995). Statistical analysis also shows correlations between hydration sites identified by crystallography and the geometry of proteins (Kuhn *et al.*, 1992). Several patterns of hydration which are associated with secondary structures have also been reported (see review, Baker, 1995). It is difficult to address the generality of these patterns of hydration water molecules in single crystals and their relevance to protein hydration in aqueous solution because of their inconsistency with results from NMR. For example the distribution of sites showing NOE cross peaks to water in BPTI does not always correlate well with the location of water molecules found in crystal structures (Otting *et al.*, 1991a).

Crystallographic and NMR Detection of Hydration Water

There are many fundamental processes and phenomena in which water seems to play a crucial role. In order to understand such processes and phenomena involving the water-protein interactions, a number of methods have been used for the characterization of hydration of proteins such as infrared spectroscopy, Raman spectroscopy, dielectric spectroscopy, NMR spectroscopy, X-ray and neutron scattering, differential scanning calorimetry, and so on (see reviews by Rupley & Carei, 1991; Gregory, 1995). Recent developments of various techniques have improved the characterization of properties of water. However, X-ray and neutron diffraction methods and NMR spectroscopy are still the only methods which can identify the location of water molecules at atomic resolution in proteins.

The crystallographic method

The crystallographic method is the most powerful method to determine the protein structures. There have been more than five thousand protein structures determined by x-ray crystallography (Protein Data Bank, Bernstein *et al.*, 1977). The number of structures determined by x-ray diffraction is rapidly growing. Direct interactions between proteins and water molecules in the interior and at the surface

of proteins are vividly manifested by the crystallography although only the positions of the oxygen atoms from water molecules are observed (Savage & Wlodawer, 1986; Saenger, 1987; Baker, 1995). X-ray crystallography observes the atomic and molecular structure as a three dimensional map of the distribution of electrons. Hence there are several difficult aspects for the identification of water molecules.

(1) Model bias

It is generally difficult to locate reliably solvent molecules in the density map because solvent molecules at the surface often have low electron densities (Savage & Wlodawer, 1986; Baker, 1995). Therefore, solvent molecules are not included in early stages of analysis, or when the resolution is limited. In addition when inaccurate or incomplete experimental phases are substituted or augmented by models, model bias will make it more difficult to interpret low density maps (Burling *et al.*, 1995). Although X-ray crystallography is a powerful method to analyze the hydration pattern on the surface, the determination of water structure around biomolecules is still a difficult task and can be biased easily during model building.

(2) Crystal contacts

Diffraction methods absolutely require a crystal for a high resolution structure. Proteins in crystals have significant contacts with neighboring molecules. For example in human lysozyme crystals with 37% solvent content 38% of the surface is involved in crystal packing. Whereas turkey egg white lysozyme crystals with 51% solvent content have 13% of the surface covered by crystal contacts (Baker *et al.*, 1983). Tighter crystal packing generally leads to better ordered crystals and the possibility of higher-resolution x-ray diffraction data. This allows better identification of water molecules. In fact crystals in which many ordered water molecules are observed often have low solvent content. For example the crystals of 2-Zn insulin crystals (Baker *et al.*, 1988), γ B-crystalline (Kumaraswamy *et al.*, 1996), and crambin (Teeter, 1993) have 30%, 36%, and 32% solvent content respectively. This suggests that for the identification of water molecules tighter packing is preferable because it results in higher resolution data but this also engenders that a larger percentage of the surface is involved in crystal packing. Thus it is impossible to characterize the

entire surface hydration shell of a protein by crystallography.

(3) Lack of dynamics

X-ray diffraction data results from time and spatial average over a large number of molecules in the crystal lattice and therefore provides an averaged view of solvation and flexibility. Two important factors of the quality of the solvent molecules are their occupancies and B factors. The occupancy refers to the fraction of time that a particular site is occupied in the crystal. The B factor, or temperature factor, in practice cover both thermal vibration and positional disorders in the crystal structures. These data might provide some idea about dynamics of solvent molecules. But it is quite difficult to derive the reliable information of dynamics from crystal structures. In addition crystal contacts may reduce the mobility of water molecules. Thus, it is difficult to answer the relevance of the observation of water molecules in crystal to hydration water in aqueous solution. The most frequently used approach is the statistical analysis of the positions of solvent molecules in crystals for the characterization of hydration (Thanki *et al.*, 1988, 1991; Kuhn *et al.*, 1992).

Neutron Scattering

In x-ray diffraction the scattering power of an atom is proportional to its number of electrons. Therefore usually the protons are not visible. In contrast some isotopes such as H and D can easily be distinguished in neutron density maps so that neutron scattering experiments can provide more reliable, precise and unambiguous localizations of H atoms in crystal structures. Neutron diffraction is often used for characterization of hydration water and verification of X-ray diffraction data (Savege & Wlodawer, 1986).

NMR spectroscopy

NMR can characterize water molecules in solution without artifacts from crystal contacts. NMR spectroscopy does not require ordering of molecules as is required in crystallographic methods. This makes a clear contrast with diffraction methods and is of advantage for the characterization of hydration around flexible side chains. There have been many attempts to characterize the hydration shell by NMR spec-

troscopy. Two major features of NMR spectroscopy have been used for the characterization of hydration molecules. One is the relaxation time of water molecules using ^1H , ^2H , and ^{17}O nuclei (Jacobson *et al.*, 1954; Koenig & Schillinger, 1969; Akasaka, 1981; Bryant, 1996). The other is the detection of NOEs for identification of water (Pitner, *et al.*, 1974; Otting & Wüthrich, 1989; Otting *et al.*, 1991a,b; Gerotheranassis, 1994).

Nuclear Relaxation

Relaxation of water molecules has long been used for characterization of water molecules by analyzing line shapes and measuring longitudinal relaxation rates (Jacobson *et al.*, 1954; Koenig & Schillinger 1969; Akasaka, 1981; Bryant, 1996). The frequency dependence of the longitudinal relaxation rate, the nuclear magnetic relaxation dispersion (NMRD), contains information about water–protein interactions and dynamics (Koenig & Schillinger 1969; Denisov & Halle, 1996; Bryant, 1996). The relaxation measurements cannot provide both the position and the dynamics of individual water molecules. The disadvantage of relaxation approach is lack of the information about the location of water molecules as compared with detection of NOEs.

Nuclear Overhauser Effect

Nuclear Overhauser effect is not only essential for the determination of three dimensional structures of proteins but also used for the identification of the location of water molecules in aqueous solution (Wüthrich *et al.*, 1996). Individual water molecules are identified via dipolar cross-relaxation between the protein protons and water protons. Detection of NOEs by using multidimensional NMR spectroscopy has identified a handful of “long-lived” water molecules in proteins and nucleic acids (see review by Kubinac & Wemmer, 1992; Gerotheranassis, 1994). In addition the sign and ratio of the cross relaxation rates of σ^{NOE} and σ^{ROE} have been used for the characterization of the dynamics of water molecules at the individual sites (Otting *et al.*, 1991a).

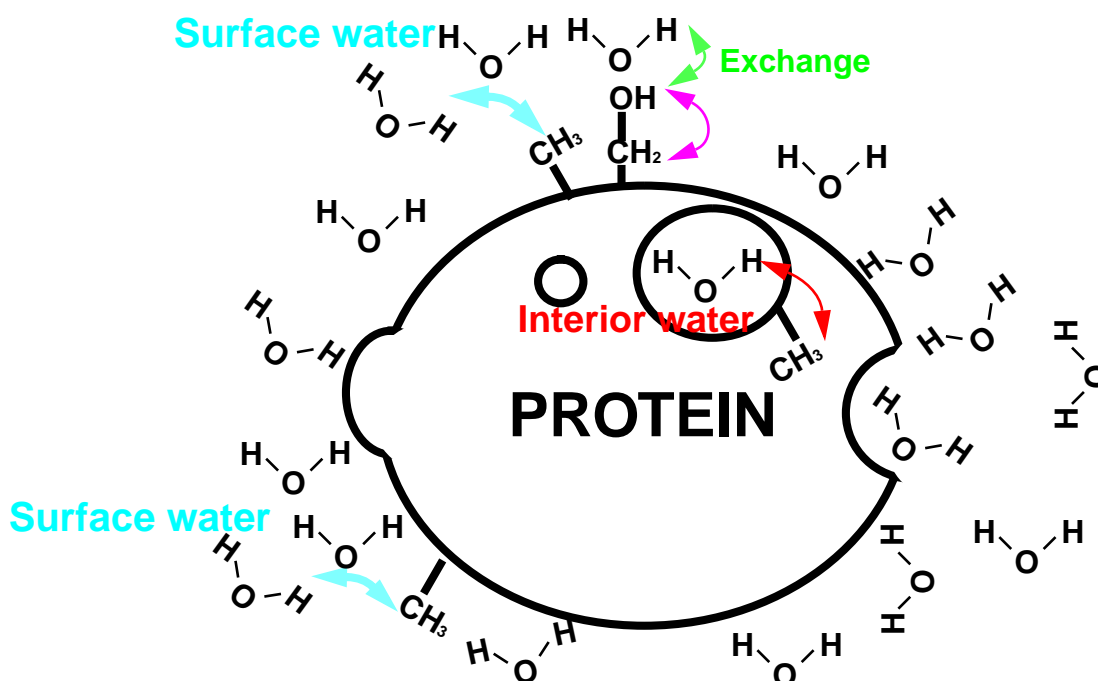


Figure 2-1: Schematic drawing of the observation of hydration water molecules by NMR. The arrows indicate NOEs originated from different interactions.

The limitation of NMR methods

NMRD can provide information on some dynamical features of water but lacks the information of the location and its atomic resolution. On the other hand detection of NOEs can provide both the location and dynamics of water molecules because NOE cross peaks between water and the protein tell proximity to the hydration site and its dynamics although it can not determine the precise positions of water molecules as determined by crystallography. Identification of water molecules detecting NOEs, however, is an attractive method to characterize the location and dynamics of the first hydration shell of a protein in aqueous solution. NOE cross peaks at the water resonance in NOESY spectra can be categorized into one of three groups. *i.e.* (1) direct NOE with water molecules (red and blue arrows in Figure 2-1). (2) NOE *via* rapidly exchanging protein protons such as hydroxyl protons and amino protons (an arrow in magenta in Figure 2-1). (3) intramolecular NOEs with protons of the protein which have chemical shifts at or close to the water resonance. The third group can be easily separated by using several NMR techniques such as isotope fil-

tering, diffusion filtering, and relaxation filtering (Grzesiek & Bax, 1993a,1993b; Wider *et al.*, 1996; Mori *et al.*, 1994). On the other hand NOEs originating from rapidly exchanging labile protons in proteins cannot be distinguished experimentally from NOEs from long-lived water molecules because they produce NOEs of the same sign as NOEs from long-lived water molecules at the water resonance due to their rapid exchanges with water. The NOEs with water molecules can also be categorized into two groups, (1) NOEs with long-lived water molecules which show negative cross relaxation rates σ^{NOE} (conventionally positive cross peaks in NOESY spectra), illustrated as a red arrow in Figure 2-1, and (2) NOEs with short-lived water molecules illustrated as a blue arrow in Figure 2-1, which show positive cross-relaxation rates σ^{NOE} (negative peaks in NOESY spectra) (Otting *et al.*, 1991a). There have been few reports on short-lived water molecules (Otting *et al.*, 1991a; Otting *et al.*, 1992; Brüschweiler *et al.*, 1995). This is presumably because weak peaks from short-lived water molecules can be covered by strong cross peaks of negative NOEs originating from long-lived water molecules and the rapidly exchanging labile protons of proteins. Thus, long-lived water molecules and rapidly exchanging labile protons are main obstacles to study the surface hydration of proteins. Therefore there has so far been no report on the complete hydration shell of a protein by NMR spectroscopy.

For the better understanding of various biologically relevant phenomena, we would like to understand the protein hydration. What do we have to know about water for the better understanding of the protein hydration? First the location of water molecules, the structure of the hydration shell, the difference of water structure around non-polar and polar groups, and the detailed hydration geometry at the interface. Second, the movement of water molecules i.e. dynamics of water molecules at the surface of proteins, and the time scale of the dynamics of water molecules. The relevant time scale for understanding a particular process, or biological function. Coupling of the water dynamics and dynamics of the macromolecules. Third, by characterizing the aforementioned features, to explain the behavior of a particular system, not only at the molecular level such as protein-substrate binding, protein folding, but also on the increasingly important mesoscopic scale such as protein aggregation, and fibrous formation.

This dissertation describes the characterization of the hydration shell of a protein by the combination of protein engineering and high field NMR spectroscopy to gain new insights for understanding protein hydration.

3. Hydration study of dh434(0–63) by NMR spectroscopy

3.1. Introduction

Despite enormous interests on protein hydration there has been no study on the complete hydration shell of a protein in aqueous solution at atomic resolution. This chapter presents the main part of the thesis and describes for the first time the complete hydration of a protein by NMR spectroscopy with a model protein, des-hydroxyl 434(0–63). The details of the preparation, characterization, structure determination of this protein, and NMR techniques are described in the following chapters 4 to 7.

3.2. Detection of surface hydration water by NMR spectroscopy

Detection of nuclear Overhauser effect (NOE) between proteins and water molecules is the only method to determine the hydration sites in proteins at atomic resolution by NMR spectroscopy, which has been extensively used for the identification of hydration water molecules in proteins (Pitner *et al.*, 1974; Otting & Wüthrich, 1989; Clore *et al.*, 1990; Forman-Kay *et al.*, 1991; Otting *et al.*, 1991a,b; Clore *et al.*, 1992; Otting *et al.*, 1992; Grzesiek & Bax 1993a,1993b; Kriwacki *et al.*, 1993; Xu *et al.*, 1993; Clore *et al.*, 1994; Gerotheranassis, 1994; Grzesiek *et al.*, 1994; Qi *et al.*, 1994; Qin *et al.*, 1994; Brüschweiler *et al.*, 1995; Ernst *et al.*, 1995a; Wang *et al.*, 1996).

In this study we have also used the detection of NOEs with water molecules for the characterization of the surface hydration of a protein, which can provide the location and dynamics of water molecules on the surface.

However, such NOEs can be originating from three different origins as discussed in Chapter 2. First is NOEs originating from long-lived water molecules with residence time of longer than 0.5 nsec such as interior water molecules illustrated by a red arrow in Figure 2-1 (Otting *et al.*, 1991a). Second is NOEs originating from short-lived water molecules with residence time of shorter than 0.5 nsec shown by

blue arrows in Figure 2-1 (Otting *et al.*, 1991a). In principle these two contributions come from the same origin, i.e. water molecules, but because of their time scale of the dynamics, can be differentiated by the difference in their sign of the cross relaxation rate σ^{NOE} (Otting *et al.*, 1991a). Third is NOEs mediated via fast exchanging labile protons in proteins, in particular, such as hydroxyl groups which produce strong NOE cross peaks in the spectra (an arrow in magenta shown in Figure 2-1). Since positive NOEs were observed for side chains of lysine and arginines in the previous reports, the most problematic amino acid types seem to be amino acids containing hydroxyl groups (Otting *et al.*, 1991a; Liepinsh *et al.*, 1992; Brüschweiler *et al.*, 1995). Usually NOEs originating from long-lived water molecules and *via* rapidly exchanging labile protons produce strong cross peaks of negative NOEs, which cover weak peaks of the opposite sign originating from short-lived water molecules and cannot be experimentally distinguished each other, which would complicate the NMR data analysis. Thus, for the observation of the complete hydration shell of a protein in aqueous solution by NMR spectroscopy the protein of no interior water molecules and no hydroxyl groups would be desirable.

3.3. Protein design for the hydration study

As a model protein, the DNA binding domain of 434 repressor was chosen for this surface hydration study by NMR spectroscopy. The DNA binding domain of 434 repressor is well-characterized both by X-ray crystallography at high resolution and by NMR spectroscopy (Mondragon *et al.*, 1989; Neri *et al.*, 1992a; Pervushin *et al.*, 1996). It is found that the crystal structure of 434(1–69) (Figure 3-11) contains no interior water molecules as found in BPTI of which interior water molecules were studied in great details and disturbs the observation of the complete surface hydration (Otting *et al.*, 1991a; Dötsch *et al.*, 1995; Denisov *et al.*, 1995a, 1996). Moreover, the DNA binding domain of 434 repressor, is highly soluble up to 20 mM in aqueous solution making it possible to gain sensitivity with highly concentrated samples. It can be also expressed and purified from *E.coli* expression system which makes it easy to design a new suitable variant for the characterization of the hydration by NMR spectroscopy. A protein, des-hydroxyl 434(0–63), in short notation, dh434(0–63), was designed which has no hydroxyl groups and is expected to have no interior

water molecule for the clear observation of the hydration shell. The details of the construction and preparation of this dh434(0–63) is described in Chapter 4 (see Appendix H). The high resolution solution structure of dh434(0–63) was determined by NMR spectroscopy and found to have a very similar backbone conformation to that of the original 434(1–63) (Chapter 6). There is no conceivable difference in the backbone conformation observed between dh434(0–63) and 434(1–63). This dh434(0–63) was used as a model protein for the characterization of the complete hydration by NMR spectroscopy. Since there is no hydroxyl groups in the protein, water-protein NOEs can be more clearly analyzed than any other proteins. This system will be the best system for the hydration study by NMR spectroscopy.

Sample preparation & NMR spectroscopy

The newly designed protein, dh434(1–63) was cloned in *E. coli*, and the uniformly $^{13}\text{C}/^{15}\text{N}$ -labelled dh434(0–63) was expressed and purified as described in Chapter 4. The NMR samples contain 3 mM protein in 20 mM potassium phosphate buffer, or 7 mM protein without buffer pH 4.8, in a mixed solvent of 90% $\text{H}_2\text{O}/10\%$ D_2O . All the final NMR experiments for the hydration studies was carried out with a 7 mM uniformly $^{13}\text{C}/^{15}\text{N}$ labelled (> 98%) dh434(0–63) without any buffer. The pH was adjusted to pH 4.8 by addition of minute HCl and NaOH. The concentration of the sample was determined by the optical density at 280 nm. All the NMR experiments for the identification of water molecules were performed using Varian Unity-plus operating at ^1H frequencies of 750 MHz with a triple resonance probe head equipped Z-gradient coil.

For the detection of NOEs between water and the protein protons, two-dimensional [^{15}N , ^1H]-HSQC relayed and [^{13}C , ^1H]-HSQC relayed NOE/ROE difference experiments. Special care was taken to suppress artifacts originating from the strong solvent magnetization at high magnetic field (Sobol *et al.*, 1998; see Chapter 7). To minimize the effect of radiation damping and demagnetizing field effect the water magnetization was carefully monitored during the whole experiment to produce artifact-free water-protein NOE/ROE spectra. In addition prior to the mixing period the proton magnetization attached to $^{13}\text{C}/^{15}\text{N}$ was destroyed by using a double filter unit to minimize the excitation due to radiation damping (Sobol *et al.*, 1998; Chapter

Table 3-1: Summary of the two-dimensional difference experiments described in Chapter 7 for the observation of hydration water molecules with 7 mM dh434(0–63) at pH 4.8 and 13°C measured at a ^1H frequency of 750 MHz.

Type	ω_1	region	mixing (msec)	total measurement time (hours)	$t_{1\text{max}}$ (msec)	$t_{2\text{max}}$ (msec)
NOE	^{13}C	aliphatic	60	61	12.8	66.7
ROE	^{13}C	aliphatic	30	52	12.8	66.7
NOE	^{13}C	aromatic	60	35	7.4	66.7
ROE	^{13}C	aromatic	30	28	7.4	66.7
NOE	^{15}N	-	60	37	32	66.7
ROE	^{15}N	-	30	40	33.3	66.7

7). The detail of the experiments were described in Chapter 7. Table 3-1 summarizes all the NMR experiments for the detection of hydration water molecules in dh434(0–63). For the controls, all the measurements were performed with the same setups but with weak pre-saturation during relaxation periods with 30~100% of the total measurement times. The control spectra contained no peaks except for a narrow region near the water resonance (see Chapter 7 for the details).

Estimation of the residence times based on the diffusion model

Otting *et al.* proposed the estimation of the residence times of the hydration water molecules based on the diffusion model (Otting *et al.*, 1991a; Ayant *et al.*, 1977). The estimation of the residence times depends on the rotational correlation time and the size of the spheres in the model (Figure 3-1A). The cross relaxation rates were calculated from the experimental data. The rotational correlation time of dh434(0–63) at 13°C was experimentally obtained as described in Chapter 6. The radius of gyration of dh434(0–63) was calculated from the 20 NMR conformers of dh434(0–63) for the three axis, $10.92 \pm 0.06 \text{ \AA}$, $10.64 \pm 0.05 \text{ \AA}$, $9.82 \pm 0.08 \text{ \AA}$ respectively using program MOLMOL (Koradi *et al.*, 1995). The radius of the protein $r^P=11 \text{ \AA}$, the rotational correlation time of the protein, $\tau_R^P=6 \text{ nsec}$, the radius of a water molecule, $r^W=2.0 \text{ \AA}$, and the rotational correlation time, $\tau_R^W=3.5 \text{ psec}$ were used for this cal-

ulation. The cross relaxation rates in the laboratory frame $\sigma^{\text{NOE}} = 6J(2\omega) - J(0)$, in the rotating frame $\sigma^{\text{ROE}} = 3J(\omega) + 2J(0)$, and their ratio $\sigma^{\text{NOE}}/\sigma^{\text{ROE}}$ were calculated by the following equations for the spectral density function $J(\omega)$ (Ayant *et al.*, 1977).

$$J_2(\omega) = \frac{N_s}{24\pi D b} \sum_{l=0}^{\infty} (2l+1)! \sum_{\lambda=0}^l \frac{1}{(2\lambda+1)! [2(l-\lambda)+1]!} \left(\frac{\rho_I}{b}\right)^{2\lambda} \left(\frac{\rho_S}{b}\right)^{2(l-\lambda)} A_{l,\lambda}(\omega)$$

Where $A_{l,\lambda}(\omega)$ and $v_{l,\lambda}(\omega)$ are defined as follows.

$$A_{l,\lambda}(\omega) = \text{Re} \left\{ \frac{1}{v_{l,\lambda}^2} \left[\frac{1}{2l+3} - \frac{l+3}{v_{l,\lambda}} \left(1 + \frac{(l+3) K_{l+5/2}(v_{l,\lambda})}{v_{l,\lambda} K_{l+3/2}(v_{l,\lambda})} \right)^{-1} \right] \right\}$$

$$v_{l,\lambda}(\omega) = \{ i\omega\tau + \tau[\lambda(\lambda+1)D_I^r + (l-\lambda)(l-\lambda+1)D_S^r] \}$$

in which $b = r^p + r^w$ (r^p and r^w are radius of the protein and water molecules respectively), $\rho^I = \rho^p$ (the distance from the center of the protein), $\rho^S = \rho^w$ (the distance from the water), $\tau = b^2/D$, D_S^r and D_I^r are the rotational diffusion coefficients ($D_S^r = 1/6\tau_w$ and $D_I^r = 1/6\tau_p$) (see the figure captions in Figure 3-1).

The average residence times are derived from $\tau_{\text{res}} = x^2/(6D)$, where $x = 4\text{\AA}$ was used as a criterion for the observation of NOEs (Otting *et al.*, 1991a). Figure 3-1 presents the numerical calculation from these equations with the experimental data of dh434(0–63). This model predicts that the sign change in NOE may be expected to occur for residence time around ~ 300 psec (Figure 3-1C). This residence times estimated from the cross relaxation rates based the diffusion model appears to depend on the shape of the molecules and the geometry of the molecules (Brüshweiler *et al.*, 1994). However, the observation of positive NOEs can be still unambiguously attributed to the short-lived water molecules (Brüshweiler *et al.*, 1994; Otting & Liepinish, 1995).

Beside of the validity of the model for the estimation of the residence times, it is worth noting that the estimated residence times derived from the experimental data have already a tendency to be longer due to the cancellation of positive NOEs by negative NOEs from neighboring protons reducing the ratio of $\sigma^{\text{NOE}}/\sigma^{\text{ROE}}$. In other words, the estimation of the residence times strongly depends on their neighboring

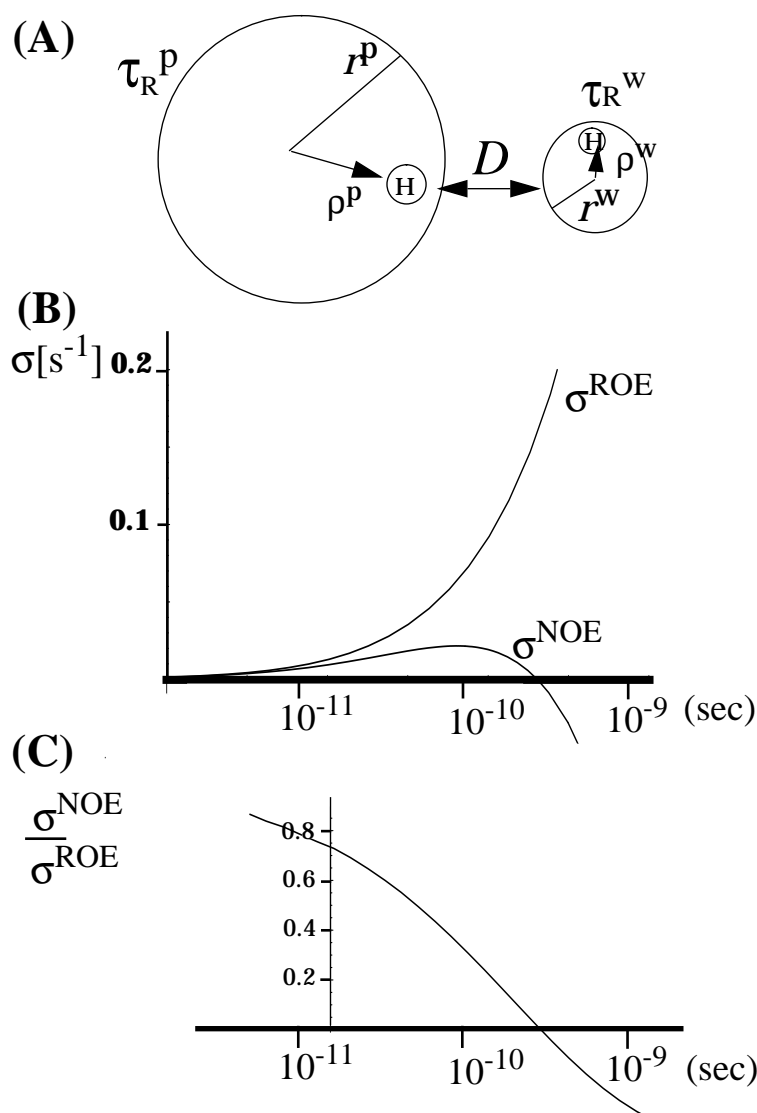


Figure 3-1: Estimation of the residence times of water molecules based on the diffusion model (Ayant et al., 1971; Otting et al., 1991). (A) Diffusion model. The radius of a protein, the correlation time of a protein, the radius of water, and the correlation time of water are presented by r^p , τ_R^p , r^w , and τ_R^w , respectively. The distance from the hydrogen atoms to the center of the radius of a protein and water are indicated by ρ^p and ρ^w (B) Plot of the estimated relaxation rates of σ^{NOE} and σ^{ROE} . (C) Plot of the ratio of σ^{NOE} and σ^{ROE} . In (B) and (C) $r^p=11 \text{ \AA}$, $\rho^p = 10 \text{ \AA}$, $\tau_R^p = 6 \text{ nsec}$, $r^w = 2.0 \text{ \AA}$, $\rho^w = 1.0 \text{ \AA}$, and $\tau_R^w = 3.5 \text{ psec}$ were used for the calculation.

protons in particular labile protons in the three-dimensional structure of dh434(0–63). It is quite difficult to assess all the contribution from the neighboring protons. Thus, only the statistical analysis could be meaningful for the estimation of the residence times.

Hydration of dh434(0–63) by NMR spectroscopy

Figure 3-2 presents the aliphatic region of the [^{13}C , ^1H]-HSQC relayed NOE/ROE difference spectra providing an impression of the quality of the spectra. Conventi-

ally a positive NOE (positive σ^{NOE}) gives rise to a negative cross peak and *vice versa*. In all the spectra presented here negative NOEs and positive NOEs are drawn by solid lines and dotted lines, respectively. It can be readily recognized that there are a few very strong cross peaks showing negative NOEs in Figure 3-2A. Those cross peaks are assumed to be originating from $\alpha\text{-NH}_3$ group at the N-terminus of dh434(0–63) which is chemically exchanging with water molecules as rapidly as hydroxyl groups (Wüthrich, 1986).

In the aliphatic region of the spectra in Figure 3-2, 123 peaks was possible to be identified, which can be used for the analysis of the hydration pattern of dh434(0–63). The identified hydration sites are mapped into a CPK-model of dh434(0–63) in Figure 3-3. The hydrogen atoms showing both positive NOEs and ROEs are colored in green. The hydrogen atoms showing positive ROEs but no NOEs are colored in light green. These hydrogen atoms are clear indications of short-lived water molecules in its proximity. The chemically exchanging groups are indicated in magenta. The hydrogen atoms in white are the hydrogen atoms which have the average solvent accessible surface of more than 5\AA^2 in the 20 NMR conformers of dh434(0–63) but for which no unambiguous evidence for the hydration could be identified mainly because of the signal overlaps. The hydrogen atoms in yellow is the hydrogen atoms showing negative NOEs, which could be originating from the long-lived water molecules, or, *via* rapidly exchanging labile protons.

Based on the obtained spectra, the individual hydration sites can be inspected in more details in the following. In particular, the difference between polar and non-polar hydrations, the geometry of the hydration sites are of special interest.

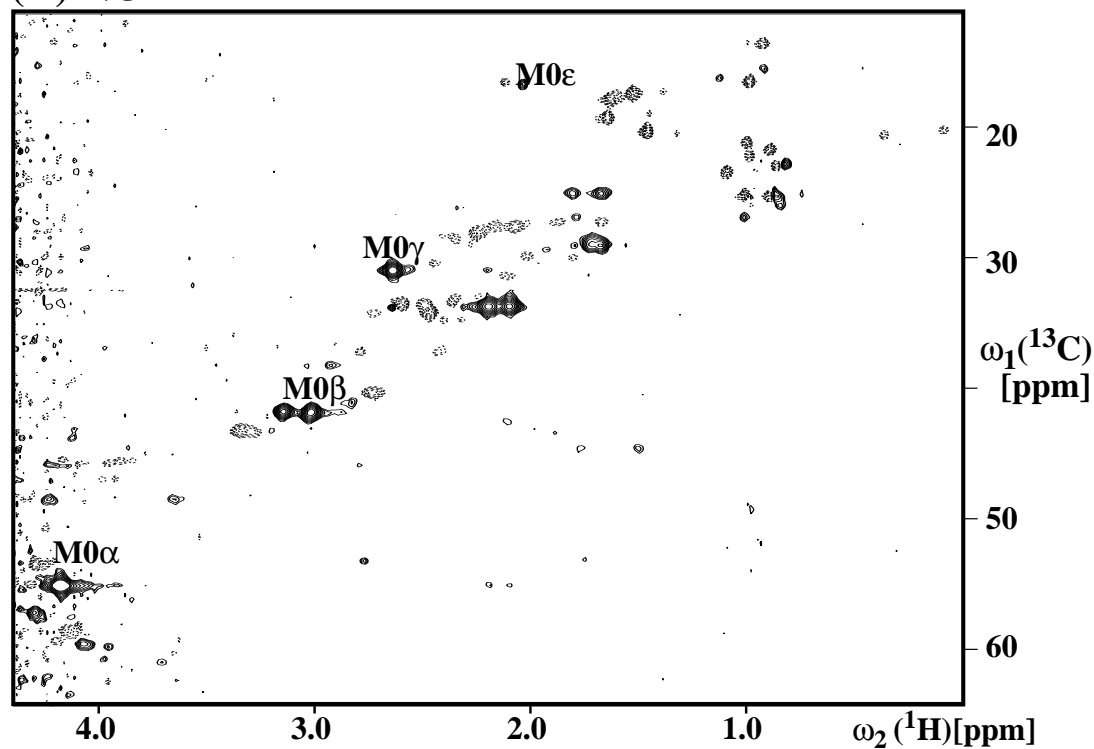
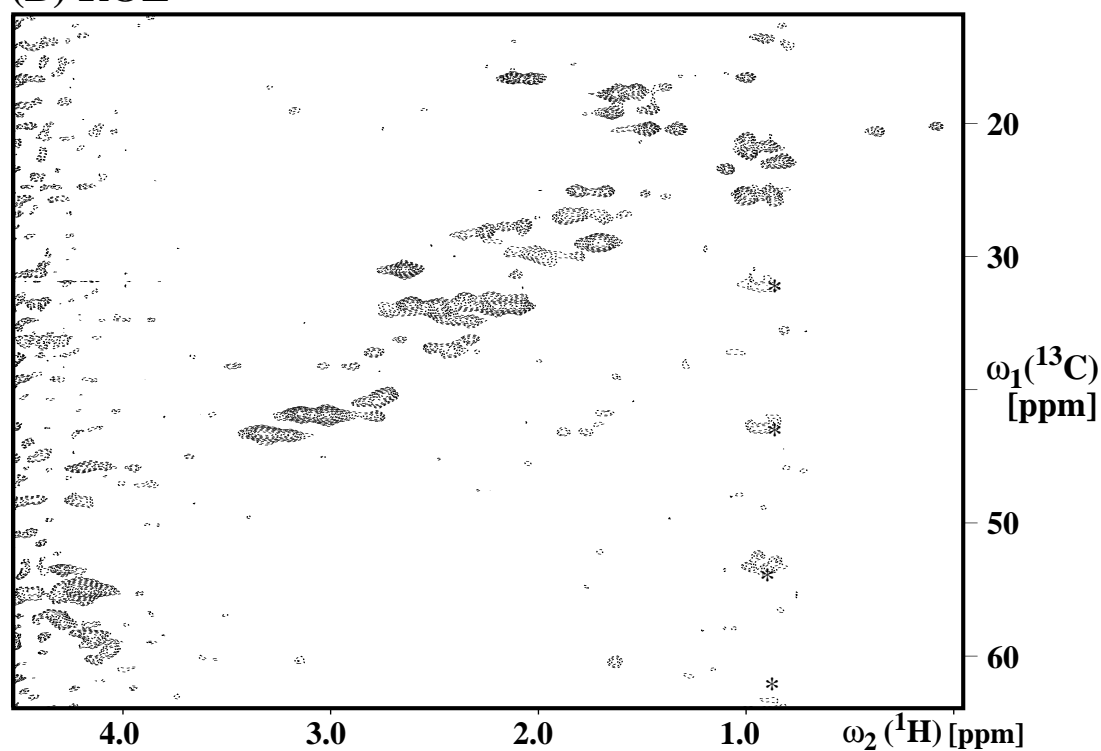
(A) NOE**(B) ROE**

Figure 3-2: Two-dimensional $[^{13}\text{C},^1\text{H}]$ -HSQC-relayed NOE/ROE difference spectra for detection of water-protein interactions with a 7 mM sample of dh434(0-63) at 13 °C and pH 4.8 measured at a ^1H frequency of 750MHz. (A) NOE spectrum with a mixing time of 60 msec. The positive peaks of Met0 are indicated. Positive and negative NOE peaks are indicated by broken and solid lines, respectively. (B) ROE spectrum with a mixing time of 30 msec.

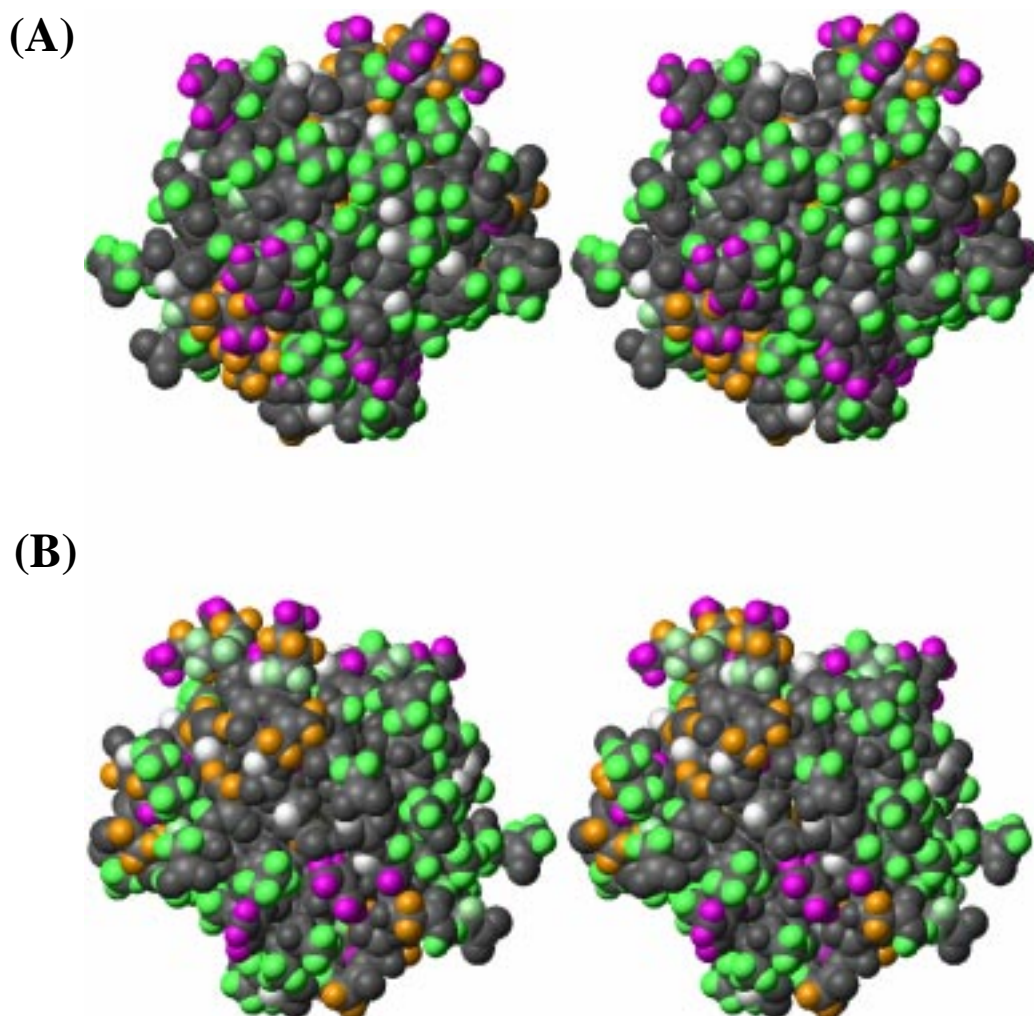


Figure 3-3: Space-filling CPK model of one of the 20 NMR structures of dh434(0–63) presenting the surface hydration. The hydration sites are indicated with the following colour coding for the hydrogen atoms: green for hydrogens showing positive NOEs and positive ROEs with the water, light green for hydrogen atoms showing no NOEs and positive ROEs with the water, orange for hydrogen atoms showing negative NOEs and positive ROEs with the water, magenta for protons showing chemical exchange with the water, and white for hydrogen atoms that have an average solvent accessible surface of more than 5 \AA^2 in the 20 NMR conformers but for which no unambiguous evidence for hydration was obtained. (A) Same orientation as Figure 6-5A. (B) The molecule was rotated about a vertical axis by 180° .

Hydration of methyl groups

In the two-dimensional ^{13}C , ^1H -HSQC relayed NOE difference spectra 28 methyl groups out of a total 45 methyl groups in the protein are observed (Figure 3-4). 4 methyl groups, $\text{I6}\delta^1$, $\text{I6}\gamma^2$, $\text{L34}\delta^2$, and $\text{L48}\delta^2$, showing negative NOEs are not located on the surface of the protein, which will be discussed below in more detail. Rapidly exchanging labile protons such as, $\alpha\text{-NH}_3$ of the N-terminus, ζNH_3 of Lys23 can be found with 5 \AA from the 3 methyl groups, $\text{M0}\epsilon$, $\text{L15}\delta^2$, and $\text{L15}\delta^1$ that show negative NOEs. Since ζNH_3 of Lys23 is in the proximity of Leu15 and shows rela-

tively strong negative NOEs with its ϵCH_2 group, this labile protons are assumed to contribute to the negative NOEs of methyl groups of Leu15. The other 21 methyl groups show positive NOEs, which is a strong indication for a short residence time of less than ~ 300 ps for the hydration water molecules based on the diffusion model (Figure 3-1). For the 20 positive NOEs with good signal to noise the average ratio of $\sigma^{\text{NOE}}/\sigma^{\text{ROE}}$ is 0.5 ± 0.3 . This might be translated into a range of estimated residence times of 10~160 ps based on the diffusion model (Ayant *et al.*, 1977; Otting *et al.*, 1991a; Figure 3-1). Some earlier NMR studies indicated mainly long-lived water molecules around methyl groups on the surface and only very few “short-lived” water molecules were found (Clare *et al.*, 1994; Qin *et al.*, 1994; Karplus & Faerman, 1994; Ernst *et al.*, 1995a; Ernst *et al.*, 1995b; Matthews *et al.*, 1995). This situation could not be supported using this optimized NMR experiments and the protein dh434(0–63) without hydroxyl groups which greatly clarifies the ambiguities in the origin of water-protein NOEs. The observation of positive NOEs indicates that the hydration water molecules around methyl groups on the surface of a protein are highly mobile which is in line with studies on small peptides that indicate “short-lived” water molecules surrounding the entire peptides (Otting *et al.*, 1991a; Otting *et al.*, 1992; Brüschweiler *et al.*, 1995).

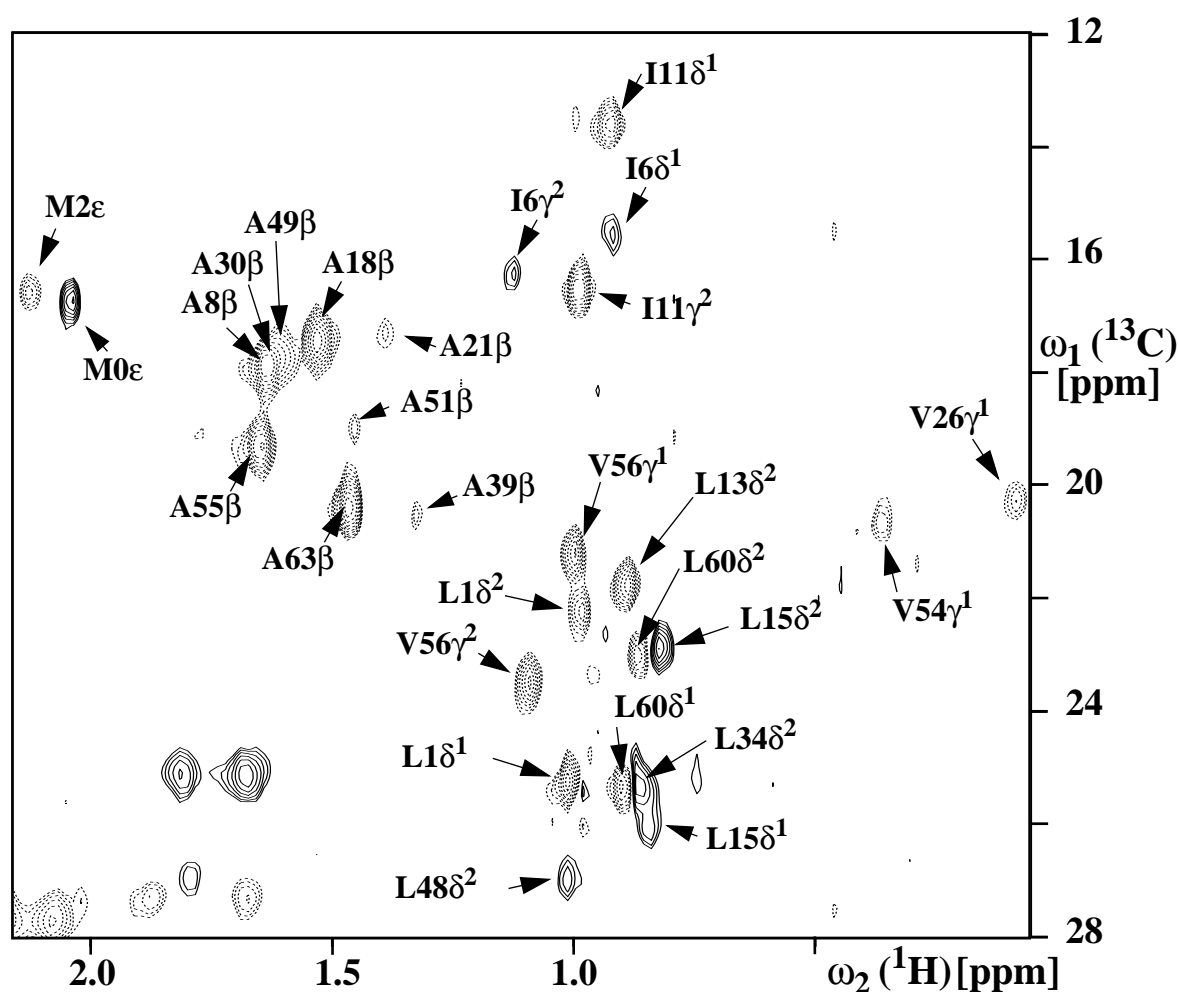


Figure 3-4: Expansion of the methyl region of the water-NOE spectrum (Figure 3-2A) with the individual assignments of the methyl resonances of dh434(0–63). Positive and negative NOEs are indicated by broken and solid lines, respectively.

Hydration of polar side chains

Polar side chains are prime candidates to form hydrogen bonds with water molecules. Hence water molecules around polar side chains might have longer residence times than around non-polar side chains. In crystal structures statistically more water molecules are found near polar residues than non-polar groups (Thanki *et al.*, 1988; Thanki *et al.*, 1991). The NMR data on dh434(0–63) reveals that the side chains of all polar amino acids class have some positive NOEs with water molecules except for the Lys residues. In the case of Lys residues, the rapidly exchanging protons in ζNH_3 groups transfer negative NOEs to the water molecules due to the fast chemical exchange covering positive NOEs (Qian *et al.*, 1993). Unambiguous posi-

tive NOEs could be assigned to the following protons: the γ protons of Glu4, Glu32, Glu47, the β protons of Asp27 and Asp57, the δ protons of Arg7, Arg43 and Arg50, the γ protons of Arg7, Arg43, and Arg50, the β protons of Arg7, Arg41, Arg43 and Arg50. The average ratio of $\sigma^{\text{NOE}}/\sigma^{\text{ROE}}$ for the methylene groups of the polar residues with positive NOE peaks is 0.4 ± 0.3 for 31 peaks. This results in the estimated residence times in the range from 20 to 230 psec (Figure 3-1C). NOEs with water molecules are observed for all the NH_2 groups of Gln and Asn except for that of the buried residue of Gln17. The NOEs of these NH_2 groups show all positive NOEs with exception of Asn36 which has negative NOEs. Those observed positive NOEs are indicative of short-lived water molecules around side chain amide groups.

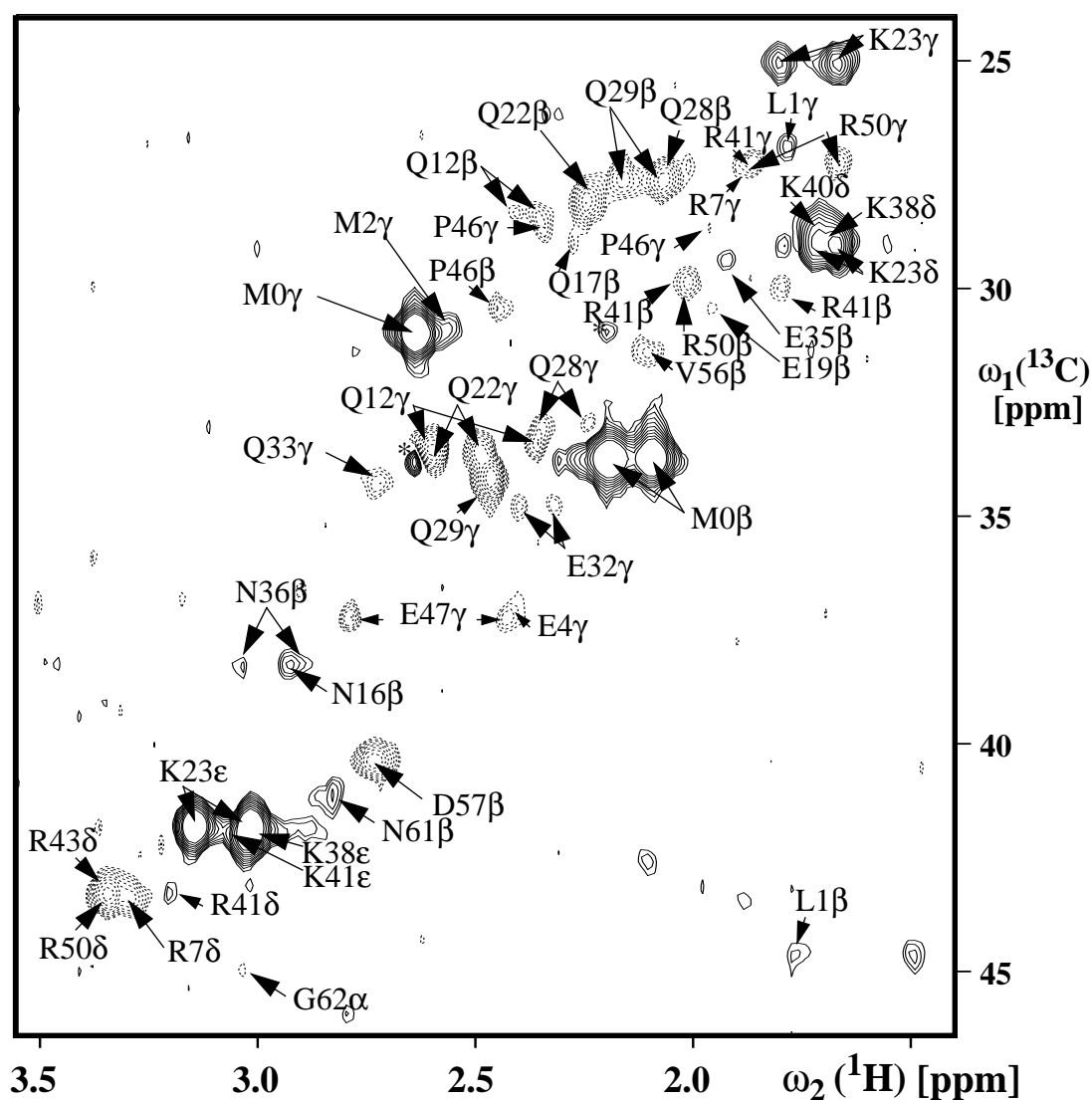


Figure 3-5: Expansion of the water-NOE spectrum (Figure 3-2A) Positive and negative NOEs are indicated by broken and solid lines, respectively.

Hydration of aromatic rings

The protein, dh434(0–63) contains only one Trp and one Phe which makes difficult to draw a general picture of hydration of aromatic rings. Positive NOEs were observed for the protons exposed on the surface, i.e. the H ϵ ¹, H δ ¹, and H ζ ² protons of Trp58 (Figure 3-6). There have been no detectable NOE or ROE with water for one H ϵ of Phe44 although it is exposed on the surface (Figure 3-6). Judging from the peak intensity in a [¹³C, ¹H]-HSQC reference spectrum, it is unlikely that the signal intensity has been dramatically reduced by the relaxation of the attached ¹³C atoms. However, the other H ϵ proton of Phe44, which is not exposed on the surface, is located near the Leu48, which shows negative NOEs. This observation might suggest that the expected positive NOEs for the H ϵ protons of Phe44 could be cancelled with negative NOEs due to ring flipping of Phe rather than the absence of hydration water molecules around aromatic rings. Thus our observation suggests that the exposed aromatic rings are equally hydrated with short-lived water molecules.

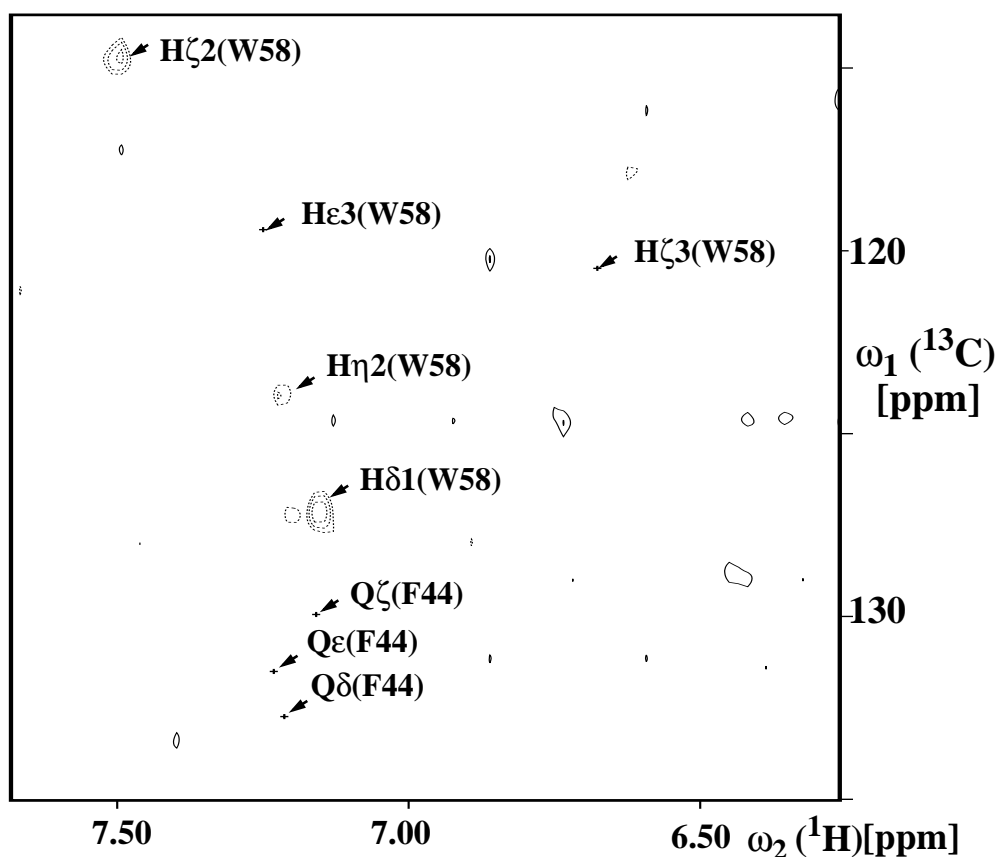


Figure 3-6: Aromatic region of the two-dimensional $[^{13}\text{C}, ^1\text{H}]$ -HSQC-relayed NOE difference spectra for the detection of water-protein interactions with a 7 mM sample of dh434(0–63) at 13 °C and pH 4.8 measured at ^1H frequency of 750MHz with a mixing time of 60 msec. The arrows and labels indicate the positions of all the peaks from the aromatic rings.

Polypeptide backbone hydration

A consistent statistical preference of hydration to the backbone atoms is found in crystal structures (Thanki *et al.*, 1988; Thanki *et al.*, 1991). Are the backbone atoms more hydrated than side chains atoms in aqueous solution? 13 α protons showing NOEs with water molecules could be identified unambiguously including all Gly, Leu13, Glu19, Val24, Gln28, Lys41, Arg50, and Ala63. The α protons showing negative NOEs are always located next to the chemically exchanging amide protons (Figure 3-12). Otherwise α protons show positive NOEs indicating of short-lived water molecules close to the backbone segments. Amide protons of proteins are in continuous chemical exchanges with water molecules which usually produce exchange peaks in both NOE and ROE spectra for fast exchanging amide groups (Wüthrich, 1986). The NOEs for the amide protons in dh434(0–63) can be divided into the three

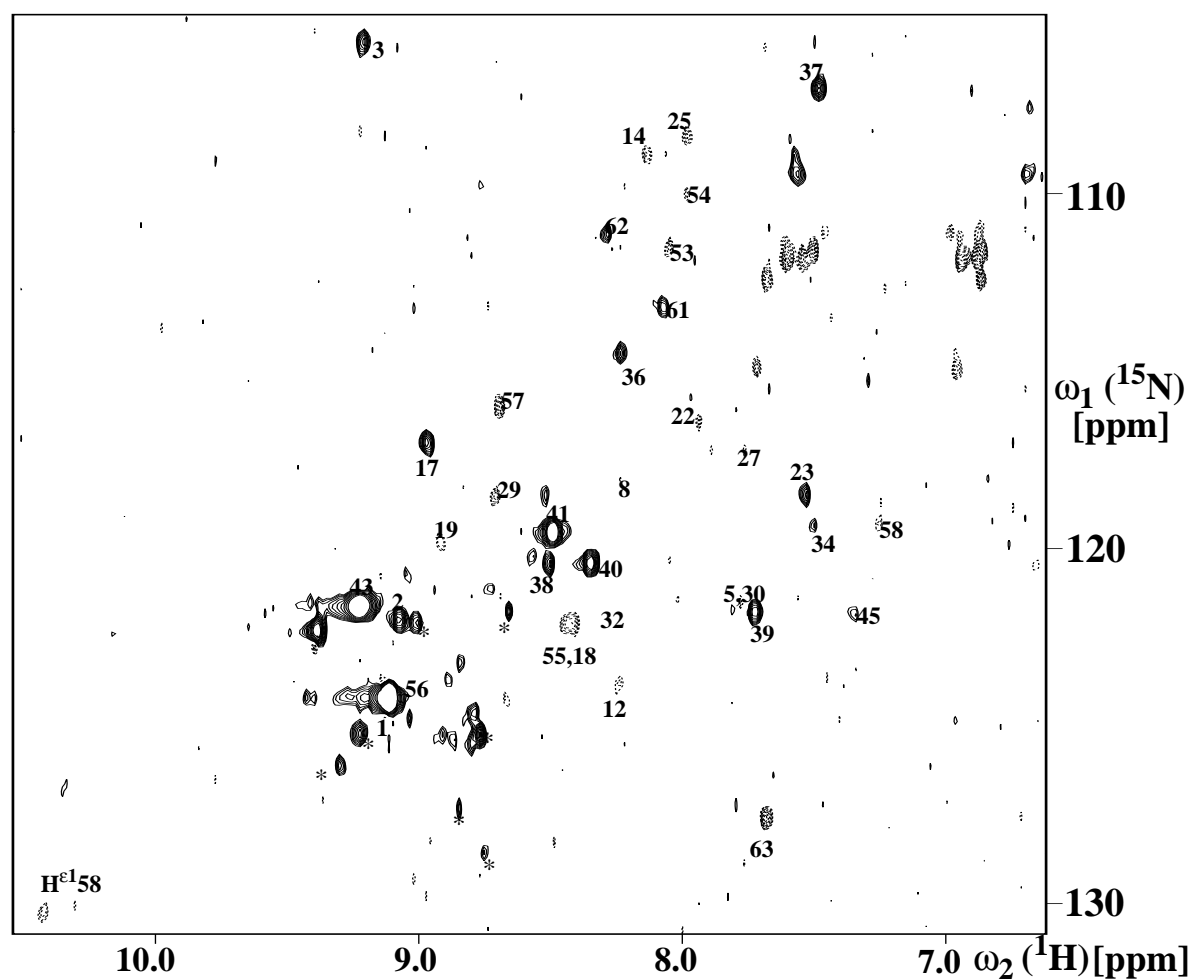


Figure 3-7: Amide region of the two-dimensional $[^{15}\text{N}, ^1\text{H}]$ -HSQC-relayed NOE difference spectra for detection of water-protein interactions with a 7 mM sample of dh434(0–63) at 13 °C and pH 4.8 measured at ^1H frequency of 750MHz with a mixing time of 60 msec. The positive and negative peaks are indicated by solid and dotted lines, respectively. The labels indicate the identified backbone amide groups. Astarisks indicates chemical exchange peaks from impurities.

classes. The first class in protons showing negative NOEs and ROEs, which are due to the predominant chemical exchange contribution. The second class has both positive NOEs and ROEs, which is a clear indication of interactions with short-lived water molecules. The third class contains protons showing negative NOEs and positive ROEs, which is indicative of long-lived water molecules or NOEs *via* chemical exchange. In spite of the chemical exchange contribution, we could identify 17 amide protons of positive NOEs for slowly exchanging amide protons (Figure 3-7; Figure 3-8). The rapidly exchanging protons of the backbone amide groups are only observed in the loop, the turns, and terminal sites of the helices in dh434(0–63). These observation suggest that there is no clear consistent difference for preferential binding of water molecules between to the backbone atoms and the side chains.

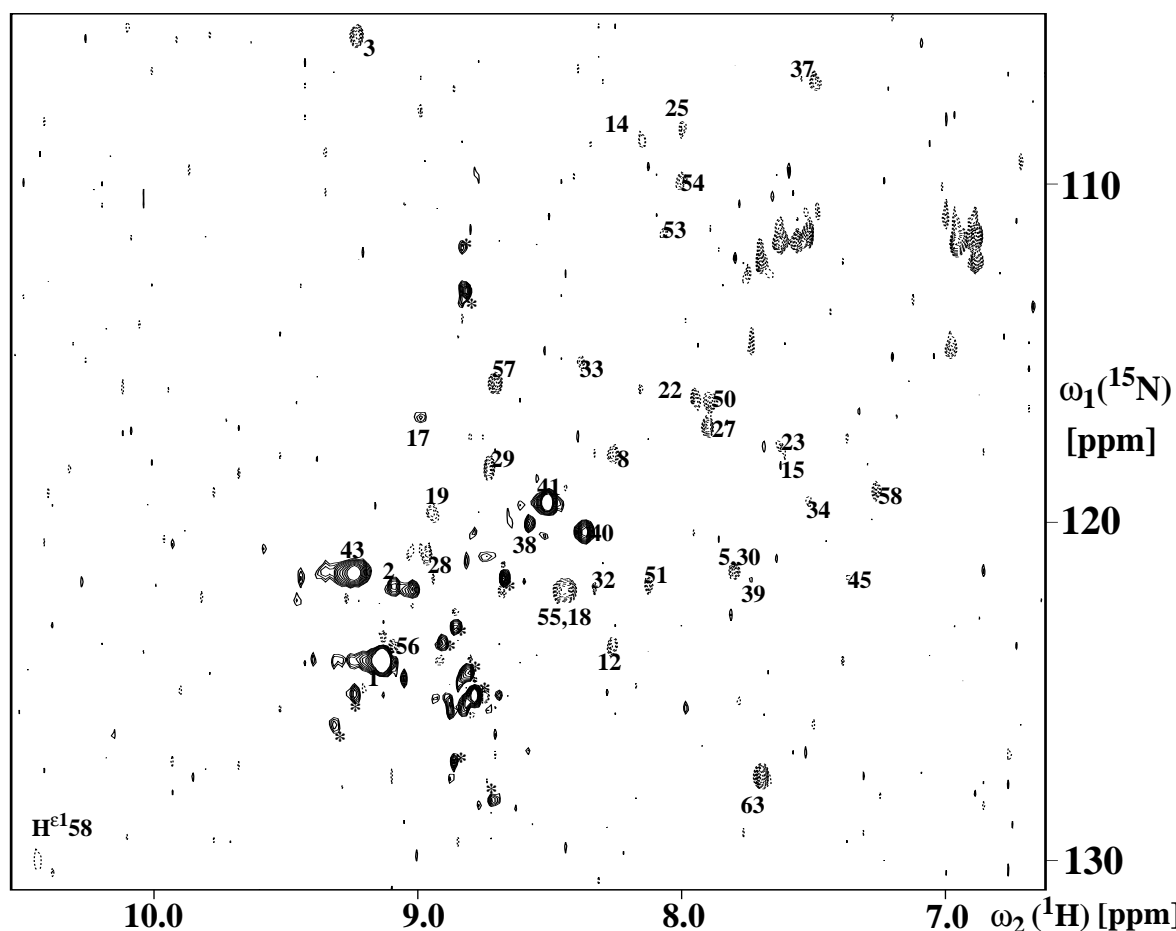


Figure 3-8: Amide region of the two-dimensional $[^{15}\text{N}, ^1\text{H}]$ -HSQC-relayed ROE difference spectra for detection of water-protein ROE with a 7 mM sample of dh434(0–63) at 13 °C and pH 4.8 measured at ^1H frequency of 750 MHz with a mixing time of 30 msec. The positive and negative peaks are indicated by solid and dotted lines, respectively. The labels indicate the identified backbone amide groups.

3.4. Discussion

Evaluation of completeness of observations on hydration

The NMR hydration measurements with dh434(0–63) allowed an almost complete characterization of the hydration shell with only a very small area for which no direct interaction with water molecules could be established (Figure 3-3). This is the first protein of which hydration has been analyzed to this extent. Figure 3-9 exhibits the extent of the observable surface hydration in detail. The protons which are on the surface but no information could be obtained unambiguously are colored in white in Figure 3-3. These sites are mainly due to overlapping of peaks. In addition some of them could not be observed presumably due to the signal attenuation by relaxation of attached ^{13}C nuclei because these peaks are observed weakly in a

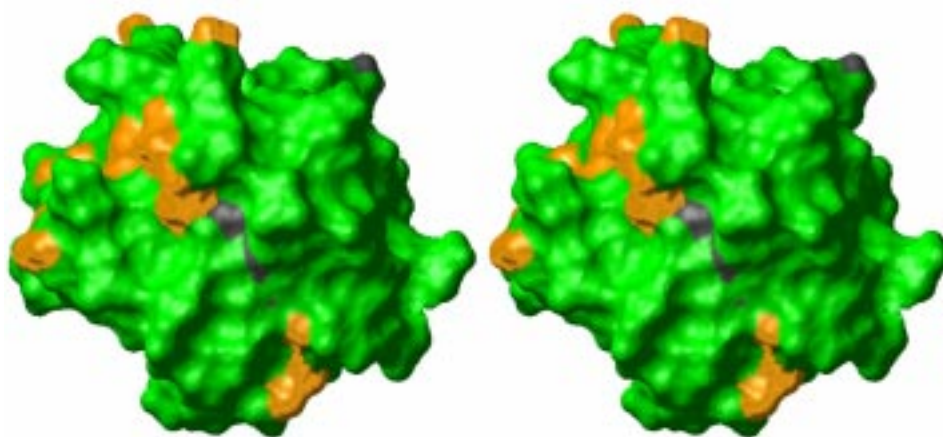


Figure 3-9: The contact surface of one of the 20 NMR conformers is shown. The surface coloured in green covers the atoms which are in a distance of 4Å from a proton identifying hydration water molecules (coloured either in green or light green in Figure 3-3 (A) and (B)). The additional surface coloured in yellow indicates the atoms which are located within a distance of 4Å from protons showing negative NOEs but not within a distance of 4Å from protons shown in green and light green in Figure 3-3 (A) and (B). The surface coloured in grey shows the area where no information on hydration could be obtained because no NOEs to water molecules could be established.

standard HSQC spectra. The surface of dh434(0–63) colored in green in the Figure 3-9 is the surface of the atoms within a distance of 4Å from the protons colored in green and light green in Figure 3-3. These atoms cover on average 88% of solvent accessible surface of the protein. Additionally the surface colored in yellow are the sites where hydration water molecules could be located although these sites are ambiguous due to its negative NOEs. Despite the sites where no information on hydration is available, total of the 98% of solvent accessible surface of a protein could be characterized by using NMR spectroscopy in this study. This is the first case in which such large area of the surface hydration has been analyzed by any kind of methods at atomic resolution in aqueous solution.

Comparison with hydration in single crystal

The first individual hydration sites were observed in single crystals by x-ray crystallography. However, the proteins in crystals have usually contacts with neighboring molecules which could prevent the observation of the complete surface hydration shell and reduce the dynamics of water molecules on the surface. Figure 3-11 presents the crystal structure of 434 repressor, 434(1–69) solved at 1.9 Å resolution with the oxygens of the identified hydration water molecules indicated by red

spheres (Mondragon *et al.*, 1989). The modeled hydrogen atoms within a distance of 5 Å from the oxygen atoms from the water molecules in the crystal structure are colored in green. These hydration sites can not cover the complete surface of the protein because of the inherent features of the diffraction methods such as crystal contacts as discussed in Chapter 2. It is found that the water molecules have more contacts with neighboring molecules in the crystal lattice of 434(1–69) (Figure 3-10). The large white area in Figure 3-11B does not have any contacts with the neighboring molecules. This observation indicates that the observed water molecules in the crystal structure are due to the geometrical restrictions in the crystal. Therefore the observation of the surface water molecules by crystallography could be artifacts from crystallographic methods rather than the actual hydration picture on the surface. In contrast NMR does not suffer from crystal contacts and does not require ordering of water molecules to be detected but can deal with a highly dynamical system.

Hence characterization by NMR spectroscopy could cover nearly the entire surface of a protein. In this study the intuitive picture of the surface of a protein in aqueous solution surrounded with water molecules and that most water molecules close to polar groups as well as non-polar groups having short residence time of less than approximately 300 psec are revealed, which is a different view of the hydration provided by crystallography.

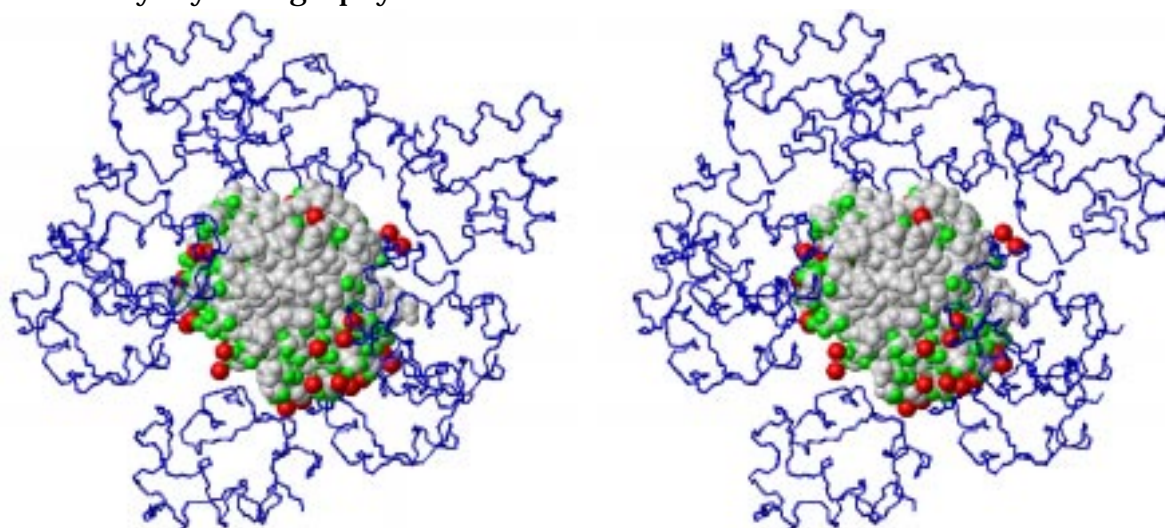


Figure 3-10: X-ray structure of 434(1–69) with neighboring molecules in the crystal lattice indicating the crystal contacts. Water molecules are shown by red spheres. A CPK space-filling model of one of molecules and the backbone atoms of the other molecules are displayed. The hydrogen atoms within a distance of 5 Å from the oxygen atoms of the hydration water molecules identified by x-ray crystallography are coloured in green. The red spheres are the oxygen atoms of the hydration water molecules.

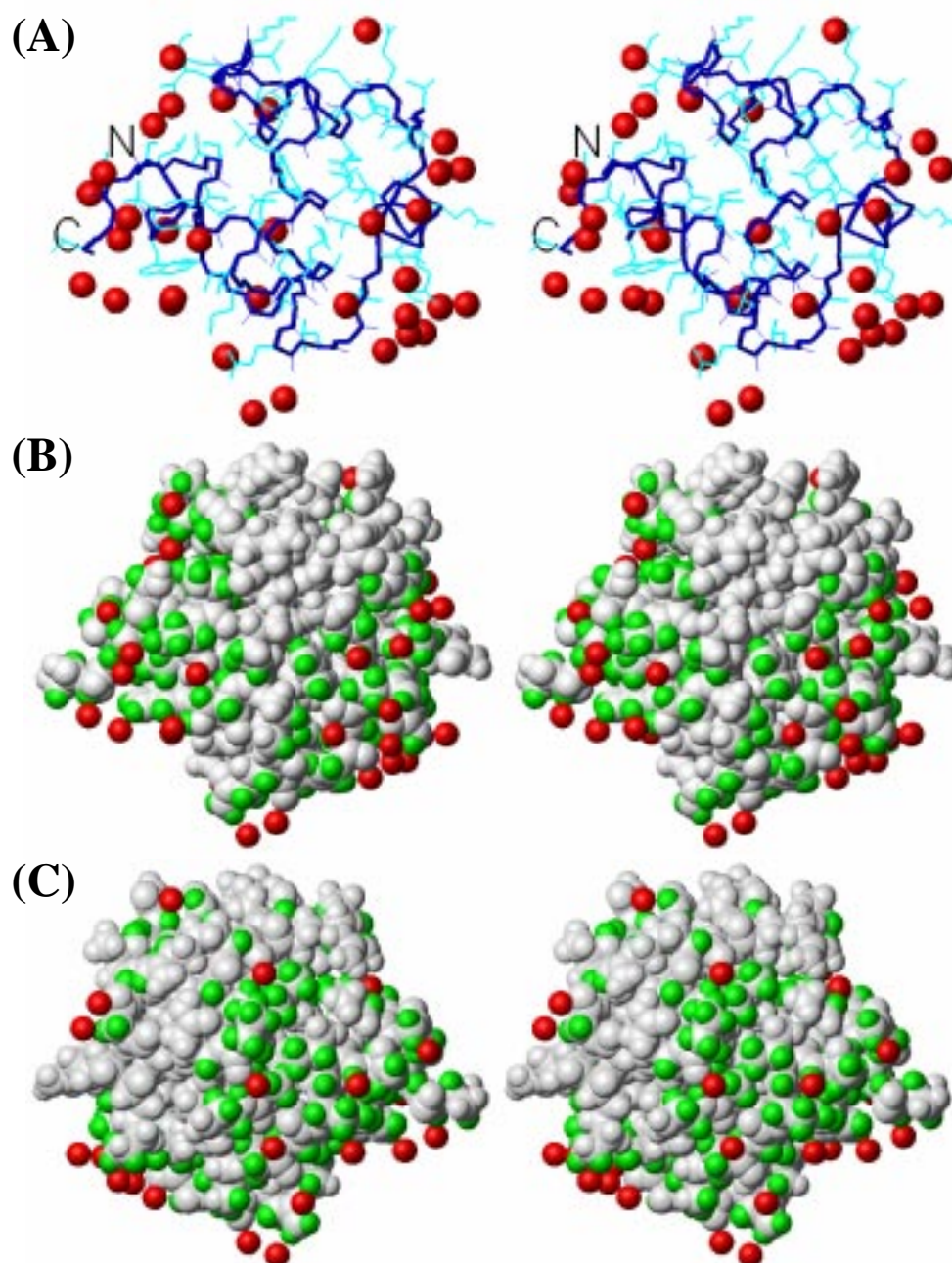


Figure 3-11: X-ray structure of 434(1–69). (A) The thick lines indicate the backbone atoms of dh434(1–69). The thin lines present the heavy atoms of the side chains. C and N denote the C- and N-terminus, respectively. Water molecules are shown by spheres. (B) CPK space-filling model of 434(1–69) illustrating the hydration water molecules by spheres. The hydrogen atoms within a distance of 5 Å from the oxygen atoms of the hydration water molecules identified by x-ray crystallography are coloured in green. The red spheres are the oxygen atoms of the hydration water molecules. (C) Same presentation as (B), with the molecules rotated around a vertical axis by 180°.

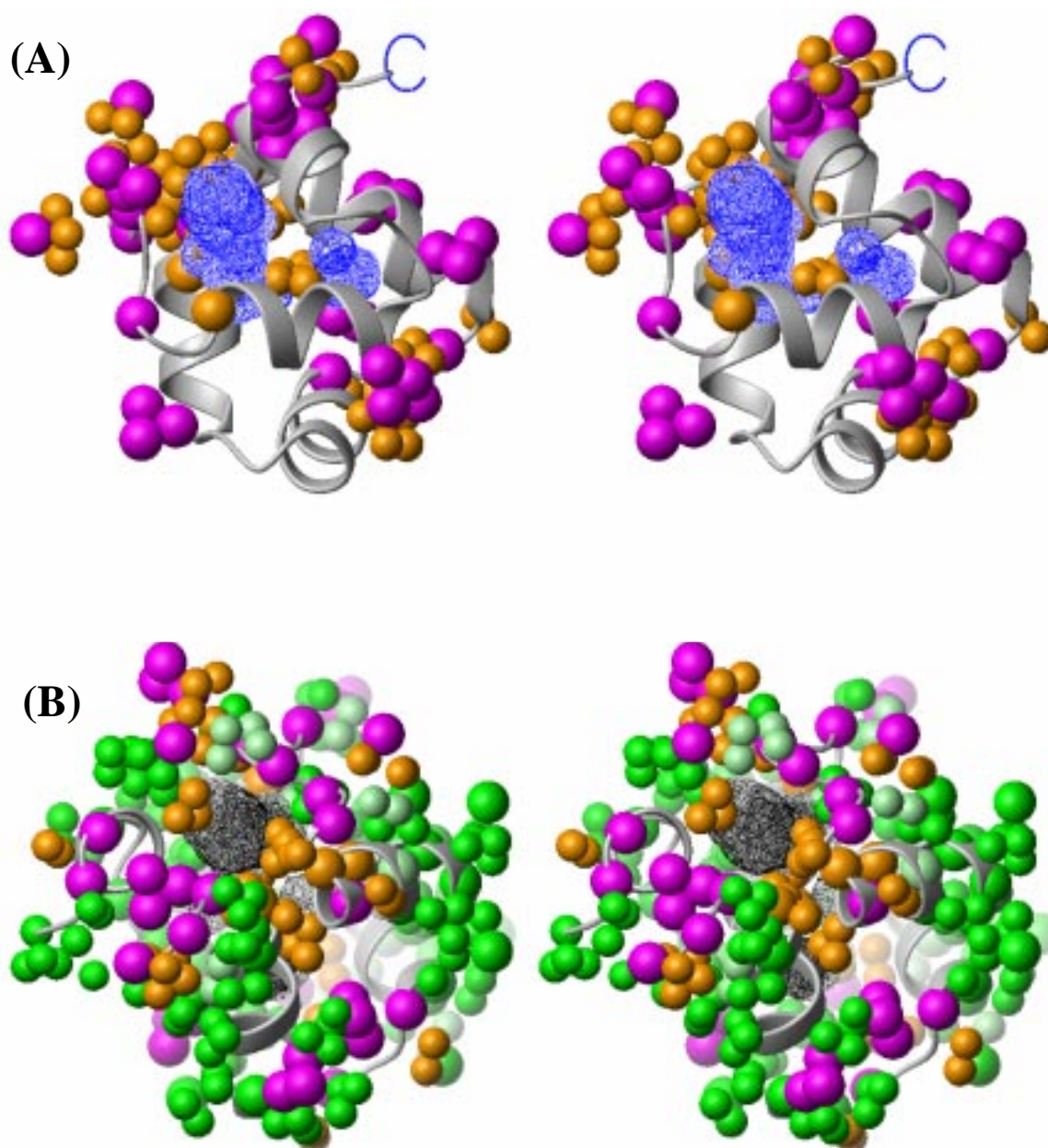


Figure 3-12: Stereo view of the backbone of dh434(0–63) with superposition of “cavities” from the 20 NMR conformers. The backbone conformation is depicted by a ribbon drawing. (A) Hydrogen atoms with negative NOEs and chemical exchanging groups are indicated by spheres, which are coloured in yellow and magenta, respectively. (B) In addition to (A) different hydration sites of protons with positive NOEs, or no NOE and positive ROEs are indicated by spheres, coloured in green and light green, respectively.

Hydration of internal cavities?

Negative NOEs observed between water and protein protons are indicative of long-lived water molecules with residence times of longer than ~300 psec (Figure 3-1C). However, such negative NOEs could also be attributed to the remaining fast exchanging labile protons in the protein. It is not possible to separate these two type

of negative NOEs experimentally. For an unambiguous identification of long-lived water molecules it is necessary to exclude the second possibility by close inspection of the three-dimensional structures. It is usually possible to conclude the presence of long-lived water molecules only if there is no labile proton within a distance of 5 Å from such protons of negative NOEs. It appears that fast exchanging labile protons are only the N-terminal α -NH₃ groups in dh434(0–63) having a similar intrinsic exchange rate as hydroxyl groups. The N-terminal α -NH₃ group produces strong negative NOEs with the α , β and γ protons of Met0 (Figure 3-2A). For all the protons showing negative NOEs there are always some labile protons such as ζ NH₃ of Lys, the N-terminal α -NH₃ group, or exchanging HN group within a distance of 5 Å in at least one of the 20 NMR conformers of dh434(0–63). This makes it difficult to identify long-lived water molecules. However, it is worth noting that there are few fast exchanging protons found within a distance of 5.0 Å from Ile6 CH₃ δ^1 and Leu48 CH₃ δ^2 , which show weak negative NOEs and are located in the core of the protein. These two sites are on the face of a cavity between the two sub-domains covered with a flexible linker (Figure 3-12). Moreover NH groups of Gly3 and Gly37 have negative NOEs and positive ROEs indicating the possibility of long-lived water molecules in their proximity (Figure 3-7; Figure 3-8). These sites are also located near the entrance of the cavities providing supports for the possibility of long-lived water near these cavities (Figure 3-12). These cavities can be seen in 10 structures out of the 20 NMR conformers of dh434(0–63). The average volume of the cavities found in this 10 NMR structures is about 30 Å³ which could accommodate a water molecule (~35 Å³). For long-lived water molecules in the interior of proteins which are integral parts of the structure, strong negative NOEs would be expected, for example as observed in bovine pancreas trypsin inhibitor (BPTI) (Otting & Wüthrich, 1989; Otting *et al.*, 1991a). However, the weak NOEs observed for these sites in dh434(0–63) might indicate low occupancy of the water molecule in the cavities implying that the water molecules could be accidentally trapped by “breathing” of the protein rather than being as an integral part of the structure. Based on these data one might speculate that a first event during unfolding includes the water molecules entering into the core of proteins between the sub-domains through a cavity or channel occasionally produced by a flexible linker. The geometry of the hydration sites appears to be more relevant to the residence time of the water molecules than amino acid types.

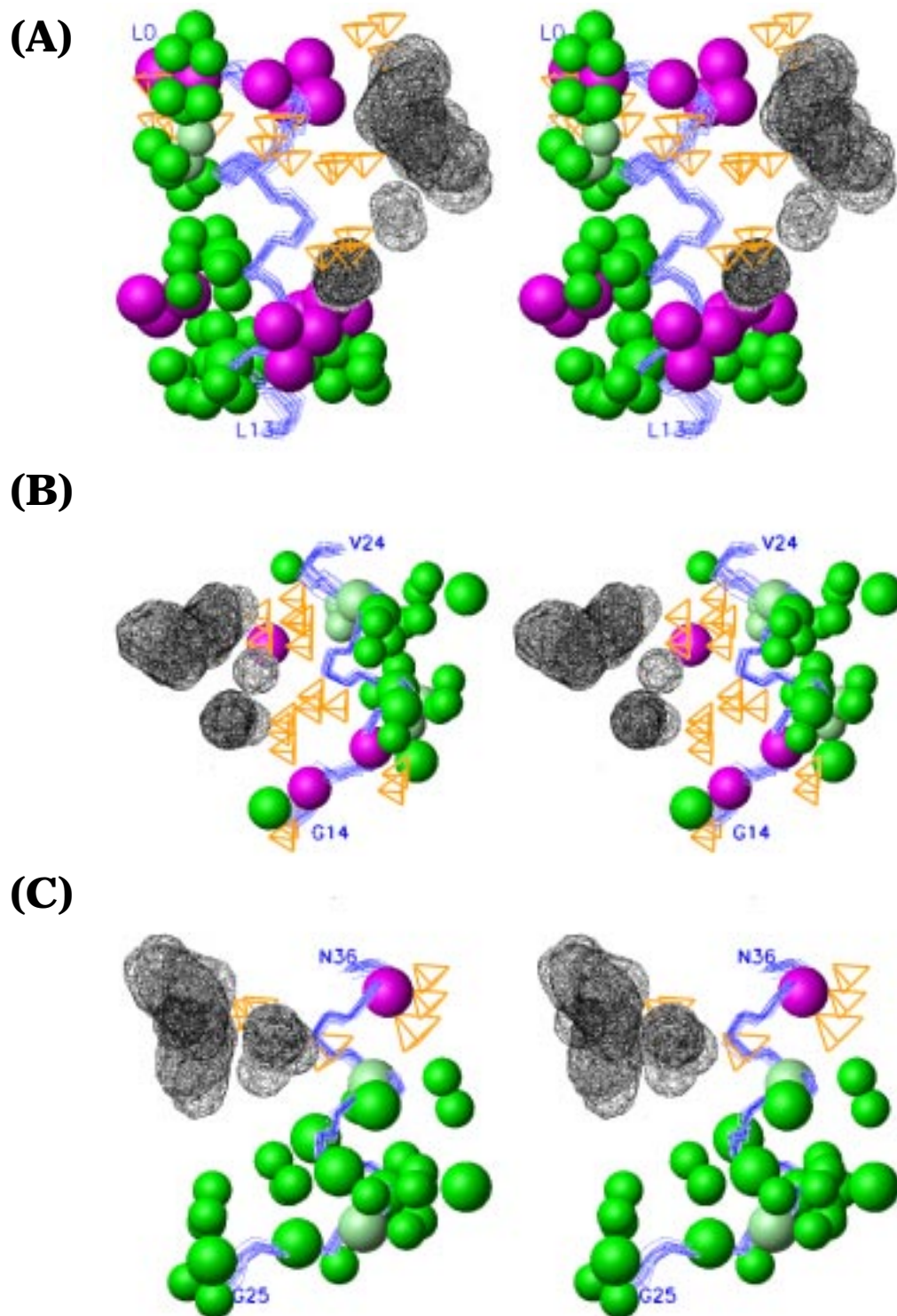


Figure 3-13:

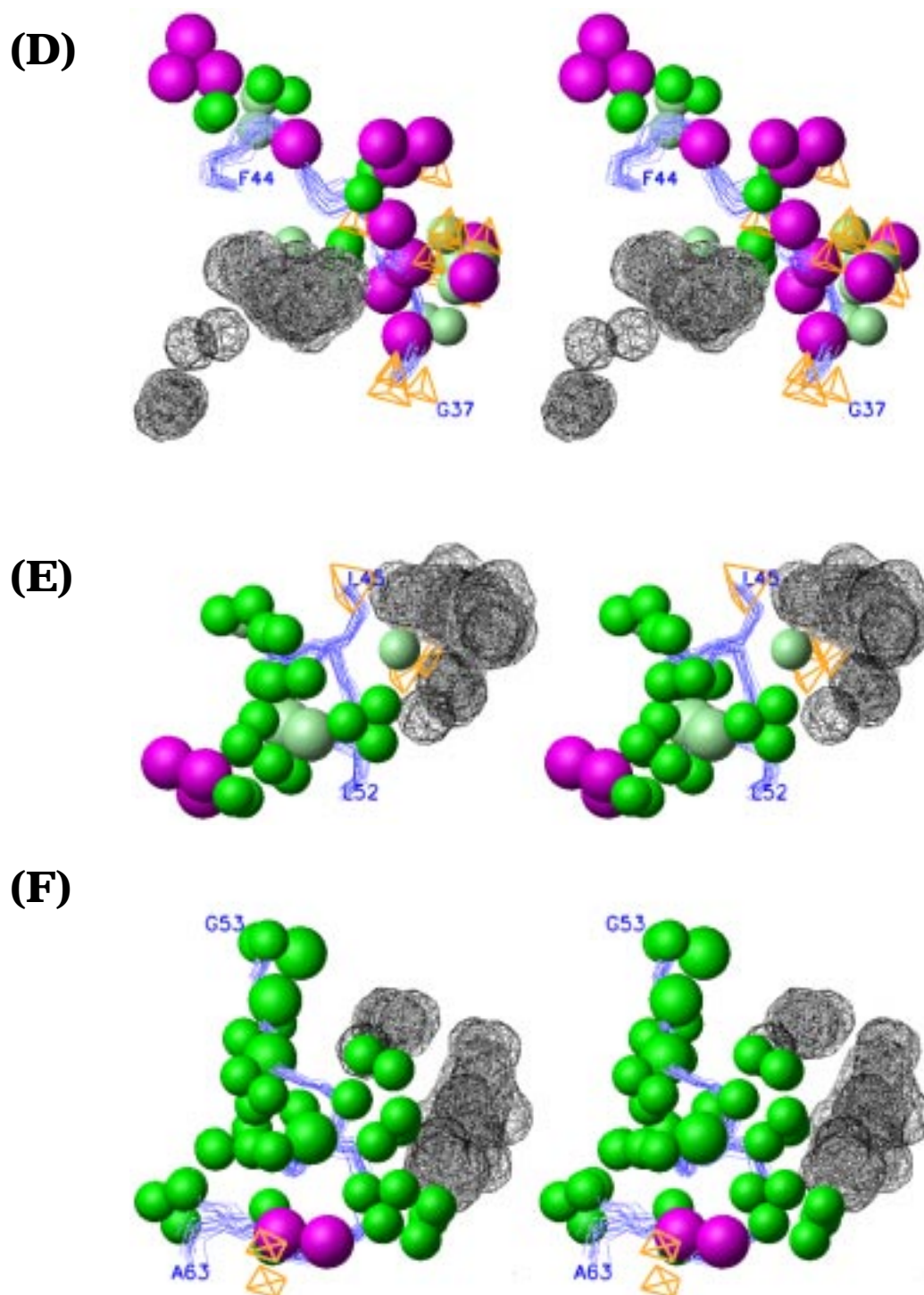


Figure 3-13: Presentation of the hydration sites. (A) residues from 0 to 13; (B) residues from 14 to 24; (C) residues from 25 to 36; (D) residues from 37 to 44; (E) residues from 45 to 52; (F) residues from 53 to 63. The protons of negative NOEs are indicated by tetrahedrons. The protons and amide groups of positive NOEs, and chemically exchanging groups are indicated by spheres. The protons exchanging with water are coloured in magenta. The cavities from the 20 NMR structures and the backbone are indicated by cages and thin lines, respectively.

Hydration in helices?

The patterns of the hydration sites in the all segments of dh434(0–63) are presented in Figure 3-13. In the crystal structures more ordered water molecules are found at the termini of helices externally or internally (Thanki *et al.*, 1991). dh434(0–63) contains five helices offering the possibility to investigate the hydration pattern of helices in aqueous solution. Figure 3-13A, B, C, E, F present the hydration of each helices. In helix I and II such N-terminal amide groups, Met2 and Gln17 show the dominant chemical exchange with water molecules in ROE spectra. N-terminal amide groups in helix III and VI tend to indicate positive NOEs, for example, Asp27 and Ala55, but dominant chemical exchanges with water molecules at the C-terminus such as Asn36, Leu61, Gly62 are observed. These amide groups in the C-terminus are involved in hydrogen bonds with the oxygen atoms of the backbones in the three-dimensional structures of dh434(0–63). In helix IV the amide group at Leu45 shows negative NOEs, which is buried in the core of the protein. The analysis of these helices does not provide any consistent view of the hydration of helices unfortunately. However, at the terminal amide groups in the helices tends to show dominantly chemical exchanges with water. This situation is the same for the loop and the turn. Moreover, it is worth to note that such backbone amide groups in the loop and in the turn show exchange peaks in ROE spectra but the none of side chain amide groups shows any exchanging peaks in ROE spectra although the backbone NH groups and side chain NH₂ groups are expected to have similar intrinsic exchange rates from the model compounds (Englander *et al.*, 1972; Bai *et al.*, 1993). The data presented here indicates faster exchange rates for the exposed NH group than the exposed side chain NH₂ groups. O-protonation mechanism, imidic acids intermediates mechanism for acid-catalyzed proton exchange of amide groups in peptides and proteins are proposed for the mechanism of proton exchange of amide groups, which require the involvement of both carbonyl oxygen and nitrogen atoms of amide groups and the accessibility of these both atoms (Tüschen & Woodward, 1985; Perrin, 1989). It is found that in crystal structures there is a water molecule forming hydrogen bonds to both the carbonyl oxygen and amide nitrogen atoms with the amide proton involved in a three-center hydrogen bond, or a internal water molecule inserting into α -helix forming a hydrogen bonded bridge between the back-

bone carbonyl oxygen and amide nitrogen atoms. This structure with three-center hydrogen bonds, an insertion of a water molecule can be regarded as folding intermediates (Sundraralingam & Sekharudu, 1989). The observation of the faster exchange for NH groups in the backbone at the termini of the helices could be relevant to this finding in crystal structure of the three-center hydrogen bond with a water because this structure appears to be an ideal situation for the O-protonation mechanism for the acid-catalyzed hydrogen exchanges of amide groups (Tüchsen & Woodward, 1985).

The question is whether there are water molecules with longer residence times near such amide groups with faster chemical exchanges, or the residence time of water molecules are associated with the exchange rates of amide groups. Although protons showing negative NOEs close to such NH groups could be found, it is not possible to separate NOEs from “long-lived water molecules, or NOEs *via* chemically exchanging labile protons with water for such amide protons because of the presence of the labile protons in its vicinity. However, the exposed side chain amide groups of dh434(0–63) in which both carbonyl oxygen and nitrogen atoms are solvent-accessible and could easily form a three-center hydrogen bonds with a water do not show predominantly chemical exchange peaks in both ROE and NOE spectra. This might imply that not only the geometry and the solvent-accessibility for water molecules but also the longer residence time of water molecules could be needed for the chemical exchange process. It is quite natural to expect water molecules transiently inserted into helices and turns to have longer residence times in aqueous solution due to their hydrogen bonds. In fact such water molecules was observed in aqueous solution by NMR spectroscopy (Clare & Gronenborn, 1992). For side chain amide groups because of the rapid motion of side chain conformation which is also coupled with water dynamics, water could not stay long enough to cause hydrogen exchanges.

One can speculate that such water molecules of long-residence times could be negative entropy effects for the solvation energy (Dunitz, 1994). Cosolute such as ions stabilizing proteins might force water molecules around the backbone amide groups in the flexible loop and turns to stay longer. In unfolded states proteins are largely flexible polypeptide chains. More ordered water molecules, *i.e.* long-lived wa-

ter molecules around the backbone amide in such large region of the polypeptide chain as well as hydrophobic hydration would contribute as large negative effects in unfolded state. This would result in the increase of the free energy of the folded protein. Such unfavorable hydration effects on the backbone in unfolded states can be a driving force of the early events of protein folding as well. This speculation is in line with the observation of preferential hydration and dehydration of proteins by cosolutes (Timasheff & Arakawa, 1989), but contradicts to a conclusion from other studies that the entropy of hydration of polar groups is negligibly small and can be ignored in protein folding (Murphy & Freire, 1992; Murphy *et al.*, 1995; Lee *et al.* 1994; Spolar & Record, 1994). However, the importance of the hydration of polar groups to the entropy of proteins is also demonstrated (Makhatadze & Privalov, 1996).

Alcohol induced helix formation could be results from the solvation effects. The low polarity of alcohol is responsible for helix induction (Tanford, 1968; Nozaki & Tanford, 1971) and increases the interaction with hydrophobic side chains with alcohol decreasing the hydration effects of side chains. This could change the balance of the entropic contribution between side chains and backbones. As a result the hydration effects on the backbone will be dominant unfavorable entropy effects. Such unfavorable entropy effects would lead to secondary structure formation such as α -helix or, β -hairpin to decrease the unfavorable hydration effects on the backbone amide groups. Alcohol-induced formation of some conformations is observed for many peptides (Sönnichen *et al.*, 1992; Dyson & Wright, 1993; Jasanoff & Fersht, 1994; Blanco *et al.*, 1994). It might be able to predict that for peptides with only polar side chain, alcohol would not necessarily induce helices when there is no specific side chains interactions. From this viewpoint it can be postulated that denaturants such as guanidium chrolide and urea would interact with the backbone amide groups dominantly to reduce the unfavorable hydration effects on the backbone amide in unfolded states by expelling water molecules resulting in keeping proteins denatured because larger volume and polar chemical structure of the denaturants would preferably interact with amide groups and could reduce the negative entropy effects from the hydration considering that water is the smallest polar liquid. This could be supported by some observations that urea and guanidium chrolide can slow proton exchange rates (Kim & Woodward, 1993; Mori *et al.*, 1997). Denaturants

might be able to induce helical conformation for polar peptides depending on the balances between the enthalpy and entropy contributions and between side chain and backbone contributions (Hagihara *et al.*, 1993; Mayr & Schmid, 1993).

It seems that solvation, or hydration effect has to be considered separately for the backbone and side chains. The entropical contribution of the backbone amide groups needs reinvestigation.

It is well anticipated that flexible loop, turn, and tails are more susceptible for proteases (Fontana *et al.*, 1986, 1989; Novotony & Bruccoleri, 1987). It is possible to speculate that necessary water molecules for catalytic reaction are delivered from the long-lived water molecules near the backbone amide groups in the turn, flexible loops. Not only the steric hindrance from the rigid parts of proteins but also hydrated water molecules in the flexible loop and turns might be responsible for the susceptibility to proteases. The investigation of the correlation between the effects of cosolutes on the susceptibilities to proteases and the backbone hydration would be beneficial for the elucidation of the functional role of hydration water molecules for enzyme activities.

Not only formation of secondary structures but also molecular association such as aggregation, precipitation can reduce such unfavorable hydration effects.

3.5. Conclusions and outlook

In this study for the first time the large part of the surface hydration shell of a protein was characterized at atomic resolution in aqueous solution by NMR spectroscopy. Based upon the this study, it might be possible to answer some questions described in Chapter 2.

Location of water molecules & structure of the hydration shell

The large part of the protons near the surface indicate the evidence of short-lived water molecules, which is also an intuitive picture of a protein surrounded by water molecules. The hydrophobic effect is often attributed to the ordering of water molecules around hydrophobic groups although the origin of the hydrophobic effects is still unclear (Frank & Evans, 1945; Klotz, 1958; Tanford, 1980). The observation of ordered water molecules around hydrophobic residues found in crystal structures is

generally used as supporting evidence for the concept of “iceberg”, or “freezing of water molecules” around hydrophobic groups (Blake, 1983; Teener, 1984). The important question is whether water molecules around hydrophobic residues are more ordered than around hydrophilic side chains in aqueous solution. Previously the observation of negative NOEs between water molecules and methyl groups was speculated as ordered pentagonal water molecules rotating with methyl groups (Ernst *et al.*, 1995a). Our observation of positive NOEs for majority of methyl groups on the surface indicates that seems not to be the case in general. If the ordering of water molecules around hydrophobic groups occurs, the life time of such ordered water molecules must be much shorter time scale of less than ~300 psec.

Difference of water structures between around the non-polar and the polar group?

The statistical analysis of the estimated life times of water molecules by NMR may indicate that the residence time of water molecules are shorter for polar groups than non-polar groups. The precision of the estimation of water molecules remains unclear. Statistical analysis of hydration patterns of proteins in single crystals indicated a predominant interaction with main chain atoms rather than side chain atoms (Thanki *et al.*, 1988; Thanki *et al.*, 1991). We have observed water-protein interactions both around the backbone atoms and around the side chain atoms indicating that both are equally hydrated. NMR does not require well ordered side chains and water molecules to detect interactions between them. Hence the less observation of hydration water molecules around side chain atoms in crystals may be a consequence of the technique rather than the representation of the actual hydration feature of proteins. From NMR data there is no clear indication of longer residence time of water molecules around the backbone atoms as compared to side chain hydration. However, the faster exchange rate for amide protons in the terminal site of helices, loop and turns than for side chain amide groups are observed, which might be an indication of water molecules with longer residence times in such regions. The relevance between residence time of water molecules and proton exchange phenomenon have to be investigate to clarify this speculation.

The detailed hydration geometry at the interface.

All the protons of negative NOEs have at least one labile proton within a distance of 5Å. Such labile protons can always be attributed to the origin of negative NOEs rather than long lived water molecules. However, a few negative not positive NOEs are observed in close proximity of the cavities found in 10 NMR conformers out of the 20 NMR conformers of dh434(0–63). Therefore such NOEs can be due to long-lived water molecules. Since short-lived water molecules are observed for both polar and non-polar groups, the residence times of water molecules are determined by geometrical features rather than preferential binding to specific amino acids types.

Movement of water molecules and dynamics of water molecules at the surface of proteins.

The indication of short-lived water molecules covered the large parts of the surface of the protein suggesting that hydration water molecules on the surface are highly mobile possibly with residence time of less than 300 psec.

Relevant time scale for understanding a particular process, or biological function and coupling of the water dynamics and dynamics of the macromolecules.

The finding of short-lived water molecules covering of the nearly entire protein is a strong indication that the dynamics of surface water molecules are also coupled with the flexible surface structure of proteins. The view of a dynamical surface structure in aqueous solution is supported by the three-dimensional structures determined by NMR methods and by NMR relaxation data (Glushko, 1972; Richarz, 1980; Wüthrich, 1996). Considering the requirement of a certain hydration for catalytic activities of enzymes, the hydration is playing an important role in providing proteins by giving a flexible matrix on the protein surface.

Do we learn enough about the protein hydration to explain the behavior of a particular system, not only at the molecular level such as protein-substrate binding, protein folding, but also on the increasingly important mesoscopic scale such as pro-

tein aggregation, fibrous formation ?

Of course, it is not possible to obtain definitive answer to such general biological process concerning hydration from a single data. However, the speculation derived from the possibility of long-lived water molecules around rapidly exchanging backbone amide protons might be relevant to such process. Long lived water molecules around the flexible backbone amide groups could be negative entropy effects, which might be one of driving forces for the formation of secondary structure, aggregation, precipitation, fibrous formation because all of such processes can reduce such unfavorable entropy effects.

Further characterization of the water-protein interaction with NOEs and proton exchange by NMR in different solvent conditions and improvement in determination of the precise residence times of water molecules would provide possibility to shed light on the effect of various cosolutes and denaturing agents on the structure and dynamics of the hydration of a protein in aqueous solution. NMR spectroscopy has great advantage for such studies because NMR spectroscopy can deal with various solution conditions. Since the difference of hydration pattern between around polar and non-polar groups could not be detected, in particular, more precise and sensitive methods for such times scale might be necessary. The relevance between hydrogen exchange and residence time of water molecules should be investigated in various conditions. It is also probably helpful to approach this problem by computer simulation such as molecular dynamics to assess the obtained experimental data. In any case such further studies will lead to the better understanding of the molecular mechanism of the hydrophobic effects, protein aggregation, protein solubility, protein folding and denaturing agents which are originating from the water structure on the proteins surface.

4. Preparation of variants of the DNA binding domain of Phage 434 repressor, 434(1–63)

4.1. Introduction

The large quantities of recombinant proteins which can be obtained from bacterial expression systems are an important prerequisite for biophysical characterization of proteins and three-dimensional structure determination. In particular, protein NMR spectroscopy has been largely developed by using isotope-labelled proteins for multidimensional heteronuclear techniques, which are usually prepared from bacterial cells. One of the difficulties in the study of the surface hydration by NMR spectroscopy is the weak signals originating from the surface hydration water molecules. Large quantities of highly soluble proteins are therefore required for the study of surface hydration of proteins to gain the sensitivity. In addition isotope-labelled proteins are of advantage to resolve the overlapping of the signals in the proton dimension. In this chapter, the construction, overproduction, and purification of variants of the DNA binding domain of 434 repressor, 434(1–63), and P22c2(1–76) by *E. coli* expression system are described.

4.2. Description of the expression system

For the overexpression of the wild-type protein and variants of 434(1–63), a promoter system based on the T7 RNA polymerase is used which is a very active enzyme and very selective for specific promoters that are rarely found in DNA unrelated to T7 DNA (Chamberlin, *et. al.*, 1970; Dunn & Studier, 1983). The bacterial host for cloning and expression is the *E. coli* B strain, BL21(F⁻, *ompT*, r_B⁻ m_B⁻), which has been extensively used here. BL21 is deficient in the *lon* protease and it also lacks the *ompT* outer membrane protease. For all expression experiments a derivative of *E. coli* BL21 was used which had been infected with the bacteriophage DE3, a λ derivative with a DNA fragment containing the *lacI* gene, the *lacUV5* pro-

moter, the beginning of the *lacZ* gene, and the gene for T7 RNA polymerase. Once a DE3 lysogen is formed, the production of T7 RNA polymerase is inducible by the addition of isopropyl- β -D-thiogalactopyranoside (IPTG). A basal level of T7 RNA polymerase in *E.coli* BL21(DE3) can be almost completely eliminated by co-transformation with plasmids encoding T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase. T7 lysozyme can be co-expressed by the plasmids pLysS (for low level of lysozyme) and pLysE (for higher level of T7 Lysozyme) both containing chloramphenicol resistance. However for expression of none of the variants of 434(1–63) and P22c2(1–76), it was necessary to reduce the basal level of T7 RNA polymerase, presumably because the proteins are not toxic to the cells. Thus *E.coli* BL21(DE3) without pLysS nor pLysE was used throughout for overproduction of all variants of 434(1–63).

The vector plasmid pT7-7 was used to express variants of 434(1–63) and P22c2(1–76) in *E.coli*. pT7-7 is a derivative of pT7-1 (Tabor & Richardson, 1985). This vector contains the T7 promoter, a polylinker sequence for cloning, the T7 gene 10, ColE1 origin, and the ampicillin resistant gene (Tabor, 1990; see also Appendix A).

4.3. Methods for mutagenesis

In vitro mutagenesis has become a standard tool for the functional analysis of proteins. In particular *in vitro* site-directed mutagenesis is a powerful method to understand contribution of each amino acid on the stability of a protein. Many mutagenesis methods have been developed for this purpose. Oligonucleotide-directed mutagenesis is one of methods to introduce mutations, insertion, and deletion at any positions. Several methods were applied for the preparation of variants of 434(1–63). The *in vitro* mutagenesis methods used in this study are the following.

PCR based site-directed mutagenesis

- Overlap extension method
- Modified overlap extension method
- Inverse PCR method
- Gene synthesis by PCR

Non-PCR based site-directed mutagenesis

- Kunkel methods
- Synthetic oligonucleotides.

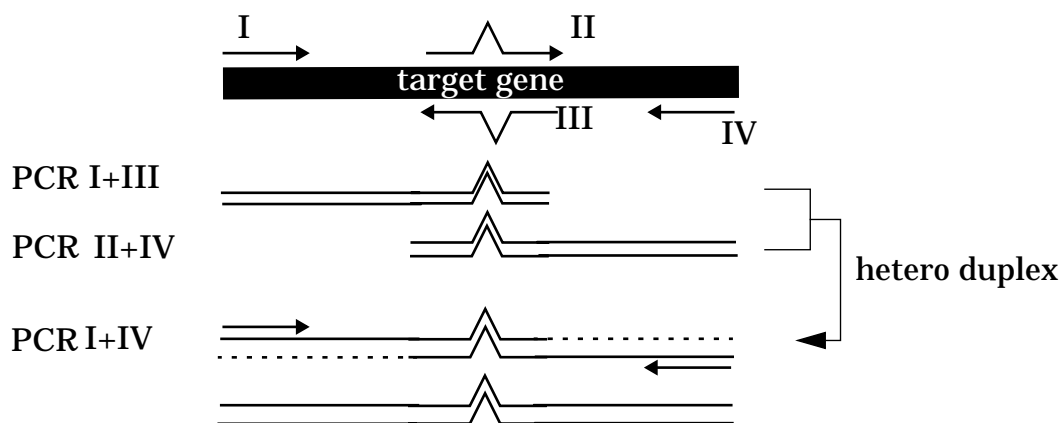


Figure 4-1: Schematic drawing of the mutagenesis procedure by PCR overlap extension method (Higuchi, 1991).

PCR based site-directed mutagenesis

PCR based site-directed mutagenesis is probably the easiest and fastest method to introduce mutations at any arbitrary site *in vitro*. PCR reactions require a few hours to introduce a series of mutations in the gene. Therefore it can be easily and quickly achieved to modify the protein.

Overlap Extension Method

Overlap extension method is an orthodox procedure for introducing a mutation with PCR (Higuchi, 1991). Briefly the procedure is as follows. One mutagenesis requires four oligonucleotides (I, II, III, and IV in Figure 4-1). One mutation of the gene can be introduced as mismatches in the synthetic oligonucleotides (II and III in Figure 4-1). The oligonucleotides II and III in Figure 4-1 have a complementary sequence to each other so that the two PCR products produced by the oligonucleotides I and III, and by the oligonucleotides II and IV can form a heteroduplex and be extended to the full length of the target gene by DNA polymerase. This PCR product can be again amplified by PCR with the two oligonucleotides I and IV. The final PCR product will contain a mutation as introduced in the synthetic oligonucleotides.

Modified Overlapping method

PCR based site-directed mutagenesis with synthetic oligonucleotides such as overlap extension method and inverse PCR method requires at least two oligonucleotides per one mutation. This is a rather expensive way to produce a large number

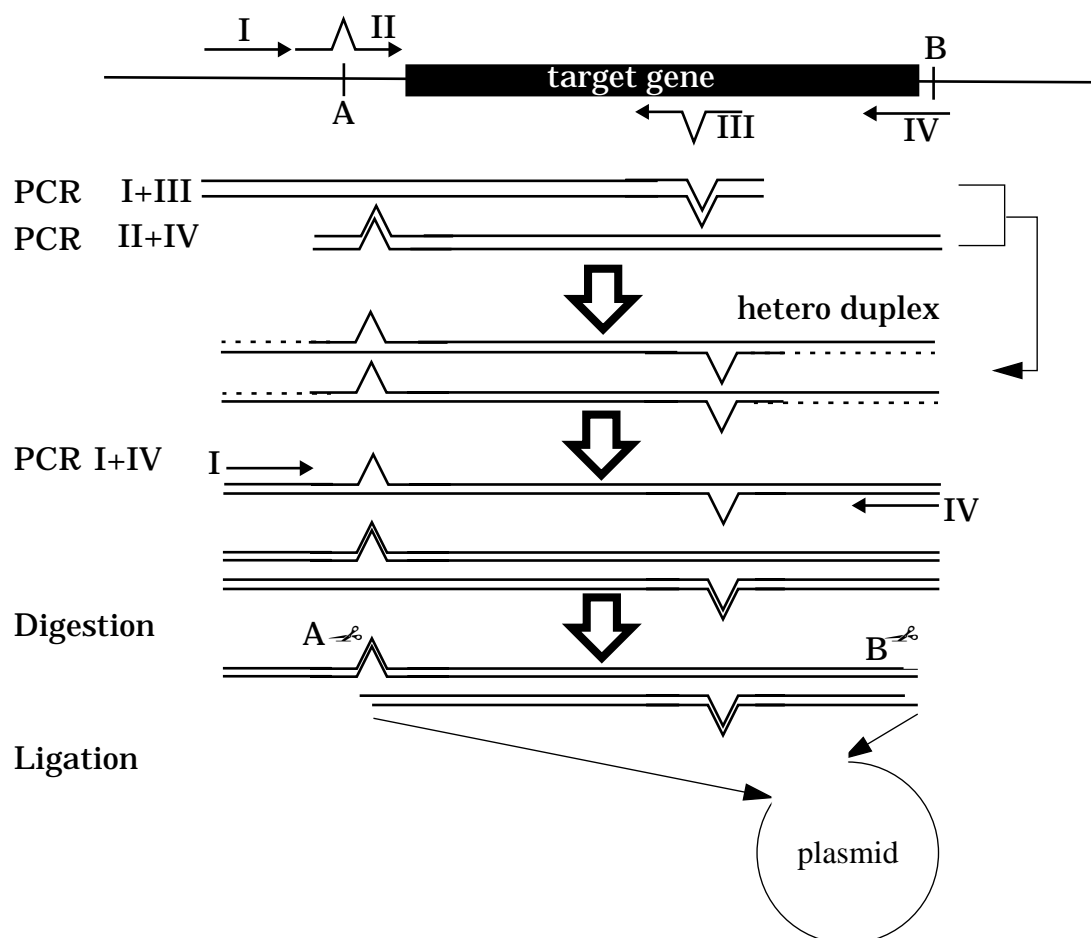


Figure 4-2: Schematic drawing of the mutagenesis procedure for modified overlap extension method (Ito *et al.*, 1991).

of mutations at many different sites, whereas the modified overlap extension method requires the minimum of four oligonucleotides for one procedure but only one oligonucleotide per one mutation (Figure 4-2) (Ito *et al.*, 1991). This method is particularly suitable for making a large number of mutants (>2) at many different positions. The procedure requires two steps as in the overlap extension method. The difference is that the final PCR product contains two PCR products (Figure 4-2). One contains the desired mutation but no mutation in the restriction site used for cloning. The other product contains no mutation in the position of the target but a mutation in the restriction site. Therefore only the PCR product having the desired mutation can be digested by the restriction enzyme and ligated into the plasmid (Figure 4-2). This method was used for the construction of the mutants at the salt-bridge of 434(1-63) (Siegal, to be published).

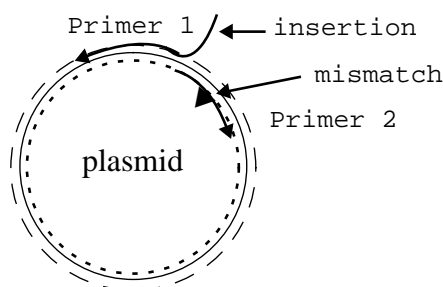


Figure 4-3: Schematic drawing of the mutagenesis procedure by inverse PCR method.

Inverse PCR

The overlap extension method and methods derived from it require two PCR steps and at least four oligonucleotides per one mutation. So-called inverse PCR method needs only one PCR step and two oligonucleotides. It is also easy to introduce deletions and insertions with this method. However, the inverse PCR method needs the complete plasmid to be amplified by DNA polymerase and blunt-end ligation (Figure 4-3). High-fidelity of DNA polymerases is necessary for this relatively long PCR. Hence *Pfu* DNA polymerase is often used for this method because of its high-fidelity (Error rate: 1.3×10^{-6}) (Flaman, *et al.*, 1994). However, the high-fidelity of DNA polymerase is usually achieved by its proofreading activity which includes a 3'-5' exonuclease activity, that may lead to a certain level of primer degradation. Therefore, the plasmid produced by inverse PCR methods tends to have deletions near the ligation sites depending on the design of the primers.

PCR based gene synthesis

Polymerase chain reaction (PCR) has become an indispensable tool for cloning and mutagenesis of the genes. It is also possible to synthesize relatively large DNA fragments by PCR (Figure 4-4) (Dillon & Rosen, 1990; Ye *et al.*, 1992; Prodromou & Pearl, 1992; Graham *et al.*, 1993; Casimiro *et al.*, 1995). There are several advantages of using gene synthesis for introducing a large number of mutations into a gene although the gene synthesis with synthetic oligonucleotides is often more costly than cloning from genomic libraries (Casimiro *et al.*, 1997). The codons can be optimized for the specific hosts in order to maximize the level of expression. The 'rare' codons that reduce the expression level can be avoided (Kane, 1995). In addition any

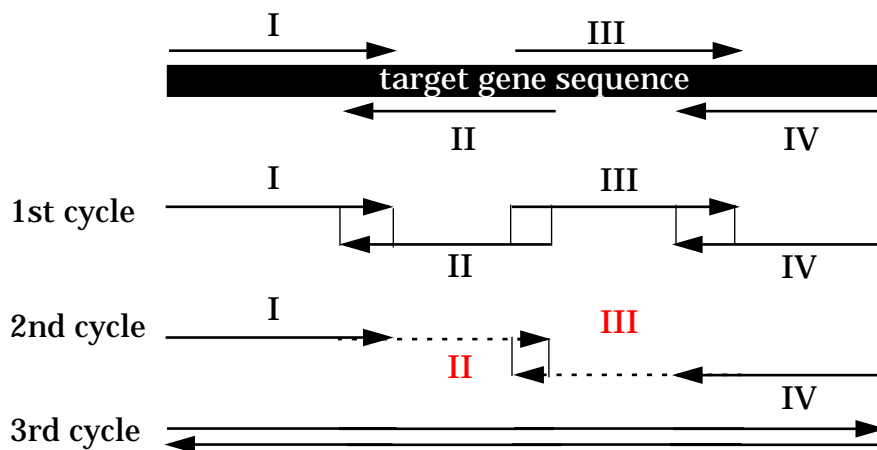


Figure 4-4: Schematic drawing of the gene synthesis by PCR.

kind of mutations, insertions, and deletions can be introduced rapidly and easily by substitution of a cassette of the oligonucleotides when single restriction sites were incorporated into the synthetic gene.

Kunkel method

The Kunkel method for site-directed mutagenesis uses uracile-containing single-stranded DNA as a template (Kunkel *et al.*, 1985). The principle is the following. Uracile-containing single-stranded DNA of the expression plasmid is prepared from M13k07 phages which had infected a *dut⁻ung⁻* *E.coli* strain that partially incorporates U instead of T into its DNA. Oligonucleotide primers containing mismatches for mutations can be hybridized to this uracile-containing single stranded DNA and extended by DNA polymerase to produce the other DNA strand. The DNA template containing uracile is then lost by delayed replication in an *E.coli* wild type host, yielding the plasmid containing the mutation.

Direct incorporation of synthetic oligonucleotides.

Direct ligation of synthetic oligonucleotides into the gene of interest is the simplest and straightforward way to introduce mutations when suitable restriction sites are available, which has been used for mutagenesis before the invention of PCR. Suitable restriction sites near the target site are necessary for the method. This is unlikely to happen very often by chance. However, unlike the other PCR based mutagenesis, the mutation does not rely on the mismatches introduced on the

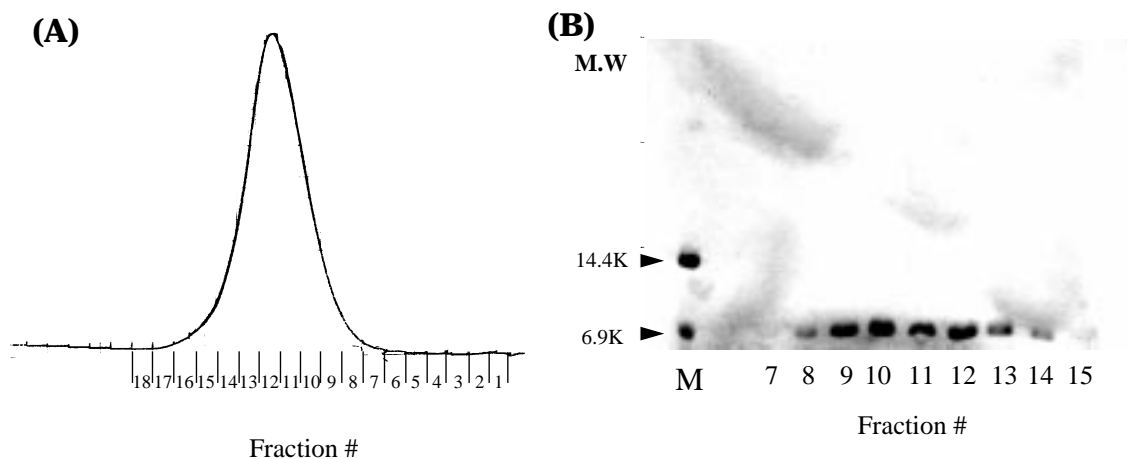


Figure 4-8:(A) Elution profile of P22/434(0–63) from the cellulose phosphate column. The optical density at 280 nm was monitored. (B) 20% SDS-PAGE analysis of the each fraction from the column.

for overproduction of the chimeric protein in *E.coli*. The protein was expressed in either LB media or the minimal media (see Materials) supplemented with $^{15}\text{NH}_4\text{Cl}$ for ^{15}N labeled samples. For expression of P22/434(0–63), *E.coli* BL21(DE3) with the plasmid p2434 were grown at 37°C until the OD_{600} reached approximately 1.0, followed by induction with IPTG at the final concentration of 0.5 mM for 6 hours. The cells were harvested and lysed by a french press at 15000 psi 2–3 times. The theoretical pI of P22/434(0–63) was calculated to 10.5 indicating a basic protein and a value close to that of 434(1–63) (10.2). The supernatant of the cell lysate after the centrifugation (17,000 rpm, SS-34 rotor, 1 hour, 4°C) was applied to an SP-Sepharose column (SP-Sepharose[®] Fast flow, Pharmacia) which was pre-equilibrated with 50 mM sodium acetate pH 4.8, 2 mM EDTA, 1 mM β -mercaptoethanol, 1 mM Pefabloc[®] SC (PENTAPHARM AG, Basel), 50 mM NaCl. P22/434(0–63) was eluted by a linear gradient from 10 mM to 500 mM NaCl in the same buffer and fractions were analyzed by SDS-PAGE. All relevant fractions were pooled and dialyzed against distilled water extensively. The solution was diluted with the initial buffer 2 times and loaded on the cellulose phosphate ion-exchanger column (Whatman P11, No. 4071010) which was pre-equilibrated with 50 mM K-phosphate pH 7.3, 50 mM NaCl. The protein was eluted by a linear gradient from 50 mM to 500 mM NaCl (Figure 4-8A). The fractions were collected and analyzed by SDS-PAGE (Figure 4-8B). Fractions containing the homogenous protein were dialyzed against distilled water extensively and lyophilized. The N-terminal amino acid sequence of the puri-

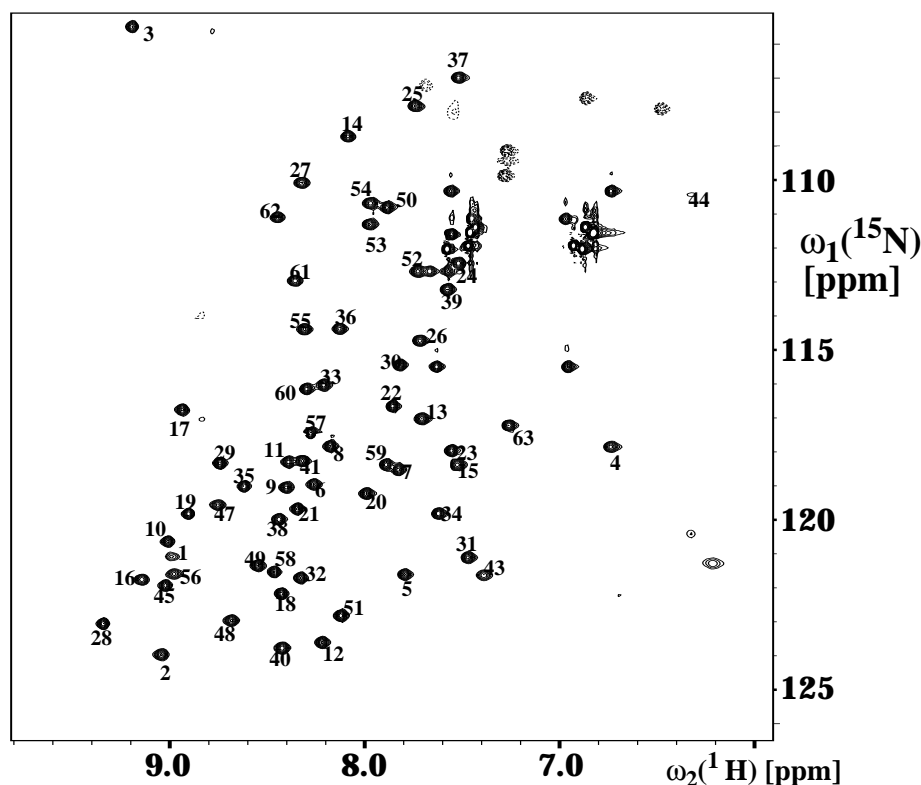


Figure 4-9: $[^{15}\text{N}, ^1\text{H}]$ -HSQC spectrum of P22/434(0–63) at pH 4.8 and 13°C in 25 mM K-phosphate buffer. All the backbone amide groups are labelled.

fied protein was analyzed by Edman degradation to confirm the purity and the sequence of the N-terminal 11 residues. It was found that P22/434(0–63) contains an additional methionine residue originating from the start codon which does not exist in 434(1–63). In Figure 4-9, the $[^{15}\text{N}, ^1\text{H}]$ -HSQC spectrum of P22/434(0–63) is shown with the backbone assignments. Many positions of the peaks in the $[^{15}\text{N}, ^1\text{H}]$ -HSQC spectrum are unchanged as compared with 434(1–63). For example, Asn16, Gln17, Gln28, Gln29, Gly37, Phe44 are virtually in the same position as found in 434(1–63) suggesting that there is no large structural changes caused after the helix exchange.

4.5. Preparation of *des-hydroxyl* 434(0–63)

The overall goal was to eliminate all hydroxyl groups from 434(1–63) for the unambiguous characterization of the surface water molecules by NMR spectroscopy (see Chapter 3). 4 serines in 434(1–63) were already successfully eliminated from 434(1–63) by making a chimeric protein with P22c2(1–76) based on the structural homology between P22c2(1–76) and 434(1–63). Interestingly the resulting chimeric protein has a higher stability than that of wild-type 434(1–63) (see Chapter 5).

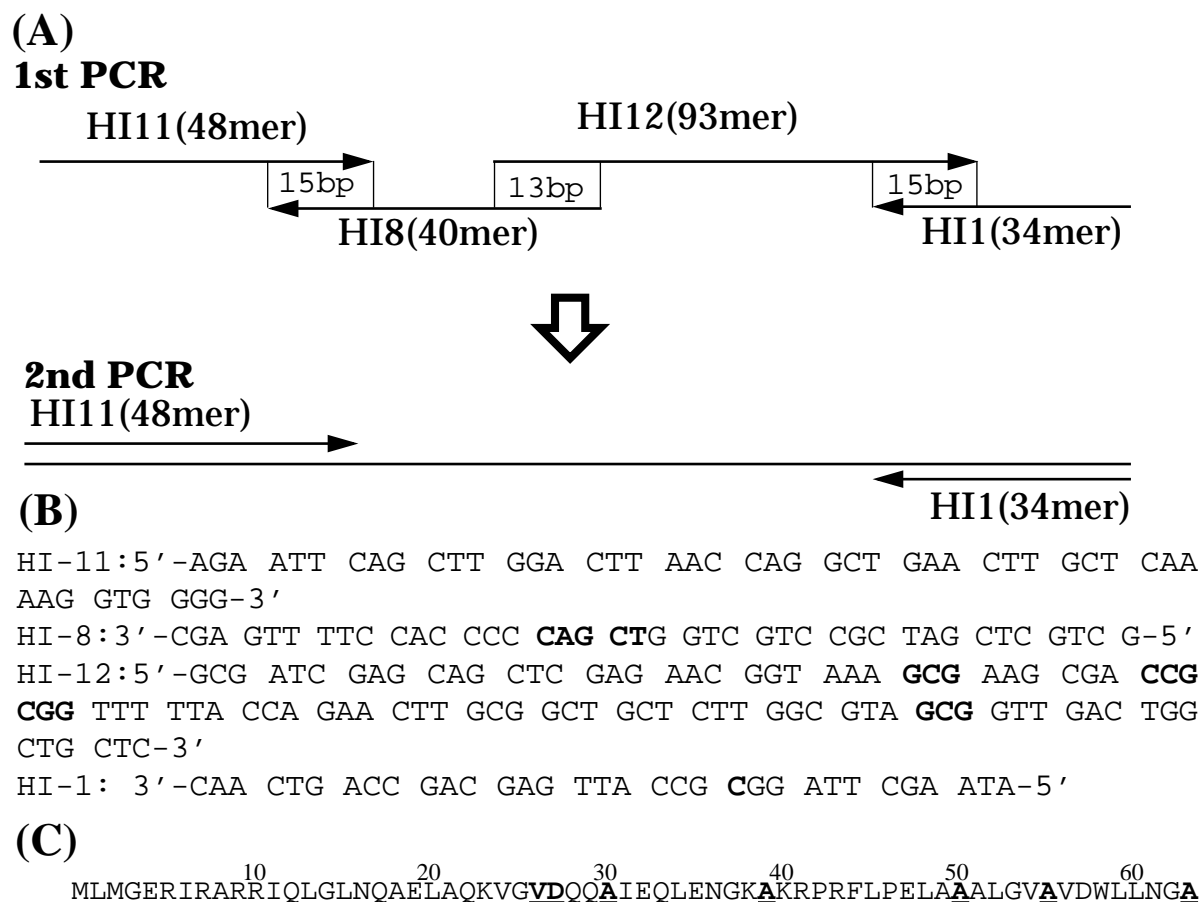


Figure 4-10:(A) Scheme for the gene construction of [R50A]-dh434(0-63) by PCR. (B) The DNA sequences of the oligonucleotides. The changes of the codons are indicated in bold. (C) The amino acid sequence of [R50A]-dh434(0-63). The substitutions of the amino acids are indicated by underlines and in bold.

Hence further mutations were decided to be introduced into this hybrid protein P22/434(0-63). P22/434(0-63) still contains 7 hydroxyl groups. It is laborious and cumbersome to replace all the residues containing hydroxyl groups one by one using oligonucleotide-directed mutagenesis because all the other serines and threonines are distributed in the primary structure (Figure 4-5). Therefore, PCR based gene synthesis was applied for the construction of the plasmid for des-hydroxyl 434(0-63).

Gene construction of des-hydroxyl 434(0-63)

PCR based gene synthesis was applied for the construction of the plasmid of dh434(0-63). Four synthetic oligonucleotides were designed (Figure 4-10). Each of the oligonucleotides contains overlapping regions (13-15 bp) as illustrated in Figure 4-10A. In a first reaction the four oligonucleotides were mixed, annealed, and elon-

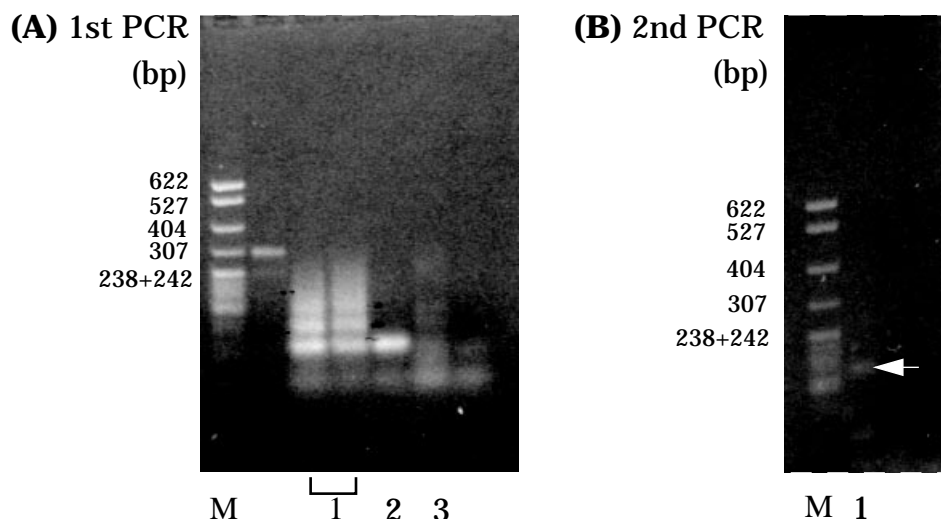


Figure 4-11:(A) PCR products after 1st PCR step. M: Marker (pBR322 DNA-*Msp*I digest) 1:PCR reaction with HI-1,HI-12, HI-11, and HI-8. 2: PCR reaction without HI-12. 3: PCR reaction without HI-8. (B) PCR products after 2nd PCR. The arrow indicates the PCR product of the correct size.

gated by Taq polymerase (Figure 4-4; Figure 4-10). The 1st reaction mixture was then used as a template and amplified with the primers HI-11 and HI-1 yielding the complete gene.

The details of the experiments are the following: All oligonucleotides were synthesized on Applied Biosystems 308B DNA synthesizer at 0.2 μ mol scale. The oligonucleotide HI-12 was purified using HPLC with a Nucleosil 300-5 C8 column. The first PCR contained 35 pmol of each of the four synthetic oligonucleotides, 2 μ l of 25 mM dNTP, 4 μ l of 100 mM $MgCl_2$ and the buffer supplied with the enzyme in the final volume of 100 μ L. The solution was kept for 10 min at 94 $^{\circ}C$, then slowly cooled down to 37 $^{\circ}C$ for 30 min and held at 37 $^{\circ}C$ for 10 min for annealing. 5 units of Taq polymerase were then added and heated to 72 $^{\circ}C$ for 3 min. The reaction was followed by 5 cycles of 1 min at 95 $^{\circ}C$, 2 min at 55 $^{\circ}C$, and 1 min at 72 $^{\circ}C$. The reaction was completed with a 5 min incubation at 72 $^{\circ}C$. The products were analyzed on a 1.2 % agarose gel (Figure 4-11A). Then 5 μ L of this reaction mixture were transferred to the second PCR tube containing 35 pmol of the two oligonucleotides HI-1 and HI-11. The final 100 μ L mixture contains 1 mM of each of dNTP, 5 units of Taq polymerase (Perkin-Elmer) in of 20 mM Tris-HCl (pH 8.4), 50 mM KCl. The second reaction involved a start at 95 $^{\circ}C$ for 1 min followed by 10 cycles of 1 min at 95 $^{\circ}C$, 2 min at 55 $^{\circ}C$ and 1 min at 72 $^{\circ}C$, and completed with a 5 min at 72 $^{\circ}C$ Figure 4-11B).

All PCR reactions were carried out on a Perkin-Elmer DNA Thermal cycler. The PCR product was purified by a PCR purification kit (QIAGEN) and then digested with EcoRI and Hind III. The digested DNA fragment was finally cloned into the plasmid p2434, in which the corresponding parts were removed by digestion with EcoRI and Hind III. *E. coli* DH5 α (Gibco BRL) was transformed with the ligation product and positive clones were identified by restriction digestion making use of a newly engineered restriction site. The DNA sequence of the complete gene was determined by the chain termination method with T7 polymerase (United States Biochemicals). This plasmid is termed pALLE. The protein from this plasmid, i.e. the 434 variant lacking all residues with hydroxyl groups, was named [R50A]-des-hydroxyl 434(0–63) (in short notation, [R50A]-dh434(0–63)) in order to differentiate the protein from the variant with the additional mutation A50R, which was termed dh434(0–63).

Expression & Purification of [R50A]-dh434(0–63)

Expression and purification procedures for [R50A]-dh434(0–63) are essentially identical to those described for 434(1–63). The purification consists of two steps. The first step is cation exchange column chromatography using SP-Sepharose Fast flow (Pharmacia). The second step is also cation exchange chromatography on phosphate cellulose P11 (Whatmann). Several modifications were applied to the previous purification method of P22/434(0–63) as the pH of the buffers and the salt conditions were slightly changed. *E. coli* BL21(DE3) transformed with pALLE was grown at 37 °C in LB medium until the OD₆₀₀ reached approximately 0.6. Cells were induced with IPTG (final concentration: 0.5 mM) for 6 hours. The cells were then harvested and lysed by a french press at 15000 psi. The supernatant from the cell lysate after the centrifugation (17000 rpm, SS34 rotor, 4 °C, 1 hour) was loaded onto the SP-Sepharose column which was pre-equilibrated with 50 mM sodium acetate pH 4.1, 1 mM EDTA, 1 mM β -mercaptoethanol, 1 mM Pefabloc SC, 10 mM NaCl. The protein was then eluted by a linear gradient from 10 mM to 500 mM NaCl in the same buffer and analyzed by SDS-PAGE. All relevant fractions were pooled, dialyzed against distilled water extensively and lyophilized. The typical yield was 10–20 mg/liter bacterial culture. The complete amino acid sequence of the purified protein was

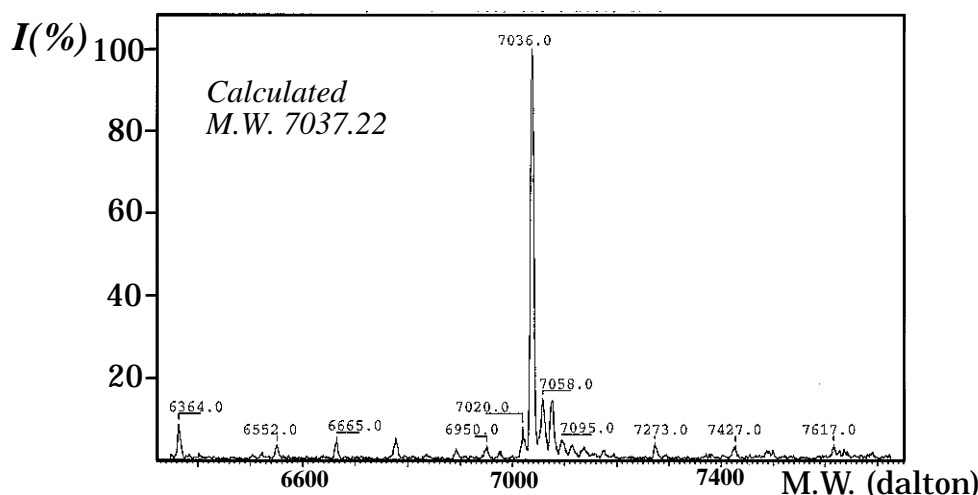


Figure 4-12: Convolution spectrum by ESI-Mass-spectroscopy of the purified [R50A]-dh434(0–63).

confirmed by Edman degradation procedure. The molecular mass of [R50A]-434(0–63) was also verified by electrospray mass spectrometry. In Figure 4-12 the mass spectrum of the recombinant protein is shown. It is in very good agreement with the expected molecular mass of 7037.22 Da (measured: 7036.0).

Light Scattering for detection of the aggregation

Unfortunately, slightly broadened lines were observed in the NMR spectra of [R50A]-dh434(0–63). Therefore the associate state of the protein was analyzed on a dynamic light scattering instrument (DyanaPro-801, Proteinsolution Inc.). In principle light scattering measurements provide several parameters indicating the state of the solution such as radius, estimated molecular mass, and polydispersity. The baseline of 1.000 and SOS (Sum of Squares) error < 5.000 should be observed in the case of 100% single size distribution (DYNAPRO-801 operator manual). The light scattering data of a 1 mM solution of [R50A]-dh434(0–63) in 20 mM K-phosphate pH 5.6 indicated a radius of 3.2 nm, a polydispersity of 1.719 nm, an estimated molecular weight of 48 kDa, and a SOS error of 16.551, suggesting significant aggregation. Addition of NaCl to a final concentration of 800 mM improved the situation yielding a radius of 1.7 nm, a polydispersity of 0.504, an estimated molecular mass of 12 kDa, a baseline of 1.001, and a SOS error of 1.696. Thus the aggregation was dramatically reduced at increased ionic strength. This can also be observed by the sharp resonance peaks in the [^{15}N , ^1H]-HSQC spectrum of [R50A]-dh434(0–63) con-

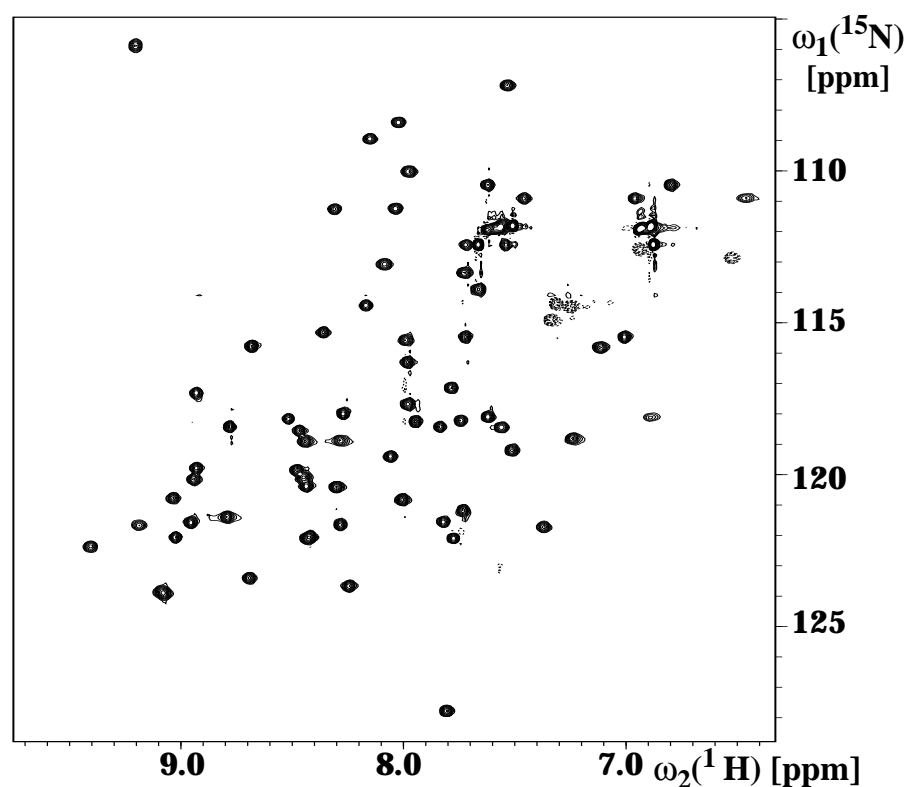


Figure 4-13: $[^{15}\text{N}, ^1\text{H}]$ -HSQC spectrum of [R50A]-dh434(0–63) at pH 5.0 and 13°C in 25 mM K-phosphate and 600 mM NaCl.

taining 600 mM NaCl (Figure 4-13).

Introduction of a Arg at position 50

It appears that [R50A]-dh434(0–63) tends to aggregate in the solution with low salt concentration as indicated by the dynamic light scattering experiments. This aggregation can be suppressed by addition of high concentration of NaCl. However, such additives could complicate the interpretation of the results of the hydration study obtained by NMR and decrease the sensitivity of NMR experiments which is a crucial point for the surface hydration study by NMR (Chapter 3). Therefore, an additional amino acid replacement was introduced in order to reduce the aggregation of [R50A]-dh434(0–63). Because of a number of hydrophobic residues such as alanines were introduced in the 434(0–63), the created hydrophobic surface might be a cause of this aggregation. Inspection of the three dimensional structure of 434(1–63) reveals that Ser50 has the highest solvent accessible surface among the positions of which side chains were replaced except for the termini (Table 6-6). A sin-

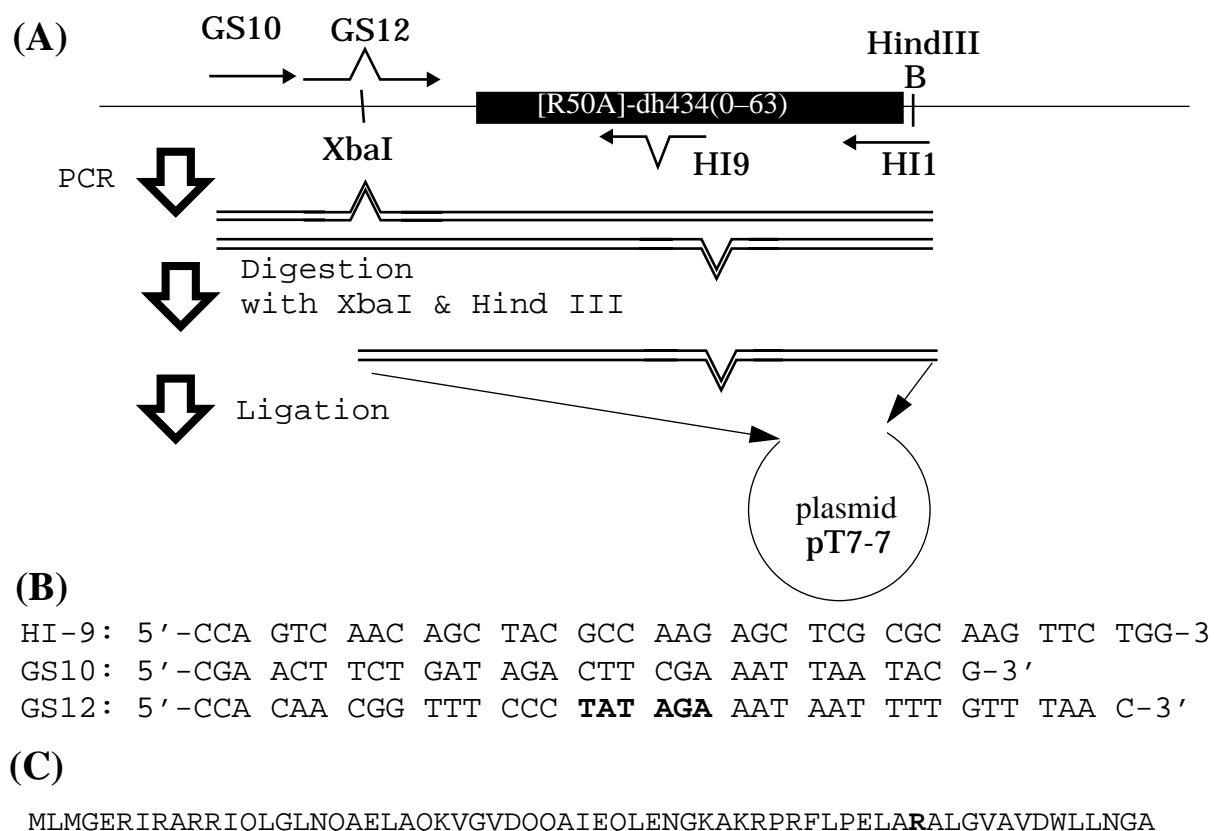


Figure 4-14: (A) Schematic drawing of the procedure for the introduction of a mutation of A50R in [R50A]-dh434(0-63) by PCR. (B) DNA sequences of the oligonucleotides used for the mutation. The changes of the DNA sequence are indicated in bold. (C) The amino acid sequence of dh434(0-63). The new mutation is indicated by underline and in bold.

gle mutation at this position was applied in order to improve the behavior of the protein based on this consideration. Arginine was chosen for this replacement because its positive charge should further facilitate the purification by ion exchange column and since a homologous protein 434Cro has also a positively charged residue at the same position (see Appendix H).

For the introduction of the mutation A50R, the modified overlap extension method was applied by using the oligonucleotides which were used for the mutations at the buried salt-bridge in 434(1-63) (Ito *et al.*, 1991; Siegal *et al.*, to be published). In Figure 4-12 and Figure 4-14 the scheme of this procedure is illustrated. This method was selected because the three oligonucleotides had already been available in the laboratory and only one more oligonucleotide had to be synthesized. For the first PCR step two reactions were prepared. The first one contained 50 pmol of each of

the oligonucleotides GS-12 and HI-1, 8 μ l of 10 mM dNTP, 8 μ l of 25 mM MgCl₂, 1 ng of the plasmid pALLE as a template, and 10 μ l of the buffer supplied with the enzyme in the final volume of 100 μ L. The other reaction contained 50 pmol each of the oligonucleotides GS-10 and HI-9. 8 μ l of 10 mM dNTP, 8 μ l of 25 mM MgCl₂, 1 ng of the plasmid pALLE as a template, and 10 μ l of the buffer supplied with the enzyme in the final volume of 100 μ L. 2.5 units of Taq polymerase were then added to each of the reaction tubes and heated to 95 °C for 1 min. The reaction was followed by 20 cycles with 1 min at 95 °C, 1 min at 50 °C and 1 min at 72 °C. The reaction was completed with a 5 min incubation at 72 °C. The resulting PCR products were analyzed on a 1.2% agarose gel and purified from a 1.2 % agarose gel by the gel extraction kit (QIAGEN) (Figure 4-15A). Then 5 μ L of the purified products were used for the second PCR step as a template. The final 100 μ L mixture for the second PCR contained 8 μ L of 10 mM dNTP, 8 μ L of 50 mM MgCl₂, 50 pmol of the two oligonucleotides GS-10 and HI-1 in 20 mM Tris-HCl (pH 8.4) and 50 mM KCl. The second PCR involved a hot start at 95 °C for 10 min followed by slow cooling to 37 °C for 30 min and kept at 37 °C for 10 min. Then 2.5 units of Taq polymerase were added and heated to 72 °C for 3 min. The reaction was followed by 10 cycles with 1 min at 95 °C, 1 min at 55 °C, and 1 min at 72 °C, and completed with a 5 min incubation at 72 °C. The PCR product was analyzed on an agarose gel (Figure 4-15B) and purified by the PCR purification kit (QIAGEN). The 307 bp DNA fragment was digested with XbaI and Hind III and ligated into the plasmid pT7-7. The DNA sequence of the modified gene in the plasmid pGFC was determined by dideoxy sequencing with T7 DNA polymerase (USB) (see Appendix D).

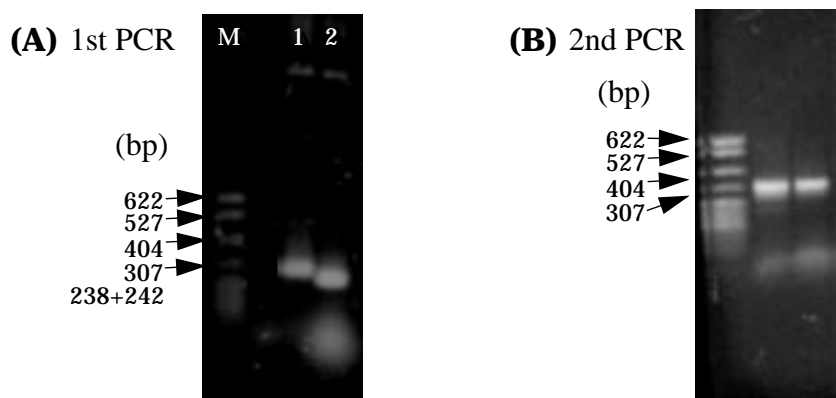


Figure 4-15: PCR products after 1st PCR step. 1: with the oligonucleotides GS-12 and HI-1. 2: with the oligonucleotides GS-10 and HI-9. M: Marker (pBR322 DNA-MspI digest)
(B) PCR products after 2nd PCR step with the two oligonucleotides, GS-10 and HI-1.

Expression & purification of dh434(0-63)

E. coli BL21(DE3) harboring plasmid pGFC was grown at 37 °C in LB-medium or the ¹⁵N-supplemented minimal medium for isotope labeling until an OD₆₀₀ of approximately 0.6 was reached, followed by induction with IPTG (final concentration: 0.5 mM) for 6 hours. The cells were harvested and lysed by a french press at 15000 psi. The extract was centrifuged at 35,000 rpm for 1 hour with an ultracentrifuge (Kontron). The supernatant was applied to SP-Sepharose column which was pre-equilibrated with 50 mM sodium acetate pH 4.1, 1 mM EDTA, 1 mM β-mercaptoethanol, 1 mM Pefabloc SC, 10 mM NaCl. dh434(0-63) was eluted by a linear gradient from 10 mM to 500 mM NaCl and fractions were analyzed by SDS-PAGE. All relevant fractions were pooled and dialyzed against distilled water extensively. This solution was diluted by the starting buffer 2 times and loaded onto the cellulose phosphate column (Whatman P11) which was previously equilibrated with 50 mM K-phosphate pH 6.1, 1 mM β-mercaptoethanol, 30 mM NaCl, and eluted by a linear gradient from 30 mM to 500 mM NaCl in the same buffer (Figure 4-16A). The fractions were analyzed by SDS-PAGE and the fractions containing the pure proteins were pooled for dialysis against water (Figure 4-16B). Typical yield of dh434(0-63) was 15–20 mg/liter bacterial culture. The complete amino acid sequence of the purified dh434(0-63) was confirmed by Edman degradation procedure. The dynamic light scattering data of a 1.8 mM solution of dh434(0-63) in 25 mM K-phosphate at

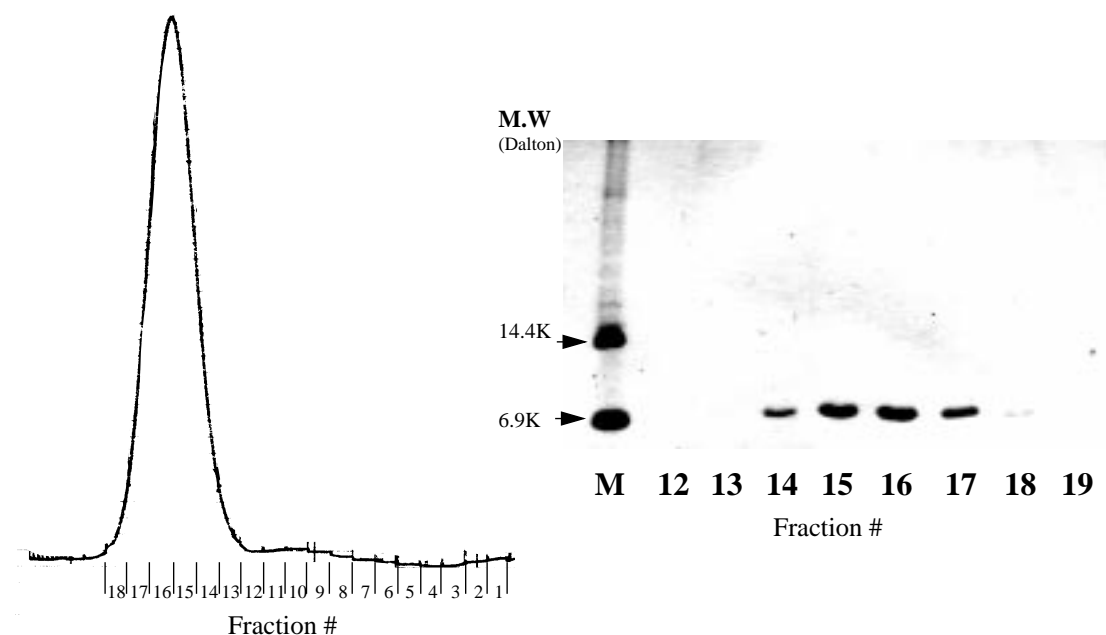


Figure 4-16: (A) Elution profile of *dh434(0–63)* from the cellulose phosphate column. The absorbance at 280 nm was monitored. (B) 20% SDS-PAGE analysis of each fraction.

pH 5.3 indicated a radius of 1.5 nm, a monodisperse protein preparation with an estimated molecular mass 9 kDa, a baseline of 1.000, and a SOS error of 0.826. Thus, the introduction of the Arg side chain at position 50 completely abolished aggregation, even at low ionic strength. The mutant protein *dh434(0–63)* was further used for the structural analysis by NMR spectroscopy (Chapter 6) and for the hydration studies (Chapter 3).

4.6. Construction, expression, and purification of the DNA binding domain of P22c2 repressor(5–68)

The DNA binding domain of phage P22c2 repressor, P22c2(1–76), is homologous to 434(1–63) but has unstructured tails at the N- and C- terminus as determined by NMR (Sevilla-Sierra *et al.*, 1994) (Figure 4-6). For comparison of its stability with that of 434(1–63) and its variants, P22c2(5–68) was constructed based on structure comparison in which the residues 1–4 and the last 8 residues were removed.

Plasmid construction, expression, and purification of P22c2(5–68)

Here the construction of the plasmid for expression of P22c2(5–68) is briefly

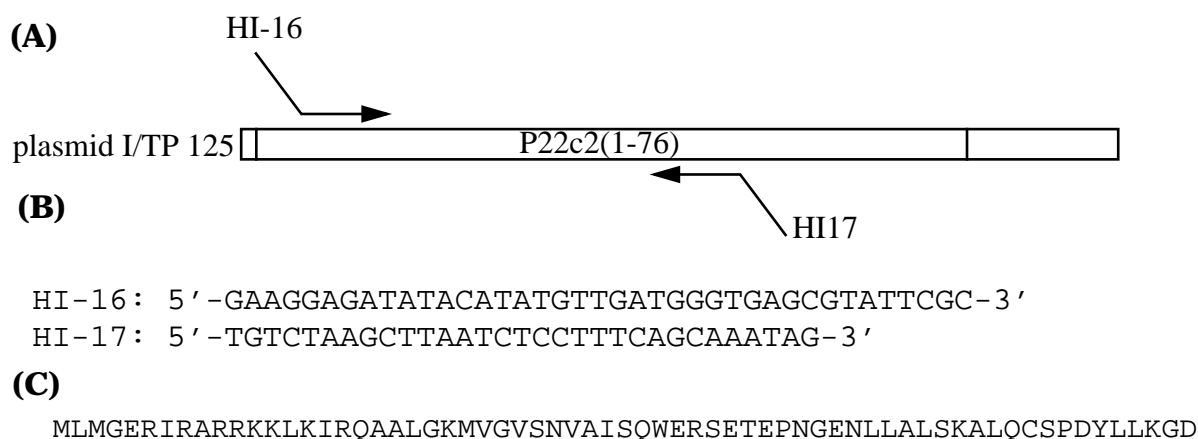


Figure 4-17: (A) Schematic drawing of the construction of P22c2(5-68). (B) DNA sequences of the primers. (C) The amino acid sequence of P22c2(5-68)

described. The plasmid I/TP 125, bearing the whole P22c2 repressor gene, was used to derive the gene of P22c2(5-68) (Sauer *et al.*, 1981). Two oligonucleotides were synthesized (Figure 4-17B). The first oligonucleotide contains the DNA sequence of residues 5-12 of P22c2(1-76). The second oligonucleotide contains the DNA sequences corresponding residues 62-68 of P22c2(1-76), a stop codon in the replace of Leu69 and a Hind III restriction site (Figure 4-18). The two oligonucleotides were used to amplify the shortened gene by PCR. The 222 bp DNA fragment made by PCR was purified with agarose gel and digested with NdeI and HindIII and ligated into the plasmid pT7-7 yielding the plasmid pT7-7/4AF. Finally *E.coli* BL21(DE3) was transformed with this plasmid for the expression of P22c2(5-68).

E.coli BL21(DE3) with plasmid pT7-7/4AF was grown at 37 °C in LB-medium until OD₆₀₀ reached approximately 0.7 and induced with IPTG (final concentration: 0.5 mM) for 5 hours. The cells were harvested and lysed by a french press at 15000 psi. The supernatant after centrifugation (SS-34, 17000 rpm, 1 hour) was applied to a SP-Sepharose column which was pre-equilibrated with 50 mM sodium acetate buffer pH 4.8 containing 1 mM EDTA, 1 mM β-mercaptoethanol, 1 mM Pefabloc SC, and 150 mM NaCl. P22c2(5-68) was eluted from the SP-Sepharose column by a linear gradient from 150 mM to 1.0 M NaCl. All relevant fractions were pooled and dialyzed against distilled water extensively. This solution was loaded onto a cellulose phosphate column (Whatmann P11) which was pre-equilibrated with 50mM K-phosphate pH 7.3, 30 mM NaCl, the protein was eluted by a linear gradient from 30

mM to 1.0 M NaCl in 50 mM K-phosphate pH 7.3 and fractions containing the pure protein were dialyzed against distilled water extensively and lyophilized. The typical yield of the protein was 50 mg/liter bacterial culture.

4.7. Overexpression and purification of an unstable 434 variant, [E35Q]-434(1–63)

Site-directed mutagenesis is a powerful method to characterize the contribution of the individual side chains to protein stability and the biological activities of proteins. However, a single mutation often results in low stability, poor expression of a mutant protein, proteolysis after translation and possibly toxicity of a protein variant. Particularly when the side chain has an important role for the stability of the native structure, variants may be poorly or not longer be expressed mainly because of the lowered stability makes the protein more susceptible to proteolytic degradation *in vivo*.

There is a single buried salt bridge in the structure of the 434 repressor between Arg10 and Glu35 (Mondragon *et al.*, 1989). A series of mutations was performed to characterize the role of this buried salt bridge (Siegal *et al.*, to be published; Perushin *et al.*, 1996). Figure 4-19 shows an SDS-PAGE analysis of the variants of both [E35Q]-434(1–63) and [R10M]-434(1–63) after the induction by IPTG for 0–3 hours. The expression of [E35Q]-434(1–63) was found dramatically diminished compared with [R10M]-434(1–63) and the wild type protein. To improve the expression yield of [E35Q]-434(1–63) a new expression vector was developed where the gene encoding for [E35Q] variant was fused to the 3' end of the wild-type 434(1–63) gene via a linker coding for a hexahistidine metal chelate affinity sequence and a methionine residue for subsequent cleavage of the fusion protein by cyanogen bromide (Figure 4-18). The fusion protein approach led to a strong increase in the yield of the expressed variant of the fusion protein level (Figure 4-19B). The expression was increased by a factor of more than 50 times compared with non-fusion [E35Q] variant judging from the band in SDS-PAGE. It was possible to purify on a Ni-NTA metal chelate affinity column despite the fact that affinity tag was introduced as a linker between 434(1–63) and [E35Q]-434(1–63) and not located at the N- or C- terminus. This histidine tag was also used for the separation of the wild type 434(1–63) from [E35Q]-434(1–63) after cleavage of the purified fusion protein by cyanogen bromide

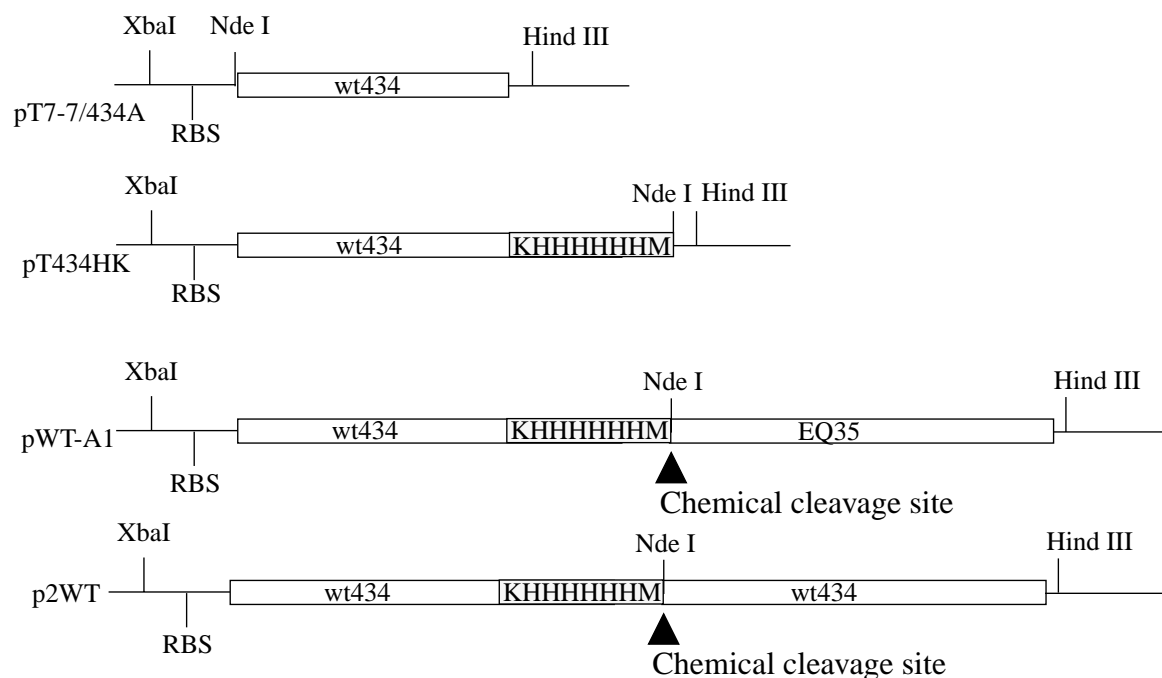


Figure 4-18: Schematic drawing of the constructed plasmids with 434(1–63).

(Figure 4-19C,D). This expression system with 434(1–63) has now been tested with the variety of proteins such as hirudin and various variant of homeodomain of LFB-1 and even with wild type 434(1–63). These results indicate that the expression system using wild type 434(1–63) as a fusion partner constitutes a generally applicable expression system for expressing proteins which are otherwise difficult to be obtained from *E.coli*. The fusion protein vector is in particular of interest for reducing the costs of preparing isotope-enriched proteins for NMR studies.

Expression and purification of [E35Q]-434(1–63)

The restriction sites in the original plasmid pT7-7/434A were modified so that a protein of interest can be cloned into the new plasmid pT434HK by using the restriction sites NdeI and Hind III (Figure 4-18). The gene coding for the variant [E35Q]-434(1–63) was then cloned into the plasmid pT434HK yielding the plasmid pWT-A1 for expression of the fusion protein. The analogous plasmid with two copies of the wild type gene was termed p2WT.

The fusion protein, 434(1–63)–[E35Q]-434(1–63) was expressed in *E.coli* BL21(DE3). Cells were grown at 37 °C, induced with IPTG at OD₆₀₀ = 1.0 to a final concentration of 1 mM and after 5 hours of induction harvested by centrifugation.

The stored cell pellet was lysed by a french press at 15000 psi. The extract was centrifuged (17000 rpm, SS34 rotor, 4 °C, 1 hour) after suspension in loading buffer (6 M GdmCl, 0.1 M NaH₂PO₄, 10 mM Tris-HCl pH 8.0). The viscous supernatant was loaded on a Ni-NTA resin (QIAGEN) for the purification and washed with the loading buffer until the absorbance of the elute at 280 nm reached a constant value. The protein was eluted with 8 M Urea, 0.1 M NaH₂PO₄, 10 mM Tris-HCl, and 250 mM imidazole at pH 8.0. The flow-through and the eluted fractions were analyzed by SDS-PAGE (Figure 4-19). The elute contains the his-tagged fusion protein with a purity of about 98% as judged by SDS-PAGE. Fraction containing the fusion protein were combined and extensively dialyzed against distilled water at 4 °C and lyophilized. The typical yield of the fusion protein 434(1–63)-[E35Q]-434(1–63) was 50 mg/liter bacterial culture.

The cleavage of the target protein from the fusion protein was achieved by chemical cleavage method with CNBr after the introduced methionine and the fact the 434(1–63) lacks methionines. 50 mg of the lyophilized fusion protein were dissolved in 4 ml of 70% (w/w) HCOOH. 5 ml (1 ml x 5) of 5 M CNBr solution in acetonitril (SIGMA) was added and incubated at room temperature for 24 hours. The reaction was stopped by adding 0.6 g of methionine and incubated at room temperature for 30 min. This reaction mixture was then dialyzed against distilled water. This solution was then diluted 3 times with the loading buffer (6 M GdmCl, 0.1 M NaH₂PO₄, 10 mM Tris-HCl pH 8.0) and again loaded onto the Ni-NTA column for the separation of the [E35Q] variant from the fusion partner and the un-cleaved fusion protein. The flow-through fractions were collected and analyzed by SDS-PAGE. The flow-through fractions contained more than 98% pure [E35Q]-434(1–63) as judged by SDS-PAGE analysis (Figure 4-19D). The final purification was performed by MonoS column (Pharmacia) equilibrated at 50 mM sodium acetate pH4.8 with a salt gradient from 50 mM to 1.0 M NaCl. The final yield of [E35Q]-434(1–63) was about 10 mg after the purification. 1 D ¹H NMR spectra of purified wild type 434(1–63) and the variant [E35Q]-434(1–63) are shown in Figure 4-20 and indicate large differences between the proteins. The poor dispersion of the signals of [E35Q]-434(1–63) as compared with 434(1–63) indicates almost complete unfolding (Figure 4-20). A further characterization of the stability of [E35Q]-434(1–63) was described in Chapter 5.

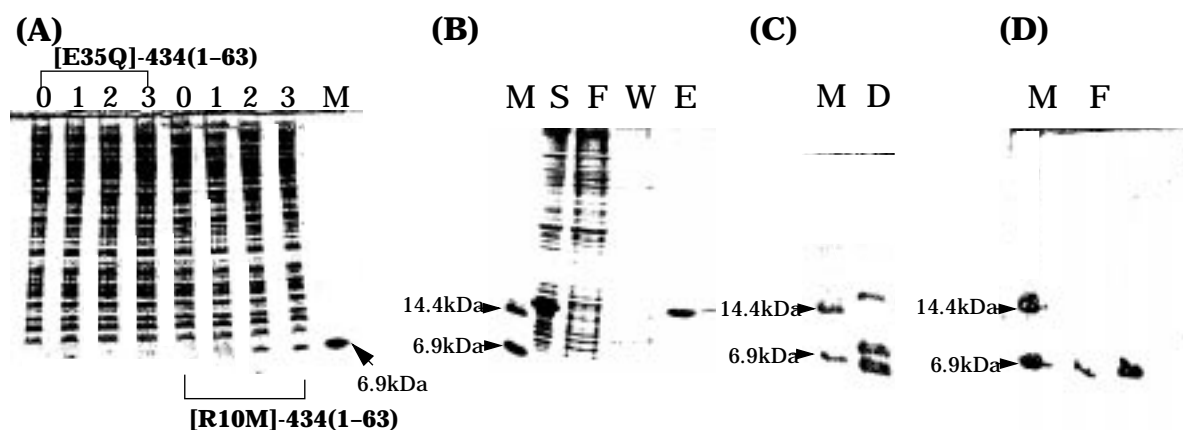


Figure 4-19: (A) Expression of [E35Q]-434(1-63) and [R10M]-434(1-63) after the induction with IPTG. 0: before induction; 1,2,3: 1, 2, 3 hours after the induction, respectively (B) Expression and purification of 434(1-63)-[E35Q]-434(1-63). M: marker, S: Supernatant, F: flow-through from the Ni-NTA column, E: Elution from the Ni-NTA column after washing with 250 mM imidazole. (C) Cleavage of the fusion protein by CNBr. (D) Purified [E35Q]-434(1-63).

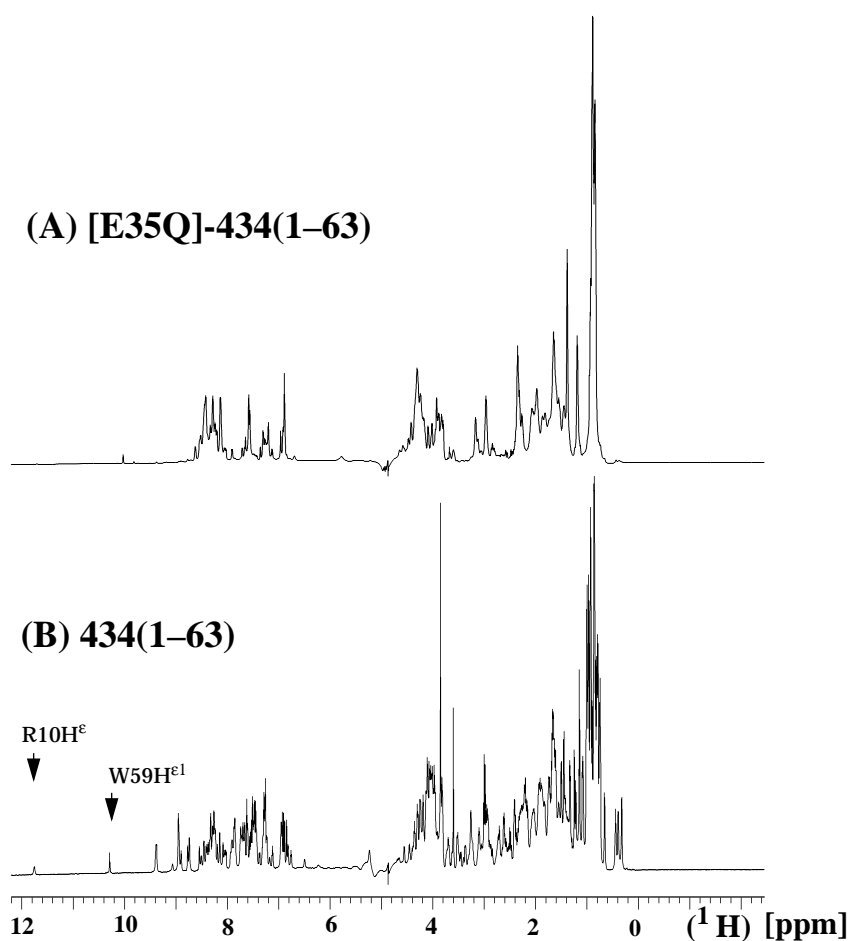


Figure 4-20: 1D ^1H NMR spectra of (A) [E35Q]-434(1-63) and (B) 434(1-63) at pH 4.8 in 50 mM K-phosphate buffer.

4.8. Materials

4.8.1. Minimum Media for isotope labeling.

Buffer solution adjusted to pH 7.0-7.4 with NaOH.

K₂HPO₄ 10.6 g/L

KH₂PO₄ 4.9 g/L

NaCl 0.5 g/L

¹⁵NH₄Cl 0.5 g/L

Minerals & Glucose #1

[D]-Glucose 4.0g/L

Biotin 1.2 mg/L (5 μM)

Vitamin B₁ 2.3 mg/L (70 μM)

MgSO₄ 123.2 mg/L (500 μM)

BH₃O₃ 6.1 mg/L (100 μM)

FeSO₄ 4.9 mg/L (35 μM)

ZnCl₂ 6.8 mg/L (50 μM)

Minerals #2

CaCl₂ 100 μM (100mM CaCl₂ stock solution 1ml/L)

Antibiotics

Ampicilin 50 μg/L

Carbenicilin 50 μg/L

4.8.2. Protein concentrations

All protein concentration of the variants of 434(1–63) was obtained by measuring the absorbance at 280 nm using $\epsilon=5550$.

5. Stabilities of variants of the DNA binding domain of phage 434 repressor

Variants of the DNA binding domain of 434 repressor, 434(1–63) were constructed and purified as described in Chapter 4. It is of interest to characterize the effects of the amino acid replacements on the stability of the protein. In this chapter the stabilities of variants were analyzed by using chemical denaturation and thermal denaturation monitoring CD (circular dichroism) signal. The mechanism of the stability changes for individual variants is discussed based on the three-dimensional structures of the proteins. This study may provide a strategy for the design of hyperstable proteins.

5.1. Thermal denaturation

CD measurement and analysis

The proteins were prepared as described in Chapter 4. 20 ~ 30 μM protein solutions were used for the CD measurements. All the CD measurements were performed with Jasco J-710 equipped with a water bath (RTE-100) using cuvettes with a path length of 1 mm. Thermal denaturation of variants of 434(1–63) containing a mutation at the position of the buried salt bridge between Arg10 and Glu35 was monitored by changes in circular dichroism at 222 nm in 2.5 °C steps. The samples were equilibrated for 1–2 minutes at each temperature and the signal was recorded for 3 min and averaged. The CD data was fitted according to the equation (1) to obtain the thermodynamical parameters. Since it is difficult to obtain a reliable value of ΔC_p from the experimental thermal denaturation curves, $\Delta C_p = 0$ was used for the curve fitting of the original data to calculate ΔH_m and T_m with the following equation (1).

$$y_{obs} = \frac{(y_N + m_N T) + (y_U + m_U T) \exp\left(\frac{\Delta H_m\left(1 - \frac{T}{T_m}\right) - \Delta C_p\left(T_m - T + T \ln \frac{T}{T_m}\right)}{RT}\right)}{1 + \exp\left(\frac{\Delta H_m\left(1 - \frac{T}{T_m}\right) - \Delta C_p\left(T_m - T + T \ln \frac{T}{T_m}\right)}{RT}\right)} \quad (1)$$

where y_{obs} is the observed circular dichroism; m_N and y_N , m_U and y_U are the slope and the intercept of the pre- and post-transition baselines of the denaturation curve, respectively; ΔH_m is the change in enthalpy at the midpoint of the denaturation temperature T_m ; ΔC_p is the difference in heat capacity between denatured and native protein; T is the absolute temperature in Kelvin; R is the gas constant. The curve fittings were performed with program Kaleida graph (Abelbeck Software). For calculation of ΔG , ΔC_p was calculated by using the value of 12 cal/mol/residues, i.e. 0.756 kcal mol⁻¹K⁻¹ (Edelhoch *et al.*, 1976; Pace *et al.*, 1990). The temperature dependence of the free energy change of denaturation ΔG is given by the following equation (Privalov, 1979).

$$\Delta G(T) = \Delta H_m\left(1 - \frac{T}{T_m}\right) - \Delta C_p\left[\left(T_m - T\right) + T \ln\left(\frac{T}{T_m}\right)\right]$$

For the analysis of the unstable mutant, [E35Q]-434(1-63), NaCl was added as stabilization reagent to measure T_m at the concentration of 2.5 M (Bowie *et al.*, 1989). For the other variants thermal denaturation curves were recorded by monitoring CD at 222 nm using the Jasco software to control the water bath with a heating rate of 50 °C/hour. The melting temperatures T_m of the variants were determined from the maximum of the first derivative.

Variants of 434(1-63) with mutations at the salt-bridge

Variants of 434(1-63) with a mutation in the buried salt bridge between Arg10 and Glu35, were constructed (Siegal *et al.* to be published; see Chapter 4). In order to estimate the contribution of each residues to the stability, the thermal stability of variants was carefully measured. Since it was found that [E35Q]-434(1-63) did not fold properly judging from CD and NMR spectra (Figure 5-1 and Figure 4-20),

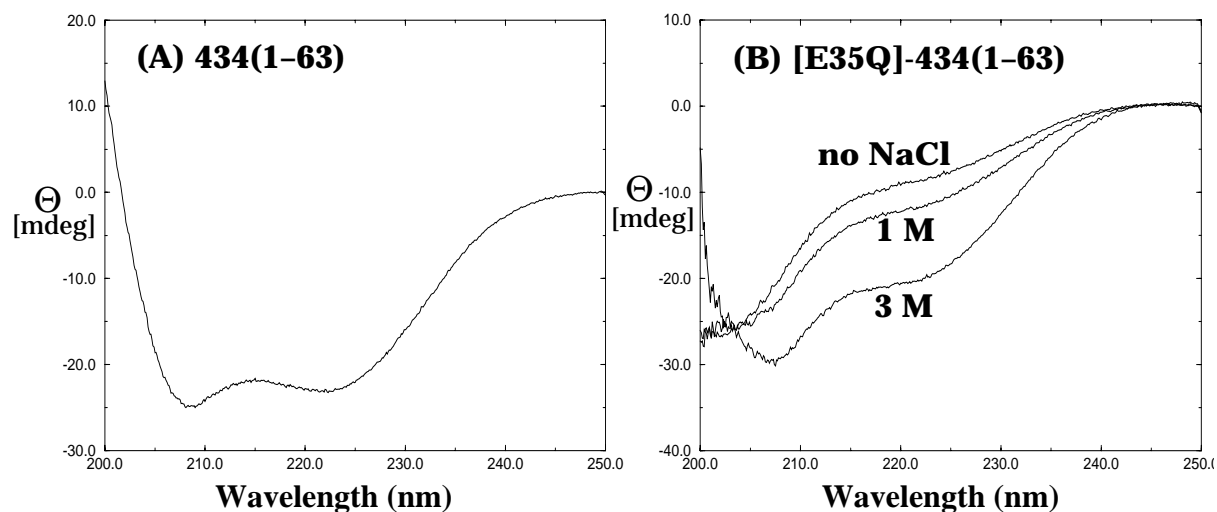


Figure 5-1: (A) CD spectrum of 20 μM 434(1-63) in 100 mM K-phosphate pH 5.0. (B) CD spectra of 16 μM [E35Q]-434(1-63) in 100 mM K-phosphate pH 5.0 with different concentrations of NaCl (0 M, 1.0 M, and 3.0 M) at 30 °C.

the thermodynamic parameters were obtained at the presence of high concentration of the salt, namely 2.5 M NaCl (Bowie *et al.*, 1989). The CD spectra recorded as different NaCl concentrations demonstrate that helical conformation is induced by the addition of the high concentration of NaCl (Figure 5-1).

For the comparison the thermal stabilities of 434(1-63), [R10M]-434(1-63), and [E35Q]-434(1-63) the thermal denaturation was also monitored at the presence of both 0 M NaCl and 2.5 M NaCl, of which thermal denaturation curves are shown in Figure 5-2.

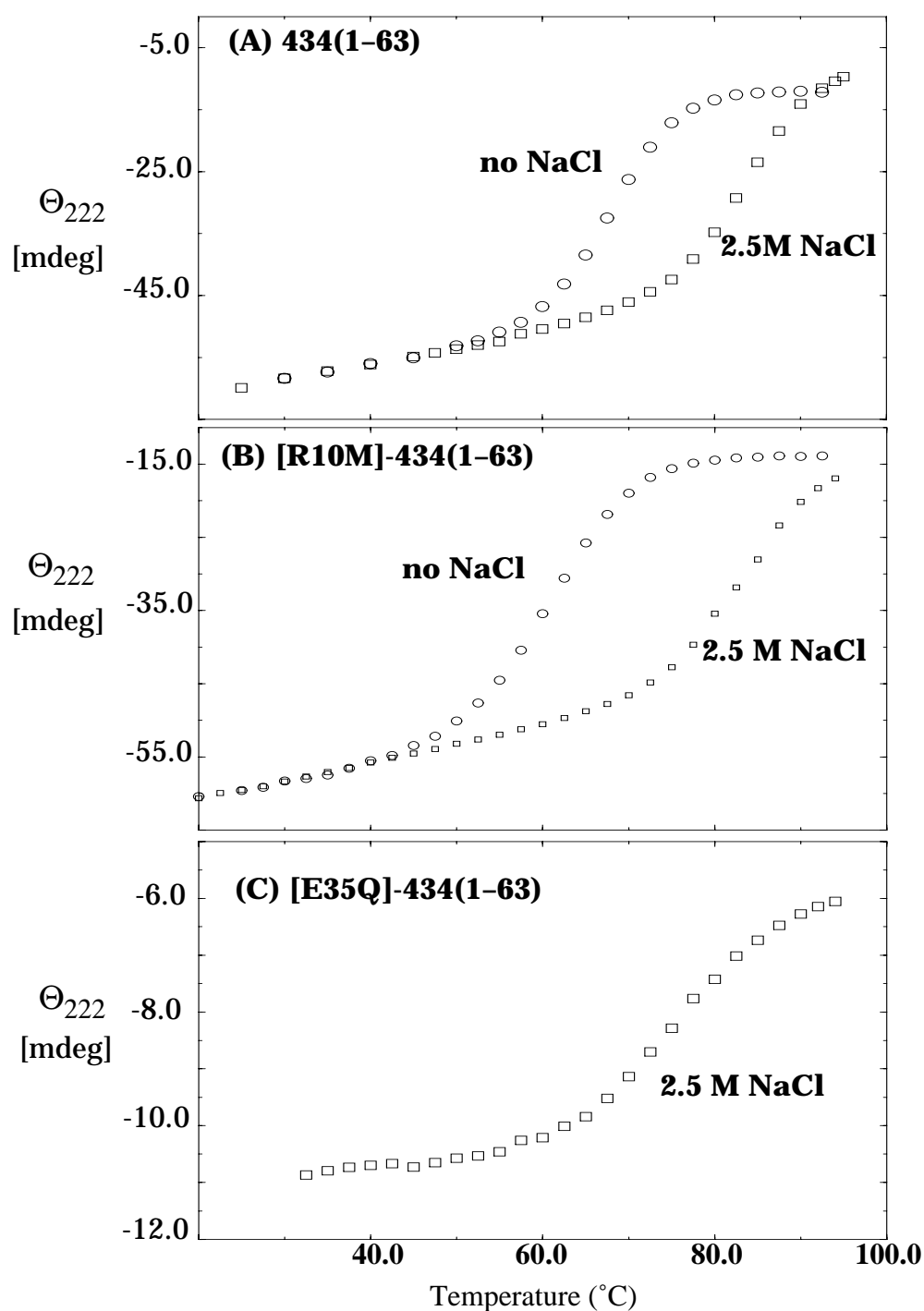


Figure 5-2: Thermal denaturation curves. (A) 45 μM wild-type 434(1-63) in 0 M NaCl and 2.5 M NaCl at pH 5.0. (B) 50 μM [R10M]-434(1-63) in 0 M NaCl and 2.5 M NaCl at pH 5.0. (C) 10 μM [E35Q]-434(1-63) with 2.5 M NaCl at pH 5.0.

Table 5-1 summarizes the thermodynamic parameters obtained by the analysis of the thermal denaturation curves of the variants of 434(1-63) at the two different concentration of NaCl. The high concentration of NaCl not only induced the helical conformation of [E35Q]-434(1-63) but also increased T_m of both 434(1-63) and [R10M]-434(1-63). However there is apparent difference in the CD spectra between

Table 5-1: Summary of the analysis of thermodynamic data for variants at the salt-bridge

Protein	T_m^a (°C)	ΔH_m^b (kcal/mol)	$\Delta G(25\text{ °C})^c$ (kcal/mol)	$\Delta G(13\text{ °C})^d$ (kcal/mol)	$\Delta(\Delta G)^e$ (kcal/mol)
434(1–63) (0M NaCl)	68.6 ± 0.1	57.0 ± 0.8	5.1 ± 0.1	5.6 ± 0.2	-
[R10M]- 434(1–63) (0 M NaCl)	61.9 ± 0.1	43.6 ± 0.7	3.2 ± 0.1	3.5 ± 0.1	-2.1 ± 0.1
434(1–63) (2.5 M NaCl)	85.1 ± 0.8	55.3 ± 2.1	5.2 ± 0.4	5.2 ± 0.6	-
[R10M]- 434(1–63) (2.5 M NaCl)	86.3 ± 0.9	52.4 ± 2.1	4.7 ± 0.4	4.6 ± 0.5	-0.6 ± 0.8
[E35Q]- 434(1–63) (2.5 M NaCl)	76.8 ± 1.7	35.6 ± 2.6	2.2 ± 0.6	1.8 ± 0.6	-3.4 ± 0.8

^a Midpoint of the thermal unfolding curve.

^b $\Delta H_m = T_m(K) \times \Delta S_m$. ΔS_m is negative of the slope of ΔG versus T at T_m in $\text{cal mol}^{-1} \text{K}^{-1}$.

^{c, d} ΔG was calculated based on $\Delta C_p = 0.76 \text{ kcal/mol}$ at the temperature 25°C and 13°C, respectively (see text).

^e $\Delta(\Delta G)$ is difference between 434(1–63) and variants of 434(1–63) in the same solution condition at temperature 13 °C.

wild-type 434(1–63) and [E35Q]-434(1–63) suggesting that the three-dimensional structures of 434(1–63) and [E35Q]-434(1–63) are not identical even at the presence of 2.5 M NaCl.

Thermal stabilities of other variants

T_m of 434(1–63), P22c2(5–68), P22/434(0–63), and dh434(0–63) are summarized in Table 5-2. Remarkable increase of T_m for a hybrid protein P22/434(0–63) was observed by 12.6 °C as compared with 434(1–63). The lowest T_m among the 4 proteins was observed for P22c2(5–68). For the detail analysis of the thermodynamic parameters of these proteins the chemical denaturation was performed as described in Chapter 5.2.

Table 5-2: Values of T_m for variants of 434(0–63) and P22c2(5–68).

Proteins	T_m
dh434(0–63)*	85.1 ± 0.9 °C
434(1–63)	68.6 ± 0.3 °C
P22/434(0–63)*	81.2 ± 0.4 °C
P22c2(5–68)	61.8 ± 0.3 °C

*residue 0 indicates an additional methionine at the N-terminus.

T_m were measured by monitoring circular dichroism at 222 nm at pH 5.0.

5.2. Chemical denaturation

CD measurement and analysis

Equilibrium chemical denaturation in GdmCl was monitored by CD (circular dichroism) at 222 nm with 20–30 μ M samples at 13.0 °C with 50 mM sodium acetate pH 4.8. 8 M GdmHCl solution (PEACE) was used for these experiments. CD signals at each concentration of GdmCl were collected for 1 min–1.5 min and averaged for the analysis of the thermodynamic parameters. GdmCl concentration [M] was determined by a refractometry after the measurements with the following equation: $[M] = 57.147(\Delta N) + 36.68(\Delta N)^2 - 91.60(\Delta N)^3$, where ΔN is the difference in refractive index between the denaturant solution and the buffer at the sodium D line. The data was fitted to a two-state transition model, as described by Pace *et al.* (1989) as following. The observed free energy of unfolding (ΔG_{obs}) and the unfolding equilibrium constant (K_{app}) are related as $\Delta G_{\text{obs}}(T) = -RT \ln K_{\text{app}} = -RT \ln \{(y_N - y)(y - y_U)\}$, where y , y_N , and y_U were the observed, native, and unfolded CD signal at a given GdmCl concentration, respectively (Santoro & Bolen, 1988). Equilibrium denaturation curves *versus* [GdmCl] concentration are shown in Figure 5-3. Difference in $[GdmCl]_{1/2}$ clearly indicates the increase of the stability for dh434(0–63) and P22/434(0–63). In Table 5-3 the thermodynamic parameters are summarized for the comparison among the variants. The m value for 434(1–63) is different from the val-

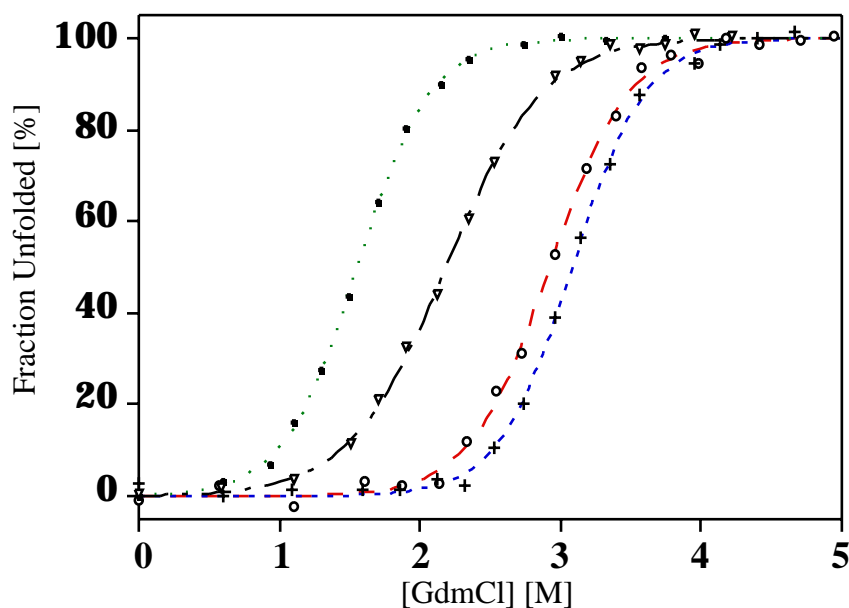


Figure 5-3: Equilibrium fraction of unfolded molecules versus [GdmCl] concentration for (∇):434(1–63), (+):P22/434(0–63), (\bullet):P22c2(5–68), and (o): dh434(0–63) in 50 mM Na-Acetate at pH 4.8, 13 °C.

Table 5-3: Thermodynamic denaturation parameters of dh434(0–63), 434(1–63), P22/434(0–63), and P22c2(5–68) at pH 4.8 and 13 °C.

Proteins	$\Delta G_{\text{eq}}(\text{H}_2\text{O})^{\text{a}}$ (kcal/mol)	$[\text{GdmCl}]_{1/2}^{\text{b}}$ (mol/l)	m^{c} (kcal \cdot l/mol 2)	$\langle m \rangle \times [\text{GdmCl}]_{1/2}$ (kcal/mol)	$\Delta \Delta G_{\text{eq}}(\text{H}_2\text{O})$ (kcal/mol)
434(1–63)	3.8 ± 0.2	2.2	1.7 ± 0.1	4.4	–
P22c2(5–68)	3.3 ± 0.1	1.6	2.1 ± 0.1	3.2	–0.5
P22/434(0–63)*	6.2 ± 0.2	3.1	2.0 ± 0.1	6.1	2.4
dh434(0–63)*	6.1 ± 0.2	2.9	2.1 ± 0.1	5.7	2.3

*residue 0 indicates an additional methionine at the N-terminus.

^{a, c} Obtained from the equation, $\Delta G = \Delta G(\text{H}_2\text{O}) - m \cdot [\text{GdmCl}]$.

^b Midpoint of the GdmHCl unfolding curve.

ues from other variants. It appears that this m value for 434(1–63) is unique feature of 434(1–63). Three times of the independent measurements were performed and always indicated the same value for 434(1–63) confirming that this is not originating the error in the experiments.

5.3. Discussion

Mutations at the salt bridge of 434(1-63)

A salt bridge between Arg10 and Glu35 is found in the interior of the three-dimensional structure of the DNA binding domain of 434 repressor (Mondragon *et al.*, 1989; Neri *et al.*, 1992a; Pervushin *et al.*, 1996). The importance of the salt-bridge for the stability of the proteins is still a matter of debate (Mrabet *et al.*, 1992; Yang & Honig, 1993; Borders *et al.*, 1994; Marqusee & Sauer, 1994; Waldburger *et al.*, 1994; Hendusch & Tidor, 1994; Tissot *et al.*, 1996). For better understanding of the role of the salt bridge in the interior of 434(1-63), variants of 434(1-63) with the mutations at the salt bridge, i.e. [R10M]-434(1-63) and [E35Q]-434(1-63) were constructed (Siegal *et al.* to be published; Pervushin *et al.*, 1996). The three-dimensional structure of [R10M]-434(1-63) has been determined by NMR and shows the same global fold with five helices as found in the structure of 434(1-63) with a minor shift of the helices (Pervushin *et al.*, 1996). A destabilization of 2.1 kcal/mol was observed for [R10M]-434(1-63) as compared with 434(1-63) from the thermal denaturation experiments (Figure 5-2 and Table 5-1). In a previous study the attempt to produce [E35Q]-434(1-63) had failed because of very low expression yields of the protein in *E.coli* presumably due to the low stability of [E35Q]-434(1-63) *in vivo* (Siegal *et al.* to be published; Figure 4-19). In this study, the unstable mutant [E35Q]-434(1-63) was successfully expressed in *E.coli* and purified by using the novel fusion expression vector as described in Chapter 4. In contrast to [R10M]-434(1-63), [E35Q]-434(1-63) shows a largely disordered structure at low ionic strength (Figure 4-20 and Figure 5-1B). This observation explains why [E35Q]-434(1-63) was not stably expressed in the cytoplasm of *E.coli* as the unfolded protein is most likely prone to degradation. It was found in the three-dimensional structure of the variant [R10M]-434(1-63) that Lys7 could also interact with Glu35, accompanied by a minor conformational changes compensating the unpaired negative charge of Glu35 (Pervushin *et al.*, 1996). It appears that in the variant [E35Q]-434(1-63) an analogous compensation by another acidic residues cannot be achieved. The apparent pKa value of 2.0 for the carboxylate of the side chain of Glu35 was obtained from the pH titration of the chemical shift of $^{13}\text{C}^\delta$ (Chapter 7). This apparent pKa shift of the side chain carboxylates of Glu35 can be translated to the contribution of this ionizable group to

the stability of the protein in pH-dependent manner as shown by the following equation,

$$\Delta\Delta G(pH) = -RT \ln \left[\frac{(1 + 10^{pH - pK_{a(f)}})}{(1 + 10^{pH - pK_{a(u)}})} \right]$$

in which $\Delta\Delta G$ is the contribution of the ionizing group to the stability of the protein, and $pK_{a(f)}$ and $pK_{a(u)}$ are the pK_a values of the acidic residue in the native and unfolded protein, respectively (Anderson *et al.*, 1990; Langsetmo *et al.*, 1991a,b). Accordingly the contribution of the side chain of Glu35 can be estimated to be about 2.7 kcal/mol at 13°C and pH 4.8 assuming the pK_a of 4.2 for Glu35 in the unfolded state. This value is more than half of free energy of folding of ΔG obtained by thermal and chemical denaturation of 434(1–63) (Table 5-1 and 5-3). Moreover, the titration of the side chain carbonyl group of Gln17 and Gln28 observed in ^{13}C chemical shift suggested the existence of a cooperative large network of hydrogen bonds between Arg10, Glu35, Gln17, and Gln28. The unsatisfied buried positive charge of Arg10 caused by replacement from Glu35 to Gln leads to the observed unfolded structure of [E35Q]-434(1–63).

It has been proven that an entire salt-bridge network can be replaced by hydrophobic residues resulting in active mutants that are significantly more stable than wild type and that simple hydrophobic interactions provide more stabilizing energy than the buried salt bridge and confer comparable conformational specificity (Waldburger *et al.*, 1995). Why does the buried salt bridge exist in proteins? There are several explanations proposed such as playing a role during protein folding (Waldburger *et al.*, 1996; Tissot *et al.*, 1996) and for the specificity (Perutz, 1989; Schirmer & Evans, 1990). Conserved salt bridges were shown to be less exposed than non-conserved ones (Schueler & Margalit, 1995). The salt-bridge found in 434(1–63) seems to be often found in homologous DNA-binding proteins with helix-turn-helix motif (Pervushin *et al.*, 1996; Appendix H). Why is the buried salt bridges conserved? Surprisingly the produced salt-bridge variant [E35Q]-434(1–63) in this study is found to be soluble up to millimolar concentration even though its three-dimensional structure is disordered, which might provide some speculations such as playing a role in a denatured state at low pH.

Comparison of the stabilities of 434(1–63), dh434(0–63), and P22c2(5–68)

A similar m value is observed for 3 homologous proteins except for 434(1–63). It is found that m values correlate with the size of the proteins (Schellman, 1978; Myers *et al.* 1995). This correlation is assumed to result from the changes in the accessible surface area (Δ ASA) upon unfolding. Inspection of the three-dimensional structures of the proteins might explain the origin of the difference in m values. Δ ASA was calculated to 4719 Å² for dh434(0–63) from the three-dimensional structure and the extended polypeptide model. This value provides expected m value of 2.0 kcal•l/mol². This is in good agreement with the experimental values of 2.1 kcal•l/mol². The structure of P22c2(5–68) can be modeled based on the structure of P22c2(1–76) by removing the residues at the N- and C- terminus and used for the calculation of Δ ASA of P22c2(5–68) resulting in Δ ASA 5002 Å². This corresponds to m value of 2.1 kcal•l/mol² according to the reported correlation (Myers *et al.*, 1995). This is also in good agreement with the experimental m value of P22c2(5–68). On the contrary Δ ASA of 434(1–63) is obtained to 4603 Å² that leads to the expected m value of 2.0 kcal•l/mol². However, the experimental m value of 1.7 kcal•l/mol² for 434(1–63) suggests a Δ ASA of about 3200 Å². This discrepancy could be because the extended peptide model as an unfolded structure is not an appropriate model in the case of 434(1–63), or because the unfolding of the protein does not fit with two-states model any more. The observation of the residual structure of 434(1–69) in 7 M urea solution might provide a support for such explanation (Neri *et al.*, 1992b). 434(1–63) in the denatured state might contain some residual structure as found in 434(1–69) in 7 M urea solution so that experimental m values might be smaller due to this residual structure.

The three-dimensional structure of dh434(0–63) reveals several small but notable differences between dh434(0–63) and 434(1–63) (see Chapter 6). An increase of hydrophobic contacts is observed in the core of dh434(0–63) (Figure 6-11). In addition the presence of a new hydrogen bond by the side chain of Arg9 connecting two sub-domains in dh434(0–63) was identified as compared with wild-type 434(1–63) (Figure 6-9 and Figure 6-10). The range of 0.5–2 kcal/mol is generally the net stabilization energy by a hydrogen bond (Fersht, 1987). Therefore the stabilization of

dh434(0–63) can be principally explained based on the formation of a new hydrogen bond. It is found that the large networks of ion-pairs exist in the three-dimensional structure of hyperthermophilic proteins as compared the structure of mesophilic counterparts suggesting the importance for hyperthermostability (Goldman, 1995). Such observation in other hyperthermophilic proteins can also support the origin of the increased stability of dh434(0–63).

A higher stability was observed for the newly constructed proteins P22/434(0–63) and dh434(0–63) compared to the wild-type protein. Interestingly three naturally occurring, structurally homologous proteins, i.e. 434(1–63), P22c2(5–68), and 434Cro(1–76) are found to have a similar $\Delta G(\text{H}_2\text{O})$ of 3.3–3.8 kcal/mol (Table 5-3; Padmanabhan *et al.*, 1996). This implies that the proteins of the 434 family are not optimized in terms of stability. Swapping of the sequences in the secondary structure among the homologous proteins in this case the first helices between P22c2(1–76) and 434(1–63) significantly improved the stability of 434(1–63). In general although a large number of data has been accumulated for understanding of protein stability (Fersht & Serrano, 1993), it is not an easy task to rationally design or modify a protein to become stable enough for application in biotechnology. The data shown here for dh434(0–63) and the hybrid protein P22/434(0–63) suggests that construction of a chimeric protein from homologous proteins may be applied in the future to generate more stable proteins.

6. Structure determination and Characterization of dh434(0–63) by NMR

High resolution structures of the proteins are indispensable for the detailed analysis of the protein hydration by NMR spectroscopy because it is necessary to inspect the three-dimensional structure carefully to separate NOEs from long-lived water molecules and NOEs *via* chemical exchange with water. Moreover, the assignments of $^{13}\text{C}/^{15}\text{N}$ atoms will be of great advantage for the separation of the signal overlaps in ^1H frequency by utilizing the dispersions in ^{13}C and ^{15}N frequencies. The resonance assignments, structure determination, characterization of dh434(0–63) by NMR spectroscopy are described in this chapter. The determined structures of dh434(0–63) are compared with wild-type 434(1–63).

6.1. ^1H resonance assignments of dh434(0–63) and secondary structure determination

Sample preparation & NMR spectroscopy

dh434(0–63) was expressed in *E. coli* and purified as described in Chapter 4. For the homonuclear experiments a NMR sample contains 5 mM dh434(0–63) in 25 mM K-phosphate buffer with a mixed solvent of 90% $\text{H}_2\text{O}/10\%$ D_2O was used. The protein samples were adjusted to pH 4.8 by the addition of minute amounts of NaOH and HCl.

The spectra were recorded on a Varian Unity-pulse 750 spectrometer and the temperature of all experiments used for the structure determination was 13.0 °C. For the assignments of ^1H resonances homonuclear 2QF-COSY (Rance *et al.*, 1983), clean-TOCSY with mixing time 70 msec (Griesinger *et al.*, 1988) and NOESY with mixing time 40 msec (Anli-Kumer *et al.*, 1980) spectra were recorded in H_2O in the pure absorption mode using the States-TPPI method (Marion *et al.*, 1988). For the collection of upper-bond $^1\text{H}-^1\text{H}$ distance constraints, a [$^1\text{H}, ^1\text{H}$]-NOESY spectra with

zero-quantum coherence suppression by systematic incrementations of mixing time in every each scans (from 35 msec to 45 msec with an incrementation of 0.454 msec) (Macura *et al.*, 1982). A data set in H₂O solution was recorded with 1024 × 2048 complex points, with $t_{1\max} = 56$ msec and $t_{2\max} = 223$ msec. Vicinal spin–spin coupling constants ${}^3J_{\text{HN}\alpha}$ were determined from the aforementioned NOESY spectrum or [¹⁵N, ¹H]-HSQC spectrum in H₂O by inverse Fourier transformation of in-phase multiplets (Szyperski *et al.*, 1992). Vicinal spin–spin coupling constants ${}^3J_{\alpha\beta}$ were obtained from an E.COSY spectrum (Griesinger *et al.*, 1985) recorded at 750 MHz proton frequency with a digital resolution of 1.6 Hz/point along ω_2 . ${}^3J_{\text{N}\beta}$ was derived from constant time 3D-HNHB experiments (Archer *et al.*, 1991; Chary *et al.*, 1991), recorded on Bruker AMX600 spectrometer by using a ¹⁵N labeled sample. This information of coupling constants was used for the inputs of program HABAS (Güntert *et al.*, 1991a).

Sequential assignments

Sequence specific ¹H n.m.r. assignments for the dh434(0–63), were obtained following the standard strategy (Wüthrich, 1986) for small proteins, using clean-TOCSY, NOESY, and 2QF-COSY. NOE upper bounds on ¹H–¹H distances were obtained from a NOESY experiment with $\tau_m = 40$ msec with selective suppression of zero-quantum coherence. The spin systems were categorized by using clean-TOCSY and 2QF-COSY. Non-degenerated alpha protons of glycine residues were easily identified in 2QF-COSY spectra. The nine alanines of dh434(0–63) were also found without ambiguity by the combination of clean-TOCSY and 2QF-COSY. Because dh434(0–63) does not contain any serines and cysteines, several AMX spin systems were readily assigned to one of aspartic acids, asparagine, or two aromatic residues. These spin systems assigned to the each amino acid types were connected by using NOESY spectra. A large part of the sequential connections was completed mainly on the basis of strong d_{NN} connections starting from the unique amino acid types such glycines and alanines. These connections are coincided with $d_{\alpha\text{N}}(i,i+3)$, $d_{\alpha\text{N}}(i,i+4)$ connections ensuring their initial assignments. All d_{NN} connection in ¹H homonuclear spectra were identified in the helices except for Glu47 due to overlapping, which was solved in ¹⁵N separated NOESY spectra. The sequential assign-

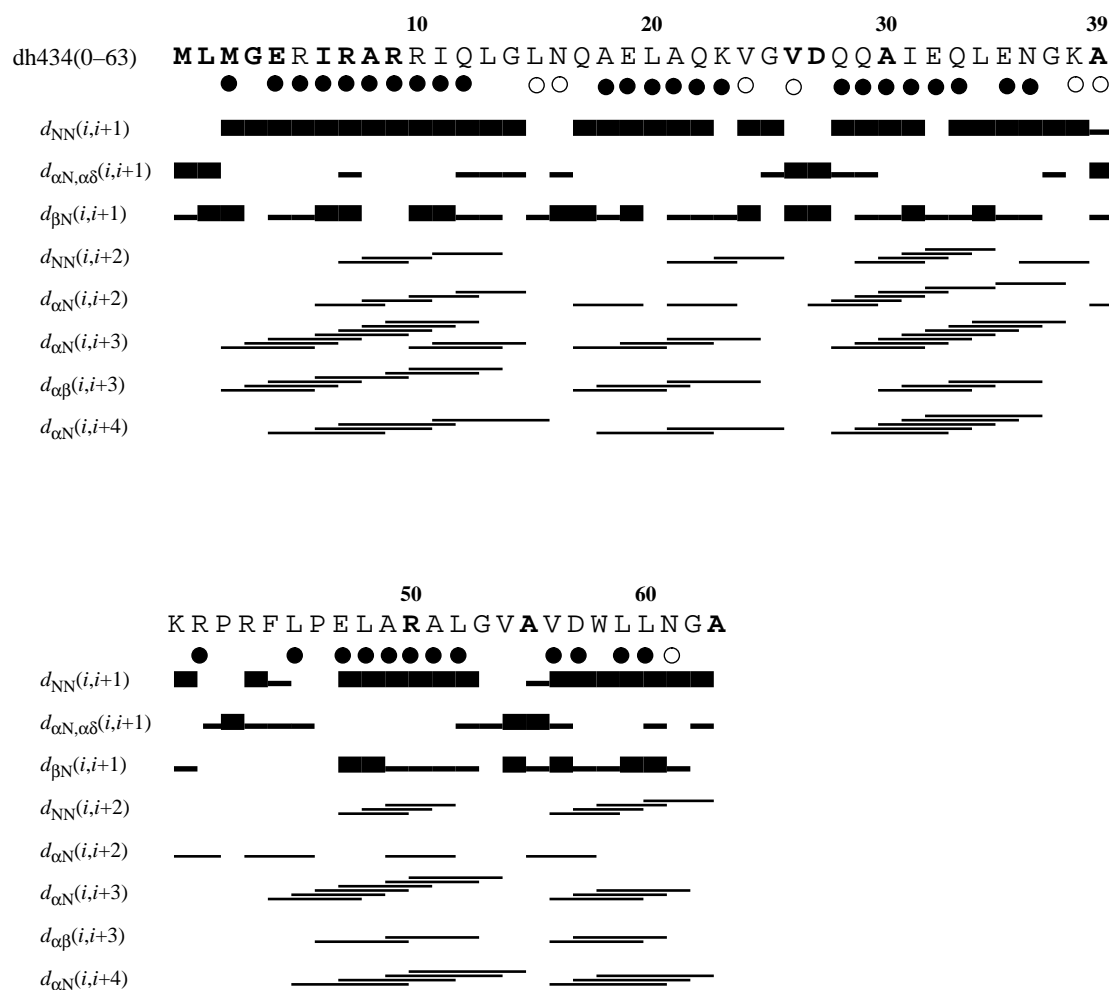


Figure 6-1: The amino acid sequence, survey of sequential and the medium-range NOE connectivities, and spin-spin coupling constants $^3J_{\text{HN}\alpha}$ of dh434(0–63). For the sequential connectivities, the thickness of the bar indicates the NOE intensities. The exchanged residues in the amino acid sequence of dh434(0–63) relative to 434(1–63) are shown in bold. Below the sequence, open and filled circles identify residues with $^3J_{\text{HN}\alpha} > 8.0$ Hz and $^3J_{\text{HN}\alpha} < 6.0$ Hz, respectively.

ments involving proline residues were completed by identifying the NOEs between δ -protons of prolines and amide protons from the preceding residues. The NOE connectivities involving sequential and medium range NOEs are summarized in Figure 6-1. The sequential assignments were also confirmed by ^{15}N edited NOESY and TOCSY spectra with ^{15}N labelled sample.

Secondary structure

Figure 6-1 provides the survey of the information obtained from NOE spectrum and $^3J_{\text{HN}\alpha}$ couplings. The successive connection of d_{NN} connectivities and the small $^3J_{\text{HN}\alpha}$ couplings suggest five helical conformations in dh434(0–63). dh434(0–63)

contains the nearly complete sets of $d_{\alpha\beta}(i,i+3)$ $d_{\alpha\beta}(i,i+3)$ indicating the presence of regular helical conformations (Wüthrich, 1984).

6.2. Solution structures of dh434(0–63)

Collection of distance and angle constraints and structure calculation

The cross peaks in the zero-quantum suppressed proton homonuclear NOESY spectrum of dh434(0–63) recorded with a mixing time of 40 msec in aqueous solution were assigned and integrated using the program XESAY (Bartels *et al.*, 1995). The peak intensities were transformed into the distance constraints using CALIBA (Güntert *et al.*, 1991). In order to obtain the maximum number of distance constraints, several rounds of structure calculations were performed with the program DYANA (Güntert *et al.*, 1991; Güntert *et al.*, 1997) and ASNO (Güntert *et al.*, 1993). In total 1236 peaks were assigned and integrated in the NOESY spectra. The total of 1026 NOE distance constraints composed of 560 long range NOEs, 94 medium range NOEs, 78 sequential and 294 intra-residues NOEs, resulted after elimination of the irrelevant constrains (Figure 6-2). The distribution of distance constraints from NOEs is quite uniform but relatively few NOEs detected for the loop between helix III and IV and the terminus. The intra-residual and sequential NOE distance constraints combined with the ${}^3J_{\text{HNH}\alpha}$, ${}^3J_{\alpha\beta}$ and ${}^3J_{\text{N}\beta}$ coupling constrains yielded a total 135 dihedral angle constraints when analyzed with the program HABAS (Güntert *et al.*, 1991). Individual assignments were obtained for 37 pairs of diastereotopic substitutes, using program HABAS (Güntert *et al.*, 1991) and GLOMSA (Güntert *et al.*, 1991). Structure calculation was performed with program DYANA starting from 50 random structures by using a standard annealing protocol with torsion angle dynamics step = 8000 (Güntert *et al.*, 1997). Energy minimization using the AMBER all-atom force field (Weiner *et al.*, 1986) in water bath was applied to the final 20 DYANA conformers with the program OPAL (Luginbühl *et al.*, 1996), which includes pseudoenergy terms for distance constraints and dihedral angle constraints (Widmer *et al.*, 1989; Billeter *et al.*, 1990a). An overview of the best 20 structures from DYANA and after minimization of the DYANA conformers by program OPAL are listed in Table 6-1. The mean energy of these 20 conformers could be reduced by about 900 kcal/mol through energy-minimization

with the program OPAL (Luginbühl *et al.*, 1996).

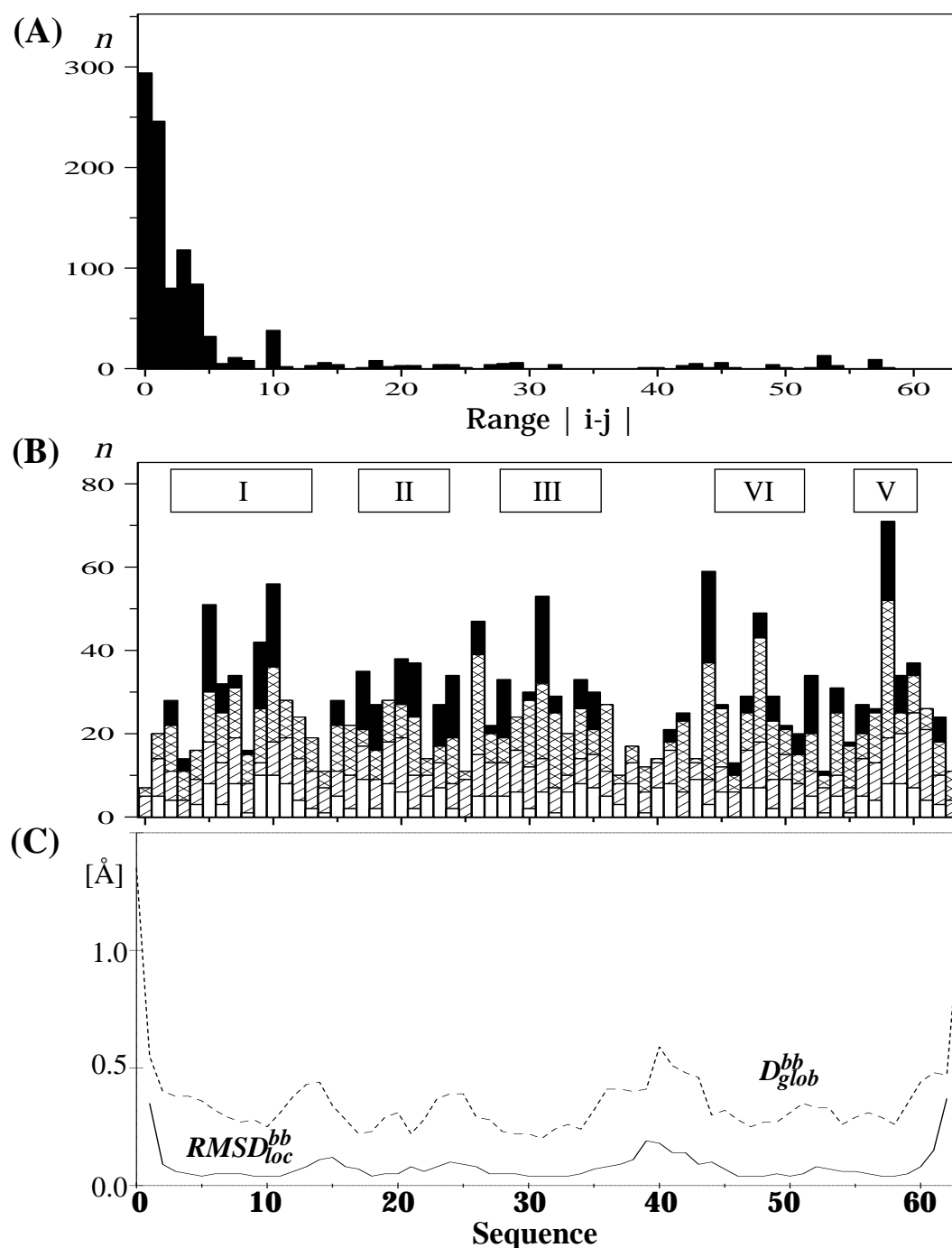


Figure 6-2: (A) Plot of the number of NOE distance constraints, n , versus their range along the amino acid sequence. (B) Plot of the number of NOE distance constraints per residue, n , versus the amino acid sequence of dh434(0-63). The constraints are specified as follows: filled, intra-residual; crosshatched, constraints between protons in sequentially neighbouring residues; vertically hatched, constraints between protons located in residues separated by 2 to 5 positions along the sequence; open, all longer-range constraints. At the top, the locations of the helices in dh434(0-63) obtained from the 20 NMR conformers are indicated. (C) Plots of the global displacement and local r.m.s.d. for the backbone-heavy atoms, C^α , N and C'.

Table 6-1. Analysis of the 20 best conformers of dh434(0–63) obtained from DYANA calculations before and after restrained minimization with the program OPAL.

Parameter	DYANA ^{a,c}	OPAL ^{b,c}
DYANA target function (Å ²) ^d	0.57 ± 0.15 (0.33 .. 0.82)	
AMBER energy (kcal/mol)	-1951 ± 58 (-2043 .. -1866)	-2852 ± 22 (-2529 .. -2333)
Van der Waals energy	439 ± 57 (339 .. 547)	-215 ± 6 (-229 .. -203)
Electrostatic energy	-2906 ± 31 (-2968 .. 2854)	-3099 ± 36 (-3151 .. -3029)
residual NOE distance constraint violations (Å)		
Number > 0.1 Å	5.00 ± 3.05 (2.00 .. 12.00)	0.20 ± 0.40 (0.00 .. 1.00)
Sum	4.71 ± 0.69 (3.68 .. 6.06)	9.49 ± 0.45 (8.73 .. 10.33)
Maximum	0.20 ± 0.07 (0.12 .. 0.40)	0.10 ± 0.01 (0.09 .. 0.11)
residual dihedral angle constraint violations (deg.)		
Number > 2.5°	0.60 ± 0.73 (0.00 .. 2.00)	0.35 ± 0.48 (0.00 .. 1.00)
Sum	9.56 ± 2.10 (6.27 .. 13.26)	31.4 ± 4.98 (22.4 .. 43.5)
Maximum	2.42 ± 0.74 (1.15 .. 3.98)	2.36 ± 0.29 (1.90 .. 2.78)

^a DYANA conformers before energy minimization.

^b DYANA conformers after restrained energy refinement with the program OPAL.

^c The numbers given are the average standard deviation calculated for the group of 20 DYANA conformers with the minimum and maximum values for the individual conformers given in parentheses.

^d The values for the DYANA target function are given with respect to all distance constraints and all dihedral angle constraints.

A ramachandran plot of the backbone dihedral angle ψ and ϕ of all the residues in the 20 energy minimized DYANA conformers of dh434(0–63) shows that almost all residues except for glycine residues indicated by crosses are in the energetically allowed regions of the ψ - ϕ plane (the region shadowed in Figure 6-3). A few residues apparently outside of this region are in the loop and the terminus, which are not well-defined and have less distance constraints as compared with the other segments of the structure of dh434(0–63) (Figure 6-2).

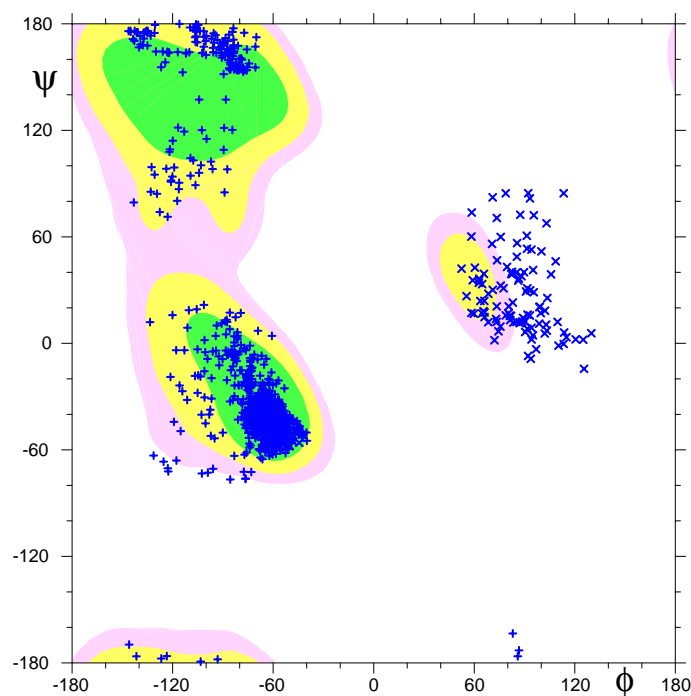


Figure 6-3: Ramachandran plot of all the 20 energy-minimized DYANA conformers of dh434(0–63). +: non-glycyl residues; ×: glycylic residues.

NMR structures of dh434(0–63)

In Figure 6-5A the backbone conformations of the final 20 energy-refined structures of dh434(0–63) are displayed after the superposition for all backbone heavy atoms which affords a visual impression of the structures and of the high quality of the structure determination. Five helices with residues 2 to 13 (helix I), 17 to 24 (helix II), 28 to 36 (helix III), 45 to 52 (helix IV), and 56 to 61 (helix V) are readily identified. A slightly disordered loop connecting helix III and helix IV can be identified as well. This global architecture of dh434(0–63) is essentially identical to the molecular architectures found in the structures of 434(1–69) and 434(1–63), solved by both X-ray and NMR despite of the 15 substitutions of the side chains (Aggarwal *et al.*, 1988.; Mondragon *et al.*, 1989; Neri *et al.*, 1992a; Pervushin *et al.*, 1996; see Figure 6-14). The r.m.s.d. values calculated for the backbone atoms and all heavy atoms of dh434(0–63) are 0.39 Å and 0.83 Å, respectively indicating the high-quality NMR structures (Table 6-2). Plots of the local backbone r.m.s.d. values and the global backbone displacement *versus* the sequence further indicates all the well-defined helices and the increase of local disorder particularly for the terminus and the loop connecting the helices III and IV (Figure 6-2C). This is also observable in the plot of

Table 6-2. RMSD values calculated for different atom selections in the solution structures of dh434(0–63).

Atoms used for comparison*	RMSD (Å) \pm standard deviation
Backbone atoms N, C $^{\alpha}$ and C' (1–63)	0.39 \pm 0.06
All heavy atoms (1–63)	0.83 \pm 0.06
Backbone atoms (17–35) [#]	0.26 \pm 0.04

* *bb* stands for the backbone heavy atoms N, C $^{\alpha}$ and C'. The numbers in parentheses indicate the residues for which the RMSD was calculated.

** Averages are given of the pairwise RMSD values between each of the 20 energy-refined DYANA conformers and the mean solution structure. RMSD values were calculated with the program MOLMOL (Koradi et al., 1995). # the helix-turn-helix motif.

the dihedral angle χ^1 , ψ , and ϕ . (Figure 6-5). The relaxation data of dh434(0–63) described in Figure 6-16 supports that these disorder in r.m.s.d. could be relevant to the conformational features rather than lack of the constraints in this region. A plot of the dihedral angle χ^1 , ψ , and ϕ observed in the 20 energy-minimized DYANA conformers of dh434(0–63) (Figure 6-5) indicates that the helices are defined in a regular helical conformation and several side chain conformations are also well defined.

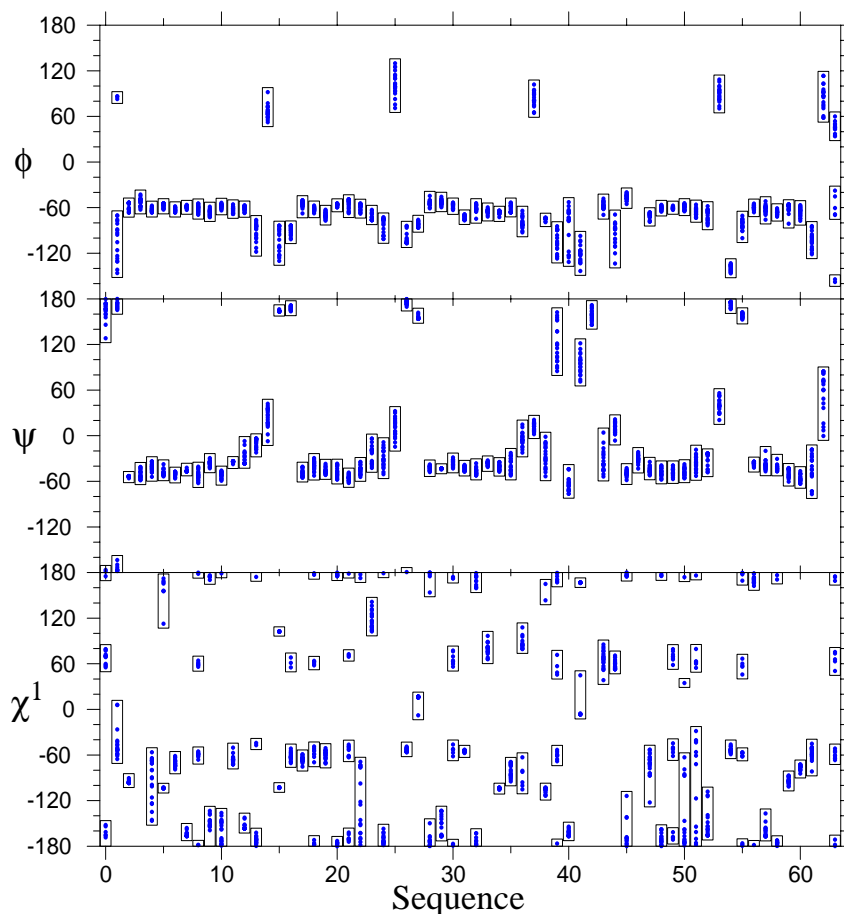


Figure 6-4: Plots of the dihedral angles ϕ , ψ , and χ^1 in the 20 NMR conformers of dh434(0-63) versus the amino acid sequence. Values for each conformer of the 20 energy minimized DYANA conformers of dh434(0-63) are represented by dots, and the ranges of these values are indicated by a bar.

Figure 6-5B affords the side chain conformations of dh434(0-63). Despite of 15 substitutions in the amino acid sequence all the substituted side chains are nicely accommodated to form the global fold of dh434(0-63), which is almost identical to that of 434(1-63) although slight rearrangements of side chain conformations have occurred as described below. The residues in the helix-turn-helix motif that make contacts with DNA bases in the major groove and the backbone of DNA such as Gln17, Gln28, Gln29, Gln33 remain unchanged in dh434(0-63). The structure of the helix-turn-helix motif is well defined and essentially identical to that of 434(1-63)(Table 6-2). Hence, from the viewpoint of the three dimensional structure one can expect that dh434(0-63) might bind the operator sites with a necessary oligomerization domains for the DNA binding.

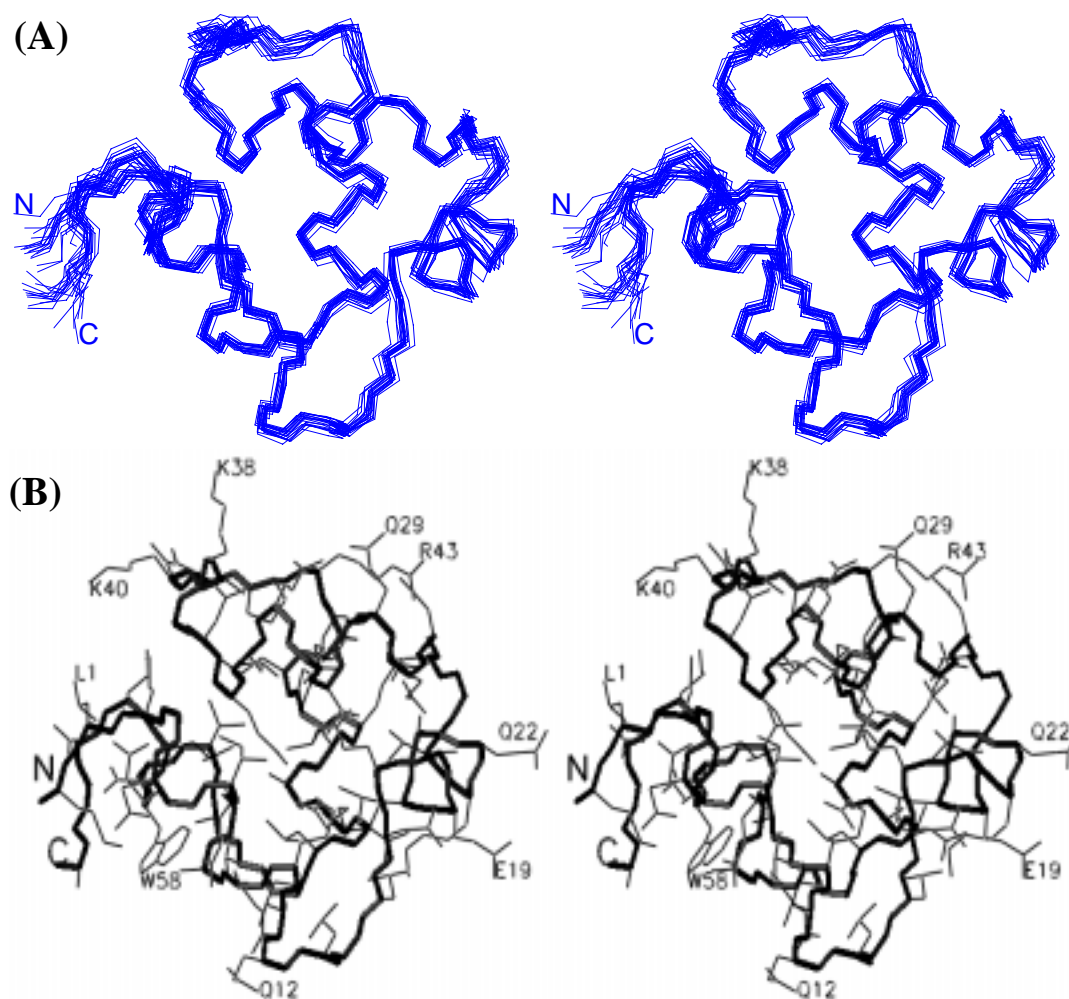


Figure 6-5: (A) Stereoview of the superposition of the backbone-heavy atoms of the energy-minimized 20 DYANA conformers of dh434(0–63). N and C termini are indicated. (B) Stereoview of one of the 20 energy-minimized NMR conformers of dh434(0–63) with side chain heavy atoms. The thick lines indicate the backbone-heavy atoms.

6.3. Comparison between dh434(0–63), 434(1–63), and P22c2(1–76)

The most interesting points to be closely analyzed are the structural changes upon these substitutions and the structural differences that had occurred when a chimeric sequence was produced and how much plasticity proteins have in order to accommodate side chains coming from another homologous protein. The superposition of the backbone atoms of the mean structures of P22c2(1–76), 434(1–63), and dh434(0–63) is shown in Figure 6-6 providing an impression of their similarities in the backbone conformations. Notable differences in the loop between the helix III and IV can be readily recognized between P22c2(1–76) and 434(1–63). This is apparently due to the insertion of one residue in P22c2(1–76), which also appears

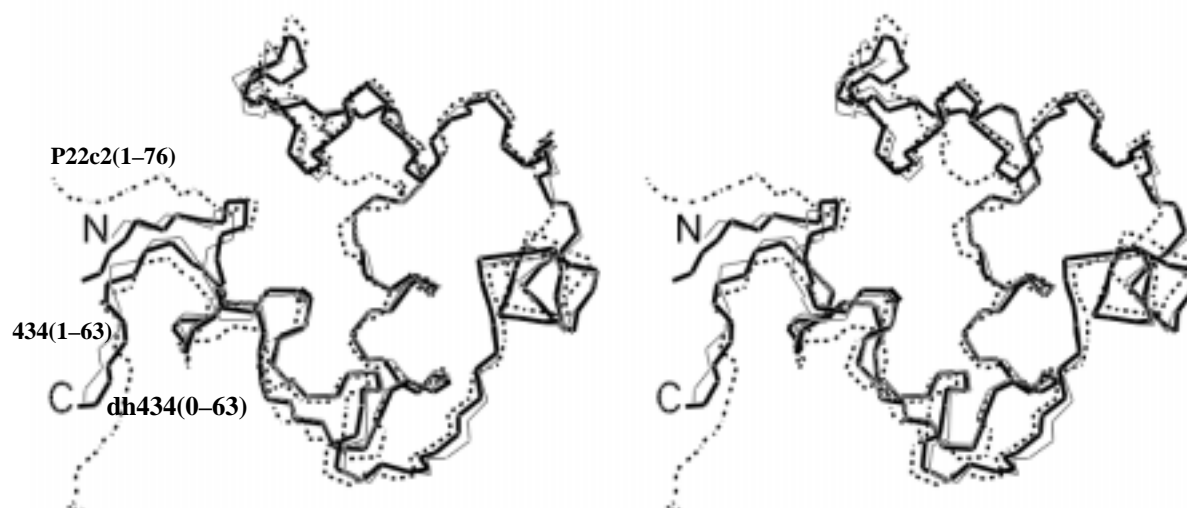


Figure 6-6: Stereoview of the backbone-heavy atoms of the three mean structures from 434(1-63), dh434(0-63), and P22c2(1-76). C^α atoms of the residues 1-63 of dh434(0-63) and 434(1-63), C^α atoms of residues 5-46, 48-68 of P22c2(1-76) were superimposed. The dotted line represents the mean structure from P22c2(1-76) (Sevilla-Sierra *et al.*, 1994). The solid thick line presents the mean structure of dh434(0-63). The mean structure of 434(1-63) is presented by the solid thin line (Pervushin *et al.*, 1996). N and C denote the N and C terminus, respectively.

as an increase of the r.m.s.d. for P22c2(1-76) (Table 6-3). Otherwise the r.m.s.d. among them are very small (Table 6-3). In particular the r.m.s.d. values with respect to 434(1-63) are comparable to the r.m.s.d. values of 0.61 Å among the 20 NMR conformers of 434(1-63) (Pervushin *et al.*, 1996), indicating that there is no difference between 434(1-63) and dh434(0-63) in the backbone conformation beyond the precision of the NMR structures.

Table 6-3: RMSD values for the superposition of three structures of 434(1-63), P22c2(1-76), and dh434(0-63).

Structures	range	dh434(0-63) RMSD (Å)**
<434(1-63)>	1-63	0.62
<434(1-63)>	17-35 [#]	0.37
<P22c2(1-76)>	5-46,48-68	1.30
<P22c2(1-76)>	21-39 [#]	0.50

**RMSDs were calculated for the backbone heavy atoms N, C^α and C' between the mean structures. 434(1-63) is the structures solved by NMR(1r63) (Pervushin et al., 1996).

P22c2(1-76) is the structures determined by NMR (1adr) (Sevilla-Sierra et al., 1994)

[#] the helix-turn-helix motif.

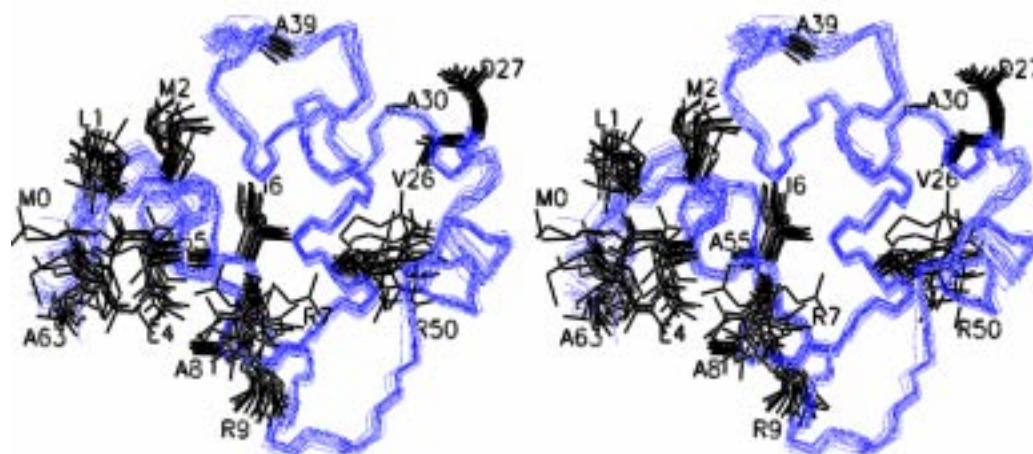


Figure 6-7: Stereoview of the superposition of the backbone atoms of the 20 energy minimized DYANA conformers of dh434(0-63) with the side chain heavy atoms of the replaced residues. The labels indicate the amino acid types and residue numbers.

Figure 6-7 shows the side chain heavy atoms of the replace amino acids and indicates the spatial locations of the amino acid substitutions in the three dimensional structures of dh434(0-63). The side chain conformations of buried residues or residues having contacts with the other parts of the proteins among the substituted side chains are of special interest to be investigated in order to detect the effects by the replacements of the side chains on the three dimensional structures as compared with 434(1-63). In Figure 6-8 the local structures of the substituted residues are

closely compared with the crystal structure of 434(1–69). The first 9 residues segment of dh434(0–63) contains the sequence of the corresponding part in the structure of P22c2(1–76) because dh434(0–63) was derived from the chimeric protein P22/434(0–63) (see Chapter 4). The backbone atoms of the residues 1–9 are superimposed in Figure 6-8A. In this segment several changes of the side chains conformation as compared with the structure of 434(1–69) can be observable. The side chains of Met2, Arg5, and Arg9 are slightly shifted as compared with the side chain conformation of 434(1–69). In particular the side chain of Arg9 is shifted and involved in the hydrogen bond with the backbone oxygen atom of Leu52, or Gly53 connecting the two sub-domains of dh434(0–63), i.e. helix I–III and the helix IV–V (Figure 6-8; Figure 6-10). Probably because of this shift of Arg9 the side chain of Arg5 is also moved and involved in the hydrogen bonds connecting the same two sub-domains (Figure 6-10). Among the other replaced residues the remarkable change is the side chain conformation of Val26. The hydroxyl group of Thr26 in the crystal structure of 434(1–69) forms a hydrogen bond with H^N of Thr27. In dh434(0–63) Thr26 was replaced with a valine and the side chain of Val26 was rotated by about 120° with respect to the χ^1 angle of Thr26 in 434(1–69) to accommodate the side chain of Val26 (Figure 6-8B). This conformational change leads to enhance van der Waals interactions with the other hydrophobic residues (Figure 6-11; Table 6-5). This replacement concludes that the side chain of Thr26 can be replaced with Val without causing any problems although this side chain in 434(1–63) is forming a hydrogen bond with the amide group of Thr27. In fact P22c2(1–76) contains Val at the corresponding position (Figure 4-5; Appendix H). This position is located in the helix-turn-helix motif which recognizes operator sites of the DNA. Therefore it might have an important role for the fine turning of the conformation of the helix-turn-helix motif although it is not tested whether dh434(0–63) is still able to bind to operator sites.

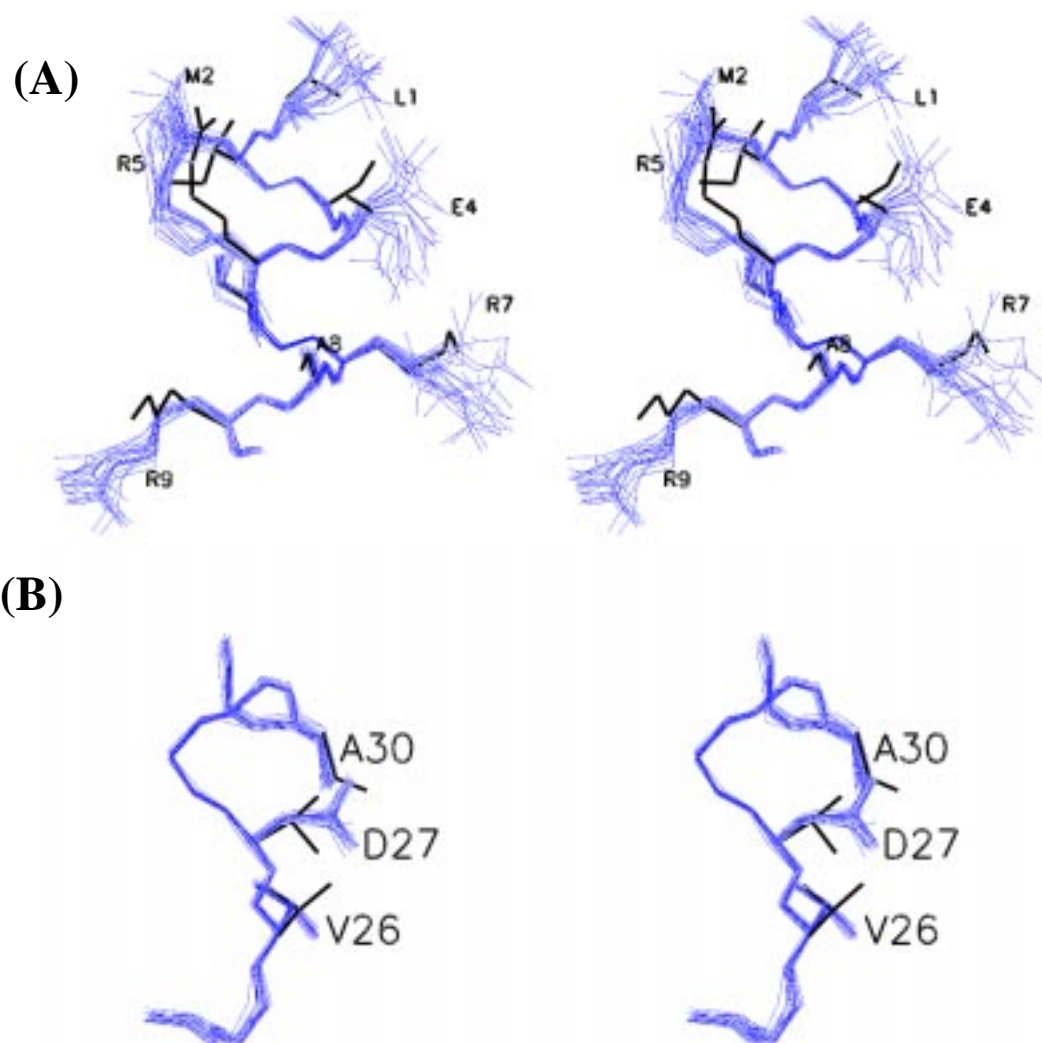


Figure 6-8: Comparison of the side chain conformations between NMR conformers of dh434(0–63) and the crystal structure 434(1–69). (A) Stereoview of the residues 1–9 of dh434(0–63) and the crystal structure of 434(1–69) with side chain heavy atoms. (B) Stereoview of the residues 24–31 with the side chain heavy atoms from residue 26, 27 and 30. The side chain heavy atoms from the crystal structure are shown by thick lines.

Hydrogen bond

Hydrogen bond is one of the most important interactions to keep the native three dimensional structure of proteins and for protein stability. In Table 6-4 the hydrogen bonds found in dh434(0–63) are summarized and compared with that of 434(1–63) and the crystal structure of 434(1–69). Among the substituted side chains in dh434(0–63) the side chains of Arg9 and Asp27 are found to be involved in new hydrogen bonds in the 20 NMR conformers of dh434(0–63). The side chain of Asp27 forms a hydrogen bond with its amide group in 11 structures out of the 20 NMR structures of dh434(0–63). In the three-dimensional structures of 434(1–63) and

434(1–69), the side chains of Ser4, Thr27, and Ser55 are forming N-caps (Mondragon *et al.*, 1989; Neri *et al.*, 1992a; Pervushin *et al.* 1996). Ser4, Thr27, and Ser55 are replaced in dh434(0–63) with Glu, Asp, and Ala, respectively. Since Glu and Asp are preferable amino acid types for N-capping (Richardson & Richardson, 1988), it was intended to keep the N-cap between Ser30 and Thr27 by the substitution to Asp in dh434(0–63). In the three-dimensional structure of dh434(0–63) only one structure out of 20 NMR structure of dh434(0–63) is found to contain this N-cap but the pH titration of dh434(0–63) indicates that the amide group of Ala30 and Asp27 were titrated with the same pKa supporting the presence of this N-cap. In dh434(0–63) Ser55 was replaced with Ala in dh434(0–63). The removal of the hydroxyl group of Ser55 should cause the loss of the hydrogen bond with the backbone amide of Trp58 (Figure 6-13). This effect could not be detectable in the 20 NMR structures of dh434(0–63) because of the short side chain but vividly observable in the chemical shift change of H^N of Trp58 which shifted to the up-field by 1.2 ppm (Figure 6-12). Moreover this effect could be detectable in H/D exchange rates as well. As shown in Figure 6-15 the most of the buried residues in the helices decreased the proton exchange rates of the amide groups. In contrast in the helix V the exchange rates did not decrease but slightly increased. This is possibly due to the loss of the N-cap in helix V. The three dimensional structure of dh434(0–63) can be divided into two sub-domains in which one sub-domain consists of the helices I–II and the other consists of the helices VI–V connected by a flexible linker (Figure 6-10). The side chain of Arg9 is found to be involved in several hydrogen bonds connecting the two sub-domains of dh434(0–63) (Table 6-4; Figure 6-10). The HE(NE)HGHH experiment can provide the correlations between H^e and H^g and the correlations between H^e and H^h through the heteronuclear three bonds $^3J_{NH}$ couplings (Figure 6-9B). It can also correlate between H^e atom and the H^δ due to the two bonds heteronuclear $^2J_{NH}$ couplings which are usually smaller than $^3J_{NH}$ for flexible arginine side chains

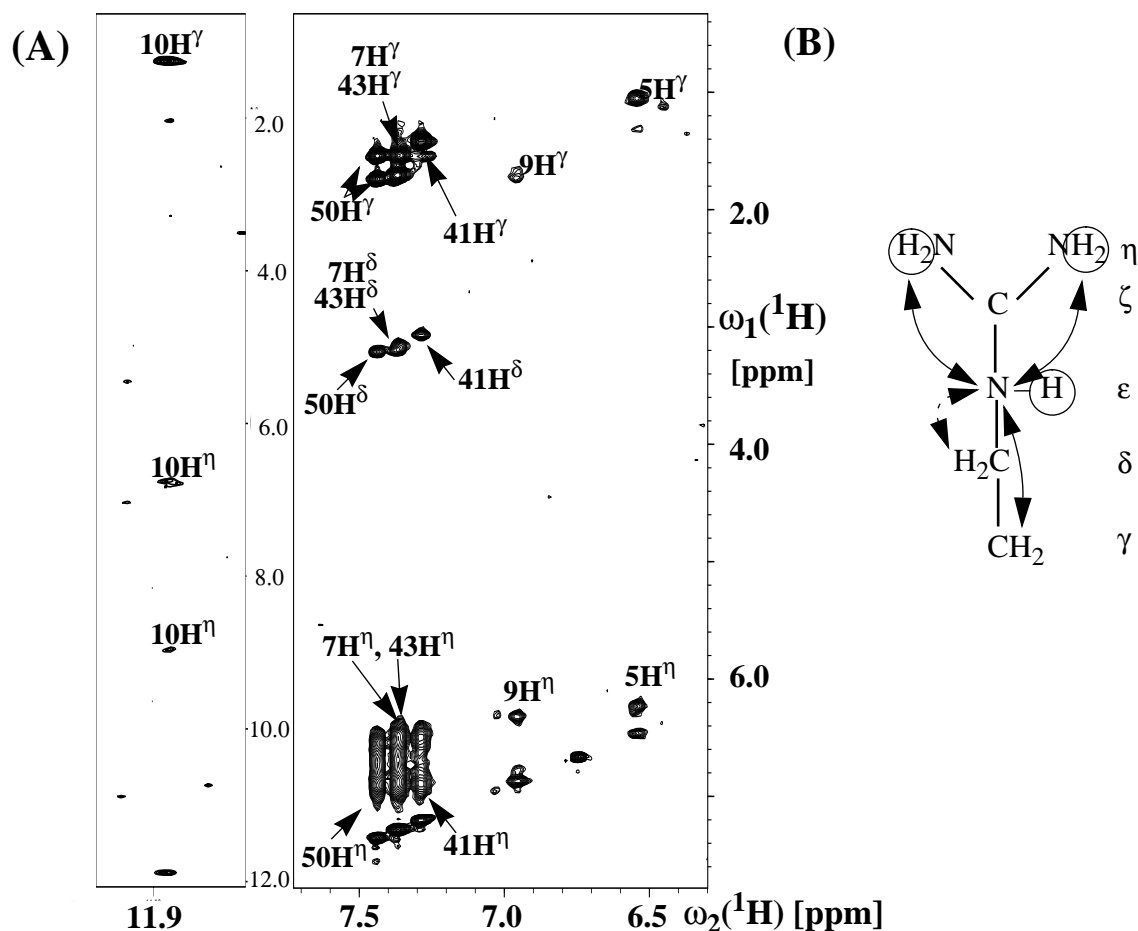


Figure 6-9: (A) Spectra of 2D HE(NE)HGHH experiment (Pellecchia et al., 1997) of 5 mM uniformly ^{15}N labeled dh434(0–63) at pH 4.8 and 13°C. (B) Schematic drawing of the scheme of the magnetization transfers in HE(NE)HGHH experiment.

presumably because of longer relaxation times. Figure 6-9A presents the spectrum of the HE(NE)HGHH experiment measured with dh434(0–63) indicating the different pattern of H^η protons for Arg5, Arg9, and Arg10 from the other Arg. In the structure of dh434(0–63) Arg10 is involved in a buried salt-bridge with Glu35 which results in the down-field shift of H^ϵ protons to 11.8 ppm and clearly shows correlations between two H^η and H^ϵ (Appendix F). For Arg5 and Arg9 the small up-field shifts of H^ϵ protons and a separate H^η peak were observed. This is very different from the other Arg on the surface such as Arg7, Arg41, Arg43, and Arg51 which show broad two peaks for H^η (Figure 6-9). In the 20 NMR structures of dh434(0–63) the hydrogen bonds between the H^η and H^ϵ of Arg5 and Leu59, and between the H^η

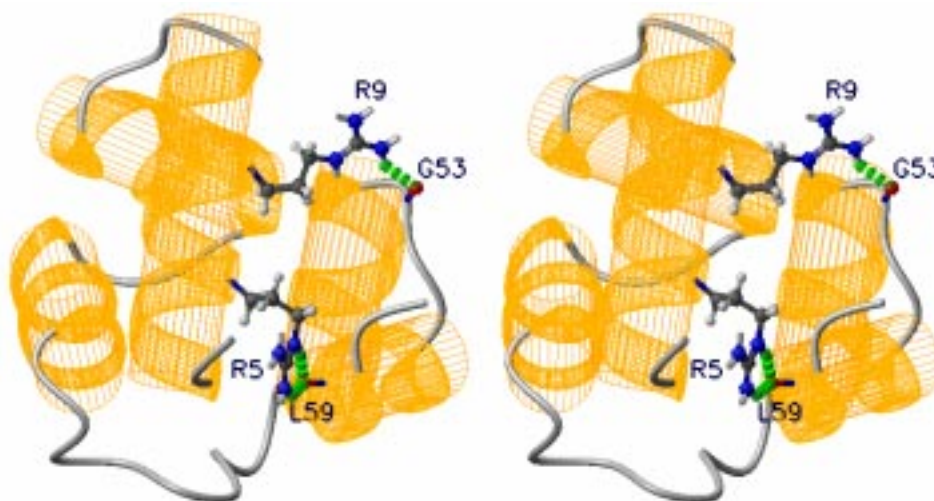


Figure 6-10: Schematic drawing of the hydrogen bonds connecting the two sub-domains of dh434(0–63). The backbone atoms are represented by ribbon drawing and cylinders. The side chains of Arg5 and Arg9 are shown by the ball and stick models. The hydrogen bonds, which connect the two sub-domain, are indicated by the dotted lines coloured in green.

and H^e of Arg9 and Leu52 or Gly53 were found (Table 6-4 and Figure 6-10). Difference in the spectrum of the HE(NE)HGHH experiment suggests that the dynamics of these side chains has dramatically changed due to these hydrogen bonds (Pellecchia *et al.*, 1996; Yamazaki *et al.*, 1995). The disappearance of the two broad peaks of H^n for Arg5, Arg9 and Arg10 suggests that the rotations around C-N bonds have been reduced. This is in good agreement with the observation of hydrogen bonds for Arg5, Arg9, and Arg10 in the three-dimensional structure of dh434(0–63). This additional hydrogen bond from Arg9 could be a factor for hyper-stability of dh434(0–63) (Mrabet *et al.*, 1992; Borders *et al.*, 1994; Chapter 5).

Table 6-4. Hydrogen bonds identified in the 20 energy-minimized DYANA conformers of dh434(0–63) and in the structures of 434(1–63) and the structure of 434(1–69) solved by x-ray.

Donor		Acceptor		dh434(0–63)	434(1–63)	434(1–69)
Arg5	HN	Leu1	O'	20	6	1
Arg5	He	Leu59	O'	15	4	1
Arg5	Hη	Leu59	O'	12	0	1
Ile6	HN	Met2	O'	20	17	1
Arg7	HN	Gly3	O'	20	19	1
Ala8	HN	Glu4	O'	20	20	1
Arg9	HN	Arg5	O'	20	20	1
Arg9	He	Leu 52	O'	10	-	-
Arg9	Hη	Leu 52	O'	10	-	-
Arg9	Hη	Gly 53	O'	10	-	-
Arg10	Hη	Glu 35	εO1	14	10	1
Arg10	Hη	Glu 35	εO2	17	15	1
Arg10	Hη	Leu 15	O'	10	3	1
Ile11	HN	Arg7	O'	20	18	1
Gln12	HN	Ala8	O'	20	19	1
Leu13	HN	Arg9	O'	11	20	1
Leu13	HN	Arg10	O'	10	0	0
Gly14	HN	Ile11	O'	15	3	1
Leu15	HN	Arg10	O'	12	10	1
Gln17	He2	Glu35	Oε	15	5	1
Gln17	He2	Glu32	Oε	14	7	0
Leu20	HN	Asn16	O'	20	18	1
Ala21	HN	Gln17	O'	20	20	1
Gln22	HN	Ala18	O'	20	20	1
Lys23	HN	Glu19	O'	17	17	1
Lys23	HZ	Ala51	O'	13	2	0
Val24	HN	Leu20	O'	12	19	1
Gly25	HN	Ala21	O'	19	14	1
Val26	HN	Ala21	O'	19	17	1
Asp27	HN	Asp27	Oδ	11	-	-
Ile31	HN	Gln28	O'	11	1	0
Glu32	HN	Gln28	O'	19	20	1
Gln33	HN	Gln29	O'	20	17	1
Leu34	HN	Ala30	O'	12	15	1
Glu35	HN	Ile31	O'	20	20	1
Asn36	HN	Glu32	O'	20	20	0
Gly37	HN	Leu34	O'	17	16	0
Lys38	HN	Gln33	O'	20	13	1
Leu45	HN	Pro42	O'	14	8	0
Glu47	HN	Phe44	O'	16	0	0
Leu48	HN	Phe44	O'	20	20	1
Ala49	HN	Leu45	O'	20	20	1
Arg50	HN	Pro46	O'	20	20	1
Ala51	HN	Glu47	O'	20	20	1
Leu52	HN	Leu48	O'	20	20	1
Gly53	HN	Ala49	O'	19	9	0
Val54	HN	Ala49	O'	20	20	1
Leu59	HN	Ala55	O'	20	17	0
Leu60	HN	Ala56	O'	20	20	1
Asn61	HN	Asp57	O'	20	20	1
Asn61	Hδ2	Asp57	Oδ	15	3	0

Hydrogen bonds are listed if the donor proton involved forms a hydrogen bond in at least 10 of the 20 energy-minimized DYANA conformers, or if a hydrogen bond was identified in the crystal structure (Mondragon et al., 1989). The number of DYANA conformer after energy minimization by OPAL is listed in which the hydrogen bond was identified. The criteria used for the identification are that the proton-acceptor distance must be less than 2.4Å and the angle between the donor-proton bond and the line connecting the acceptor and donor atoms is less than 35°. This criterion is sufficiently loose to detect bifurcated hydrogen bonds.

van der Waals interaction

Hydrophobic residues such as valines and alanines are newly introduced into 434(1–63). In particular alanine was used as a first choice for the substitution of serine residues. Therefore the introduction of hydrophobic residues may lead to the changes in the van der Waals contacts between the hydrophobic residues. In Table 6-5 the observed van der Waals contacts involving the newly introduced or replaced hydrophobic residues are compared with those of 434(1–63). Apparently several new van der Waals contacts involving Val26 emerged as compared with 434(1–63). The side chain of Val26 interacts with the side chains of Val24, Ala21, Ala30 and Phe44 in dh434(0–63). The large up-field shift to 0.05 ppm of the methyl group of Val26 was observed (Appendix F). This up-field shift can be attributed to the ring current effect of Phe44 located in the vicinity of the methyl group of Val26 providing an evidence for the interaction between the aromatic ring of Phe44 and the methyl group of Val26. A contact between Ala30 and Phe44 has also newly emerged as compared with 434(1–63). The side chain of Leu34 has contacts with the newly introduced Ala39 and also located in the proximity of Phe44. These substitutions from Ser30 to Ala30 and from Thr39 to Ala39 resulted in the reduction of the solvent accessible area of each residues by 10% (Table 6-6). In total these contacts from the newly introduced hydrophobic residues enhance the hydrophobic core between two sub-domains in the structure of dh434(0–63) (Figure 6-11). These enhancements of the van der Waals contacts in the interior of dh434(0–63) could contribute to the hyper-stability of dh434(0–63) as compared with 434(1–63) (Chapter 5).

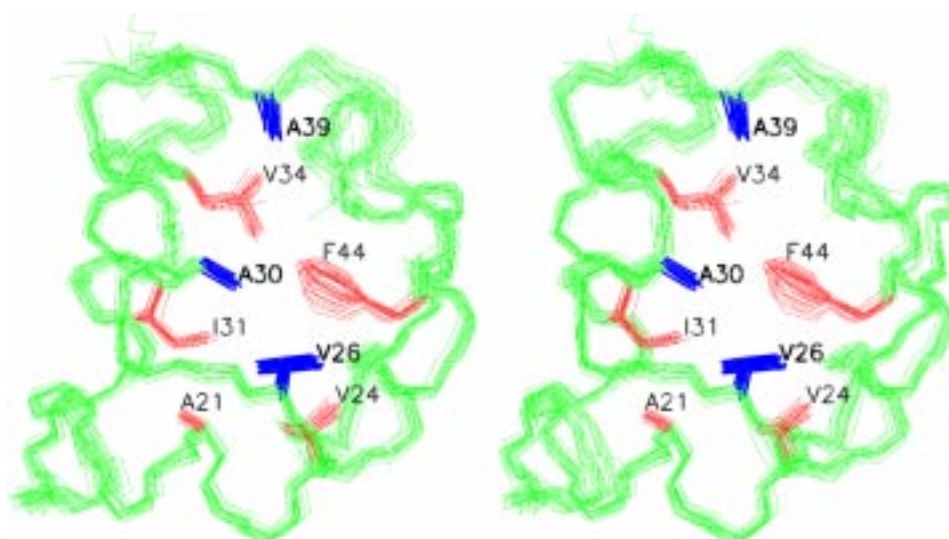


Figure 6-11: Stereoview of the backbone atoms of dh434(0–63) with the side chains involved in the newly introduced van der Waals contacts. The side chain heavy atoms from the substituted residues are indicated by thick lines. The others are shown by thin lines.

Table 6-5: Analysis of newly introduced van der Waals contacts for hydrophobic residues in dh434(0–63)

Residues & Atoms			Number of structures	
			dh434(0–63)	434(1–63)
Ile6	C γ^1	L59 C δ^2	18	17
	C δ^1	L34 C δ^2	20	5
Val26	C γ^2	A21 C β	20	3
	C γ^1	V24 C γ	11	18
	C β	A30 C β	13	–
	C γ^1	F44 C β	10	–
	C γ^1	F44 C γ	20	–
	C γ^1	F44 C δ	18	–
	C γ^1	F44 C ϵ	13	–
	C γ^2	I31 C δ^1	20	0
Ala30	C β	F44 C ζ	17	–
	C β	F44 C ϵ	13	–
Ala39	C β	L34 C γ	16	–
	C β	L34 C δ^2	15	–
Ala63	C β	W58 C δ^1	11	–

The new van der Waals contacts involved in the substituted residues are listed when more than 10 structures out of the 20 energy-minimized DYANA conformers have contacts with carbons distance < 4.0 Å. For 434(1–63), the coordinates from pdb-entry 1r63 was used.

Table 6-6. Summary of solvent-accessible surface area at the substituted residues

Residues	Solvent accessibility(%)	
	dh434(0–63)	434(1–63)
Leu1	47.6	51.8 (Ser)
Met2	11.8	17.1 (Ile)
Gly3	4.3	15.8 (Ser)
Glu4	33.9	30.0 (Ser)
Ile6	1.6	0.7 (Val)
Arg7	34.9	32.9 (Lys)
Ala8	25.0	27.3 (Ser)
Arg9	16.6	15.3 (Lys)
Val26	11.9	15.1 (Thr)
Asp27	35.6	37.4 (Thr)
Ala30	3.1	12.1 (Ser)
Ala39	10.5	20.8 (Ser)
Arg50	57.1	40.9 (Ser)
Ala55	28.6	27.7 (Ser)
Ala63	58.1	59.5 (Thr)

The average solvent-accessible surface area was calculated by using MOLMOL (Richmond, 1984; Koradi et al., 1995).

Comparison of the chemical shifts between dh434(0–63) and 434(1–63)

Chemical shifts are usually very sensitive to changes in structures and dynamics of proteins. The chemical shifts of H^α and H^N from dh434(0–63) and 434(1–63) are compared in Figure 6-12. A large down-field shift of 1.46 ppm was observed for H^N of Gly3. This down-field shift could be from the effect by forming a hydrogen bond but there is no new hydrogen bond between the other side chains and H^N of Gly3 in the 20 NMR structure of dh434(0–63). Therefore it is unclear and intriguing what cause this down-field shift of H^N of Gly3. It could be due to a hydrogen bond to a water molecule with amide group of Gly3 since the amide group of Gly3 indicate positive ROE and negative NOE (Figure 3-7; Figure 3-8). H^N of Trp58 also shows a large up-field shift of 1.4 ppm even though there is no substitution of amino acids in the neighboring residues. However, the H^N of Trp58 is involved in a hydrogen bond with Ser55 in 434(1–63) (Figure 6-13). Therefore this up-field shift can be attributed

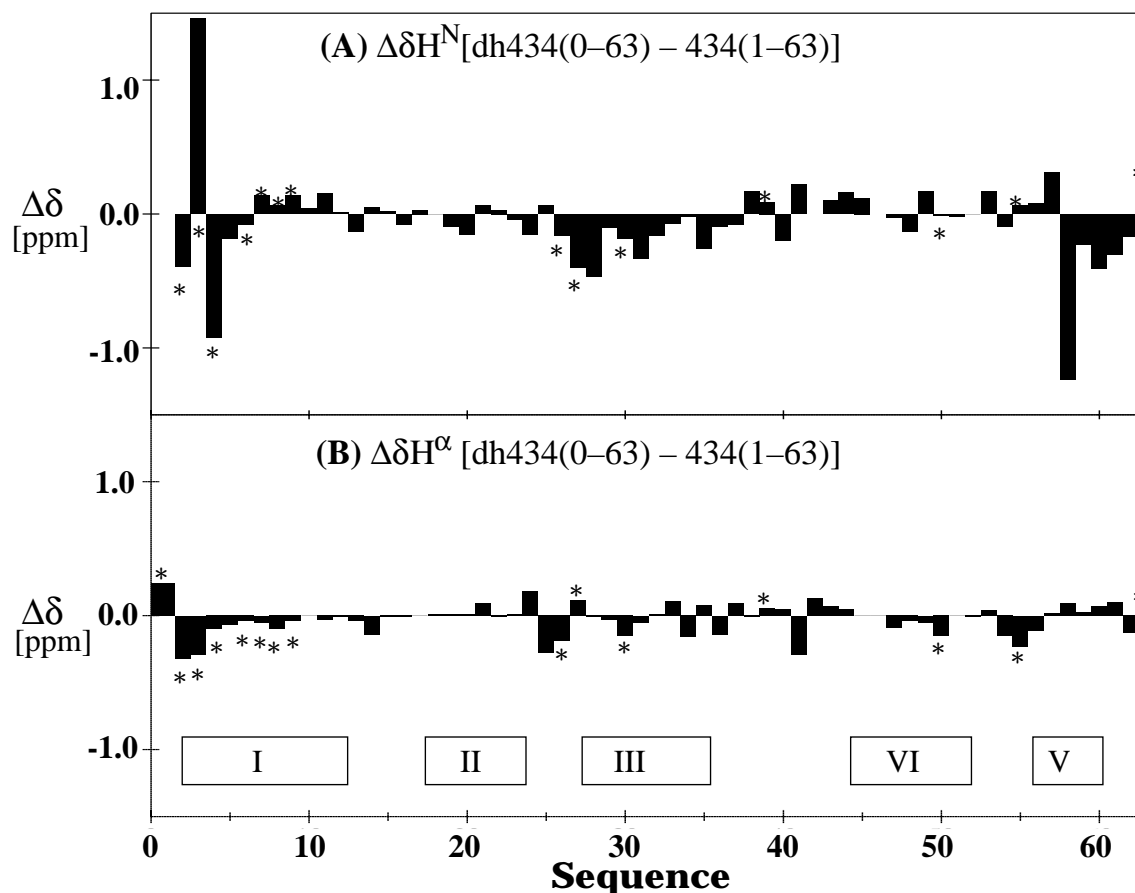


Figure 6-12: Comparison of the chemical shifts between dh434(0-63) and 434(1-63) at pH 4.8 and 13°C. (A) Backbone amide protons. (B) Alpha protons. At the bottom, the helical segments are indicated. The positions of the amino acid substitutions are indicated by asterisks.

to the removal of the hydroxyl group of Ser55 which is forming N-cap in the structure of 434(1-69). Interestingly, relatively large up-field shifts are also observed for the H^N protons of the helix V such as Leu59, Leu60, and Asn61 which are not located in a close distance from any replaced residues (Figure 6-7). This observation suggests that the removal of the N-capping residue affects the chemical shifts of the entire helix V. This can also be supported by the slightly faster H/D exchange rates in the helix V. The N-cap of helix V has influenced on the entire helix V in a cooperative manner. Compared with the chemical shift of H^N protons, H^α chemical shift changes are not significant, which is assumed to be because of no large changes in the secondary structures because H^α protons are usually sensitive to the backbone conformations (Wishart *et al.*, 1991). This is in agreement with the three dimensional structure of dh434(0-63) calculated from the distance and angle constraints obtained from the other NMR parameters. Large differences of H^α chemical shifts are

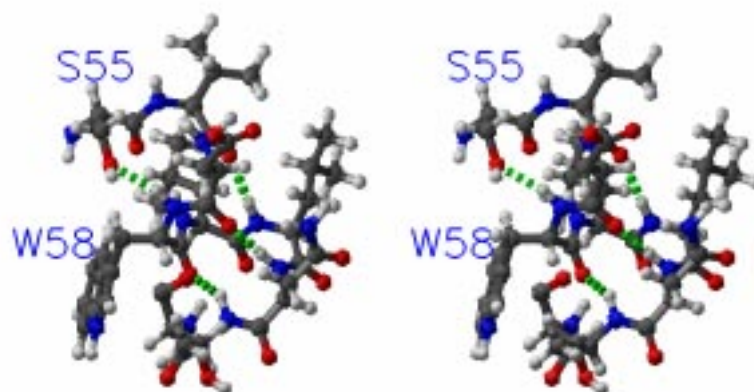


Figure 6-13: Stereoview of a segment of residues 55–63 from the crystal structure of dh434(1–69). The dotted lines indicate hydrogen bonds observed in the crystal structure of dh434(1–69) in this segment.

observed mainly for the substituted residues because of the effects of the amino acid types rather than conformational changes. However, comparably large changes are also observed for several residues in the turn and the loop (Figure 6-12).

6.4. ^{15}N relaxation measurements of dh434(0–63)

Nuclear spin relaxation measurements of the backbone ^{15}N atoms in aqueous solution with uniformly ^{15}N -enriched proteins has become a popular approach for the characterization of global protein motions and internal backbone mobility (Peng & Wagner, 1994). It is desirable to obtain experimentally the rotational correlation time of dh434(0–63) for the estimation of the residence time of water molecules based on the diffusion model proposed by Ayant *et al.*, 1978 and Otting *et al.*, 1991a. In this section the correlation time of dh434(0–63) was experimentally determined by measuring the relaxation time of the backbone ^{15}N atoms.

Sample preparation & NMR measurements

Uniformly ^{15}N -labeled dh434(0–63) was expressed and purified as described in Chapter 4. The NMR sample for the relaxation measurements contained 5 mM protein in 25 mM K-phosphate at pH 4.8 in a mixed solvent of 90% H_2O /10% D_2O .

All measurements were performed at 13°C, using Varian Unity-plus 400 spectrometers operating at ^1H frequencies of 400 MHz with a triple resonance probe head equipped with Z gradient coil. The pulse scheme used for measurements of

$T_1(^{15}\text{N})$ has been adapted from Farrow *et al.* (1994) by replacing the H^{N} -selective 180° proton pulse in the carrier position for the suppression of the cross-polarization. For the relaxation times, T_1 , a series of 11 spectra was measured with the following relaxation delays: $t_{rel}=1, 6, 16, 36, 76, 146, 196, 296, 396, 496,$ and 596 msec. The T_2 measurements was performed using CPMG-type sequence with sensitivity enhancement scheme. A series of 10 spectra was recorded with the following the relaxation delays: $t_{rel} = 15.8, 31.6, 47.5, 63.3, 79.1, 110.8, 142.4, 189.9, 237.4,$ and 316.5 msec. $^{15}\text{N}\{^1\text{H}\}$ steady-state heteronuclear Overhauser effects were determined as described by Farrow *et al.* (1994). The total measurements time for the two spectra was 48 hours at ^1H frequency of 400 MHz. NOE values were determined as the ratio of corresponding cross-peaks volumes measured from the spectra acquired with and without ^1H saturation during the recycle delay. All data sets were processed using the program PROSA (Güntert *et al.*, 1992), and peak picking and volume integration were performed with the program XEASY (Bartels *et al.*, 1995).

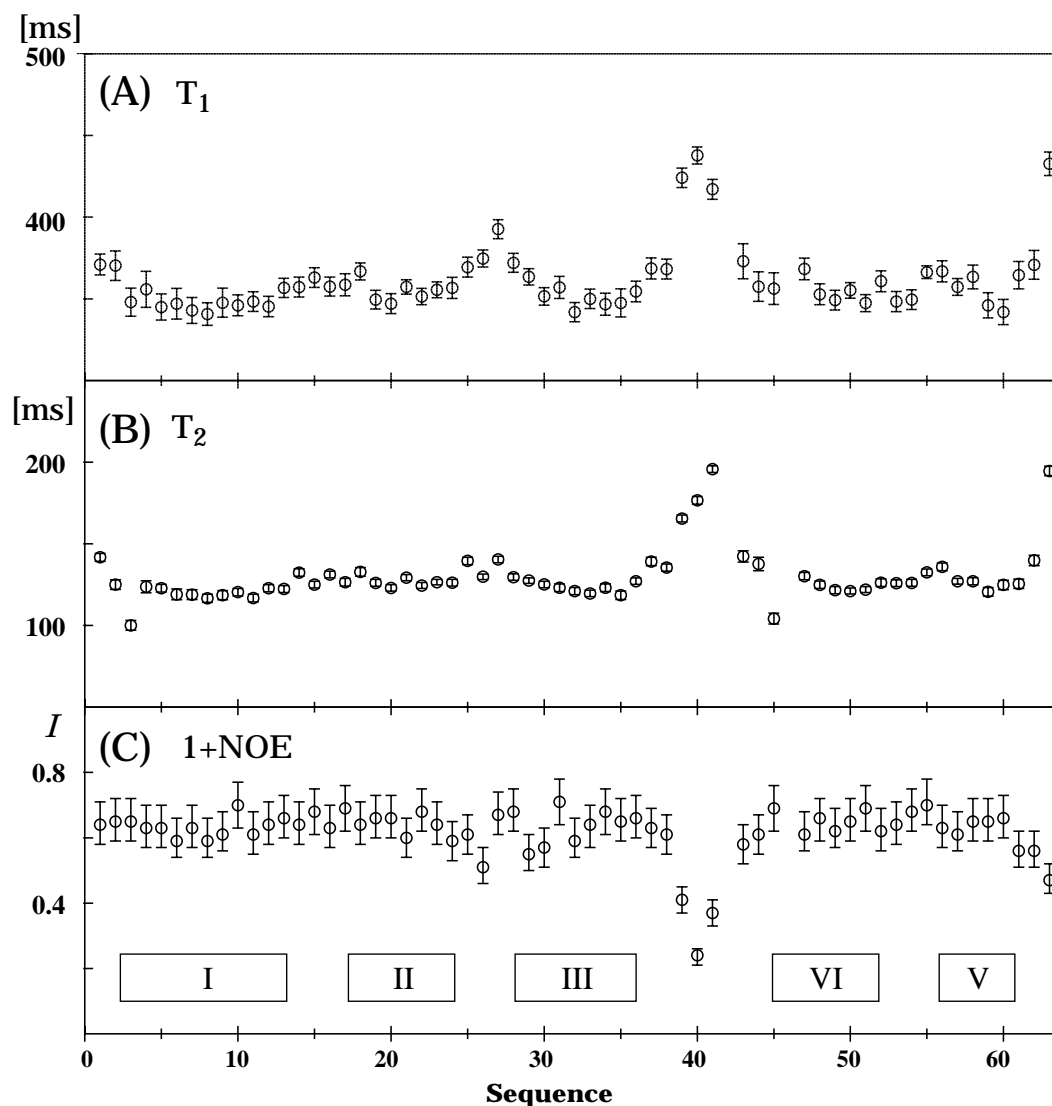


Figure 6-14: Plots of ^{15}N relaxation data obtained from a sample of 5mM uniformly ^{15}N labelled dh434(0–63) containing 25 mM K-phosphate buffer at pH 4.8 and 13.0°C measured at a ^1H frequency of 400MHz. (A) T_1 relaxation (B) T_2 relaxation (C) $^{15}\text{N}\{^1\text{H}\}$ -NOE, versus the amino acid sequence. At the bottom the positions of the helices are indicated.

Experimental T_1 , T_2 and NOE data & Global correlation times from a spherical top model

The T_1 and T_2 relaxation rates of individual ^{15}N atoms were determined by nonlinear least squares fitting of the magnetization decay curves to a single exponential function (Figure 6-14). $\langle T_1/T_2 \rangle = 2.77 \pm 0.07$ was found at the ^1H frequency of 400 MHz when disregarding the residues for which T_1/T_2 is outside of the standard deviation. This results in an apparent global correlation time $\tau_c = 6.2 \pm 0.2$ nsec for dh434(0–63). Figure 6-14 indicates the increases of T_1 and T_2 in the loop region

and the C-terminus of dh434(0–63). This is a very similar feature to the relaxation data obtained from 434(1–63) in which the increase in T_1 and T_2 was attributed to the increased local mobility indicating that not only three dimensional structures but also dynamical feature of dh434(0–63) is very similar to that of wild-type 434(1–63) (Luginbühl *et al.*, 1997).

6.5. H/D exchange rates of dh434(0–63) and 434(1–63)

Information about the structure and conformational dynamics of proteins is contained in hydrogen exchange behavior of amide groups. H/D exchange rates of amide groups have been used for the characterization of hydrogen bonds, protein folding process, protein stability, and protein-protein and protein-ligand interactions. In this section the characterization of dh434(0–63) by measuring H/D exchange rates of dh434(0–63) is described.

Sample preparation & NMR measurements

Each of 1 mM solution of the uniformly ^{15}N -labeled dh434(0–63) and the uniformly ^{15}N labeled 434(1–63) were used for the measurements of H/D exchange rates. The sample dissolved in H_2O at pH 4.8 in 25 mM K-phosphate buffer was lyophilized. H/D exchange rates of the slowly exchanging amide protons of dh434(0–63) were obtained from a series of 2D- ^{15}N , ^1H -COSY (Bodenhausen & Ruben, 1980) spectra recorded immediately after the fully protonated ^{15}N labelled lyophilized dh434(0–63) dissolved in D_2O . All measurements were performed at 13°C , using Varian Unity-plus 400 spectrometer. All data sets were processed using the program PROSA (Güntert *et al.*, 1992), and peak picking and volume integration were performed with the program XEASY (Bartels *et al.*, 1995). The exchange rates were determined by least square fitting the cross-peak intensities to an exponential decay function.

Experimental H/D exchange rates

Plots of the H/D exchange rates of dh434(0–63) and 434(1–63) were presented *versus* the amino acid sequence in the Figure 6-15 together with a plot of the solvent accessible surface of dh434(0–63). Overall, slow amide proton exchange rates are co-

incident with the positions of the helices. The slower amide proton exchange rates for dh434(0–63) as compared with 434(1–63) correlate well with the higher precision of the determined NMR structures of dh434(0–63). The significant differences in the exchange rates between dh434(0–63) and 434(1–63) can be readily recognized in the helix I. This could be attributed to the effects originating from the differences in the amino acid sequences due to the neighboring effects (Bai *et al.*, 1993). Amide protons of the exposed residues in the helix II, III and IV show little difference in the H/D exchange rates between dh434(0–63) and 434(1–63) unless there is any side chain substitutions at the positions (Figure 6-15). In contrast the exchange rates of the residues with smaller solvent accessible surface areas are largely reduced by more than factor of one (Figure 6-12). This observation suggests that the opening rate of the protein was largely reduced. This is in good agreement with the data of the stability of dh434(0–63) (Chapter 5). Whereas the exchange rates in the helix V of dh434(0–63) are not slower than that of 434(1–63). Even slightly faster exchange rates could be observed for some residues in this helix although there is no amino acid substitution introduced in the helix V. This can be explained based on the fact that the N-cap formed by Ser55 in 434(1–63) was removed (Figure 6-13). The effect of the removal of N-cap was observed in the $^1\text{H}^{\text{N}}$ chemical shifts in the helix V as well (Figure 6-12). This observation may support the local destabilization of the helix V. Overall the H/D exchange rates of amide groups of dh434(0–63) indicates a more compact structure than that of 434(1–63). This strongly supports that dh434(0–63) does not have a “molten globule” structure that was often found in *de novo* designed proteins, or artificial proteins (Shang *et al.*, 1994; Betz *et al.*, 1993; Betz *et al.*, 1995).

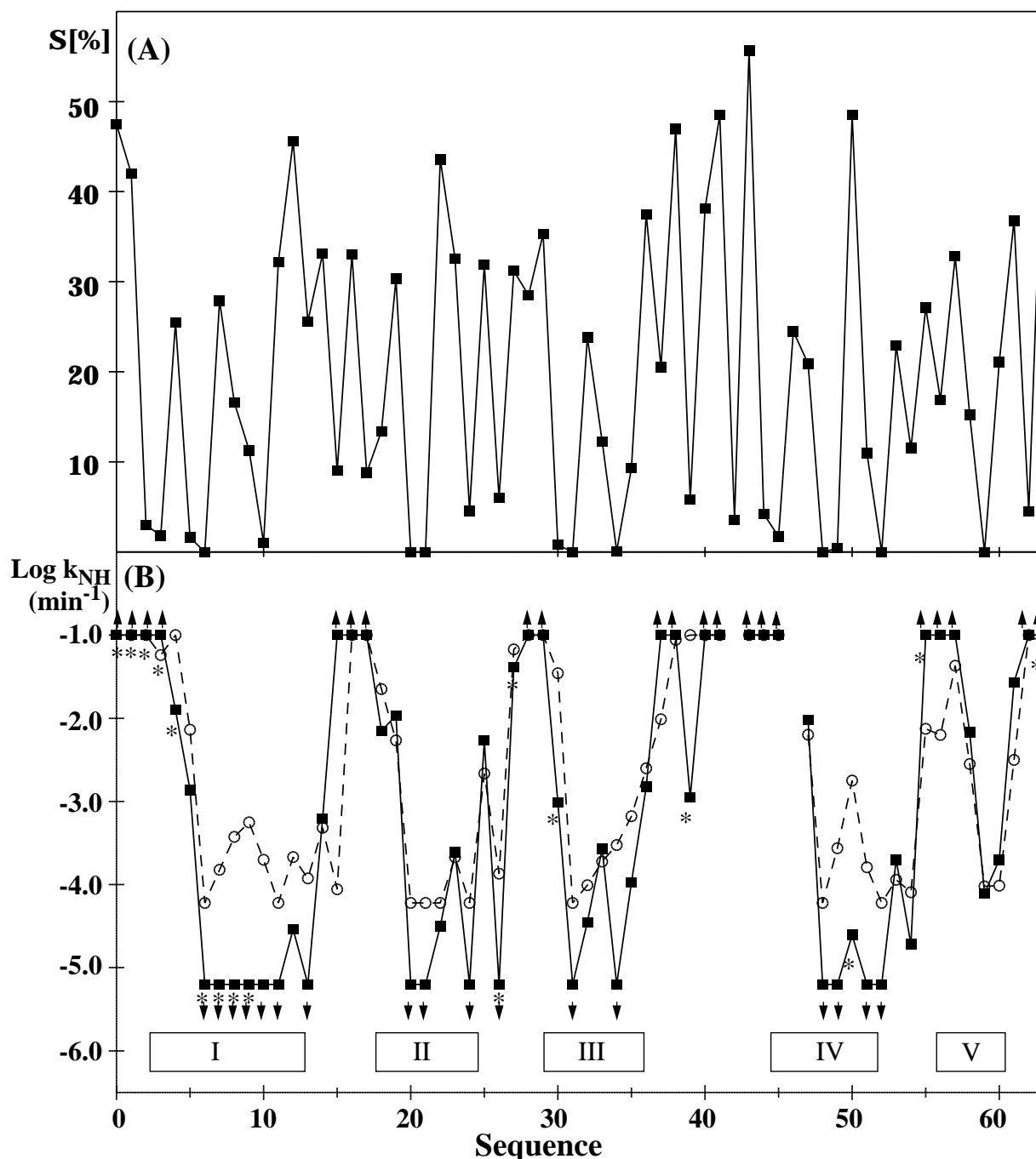


Figure 6-15: (A) Plot of the average solvent-accessible surface area of dh434(0–63) (Richmond et al. 1984). (B) Plots of the rates of exchange with solvent $^2\text{H}_2\text{O}$ for the backbone amide protons of dh434(0–63) (solid lines and filled squares) and 434(1–63) (dotted lines and open circles) at pH 4.8 and 13.0 °C. At the bottom, the locations of the helices in the NMR structure are indicated. Asterisks indicate the positions of the amino acid substitutions as compared with 434(1–63).

6.6. ^{13}C and ^{15}N resonance assignments of dh434(0–63)

$^{13}\text{C}/^{15}\text{N}$ doubly labeled samples are very useful to provide additional frequencies to resolve the overlapping in ^1H frequency. It has become a powerful approach

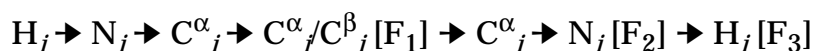
to resolve the overlapping using isotope-labelled samples to determine three dimensional structures by NMR. It has also been used for the identification of water molecules by separating the overlapping using ^{13}C and ^{15}N frequencies (Grzesiek & Bax, 1993a,b; Kriwacki, 1993). However, there is only homonuclear experiments which have been used for the identification of the short-lived water molecules. This is presumably because the observation of short-lived water molecules requires the sensitivity hence often used the high concentration (20 mM ~100 mM) of the protein or peptides solution (Otting *et al.*, 1991a; Liepinish *et al.*, 1992; Brühweiler *et al.*, 1995). This high concentration of proteins is very demanding requirement for an isotope-labelled sample from the viewpoint of the cost. The disadvantage with unlabeled samples in the hydration studies is that the separation of the signals is often achieved by the homonuclear 3D experiments such as NOESY-TOCSY experiment in which TOCSY scheme might produce undesired effects disturbing the interpretation of the spectra.

Sample preparation & NMR spectroscopy

$^{13}\text{C}/^{15}\text{N}$ doubly labeled sample of dh434(0–63) was prepared as described in Chapter 4. 1 mM and 5 mM protein solution of dh434(0–63) in 90%/10% $=\text{H}_2\text{O}/\text{D}_2\text{O}$ in 25 mM K-phosphate buffer are used for the measurements. pH of the samples was adjusted to 4.8 with minute amounts of HCl and NaOH. 2D-, 3D- HCCH-COSY (Ikura *et al.*, 1991), HC(CO)NH-TOCSY (Logan *et al.*, 1991), HNCACB (Wittekind *et al.*, 1993) experiments were performed for the resonance assignments of ^{13}C , ^{15}N of dh434(0–63). All experiments were performed at 13°C, using Varian Unity-plus 400, 750, or Bruker DMX 600 spectrometers.

Assignments of C^α , C^β atoms

3D HNCACB experiment can provide the information of the chemical shifts of $\text{C}^\alpha/\text{C}^\beta$ (Wittekind *et al.* 1993). In the HNCACB experiment, the magnetization transfer follows the following steps.



In addition to this magnetization pathway there is a simultaneous magnetization transfer mediated by the two-bond $^2\text{J}(\text{N}_i-\text{C}^\alpha_{i-1})$ coupling as following.

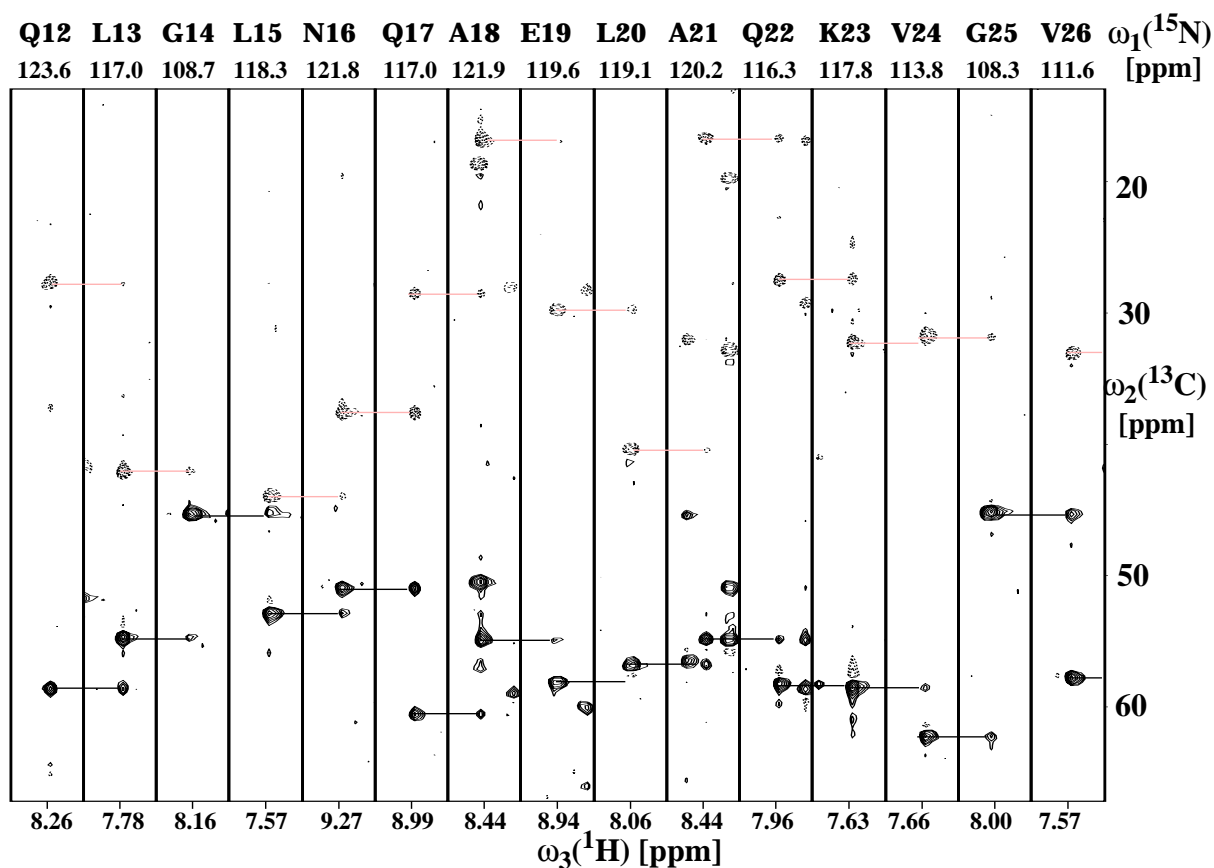
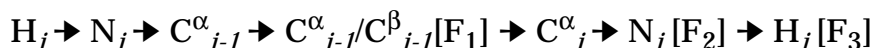


Figure 6-16: Strips from the 3D HNCACB spectrum (Wittekind et al., 1993) of 1mM dh434(0–63) at pH 4.8 and 13°C with 25mM K-phosphate buffer measured on DMX600 illustrating the sequential connectivities for the residues 24–31. The sequential connectivities are shown by solid lines for C^α and by dotted lines for C^β . The positive peaks for C^α atoms and negative peaks for C^β atoms are shown by solid lines and dotted lines, respectively.



These magnetization transfer pathways provide both intra-residual correlation of C^α/C^β , N_i , H_i and the inter-residual correlation of $C^\alpha_{i-1}/C^\beta_{i-1}$, N_i , H_i . In Figure 6-16 the typical strips from the spectrum of HNCACB experiments with dh434(0–63) are indicated. For each strips the stronger positive/negative peaks correspond to the intra-residue correlation between C^α/C^β and the amide ^{15}N , ^1H . The weaker pairs of the peaks correspond to the correlations between the $C^\alpha_{i-1}/C^\beta_{i-1}$ of the preceding residue and the amide ^{15}N , ^1H which is due to the two-bond $^2J(N_i-C^\alpha_{i-1})$ coupling. Because of the different magnetization transfer pathways in the HNCACB experiment the C^α and C^β have the opposite sign of the peaks which can be used to distinguish between C^α and C^β atoms. In this HNCACB experiment all the C^α and C^β chemical shifts were identified except for the preceding residues of proline residues because of the lack of the ^{15}N , ^1H correlations. By this experiment, all the C^α and C^β of Gly,

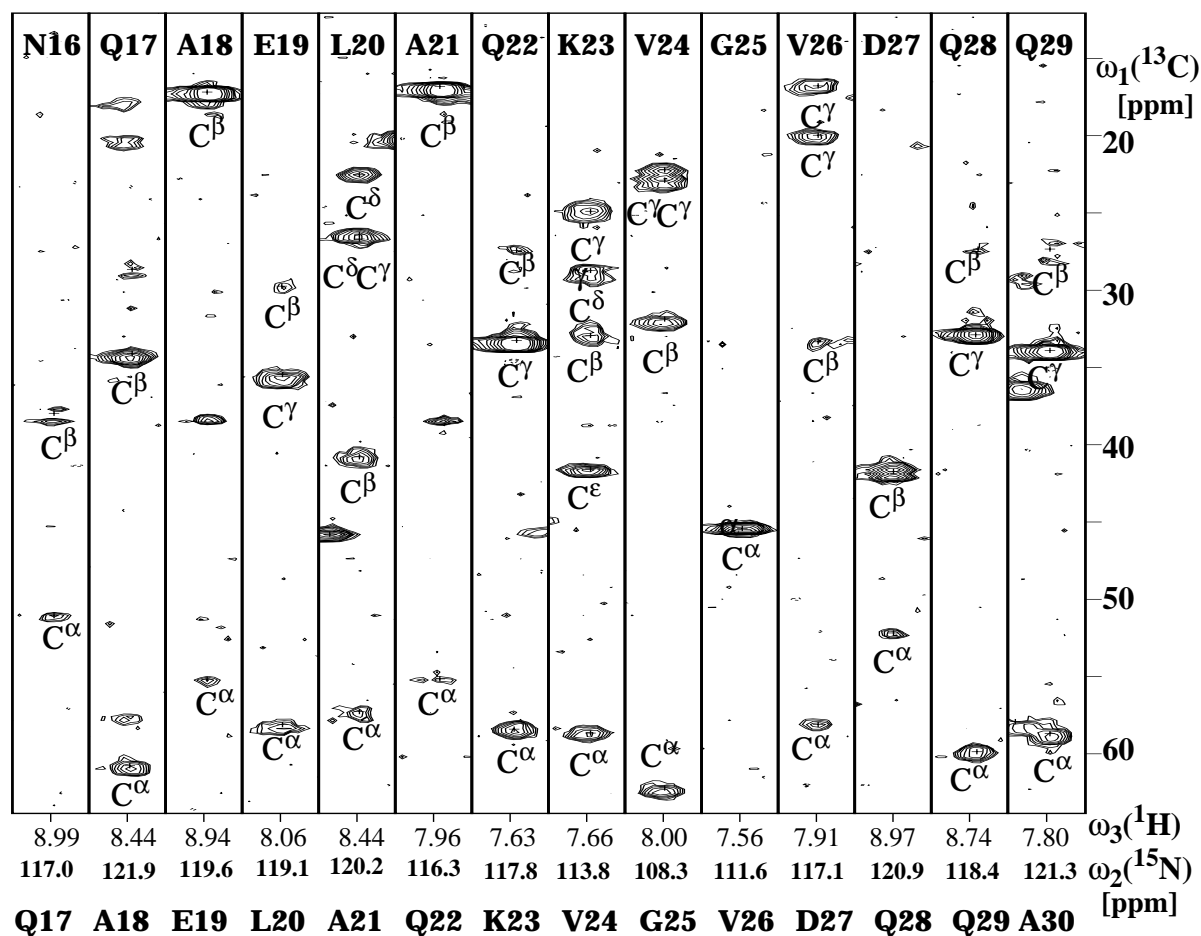
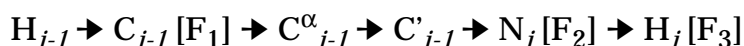


Figure 6-17: Strips from the 3D HC(CO)NH-TOCSY spectrum (Logan *et al.*, 1993) of 5mM ^{13}C , ^{15}N doubly labeled dh434(0–63) at pH 4.8, 13°C measured on UNITY+ 400 at the ^1H frequency of 400 MHz. The strips are illustrating the residues 17–30 containing all the peaks originated from the preceding residues' (labeled at the top) carbon atoms.

Ala, Asn, Asp, Trp, and Phe of dh434(0–63) were identified.

Assignments of ^{13}C atoms of longer side chains

From the HNCACB experiment only the information of $\text{C}^\alpha/\text{C}^\beta$ was obtained. It is necessary to use additional spectra for the assignment of ^{13}C atoms of longer side chains. For this purpose, 3D-HC(CO)NH-TOCSY spectrum was recorded (Logan *et al.*, 1993). In this experiment the magnetization transfer follows the following steps.



This magnetization transfer provides the correlation of all ^1H attached ^{13}C atoms of the preceding residue C_{i-1} , and N_i , H_i from the backbone amide groups. Fig-

ure 6-17 presents the typical strips from the HC(CO)NH-TOCSY spectrum of dh434(0–63).

Each strips contains the information of all the chemical shifts of ^{13}C atoms of the preceding residues except for the aromatic and carbonyl carbons. Combined with the data from the HNCACB experiment, all the C^α , C^β , C^γ atoms of Gln, Glu, Met, and Val were assigned unambiguously. For the other type of amino acids all the ^{13}C chemical shifts of the side chains except for 2 residues from the preceding residues of proline residues were obtained but it was necessary to assign these chemical shifts to the individual carbon atoms. For this purpose additional spectra. 2D- and 3D-HCCH-COSY spectra were recorded on Unity plus 750 or 400. Based on the ^1H assignment obtained from the analysis of ^1H homonuclear experiments, the HCCH-COSY and [^{13}C , ^1H]-HSQC spectra were used for the assignments of the rest of ^{13}C atoms. Since the methyl group of a methionine does not have any C-C bond connection, the assignments of ϵCH_3 in methionines could not be confirmed by the HCCH-COSY spectra. However the assignments of the ^{13}C atoms of the methyl groups were derived from ^{13}C edited ct-HSQC spectra based on the ^1H assignment of methionines obtained from the ^1H homonuclear experiments. dh434(0–63) contains the two aromatic groups, the assignments of the aromatic rings were completed in the same manner with HSQC and HCCH-COSY spectra. For the assignments of the preceding residues of the proline, HCCH-COSY and HCCH-TOCSY spectra were analyzed. The obtained assignments of ^{13}C and ^{15}N atoms are summarized in APPENDIX G, which were used for the analysis of the hydration spectra in Chapter 3. Figure 6-18 presents all the complete assignments of the backbone amide groups of dh434(0–63) except for the N-terminus in the [^{15}N , ^1H]-HSQC spectrum. These backbone assignments were used for ^{15}N relaxation and H/D exchange rates measurements.

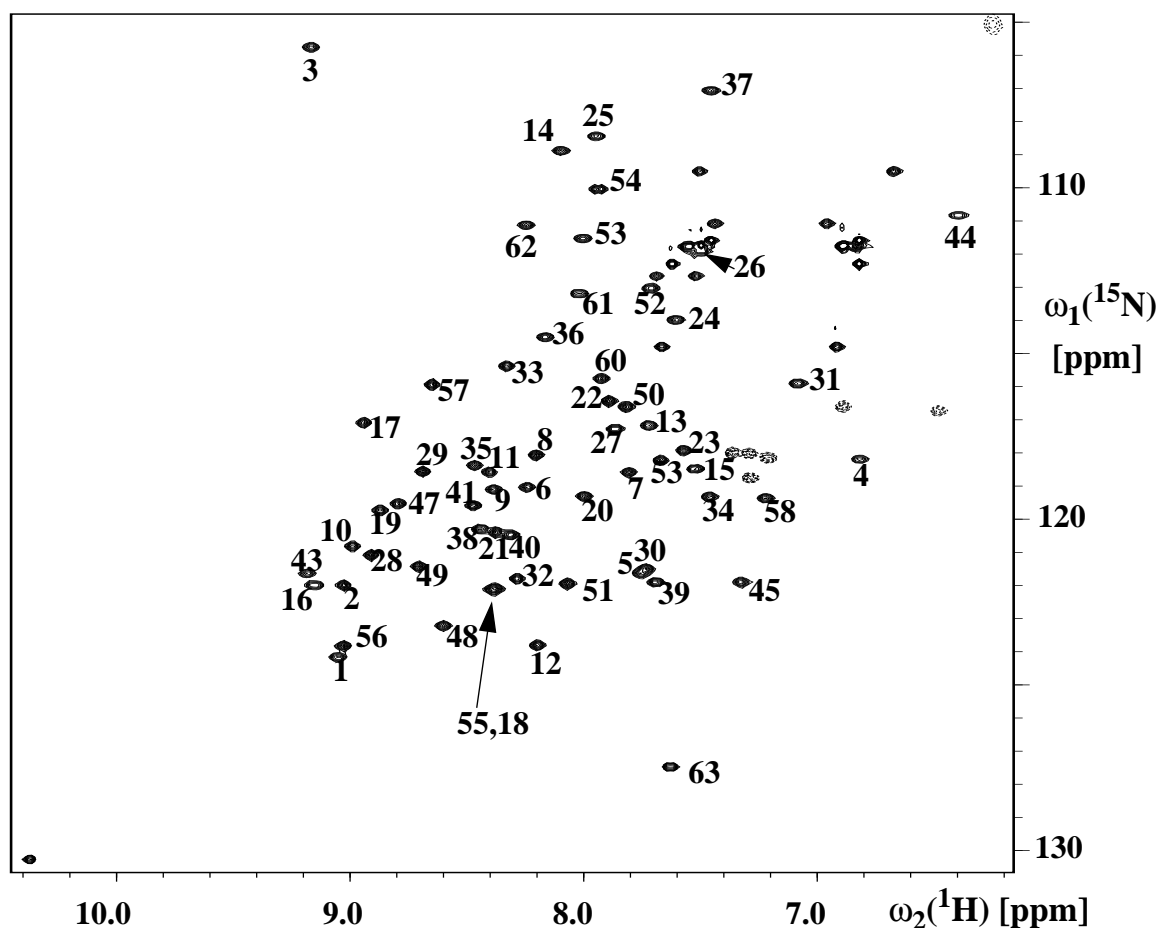


Figure 6-18: $[^{15}\text{N}, ^1\text{H}]$ -HSQC spectrum of 5 mM dh434(0–63) at pH 4.8 and 13°C with the assignments of the backbone amide groups. The spectra was recorded on Unity+ 400. The peaks with dotted lines are the folded peaks from the side chain N^ϵ of Arg residues in dh434(0–63).

Chemical shifts index

Chemical shifts such as H^α , C^α , and C^β atoms are found to correlate with the secondary structures of proteins (Wishart *et al.*, 1991; Spera & Bax, 1991). Such correlations have been used for the identification of the secondary structures (Wishart *et al.*, 1994; Luginbühl *et al.*, 1995). Figure 6-19 presents the chemical shift differences from the values in random coli peptides for H^α , C^α , and C^β , and CSI (Chemical Shift Index proposed by Wishart *et al.*, 1994). The down-field shifts of C^α atoms, the up-field shifts of H^α atoms are in good agreement with the positions of the helices in the NMR structures of dh434(0–63). The tendency of the up-field shifts of C^β atoms also weakly correlates the positions of the helices. These observation provide supports for the correct ^1H , ^{13}C , resonance assignments of dh434(0–63) and the calculated NMR structures.

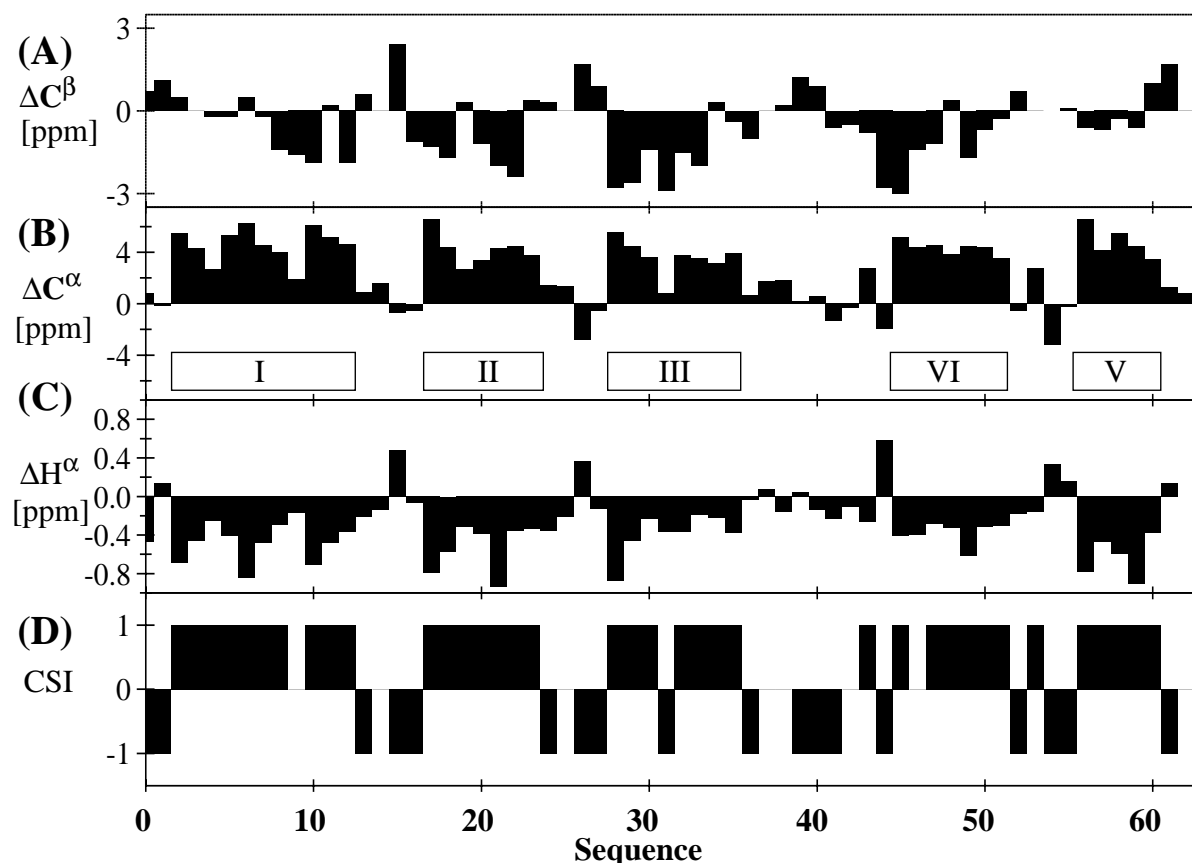


Figure 6-19: Chemical shifts differences between the measured values of dh434(0–63) and the random coil values, (A) $^{13}\text{C}^\beta$, (B) $^{13}\text{C}^\alpha$, and (C) $^1\text{H}^\alpha$, versus the amino acid sequence. The positions of the helices are indicated. (D) Chemical shift index (CSI) (Wishart *et al.*, 1994).

6.7. Stereospecific assignments of methyl groups by fractionally 10% ^{13}C labeling

Assignments of the diastereotopic methyl groups have large impact on the precision of the structures determined by NMR spectroscopy (Güntert *et al.*, 1991a). Biosynthetically fractionally labeling procedure is the utmost reliable method for the assignment of diastereotopic methyl groups in valine and leucine because of the use of the biosynthetic pathway (Senn *et al.*, 1989; Neri *et al.*, 1989). In this section the methyl groups of valines and leucines are assigned by using 10% fractionally labeled samples.

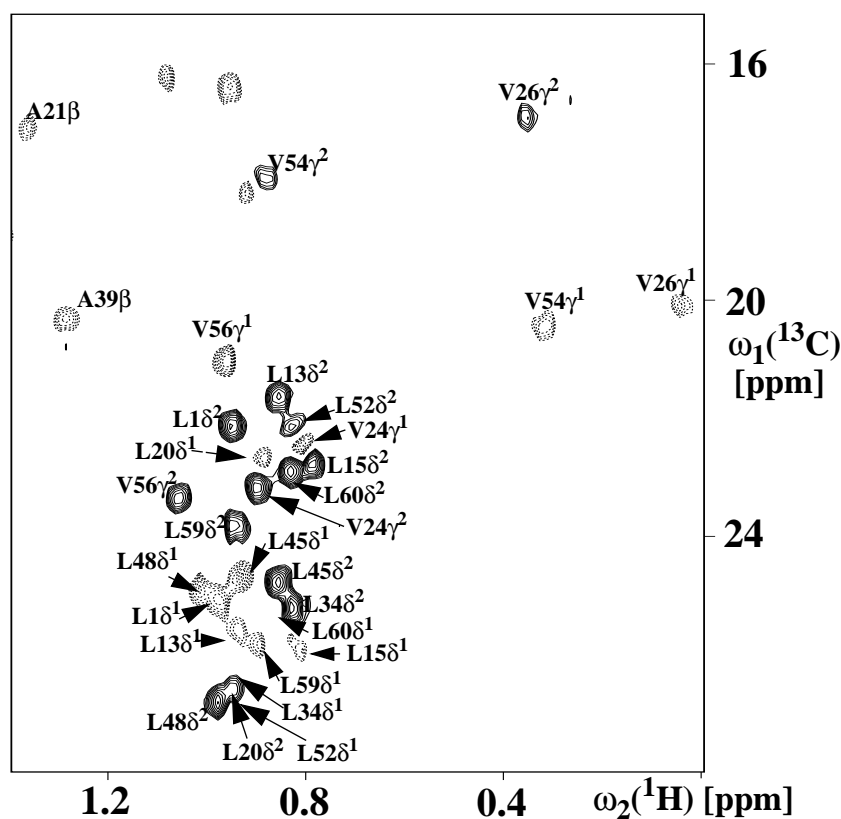


Figure 6-20: Part of *ct*- ^{13}C , ^1H -HSQC spectrum of the 10% fractionally labeled dh434(0–63) for the stereospecific assignments of the methyl groups in leucines and valines. The duration of the constant time was set to 27 msec so that the prochiral pro-(S) methyl groups of leucine and valine were of opposite sign relative to the pro-(R) methyl group. Negative peaks are indicated by dotted lines.

Sample preparation and NMR spectroscopy

10% fractionally ^{13}C labeled sample was prepared from the M9 culture supplemented with 90% /10% = $[\text{D}]-^{12}\text{C}_6- / ^{13}\text{C}_6-$ Glucose (Neri *et al.*, 1989). Based on ^1H and ^{13}C resonance assignments obtained from the other experiments described above, the stereospecific assignments of the methyl groups of leucines and valines of dh434(0–63) were completed using *ct*- ^{13}C , ^1H -HSQC and ^{13}C , ^1H -HSQC spectra. Figure 6-20 presents *ct*- ^{13}C , ^1H -HSQC spectrum of fractionally 10% labelled dh434(0–63) with constant time, 27 msec so that the prochiral pro-(S) methyl groups of leucine (δ^2) and valine (γ^2) indicated by solid lines were of opposite sign relative to the pro-(R) methyl group (γ^1 and δ^1) (Hiroaki *et al.*, 1996). The stereospecific assignments of the diastereotopic methyl groups of valines and leucines of dh434(0–63) are also summarized in APPENDIX G.

6.8. Empirical correlation between $^{13}\text{C}_\gamma$ chemical shifts and conformations in valines.

There has been an attempt to correlate the chemical shifts of the diastereotopic methyl groups and the conformation in which among homologous protein the similar values at the same position among the homologous protein are observed but no clear correlation has been identified (Neri *et al.*, 1990). So far there are few further investigations between the chemical shifts of side chains and the conformations reported. In this study the three homologous proteins with the high resolution NMR structures could be investigated to find a correlation between the chemical shifts and the conformation.

Valine is an unique amino acid type because valine residues is special in having the greatest predominance of just one rotamer in highly refined protein structures i.e. one of the three conformers either gauche⁺, gauche⁻, or trans (Ponder & Richard 1987). In addition to this preferable conformation, the relatively short side chain of valine of which chemical shifts could be influenced by the carbonyl group and the backbone conformation in contrast to leucine residue.

Table 6-7: Summary of ^{13}C chemical shift differences for diastereotopic methyl groups of valines

residue #	dh434(0–63) $\delta(^{13}\text{C}\gamma^1 - ^{13}\text{C}\gamma^2)$ χ^1 angle (Ψ angle)	434(1–69) ^b $\delta(^{13}\text{C}\gamma^1 - ^{13}\text{C}\gamma^2)$ χ^1 angle (Ψ angle)	P22c2(1–76) ^c $\delta(^{13}\text{C}\gamma^1 - ^{13}\text{C}\gamma^2)$ χ^1 angle (Ψ angle)
6	–	– 4.6 173.2 ± 8.3 (–53.6 ± 4.0)	–
24(28 ^a)	– 0.7 188.3 ± 6.2 (–29.8 ± 12.0)	– 2.5 174.2 ± 4.7 (–30.8 ± 5.2)	– 0.6 180.1 ± 4.1 (–50.5 ± 10.7)
26(30 ^a)	+3.2 –53.4 ± 2.3 (175.9 ± 2.9)	–	+3.0 –57.4 ± 2.9 (173.4 ± 6.7)
29(33 ^a)	–	–	–1.5 171.4 ± 2.2 (–51.1 ± 3.7)
54	+ 2.5 –53.6 ± 3.13 (173.2 ± 3.3)	+ 2.5 –68.6 ± 8.5 (173.2 ± 4.1)	–
56	– 2.3 171.0 ± 4.8 (–36.5 ± 2.2)	– 1.5 165.8 ± 8.9 (–45.1 ± 2.2)	–

^anumbering for P22c2(1–76).

^bfrom Neri *et al.*, 1989 and Neri *et al.*, 1992a; ^cfrom Neri *et al.*, 1990 and Sevilla-Sierra *et al.*, 1994.

Table 6-7 summarizes the chemical shifts difference between the two diastereotopic methyl groups of a valine, i.e. $\Delta(\delta^{13}\text{C}\gamma^1 - \delta^{13}\text{C}\gamma^2)$, and the χ^1 angles for 434(1–69), P22c2(1–76) and dh434(0–63). Here only the chemical shift differences between diastereotopic methyl groups are analyzed which simplify the errors originating from the calibration and differences in the solvent and temperature conditions. It can be readily recognized that the values of the valines in the corresponding position among the three protein are very similar as noticed before (Neri *et al.*, 1990). In particular it is worth noting that the value for Val26 in dh434(0–63) has a close value to that of P22c2(1–76) although ^1H resonances of the methyl group of Val26 were shifted by the strong ring current effect from Phe44 which does not exist in P22c2(1–76). This might suggest that the ^{13}C chemical shift difference of the methyl groups in valine could be largely determined by the conformation of valine residues rather than the effects from its neighboring groups in the vicinity. In addition to this observation Table 6-7 indicates a pattern that the residues having negative $\Delta(^{13}\text{C}\gamma^1 - ^{13}\text{C}\gamma^2)$ have a value of around 180° for the χ^1 angle. In contrast the residues having

positive $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$ have a value around -60° for the χ^1 angle despite of the positions in the amino acid sequence. This observation indicates the correlation between $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$ and χ^1 angles in valine residues and tempts to analyze other proteins. Figure 6-21 presents that the results from 11 different proteins including β -sheet proteins, which are 434(1-69)(5), P22c2(1-76) (4), dh434(0-63) (4), Cyclophilin A (8), lck-SH3 (3), E2-DBD (5), SCP-2 (7), P14A(7), dnaJ(2), HSF(8), HIV protease complex (7) (the number of valine residues is indicated in the parenthesis) (Neri *et al.*, 1989; Neri *et al.*, 1990; Sevilla-Sierra *et al.*, 1994; Ottiger *et al.*, 1997; Hiroaki *et al.*, 1996; Liang *et al.*, 1996; Lopez *et al.*, to be published; Fernández *et al.*, 1997; Pellecchia *et al.*, 1996; Damberger *et al.*, 1995; Yamazaki *et al.*, 1996). For P14A, dnaJ, HSF, and HIV protease complex 10% fractional labelling was not applied. NOEs and couplings were used for the stereospecific assignments instead. The pattern for the 11 proteins appears to be more complicated than the expected pattern from the three homologous proteins. It appears that side chain conformations of valines in helices are limited to the trans conformation due to steric hindrance (Figure 6-21A). These valine residues in helices clearly indicate the tendency of negative values for $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$ as observed in the variants of 434(1-63). $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$ in 19 residues of this category is found to be -1.5 ± 1.1 ppm (Figure 6-21B). This standard deviation is rather large but a very similar value found in the correlation between $^{13}\text{C}^\alpha/^{13}\text{C}^\beta$ chemical shifts and the secondary structures (Spera & Bax, 1991). The residues having positive values of $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$ in the three homologous proteins are all in the gauche⁻ conformation. From 11 proteins including these three homologous proteins the valines in gauche⁻ conformation indicates positive value $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$, $+2.3 \pm 1.1$ ppm (Figure 6-21B).

This two distinct patterns were clearly identified in the three homologous proteins, 434(1-63), P22c2(1-76), and dh434(0-63) (Table 6-7). However, analysis of the other proteins shows that the valines having negative values of $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$ are not necessarily in trans conformation when they are not in the helical conformation, i.e. $30^\circ < \Psi < 90^\circ$ (Figure 6-21C). The valines of trans side chain conformation in the backbone β conformation, i.e. $90^\circ < \Psi < 200^\circ$ indicate values around zero, i.e. 0.0 ± 0.8 ppm. This observation complicates the interpretation of $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$ because the range of $\Delta(^{13}\text{C}^{\gamma 1}-^{13}\text{C}^{\gamma 2})$ for this category overlaps both the range of α -trans and β -gauche⁻ conformations.

However, these two distinct patterns of α -trans and β -gauche⁻ would be still useful information as a guideline for verification of stereospecific assignments of the methyl groups of valines in particular when fractional samples are not available. This observed correlation implies the potential use of the ^{13}C chemical shifts for three dimensional structure determination of proteins. It may be tempting to use the chemical shifts information of $^{13}\text{C}^\gamma$ of valines for the stereospecific assignments of the diastereotopic methyl groups particularly for the α conformation, it should be noted that the basis of $^{13}\text{C}^\gamma$ data presented in this study is probably too small to use for the stereospecific assignments. The increase of the data will certainly provide a reliable guidance for the assignment verification of valine residues and structure refinements. In addition the increase of the data will also provide the more precise correlations including ϕ angle and make it possible to produce the chemical shifts difference surface depending on ψ , χ^1 angle even including ϕ angle. For this purpose it is necessary to accumulate the information of both high resolution structures and the precise chemical shifts, and possibly with the information of the side chain dynamics. It may be also possible to obtain the information of the side chain conformation from the $\Delta(^{13}\text{C}^{\gamma 1} - ^{13}\text{C}^{\gamma 2})$ chemical shifts of valine residues combined with the backbone conformation that can also be obtained from $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$ chemical shifts. This data gained from the three homologous proteins will be of particular interest in cases when the data of NOEs and J coupling are limited such as a case of larger proteins and unfolded proteins.

Very recently Oldfield group reached a same conclusion from *ab initio* calculation which predicts this correlation and shows a good correlation with experimental data (Havlin *et al.*, 1997; Pearson *et al.*, 1997).

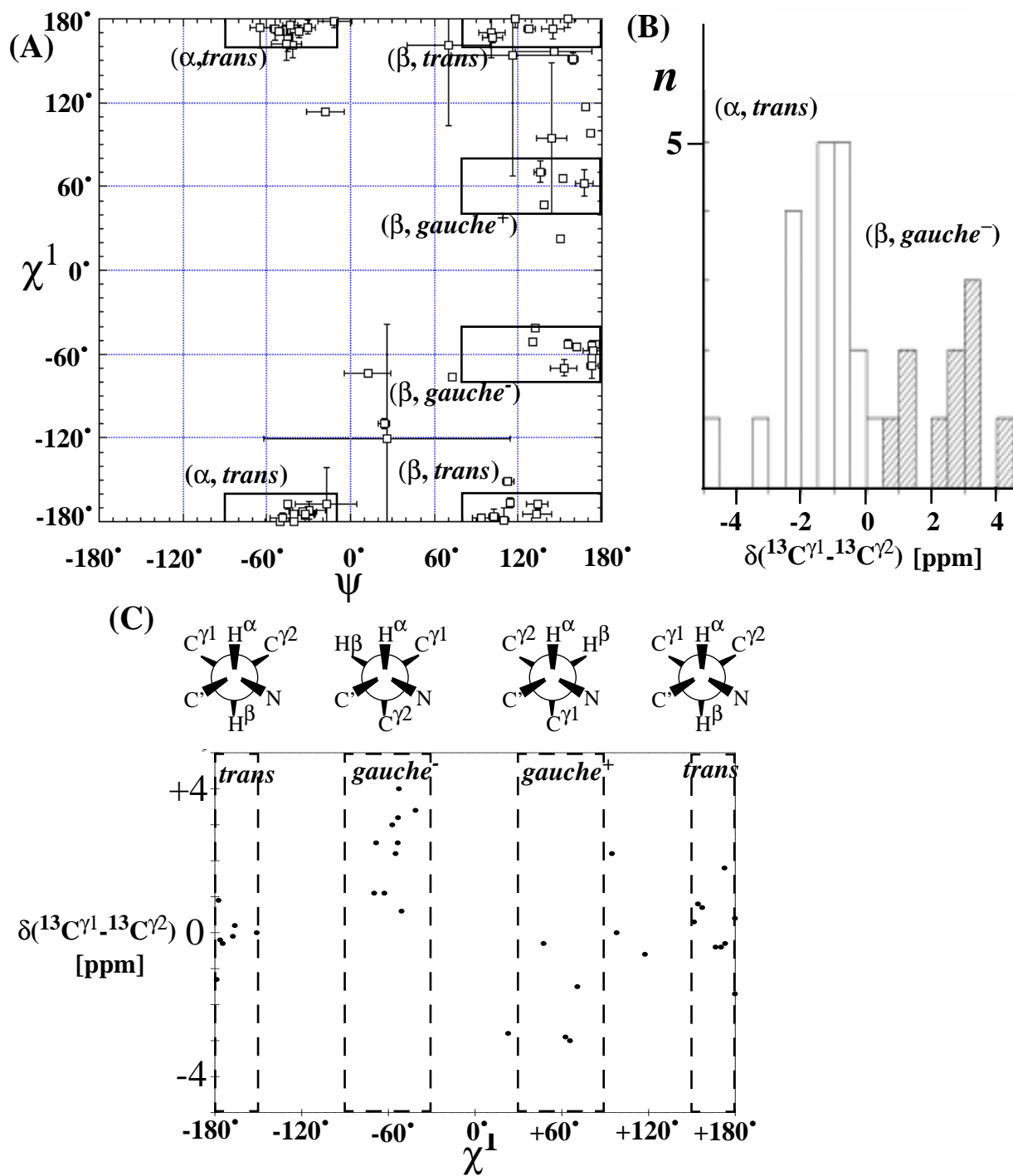


Figure 6-21: (A) The distribution of the χ^1 angle and ψ angle of valines from 11 different proteins. Based on χ^1 and ψ angles, the distribution can be divided into the 4 categories, $(\alpha, trans)$, $(\beta, gauche^+)$, $(\beta, gauche^-)$, and $(\beta, trans)$. The error bars are obtained from the NMR structures when there are more than one available structures. (B) Histogram showing the $\delta(^{13}C^{\gamma 1} - ^{13}C^{\gamma 2})$ distribution in $(\alpha, trans)$ and $(\beta, gauche^-)$ conformations. (C) Plot of the $\delta(^{13}C^{\gamma 1} - ^{13}C^{\gamma 2})$, versus χ^1 angles for residues in β conformations ($90^\circ < \psi < 200^\circ$). The regions of $gauche^+$, $gauche^-$ and $trans$ were indicated with dotted boxes.

6.9. Summary of Chapter 6

In Chapter 6 the high resolution structures of dh434(0–63) was determined by NMR spectroscopy, which provides the basis for the analysis of the hydration study of dh434(0–63) described in Chapter 3 and for understanding the mechanism of the hyper-stability of dh434(0–63) described in Chapter 5. The ^{13}C and ^{15}N resonance assignments of dh434(0–63) described in this chapter also provide the necessary information for the analysis of the spectra obtained from the NMR experiments for the hydration study. These resonance assignments also verified the structures calculated from the NMR data of dh434(0–63). H/D exchange rates of amide groups of dh434(0–63) supports the hyper-stability observed from the thermal and chemical denaturation of dh434(0–63) indicating the stable compact globular structure of dh434(0–63).

7.NMR techniques

7.1. Pulse scheme for 'artifact-free' water-protein NOEs detection

NMR spectroscopy has become a powerful method to characterize hydration of macromolecules (Chapter 2). A number of two- and three- dimensional NMR experiments have been developed to investigate the interactions between water and proteins. These pulse schemes for the observation of the interaction between water and proteins can be usually divided into two blocks. The first part is to separate water resonance from the other resonances. For this purpose, ^1H homonuclear NOESY, selective pulse, selective excitation of water resonance by radiation dumping, diffusion filtering, relaxation properties, or isotope filtering have been applied (Otting & Wüthrich, 1989; Otting *et al.*, 1991a,b; Kriwacki *et al.*, 1993; Grzesiek & Bax, 1993a,1993b; Mori *et al.*, 1994; Dötsch & Wider, 1995; Otting & Liepinsh, 1995b; Dalvit & Hommel, 1995; Wider *et al.*, 1996; Mori *et al.*, 1996). The second part of the pulse scheme is a relay step for the separation of signals in the proteins. Either ^1H homonuclear TOCSY or heteronuclear correlation schemes have been used for this purpose (Otting *et al.*, 1991b; Grzesiek & Bax, 1993a, 1993b). In our studies of hydration of dh434(0–63) the protein was $^{13}\text{C}/^{15}\text{N}$ doubly labeled, so that we could use heteronuclear correlation steps to separate overlapping signals.

Figure 7-1 presents a pulse scheme for the detection of water–protein NOE making use of radiation dumping as a selective pulse (Otting & Liepinish 1995b; Widar *et al.*, 1996). Two different experiments have been recorded to obtain water–protein NOEs. In the experiment denoted with (I) in Figure 7-1 the pulse field gradient pulse was placed at the end of the mixing time so that the water magnetization was rotated back to +Z axis due to radiation dumping during the mixing time and all the other transverse magnetization was dephased at the end of mixing time by a strong gradient pulse (G3). In the second experiment as denoted with (II) in Figure 7-1 the radiation dumping was prevented by applying a strong field gradient pulse (G3) at the beginning of the mixing time and a weak gradient pulse during the entire mix-

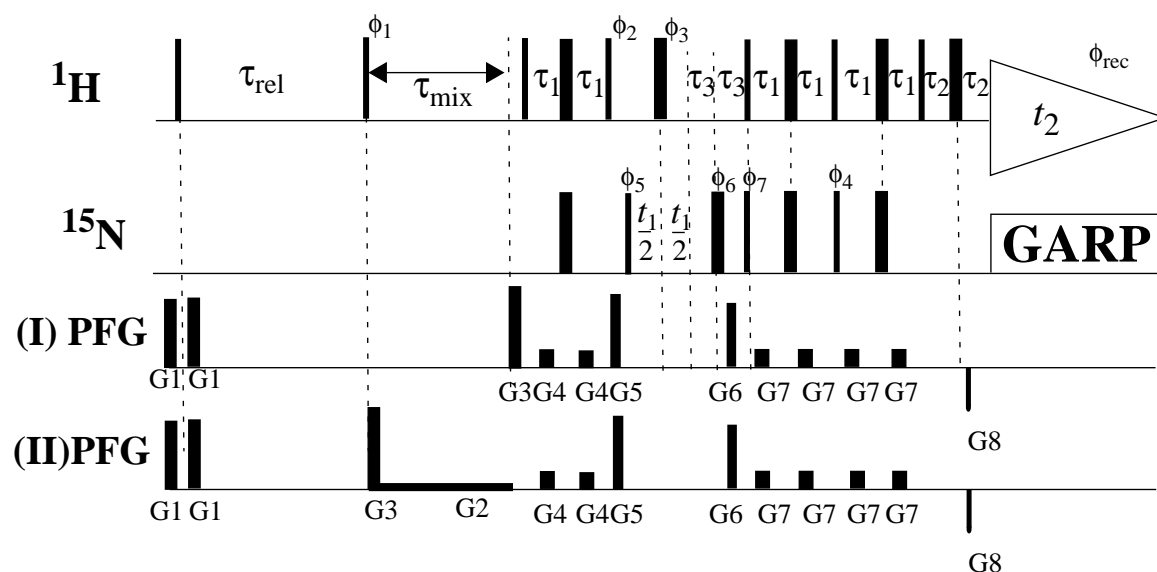


Figure 7-1: Pulse scheme for the detection of water-protein NOEs using radiation damping as a water selective pulse. (I) and (II) denote the two experiments to compute subtraction spectrum to produce water-protein NOEs. Where no rf-phase is marked, the pulse is applied along x , otherwise the phases are indicated above the pulses. The delays are set to the following values: $\tau_{rel} = 2.0$ sec; $\tau_1 = 2.25$ msec; $\tau_2 = 0.4$ msec; $\tau_3 = 2.75$ msec. The duration and the amplitude for the gradient pulses are the following; $G1 = 1$ msec and 25 G/cm, $G2 = 0.5$ G/cm; $G3 = 2$ msec and 30 G/cm, $G4 = 0.5$ msec and 14 G/cm, $G5 = 0.5$ msec and 18 G/cm, $G6 = 1.25$ msec and 30 G/cm, $G7 = 1.5$ msec and 15 G/cm, $G8 = 0.125$ msec and 29 G/cm. All the gradient pulses are rectangular pulses. The phase cycling for the pulse scheme is the following; $\phi_1 = x, x, -x, -x$; $\phi_2 = y$; $\phi_3 = x, x, -x, -x$; $\phi_4 = y$; $\phi_5 = x$; $\phi_6 = 4(x), 4(y), 4(-x), 4(-y)$; $\phi_7 = x$; $\phi_{rec} = x, -x, x, -x, -x, x, -x, x, -x, x$. Quadrature detection in t_1 is accomplished by altering the phase ϕ_5 according to States-TPPI. τ_{mix} indicates the mixing time.

ing time. As a consequence all transverse magnetization including water magnetization was dephased by the gradient pulse. The spectrum resulting from the difference between two experiments must therefore contain water-protein NOEs, exchange peaks with water and, possibly H^N-H^α NOEs deriving from H^α that resonate at a close chemical shift to the water chemical shift.

Figure 7-2 shows such a spectrum recorded with a sample of 5 mM uniformly ^{15}N labeled dh434(0–63). The experiment as described in Figure 7-1 was measured at two different 1H frequencies of 750 MHz and 400 MHz. The spectrum should contain only water-protein NOEs, exchange peaks with water and, possibly, H^N-H^α NOEs in case that H^α protons have the same chemical shift of water. Surprisingly the spectrum measured at the 1H frequency of 750 MHz contains many cross peaks unexpectedly corresponding to all amide backbone groups of dh434(0–63) including some buried amide groups. For example, the arrows in Figure 7-2A indicate the positions of the backbone amide resonances of the residues with less than 5% of the solvent accessible surface area. This observation could lead to a hypothesis that

dh434(0–63) might have non-compact structures such as ‘molten globule structure’ in which water is penetrating into the deep core of the molecule because this could arise from a number of substitution of amino acids introduced in wild-type 434(1–63) (Chapter 4). Contrarily the other data such as H/D exchange rates, thermal and chemical denaturations suggest the more compact and stable structure of dh434(0–63) as compared with the wild-type 434(1–63) as described in Chapter 5 and Chapter 6. Therefore the cause of unexpected peaks at ^1H frequency of 750 MHz must originate from the experimental scheme not from the feature of the protein. In addition the spectrum measured with the same pulse scheme at ^1H frequency of 400 MHz is rather different from one measured at ^1H frequency of 750 MHz. In particular the absence of the peaks from buried residues indicates that the data at ^1H frequency of 400 MHz are much more reasonable (Figure 7-2B). This observation suggests that there are undesirable effects in the pulse scheme which would produce some artifacts in the spectrum only at the high magnetic field.

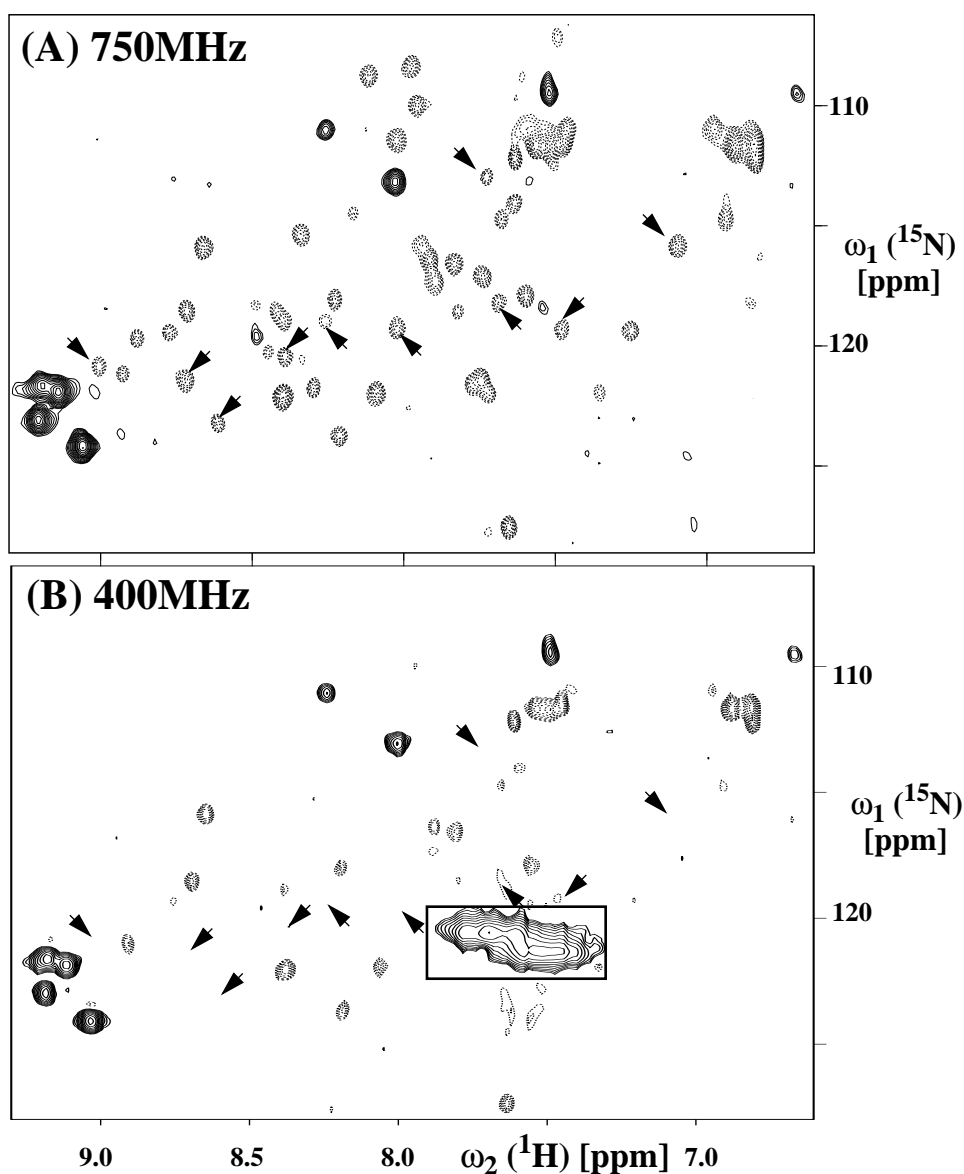


Figure 7-2: Spectra of 5mM uniformly ^{15}N labelled dh434(0–63) measured by the pulse scheme for the detection of water-protein NOE described in Figure 7-1. (A) At ^1H frequency of 750MHz with mixing time of 70 msec. (B) At ^1H frequency of 400MHz with mixing time of 120 msec. The arrows indicate the residues of less than 5% of solvent accessible surface. The peaks indicated by a box are the folded signals originated from side chains due to the different spectrum width along ω^1 frequency. Positive (negative NOEs) and negative (positive NOEs) peaks are shown by solid and dotted lines, respectively.

In fact the extensive studies as described in Sobol *et al.* 1998 have demonstrated that radiation damping field, despite its small magnitude, is able to disturb magnetization of nuclei resonating 2–3 kHz away from the solvent signal. The amplitude of such disturbances can be of the order of 1% of the nuclear magnetization and in particular of importance for the detection of weak water-protein NOEs such as NOEs originating from the surface hydration water. This effect was vividly manifested in the spectrum at ^1H frequency of 750 MHz (Figure 7-2A). Since radiation

damping is a well known effect in high resolution solution NMR, in which solvent magnetization is rotated back to the external magnetic field direction by the strong coupling of the magnetization with the resonant circuit of the detection system, the effect becomes more and more important at higher static magnetic field strength and with improvements in the probe quality factors Q .

To overcome this problem at high magnetic field, a new pulse scheme was developed (Figure 7-3). The pulse scheme was designed so as to suppress the undesired excitation by radiation damping and to reduce the radiation damping and demagnetization field effect through the entire pulse scheme. The scheme starts with two gradient with a 90 degree ^1H pulse surrounded with two gradient pulses (G1) in order to keep the initial magnetization constant at the beginning of the pulse sequence. The relaxation delay (τ_{rel}) of 2.5 sec ensures the recovery of the magnetization of water and proteins. The double filter unit follows the relaxation delay, which eliminates all proton magnetization attached to either ^{13}C or ^{15}N nuclei (Gemmecker *et al.*, 1993). In the first experiment denoted with (I) in Figure 7-3 during the double filter unit the water magnetization was refocused by 180 degree ^1H pulse and two equivalent gradient pulses (G2) and restored to +Z axis by a selective water pulse at the point *a* in Figure 7-3. All the transverse components were dephased by a Z-gradient pulse (G3) after this selective pulse. At this point only the water magnetization is aligned to +Z axis. All the other magnetization from the protein should be dephased by the gradient pulses including the protons under and near the water resonance. In the second experiments denoted with (II) all the transverse magnetization including water at the point *a* was dephased by a strong gradient pulse (G3). The both scans after the mixing time were followed by a heteronuclear relay step, i.e. either [^{15}N , ^1H]-, or [^{13}C , ^1H]-HSQC with sensitivity enhancement schemes (Kay *et al.*, 1992; Schueler *et al.*, 1994). The subtraction of these two experiments produces the spectrum containing water-protein NOEs, exchange peaks with water. During the HSQC-relay step, the gradient pulses were placed carefully during all the delay periods to dephase water magnetization and to suppress radiation damping and demagnetization effect which could also produce artifacts in the spectrum. The control experiments of the pulse scheme were performed to obtain artifacts-free spectra by following the test procedure that include a cancellation test for the relay step, a difference test with presaturation of the sol-

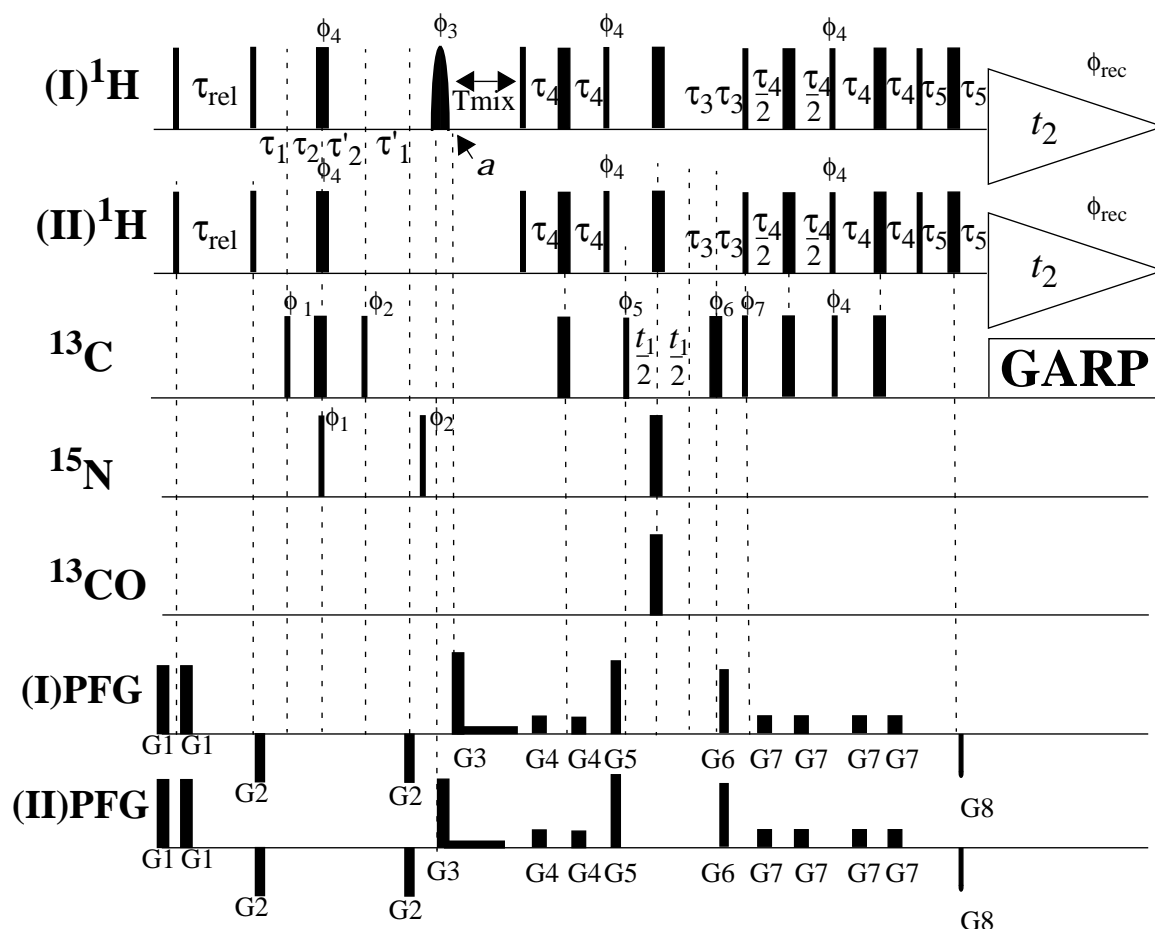


Figure 7-3: Pulse scheme for “artifact-free” two-dimensional $[^{13}\text{C}, ^1\text{H}]$ -HSQC difference spectrum. 90° and 180° pulses are indicated by thin and thick vertical bars, respectively. The sinc shaped soft pulse is indicated by a filled curved shape. The duration of the soft pulse was 3.7msec. Where no rf-phase is marked, the pulse is applied along x , otherwise the phases are indicated above the pulses. The delays are set to the following values: $\tau_{\text{rel}}=2.5$ sec; $\tau_1=3.9$ msec; $\tau_1'=3.4$ msec; $\tau_2=1.6$ msec; $\tau_2'=2.0$ msec; $\tau_3=0.7$ msec; $\tau_4=1.7$ msec; $\tau_5=0.4$ msec. The duration and the amplitude for the gradient pulses are the following; $G1=2$ msec and 16 G/cm, $G2=0.5$ msec and -8 G/cm, $G3=2$ msec and 15 G/cm, $G4=0.5$ msec and 7 G/cm, $G5=0.5$ msec and 15 G/cm, $G6=0.5$ msec and 9 G/cm, $G7=0.5$ msec and 2 G/cm, $G8=0.125$ msec and -16 G/cm. All the gradient pulses are rectangular pulses except for $G8$ which has a sine-bell shape. The phase cycling for the pulse scheme is the following; $\phi_1 = x, x, -x, -x$; $\phi_2 = 4(x), 4(-x)$; $\phi_3 = -x$; $\phi_4 = y$; $\phi_5 = x$; $\phi_6 = 4(x), 4(y), 4(-x), 4(-y)$; $\phi_7 = x$; $\phi_{\text{rec}} = x, -x, x, -x, -x, x, -x, x$. Quadrature detection in t_1 is accomplished by altering the phase ϕ_5 according to States-TPPI. T_{mix} indicates the mixing time.

vent signal and an experiment with short mixing time (<5 msec) or even build-up curves of the solvent-protein NOE. Thus the artifacts-free water-protein NOE/ROE spectra with the pulse scheme described in Figure 7-3 are shown in Chapter 3 and analysed for the detailed investigation of water-protein interactions in dh434(0–63).

7.2. The 2D NMR Experiments $H(C)CO_2$ and $HCCO_2$

Side chain carboxylates of Asp and Glu residues are frequently involved in active sites of proteins, in salt bridges, and in hydrogen bonds. The identification and the determination of their pK_a values is of importance to understand the mechanisms of catalytic reactions as well as the contributions of these groups to the stability of proteins by formation of salt bridges and hydrogen bonds, and to characterize transient hydrogen bonding interactions on the protein surface (Anderson *et al.*, 1990; Szyperski *et al.*, 1994; Jeng & Dyson, 1996). Apparent pK_a shift of ionizable groups can be used to assess the contribution of each ionizable group to the stability of the folded state (Anderson *et al.*, 1990). A buried salt-bridge can be found in the DNA binding domain of 434 repressor, 434(1–63) (Mondragon *et al.*, 1989; Neri *et al.*, 1992; Pervushin *et al.*, 1996). Determination of pK_a of side chain carboxylates in the salt-bridge of 434(1–63) is of interest to assess the contribution of the buried salt-bridge to the stability of the protein because it is generally still unclear what contribution such buried salt bridge interactions per se make to the stability of proteins (Mrabet *et al.*, 1992; Yang & Honig, 1993; Hendsch & Tidor, 1994; Borders *et al.*, 1994; Marqusee & Sauer, 1994; Waldburger *et al.*, 1995; Tissot *et al.*, 1996). pH dependence of the ^{13}C chemical shifts of the side chain carbonyl group can be one of powerful methods to determine the individual pK since the chemical shifts is very sensitive to the ionization states. Recently, 2D ct- $H(CA)CO$ adapted to detect side chain ^{13}CO resonance has been used to determine pK_a values for the carboxylates in $^{13}C/^{15}N$ -labelled proteins (Yamazaki *et al.*, 1993; Jeng & Dyson, 1996; Oda *et al.*, 1994; Wang *et al.*, 1996; Qin *et al.*, 1996).

Figure 7-4 presents the pulse scheme of the 2D $H(C)CO_2$ and the 2D $HCCO_2$ experiments, which designed to eliminate cross peaks originating from the side chain $-CH_2-CO-NH_2$ moieties likewise occurs for the backbone fragments $-CH-CO-NH-$. This filtration of cross peaks is achieved by the delay τ_4 , which is set to 32 msec, and pulses for ^{15}N atoms in Figure 7-4. Such delay and pulses convert the transferred magnetization on ^{13}CO in $-CH_2-CO-NH_2-$, $-CH-CO-NH-$ moieties into the magnetization, which can not be refocused to observable magnetization by the successive pulses. In addition the gradient pulse after the 90° pulse for ^{15}N atoms

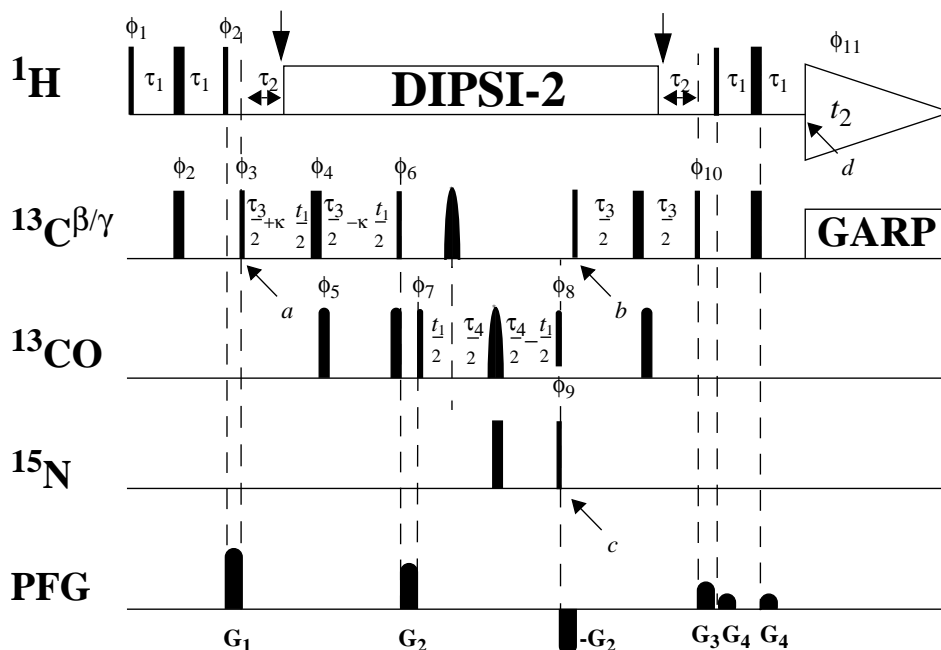


Figure 7-4: Experimental scheme of the 2D HCCO_2 and 2D H(C)CO_2 experiment. 90° and 180° pulses are indicated by thin and thick vertical bars, respectively. For the present project we used the following conditions: The carrier for the first three ^1H pulses is placed on the solvent line at 4.9 ppm; then it is switched to 2.5 ppm for the duration of the DIPSI-2 sequence (Shaka et al., 1988), and subsequently it is switched back to 4.9 ppm (the two vertical arrows indicate the time points at which the offset is switched). The $^{13}\text{C}^{\beta/\gamma}$ and ^{15}N carriers are set to 40 ppm and 110 ppm, respectively. The 90° pulse lengths are 12 μs for ^1H , 40 μs for ^{15}N , 47 μs for $^{13}\text{C}^{\beta/\gamma}$, and 92 μs for ^{13}CO . The corresponding 180° pulses are applied with the same power. The pulses on ^{13}CO have the shape of a sinc centre lobe applied in a phase-modulated manner with the centre of excitation at 180 ppm, and they have been optimized to avoid excitation of $^{13}\text{C}^{\beta/\gamma}$ magnetization. A ^{13}CO RE-BURP pulse (Geen & Freeman, 1991) of 400 μs and 15.7 KHz peak amplitude and a $^{13}\text{C}^{\beta/\gamma}$ I-BURP2 pulse (Geen & Freeman, 1991) of 400 μs and 12.4 KHz peak amplitude are applied during the ^{13}CO evolution period (these two pulses are identified with asterisks). The duration and amplitudes of the sine-bell-shaped PFGs are 1 ms and 60 G/cm for G_1 , 1 ms and 40 G/cm for G_2 , 1 ms and 20 G/cm for G_3 , and 200 μs and 8 G/cm for G_4 . The delays are $\tau_1 = 3.0$ ms, $\tau_2 = 1.6$ ms, $\tau_3 = 7.2$ ms, and $\tau_4 = 32$ ms. κ represents a scaling factor for the chemical shift evolution of $^{13}\text{C}^{\beta/\gamma}$ (see text). The 2D H(C)CO_2 experiment is recorded by setting $\kappa = 0$ (see text). The phase cycling for the pulse scheme is: $\phi_1 = y$; $\phi_2 = x, -x$; $\phi_3 = x$; $\phi_4 = 8x, 8y, 8(-x), 8(-y)$; $\phi_5 = 8x, 8(-x)$; $\phi_6 = 4x, 4(-x)$; $\phi_7 = x, x, -x, -x$; $\phi_8 = x, -x$; $\phi_9 = x$; ϕ_{10} (receiver) = $2x, 4(-x), 2x, 2(-x), 4x, 2(-x)$; where no rf-phase is marked, the pulse is applied along x . Quadrature detection in t_1 is accomplished by altering the phase ϕ_7 according to States-TPPI. To shift the apparent $^{13}\text{C}^{\beta/\gamma}$ carrier position to 48.5 ppm in the 2D HCCO_2 experiment, the phase ϕ_3 is incremented in 40° steps according to TPPI. A DIPSI-2 sequence with $rf = 1.2$ kHz is used to decouple ^1H during the heteronuclear magnetization transfer, and a GARP sequence (Shaka et al., 1985) with $rf = 1.0$ kHz is employed to decouple $^{13}\text{C}^{\beta/\gamma}$ during t_2 .

dephase the transverse magnetization at the point *c*. Therefore cross peaks only originating from Asp and Glu appear in the spectrum of the 2D H(C)CO_2 (Figure 7-5B). Moreover, elimination of signals arising from Asn and Gln residues notably reduces the total number of cross-peaks, making this experiment ideally suited for use

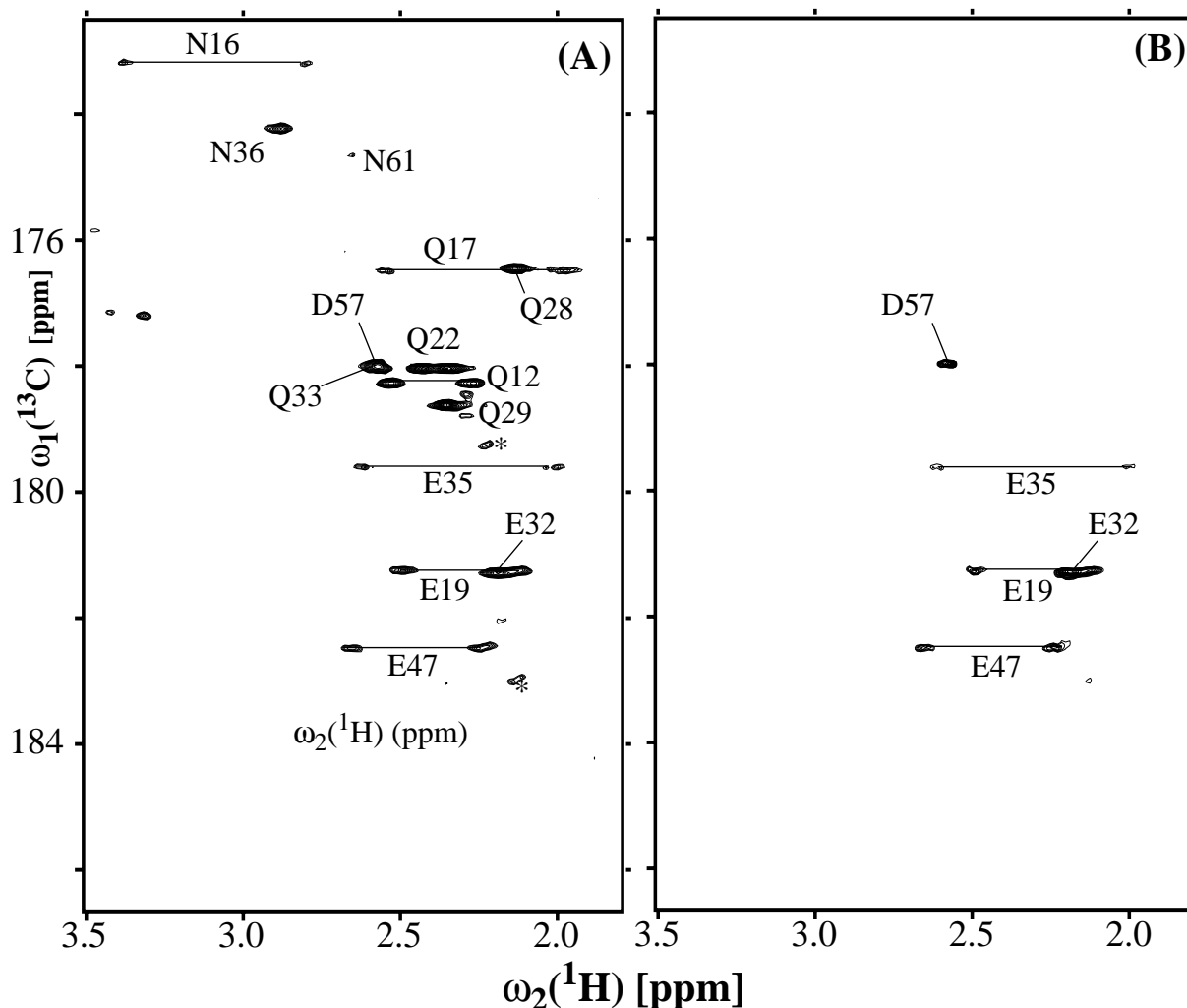


Figure 7-5: Spectra obtained by the pulse scheme described in Figure 7-4, with a 1 mM sample of uniformly $^{13}\text{C}/^{15}\text{N}$ -labelled 434(1–63) (solvent D_2O , $\text{pD}^* = 7.0$, $T = 13^\circ\text{C}$). The spectra were recorded on a Bruker AMX 600 spectrometer operating at 600 MHz ^1H frequency. (A) spectrum of 2D H(CA)CO . (B) Spectrum of 2D H(C)CO_2 . The final data size of the experiments in (A) and (B) was doubled in t_1 by linear prediction. The assignments of side chain carbonyl groups are indicated. Asterisks indicate cross-peaks arising from an impurity.

of the projection technique which in addition provides the $^{13}\text{C}^{\beta/\gamma}$ chemical shifts (Szyperski *et al.*, 1993a; Szyperski *et al.*, 1993b). Such experiment can be presented as 2D HCCO_2 experiments by setting the κ to non-zero values, which provides $^{13}\text{C}^{\beta/\gamma}$ chemical shifts encoded as splitting of peaks. This information could be useful for the assignments of the side chain carboxylates.

The 2D H(C)CO_2 and the 2D HCCO_2 experiments described here greatly reduce the problem of spectral overlap when compared with 2D ct-H(CA)CO (Figure 7-5) and thus circumvent the necessity to produce selectively $^{13}\text{CO}_2$ labelled samples (Anderson *et al.*, 1990; Anderson *et al.*, 1993; McIntosh *et al.*, 1996). Since it also

provides correlations with $C^{\beta/\gamma}$, 2D $HCCO_2$ is most informative for obtaining the resonance assignments, whereas 2D $H(C)CO_2$ is the suitable experiment for pH-titrations of the carboxylate groups. Normally, there is hardly any interference between the resonances of the side chain carboxylates of Asp and Glu and the C-terminus.

The spectra of 2D ct-H(CA)CO adapted for side chain correlation and 2D $H(C)CO_2$ experiments of a 1 mM ^{13}C , ^{15}N -doubly-labeled 434(1–63) solution are shown in Figure 7-5 at 13°C. In the 2D ct-H(CA)CO spectrum in Figure 7-5A all correlations between $^1H^{\beta}$ and $^{13}CO^{\gamma}$ of Asp57, Asn16, Asn36 and Asn61 as well as the correlations between $^1H^{\gamma}$ and $^{13}CO^{\delta}$ of the Glu residues in positions 19, 32, 35, and 47 and the Gln residues 12, 17, 22, 28, 29, and 33 are observed. In the 2D $H(C)CO_2$ spectrum in Figure 7-5B only the cross peaks corresponding to Asp and Glu are selected (weak cross peaks from the C-terminal residue were observed outside of the spectral region shown). In particular, the assignment of the side chain carboxyl carbon of Asp57 in the 2D ct-H(CA)CO spectrum is ambiguous because of overlap with the resonances of Gln 33 (Figure 7-5A). This overlap, and thus the assignment ambiguity, is neatly resolved in Figure 7-5B. With use of 2D $H(C)CO_2$ experiment pK of side chain carboxylates were determined in 434(1–63). pK 4.4 for Glu19, pK 4.1 for Glu32, pK 2.0 for Glu35, pK 4.1 for Glu47, pK 3.4 for Asp56 could be determined by the pH titration using $^{13}C^{\gamma\delta}$ chemical shifts, which generally agree with data obtained from 1H chemical shifts (Dötsch, 1994). These determined pK values could be used for analysis of the contribution to the protein stability (Chapter 5). The newly designed 2D $H(C)CO_2$ experiment simplifies the spectrum which makes the procedure of pH titration easier. pH titration by measuring a series of the spectra is a simple but still tedious procedure. Improvement in the procedure which can overcome this tedious adjustment of pH in pH titration such as manual adjustment of pH values is of special interest for the next development of the procedure.

8. References

- Aggarwal, A.K., Rodgers, D.W., Drottar, M., Ptashne, M. & Harrison, S.C. (1988) Recognition of a DNA operator by the repressor of phage 434. A View at high resolution. *Science* **242**, 899–892.
- Akasaka, K. (1981) Longitudinal relaxation of proton under cross saturation and spin diffusion. *J. Mag. Reson.* **45**, 337–343.
- Anderson, D.E., Wayne, J.B. & Dahlquist, F.W. (1990) pH-Induced denaturation of proteins: A single salt bridge contributes 3–5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**, 2403–2408.
- Anderson, D.E., Lu, J., Macintosh, L. & Dahlquist, F.W. (1993) In *NMR of Proteins*. MacMillan. pp. 258.
- Anil-Kumar, Ernst, R.R. & Wüthrich, K. (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Comm.* **95**, 1–6.
- Archer, S., Ikura, M., Torchia, D.A. & Bax, A. (1991) An alternative 3D NMR technique for correlating backbone ^{15}N with side chain H^β resonances in larger proteins. *J. Magn. Reson.* **95**, 636–641.
- Ayant, Y., Belorizky, E., Fries, P. & Rosset, J. (1977) Effet des interactions dipolaires magnétiques intermoléculaires sur la relaxation nucléaire de molécules polyatomiques dans les liquides. *le journal de physique* **38**, 325–337.
- Badger, J., Kapulsky, A., Caspar, D.L. & Korszun, R. (1995) Neutron diffraction analysis of the solvent accessible volume in cubic insulin crystals. *Nat. Struct. Biol.* **2**, 77–80.
- Bai, Y., Milne, J.S., Mayne, L. & Englander, S.W. (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* **17**, 75–86.
- Baker, E.N., Blundell, T.L., Cutfield, J.F., Cutfield, S.M., Dodson, E.J., Dodson, G.G., Hodgkin, D.M., Hubbard, R.E., Isaacs, N.W., Reynolds, C.D. Sakabe, K., Sakabe, N. & Vijayan, N.M. (1988) The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Phil. Trans. Roy. Soc.* **B319**, 369–456.
- Baker, E.N. & Hubbard, R.E. (1984) Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **44** 97–179.
- Baker, E.N. (1995) Solvent interactions with proteins as revealed by X-ray crystallographic studies *Protein-solvent interactions*, Gregory (ed.) Marcel Dekker, Inc. pp143–189.
- Bartels, C., Xia, T., Billeter, M., Güntert, P. & Wüthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **5**, 1–10.
- Ben-Naim, A. (1974) *Water and aqueous solutions: introduction to a molecular theory*. Plenum Press, New York.
- Benveniste, J. (1988) Benveniste on the Benveniste affair. *Nature* **335**, 759.
- Benveniste, J. (1991) Defence of diluted water. *Nature* **353**, 787.
- Benveniste, J., Duco, B. & Spira, A. (1994) Memory of water revisited. *Nature* **370**, 322.
- Berndt, K.D., Beunink, J., Schröder, W. & Wüthrich, K. (1993) Designed replacement of an internal hydration water molecule in BPTI: structural and functional implications of a Gly-to-Ser mutation. *Biochemistry* **32**, 4564–4570.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer, E.E. Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
- Betz, S.F., Raleigh, D.P. & DeGrado, W.F. (1993) *De novo* protein design: from molten globules to native-like states. *Curr. Opin. Struct. Biol.* **3**, 601–610.
- Betz, S.F., Bryson, J.W. & DeGrado, W.F. (1995) Native-like and structurally characterized designed α -helical bundles. *Curr. Opin. Struct. Biol.* **5**, 457–463.

- Billeter, M., Schaumann, T., Braun, W. & Wüthrich, K. (1990) Restrained energy refinement with two different algorithms and force fields of the structure of the α -amylase inhibitor tendamistat determined by NMR in solution. *Biopolymers* **29**, 695–706.
- Billeter, M., Güntert, P., Luginbühl, P. & Wüthrich, K. (1996) Hydration and DNA recognition by homeodomains. *Cell* **85**, 1057–1065.
- Blake, C.C., Pulford, W.C. & Artymiuk, P.J. (1983) X-ray studies of water in crystals of lysozyme. *J. Mol. Biol.* **167**, 693–723.
- Blanco, F.J., Jimenez, M.A., Pineda, A., Rico, M., Santoro, J. & Nieto, J.L. (1994) NMR solution structure of the isolated N-terminal fragment of protein-G B1 domain. Evidence of trifluoroethanol induced native-like beta-hairpin formation. *Biochemistry* **33**, 6004–6014.
- Blundell, T.L. & Johnson, L.E. (1976) *Protein Crystallography*. Academic press, London.
- Blundell, T.L., Barlow, D., Borkakoti, N. & Thornton, J. (1983) Solvent-induced distortions and the curvature of alpha-helices. *Nature* **306**, 281–283.
- Bodenhausen, G. & Ruben, D. (1980) Natural abundance ^{15}N NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Letters* **69**, 185–188.
- Borders, C.L. Jr., Broadwater, J.A., Bekeny, P.A., Salmon, J.E., Lee, A.S., Eldridge, A.M. & Pett, V.B. (1994) A structural role for arginine in proteins: multiple hydrogen bonds to backbone carbonyl oxygens. *Protein Sci.* **3**, 541–548.
- Bowie, J.U. & Sauer, R.T. (1989) Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry* **28**, 7139–7143.
- Brunne, R.M., Liepinsh, E., Otting, G., Wüthrich, K. & van Gunsteren, W.F. (1993) Hydration of protein. A comparison of experimental residence times of water molecules solvating the bovine pancreatic trypsin inhibitor with theoretical model calculations. *J. Mol. Biol.* **231**, 1040–1048.
- Brüschweiler, R. & Wright, P.E. (1994) Water self-diffusion model for protein-water NMR cross relaxation. *Chem. Phys. Lett.* **229**, 75–81.
- Brüschweiler, R., Morikis, D. & Wright, P.E. (1995) Hydration of the partially folded peptide RN-24 studied by multidimensional NMR. *J. Biomol. NMR* **5**, 353–356.
- Bryant, R.G. (1996) The dynamics of water–protein interactions. *Annu. Rev. Biophys. Biomol. Struct.* **25**, 29–53.
- Buckle, A.M., Cramer, P. & Fersht, A.R. (1996) Structural and energetic responses to cavity-creating mutations in hydrophobic cores: observation of a buried water molecule and the hydrophilic nature of such hydrophobic cavities. *Biochemistry* **30**, 4298–4305.
- Bundi, A. & Wüthrich, K. (1977) ^1H NMR titration shifts of amide proton resonances in polypeptide chains. *FEBS Lett.* **77**, 11–14.
- Burling, F.T., Weis, W.I., Flaherty, K.M. & Brünger, A.T. (1996) Direct observation of protein solvation and discrete disorder with experimental crystallographic phases. *Science*. **271**, 72–77.
- Cacace, M.G., Landau, E.M. & Rasmsden, J.J. (1997) The Hofmeister series: salt and solvent effects on interfacial phenomena. *Q. Rev. Biophys.* **30**, 241–277.
- Casimiro, D.R., Wright, P.E. & Dyson, H.J. (1997) PCR-based gene synthesis and protein NMR spectroscopy. *Structure* **5**, 1407–1412.
- Chamberlin, M., McGrath, J. & Waskell, L. (1970) New RNA polymerase from *Escherichia coli* infected with bacteriophage T7. *Nature* **228**, 227–231.
- Chary, K.V.R., Otting, G. & Wüthrich, K. (1991) Measurement of small heteronuclear ^1H - ^{15}N coupling constants in ^{15}N -labeled proteins by 3D HNNHAB-COSY. *J. Magn. Reson.* **93**, 218–224.
- Chyan, C.L., Wormald, C., Dobson, C.M., Evans, P.A. & Baum, J. (1993) Structure and stability of the molten globule state of guinea-pig alpha-lactalbumin: a hydrogen exchange study. *Biochemistry* **32**, 5681–5691.
- Clore, G.M., Bax, A., Wingfield, P.T. & Gronenborn, A.M. (1990) Identification and localization of bound internal water in the solution structure of interleukin 1β by heteronuclear three-dimensional ^1H rotating-frame Overhauser ^{15}N - ^1H multiple quantum coherence NMR spec-

- troscopy. *Biochemistry* **29**, 5671–5676.
- Clore, G.M. & Gronenborn, A.M. (1992) Localization of bound water in the solution structure of the immunoglobulin binding domain of streptococcal protein G. Evidence for solvent-induced helical distortion in solution. *J.Mol.Biol.* **223**, 853–856.
- Clore, G.M., Bax, A., Omichinski, J.G. & Gronenborn, A.M. (1994) Localization of bound water in the solution structure of a complex of the erythroid transcription factor GATA-1 with DNA. *Structure* **2**, 89–94.
- Collins, K.D. & Washabaugh, M.W. (1985) The Hofmeister effect and the behaviour of water at interface. *Q. Rev. Biophys.* **18**, 323–422.
- Dalvit, C. & Hommel, U. (1995) New pulsed field gradient NMR experiments for the detection of bound water in proteins. *J. Biomol. NMR* **5**, 306–310.
- Damberger, F.F., Pelton, J.G., Liu, C., Cho, H., Harrison, C.J., Nelson, H.C. & Wemmer, D.E. (1995) Refined solution structure and dynamics of the DNA-binding domain of the heat shock factor from *Kluyveromyces lactis*. *J. Mol. Biol.* **254**, 704–719.
- Das, G., Hickey, D.R., McLendon, D., McLendon, G. & Sherman, F. (1989) Dramatic thermostability of yeast iso-1-cytochrome C by an asparagine → isoleucine replacement at position 57. *Proc. Natl. Acad. Sci. USA.* **86**, 496–499.
- Davenas, E., Beauvais, F., Amara, J., Oberbaum, M., Robinzon, B., Miadonna, A., Tedeschi, A., Pomeranz, B., Fortner, P., Belon, P., Sainte-Laudy, J., Poitevin, B. & Benveniste, J. (1988) Human basophil degranulation triggered by very dilute antiserum against IgE. *Nature* **333**, 816–818.
- De Anda, J., Poteete, A.R. & Sauer, R.T. (1983) P22c2 repressor. Domain structure and function. *J. Biol. Chem.* **258**, 10536–10542.
- Denisov, V.P. & Halle, B. (1996) Protein hydration dynamics in aqueous solution. *Faraday Discuss.* **103**, 227–244.
- Denisov, V.P. & Halle, B. (1995) Protein hydration dynamics in aqueous solution: a comparison of bovine pancreatic trypsin inhibitor and ubiquitin by oxygen-17 spin relaxation dispersion. *J. Mol. Biol.* **245**, 682–697.
- Dillon, P.J. & Rosen, C.A. (1990) A rapid method for the construction of synthetic genes using the polymerase chain reaction. *BioTechniques* **9**, 298–300.
- Dötsch, V. (1994) Charakterisierung der Wechselwirkungen zwischen Proteinen und Lösungsmittelmolekülen. Diss. ETH Nr. 10856.
- Dötsch, V., Wider, G., Siegal, G. & Wüthrich, K. (1995a) Interaction of urea with an unfolded protein. The DNA-binding domain of the 434-repressor. *FEBS Lett.* **366**, 6–10.
- Dötsch, V. & Wider, G. (1995b) Exchange rates of internal water molecules in proteins measured using pulsed field gradients. *J. Am. Chem. Soc.* **117**, 6064–6070.
- Dunn, J.J. & Studier, F.W. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **166**, 477–535.
- Dunitz, J.D. (1994) The entropic cost of bound water in crystals and biomolecules. *Science* **264**, 670.
- Dyson, H.J. & Wright, P.E. (1993) Peptide conformation and protein folding. *Curr.Opin. Struct. Biol.* **3**, 60–65.
- Edelhoch, H. & Osborne, J.C. (1976) The thermodynamic basis of the stability of proteins, nucleic acids, and membranes. *Adv. Protein Chem.* **30**, 183–250.
- Eisenberg, D. & Kauzmann, W. (1969) *The Structure and Properties of Water*, Oxford University press, New York.
- Eisenberg, D. & McLachlan, A.D. (1986) Solvation energy in protein folding and binding. *Nature* **319**, 99–203.
- Englander, S.W., Downer, N.W. & Teitelbaum, H. (1972) Hydrogen exchange. *Annu. Rev. Biochem.* **41**, 903–924.
- Englander, S.W. & Kallenbach, N.R. (1989) Hydrogen exchange and structural dynamics of protein and nucleic acids. *Q. Rev. Biophys.* **22**, 521–655.

- Ernst, J.A., Clubb, R.T., Zhou, H.X., Gronenborn, A.M. & Clore, G.M. (1995a) Demonstration of positionally disordered water within a protein hydrophobic cavity by NMR. *Science* **267**, 1813–1817.
- Ernst, J.A., Clubb, R.T., Zhou, H.X., Gronenborn, A.M. & Clore, G.M. (1995b) Use of NMR to detect water within nonpolar protein cavities. *response. Science*. **270**, 1848–1849.
- Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D. & Kay, L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* **33**, 5984–6003.
- Fersht, A.R. (1987) The hydrogen bond in molecular recognition. *Trends. Biochem. Sci.* **12**, 301–304.
- Fersht, A.R. & Serrano, L. (1993) Principles of protein stability derived from protein engineering experiments. *Curr. Opin. Struct. Biol.* **3**, 75–83.
- Finer-Moore, J.S., Kossiakoff, A.A., Hurley, J.H., Earnest, T. & Stroud, R.M. (1992) Solvent structure in crystals of trypsin determined by X-ray and neutron diffraction. *Proteins* **12**, 203–222.
- Flaman, J.M., Frebourg, T., Moreau, V., Charbonnier, F., Martin, C., Ishioka, C., Friend, S.H. & Iggó, R. (1994) A rapid PCR fidelity assay. *Nucleic Acids Res.* **22**, 3259–3260.
- Fontana, A., Fassina, G., Vita, C., Dalzoppo, D., Zamai, M. & Zambonin, M. (1986) Correlation between sites of limited proteolysis and segmental mobility in thermolysin. *Biochemistry* **25**, 1847–1851.
- Fontana, A., Vita, C., Dalzoppo, D. & Zambonin, M. (1989) Limited proteolysis as a tool to detect structure and dynamics features of globular proteins: studies on thermolysin. In *Methods in protein sequence analysis. Wittman-Liebold, B. (ed.)* pp 315–324. Springer Verlag, Berlin.
- Forman-Kay, J.D., Gronenborn, A.M., Wingfield, P.T. & Clore, G.M. (1991) Determination of the positions of bound water molecules in the solution structure of reduced human thioredoxin by heteronuclear three-dimensional nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **220**, 209–216.
- Frank, H.S. & Evans, M.W. (1945) Free volume and entropy in condensed systems. *J. Chem. Phys.* **13**, 507–527.
- Franks, F. (ed.) (1972-83) *Water: A comprehensive Treatise*, Volume 1-6, Plenum Press, New York.
- Geen, H. & Freeman, R. (1991) Band-selective radiofrequency pulses. *J. Mag. Reson.* **93**, 93–141.
- Gemmecker, G., Jahnke, W. & Kessler, H. (1993) Measurement of fast proton exchange rates in isotopically labelled compounds. *J. Am. Chem. Soc.* **115**, 11620–11621.
- Gerothanassis, I.P. (1994) Multinuclear and multidimensional NMR methodology for studying individual water molecules bounds to peptides and proteins in solution: principles and applications. *Prog. Nucl. Magn. Reson. Spectrosc.* **26**, 171–237.
- Glushko, V., Lawson, P.J. & Gurd, F.R. (1972) Conformational states of bovine pancreatic ribonuclease A observed by normal and partially relaxed carbon 13 nuclear magnetic resonance. *J. Biol. Chem.* **247**, 3176–3185.
- Goldman, A. (1995) How to make my blood boil. *Structure* **3**, 1277–1279.
- Graham, R.W., Atkinson, T., Kilburn, D.G., Miller, R.C. Jr. & Warren, R.A. (1993) Rational design and PCR-based synthesis of an artificial Schizophyllum commune xylanase gene. *Nucleic Acids Res.* **21**, 4923–4928.
- Gregory, R.B. (ed.) (1995) *Protein-solvent interactions*. Marcel Dekker, Inc.
- Griesinger, C., Sørensen, O.W. & Ernst, R.R. (1985) Two-dimensional correlation of connected NMR transitions. *J. Am. Chem. Soc.* **107**, 6394–6396.
- Griesinger, C., Otting, G., Wüthrich, K. & Ernst, R.R. (1988) Clean-TOCSY for ^1H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110**, 7870–7872.
- Grzesiek, S. & Bax, A. (1993a) The importance of not saturating water in protein NMR. Applications to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**, 12593–12594.
- Grzesiek, S. & Bax, A. (1993b) Measurements of amide proton exchange rates and NOEs with water

- in $^{13}\text{C}/^{15}\text{N}$ -enriched calineurin B. *J. Biomol. NMR* **3**, 627–638.
- Grzesiek, S., Bax, A., Nicholson, L.K., Yamazaki, T., Wingfield, P.T., Stahl, S.J., Eyermann, C.J., Torchia, D.A., Hodge, C.N., Lam, P.Y.S., Jadhav, P.K. & Chang, C.-H. (1994) NMR evidence for the displacement of conserved interior water molecule in HIV protease by non-peptide cyclic urea based inhibitor. *J. Am. Chem. Soc.* **116**, 1581–1582.
- Güntert, P., Braun, W. & Wüthrich, K. (1991a) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517–530.
- Güntert, P. & Wüthrich, K. (1991b) Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J. Biomol. NMR* **1**, 447–456.
- Güntert, P., Dötsch, V., Wider, G. & Wüthrich, K. (1992) Processing of multi-dimensional NMR data with the new software PROSA. *J. Biomol. NMR* **2**, 619–629.
- Güntert, P., Berndt, K.D. & Wüthrich, K. (1993) The program ASNO for computer-supported collection of NOE upper distance constraints as input for protein structure determination. *J. Biomol. NMR* **3**, 601–606.
- Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298.
- Hagen, S.J., Hofrichter, J., Szabo, A., & Eaton, W.A. (1996) Diffusion-limited contact formation in unfolded cytochrome c: estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. USA.* **15**, 11615–11617.
- Havlin, R. H., Le, H., Laws, D.D., deDios, A.C. & Oldfield, E. (1997) An *ab initio* quantum chemical investigation of carbon-13 NMR shielding tensors in glycine, alanine, valine, isoleucine, serine, and threonine: comparisons between helical and sheet tensors, and the effects of χ^1 on shielding. *J. Am. Chem. Soc.* **119**, 11951–11958.
- Harrison, S.C. & Aggarwal, A.K. (1990) DNA recognition by proteins with the helix-turn-helix motif. *Annu. Rev. Biochem.* **59**, 933–969.
- Haruyama, H., Qian, Y.Q. & Wüthrich, K. (1989) Static and transient hydrogen-bonding interactions in recombinant desulfatohirudin studied by ^1H nuclear magnetic resonance measurements of amide proton exchange rates and pH-dependent chemical shifts. *Biochemistry* **28**, 4312–4317.
- Hamada, D. & Goto, Y. (1997) The equilibrium intermediate of beta-lactoglobulin with non-native alpha-helical structure. *J. Mol. Biol.* **269**, 479–487.
- Hagihara, Y., Aimoto, S., Fink, A.L. & Goto, Y. (1993) Guanidine hydrochloride-induced folding of proteins. *J. Mol. Biol.* **231**, 180–184.
- Hendsch, Z.S. & Tidor, B. (1994) Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* **3**, 211–226.
- Hiroaki, H., Klaus, W. & Senn, H. (1996) Determination of the solution structure of the SH3 domain of human p56 Lck tyrosine kinase. *J. Biomol. NMR.* **8**, 105–122.
- Higuchi, R. (1991) Recombinant PCR. In *PCR protocols: A guide to methods and application.* **22**, 177–183.
- Hofmeister, F. (1888) On the understanding of the effect of salts. Second report. On regularities in the precipitating effect of salts and their relationship to their physiological behavior. *Nauym-Schmiedebergs Archiv für Experimentelle Pathologie und Pharmakologie (Leipzig)* **24**, 247–260.
- Hughson, F.M., Wright, P.E. & Baldwin, R.L. (1990) Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249**, 1544–1548.
- Ikura, M., Kay, L.E. & Bax, A. (1991) Improved three dimensional ^1H - ^{13}C - ^1H correlation spectroscopy of a ^{13}C -labeled protein using constant-time evolution. *J. Biomol. NMR* **1**, 299–304.
- Ito, W., Ishiguro, H. & Kurosawa, Y. (1991) A general method for introducing a series of mutations into cloned DNA using the polymerase chain reaction. *Gene* **102**, 67–70.

- Jacobson, B., Anderson, W.A. & Arnold, J.T. (1954) A proton magnetic resonance study of the hydration of deoxyribonucleic acid. *Nature* **173**, 772–773.
- James, M.N. & Sielecki, A.R. (1983) Structure and refinement of penicillopepsin at 1.8 Å resolution. *J. Mol. Biol.* **163**, 299–361.
- Janin, J. & Chothia, C. (1990) The structure of protein-protein recognition sites. *J. Biol. Chem.* **25**, 16027–16030.
- Jasanoff, A. & Fersht, A.R. (1994) Quantitative determination of helical propensities from trifluoroethanol titration curves. *Biochemistry* **33**, 2129–2135.
- Jeffrey, G.A. & Saenger, W. (1991) *Hydrogen bonding in biological structures*. Springer-Verlag, Berlin.
- Jeng, M.F. & Dyson, H. J. (1996) Direct measurement of the aspartic acid 26 pKa for reduced *Escherichia coli* thioredoxin by ¹³C NMR. *Biochemistry* **35**, 1–6.
- Jeng, M.F. & Englander, S.W. (1991) Stable submolecular folding units in a non-compact form of cytochrome c. *J. Mol. Biol.* **221**, 1045–1061.
- Jeng, M.F., Englander, S.W., Elove G.A., Wand, A.J. & Roder, H. (1990) Structural description of acid-denatured cytochrome c by hydrogen exchange and 2D NMR. *Biochemistry* **29**, 10433–10437.
- Kay, L.E. Keifer, P. & Saarinen, T. (1992) Pure Absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivities. *J. Am. Chem. Soc.* **114**, 10663.
- Kane, J.F. (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* **6**, 494–500.
- Karplus, P.A. & Faerman, C. (1994) Ordered water in macromolecular structure. *Curr. Opin. Struct. Biol.* **4**, 770–776.
- Kauzmann, W. (1959) Some factors in the interaction of protein denaturation. *Adv. Protein Chem.* 1–63.
- Kim, K.S. & Woodward, C. (1993) Protein internal flexibility and global stability: effect of urea on hydrogen exchange rates of bovine pancreatic trypsin inhibitor. *Biochemistry* **32**, 9609–9613.
- Klotz, I.M. (1958) Protein hydration and behavior. *Science* **128**, 815–822.
- Koenig, S.H. & Schillinger, W.E. (1969) Nuclear magnetic relaxation dispersion in protein solutions. *J. Biol. Chem.* **244**, 3283–3289.
- Koradi, R., Billeter, M. & Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.
- Kriwacki, R.W., Hill, R.B., Flanagan, J.M., Caradonna, J.P., & Prestegard, J.H. (1993) New NMR methods for the characterization of bound waters on macromolecules. *J. Am. Chem. Soc.* **115**, 8907–8911.
- Kubinec, M.G. & Wemmer, D.E. (1992) Solution NMR of water on protein surfaces. *Curr. Opin. Struct. Biol.* **2**, 828–831.
- Kuhn, L.A., Siani, M. A., Pique, M.E., Fisher, C.L., Getzoff, E.D. & Tainer, J.A. (1992) The interdependence of protein surface topography and bound water molecules revealed by surface accessibility and fractal density measures. *J. Mol. Biol.* **228**, 13–22.
- Kumaraswamy, V.S., Lindley, P. F., Slingsby, C., & Glover, I.D. (1996) An eye lens protein-water structure: 1.2 Å resolution structure of γ B-crystallin at 150K. *Acta Cryst.* **D52**, 611–622.
- Kunkel, T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA.* **82**, 488–492.
- Kuntz, I.D. & Kauzmann, W. (1974) Hydration of proteins and polypeptides. *Adv. Protein Chem.* 239–345.
- Langsetmo, K., Fuchs, J.A. & Woodward, C. & Sharp, K.A. (1991a) Linkage of thioredoxin stability to titration of ionizable groups with perturbed pKa. *Biochemistry* **30**, 7609–7614.
- Langsetmo, K., Fuchs, J.A. & Woodward, C. (1991b) The conserved, buried aspartic acid in oxidized *Escherichia coli* thioredoxin has a pKa of 7.5. Its titration produces a related shift in global

- stability. *Biochemistry* **30**, 7603–7609.
- Leberman, R. & Soper, A.K. (1995) Effect of high salt concentration on water structure. *Nature* **378**, 369–366.
- Lee, K.H., Xie, D., Freire, E. & Amzel, L.M. (1994) Estimation of changes in side chain configurational entropy in binding and folding: general methods and application to helix formation. *Proteins* **20**, 68–84.
- Levitt, M., & Park, B.H. (1993) Water: now you see it, now you don't. *Structure* **1**, 223–226.
- Liang, H., Petros, A.M., Meadows, R.P., Yoon, H.S., Egan, D.A., Walter, K., Holzman, T.F., Robins, T. & Fesik, S.W. (1996) Solution structure of the DNA-binding domain of a human papilloma-virus E2 protein: evidence for flexible DNA-binding regions. *Biochemistry* **35**, 2095–2103.
- Liepinsh, E., Otting, G. & Wüthrich, K. (1992) NMR spectroscopy of hydroxyl protons in aqueous solutions of peptides and proteins. *J. Biomol. NMR* **2**, 447–465.
- Liepinsh, E., Rink, H., Otting, G. & Wüthrich, K. (1993) Contribution from hydration of carboxylate groups to the spectrum of water-polypeptide proton-proton Overhauser effects in aqueous solution. *J. Biomol. NMR* **3**, 253–257.
- Liepinsh, E. & Otting, G. (1995) Selective excitation of intense solvent signals in the presence of radiation damping. *J. Biomol. NMR* **5**, 420–426.
- Logan, T.M., Olejniczak, E.T., Xu, R.X. & Fesik, S.W. (1993) A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. *J. Biomol. NMR* **3**, 225–231.
- Luginbühl, P., Szyperski, T. & Wüthrich, K. (1995) Statistical basis for the use of $^{13}\text{C}^\alpha$ chemical shifts in protein structure determination. *J. Magn. Reson. B* **109**, 229–233.
- Luginbühl, P., Güntert, P., Billeter, M. & Wüthrich, K. (1996) The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR* **8**, 136–146.
- Luginbühl, P., Pervushin, K., Iwai, H. & Wüthrich, K. (1997) Anisotropic molecular rotational diffusion in ^{15}N spin relaxation studies of protein mobility. *Biochemistry* **36**, 7305–7312.
- Macura, S., Wüthrich, K. & Ernst, R.R. (1982) Separation and suppression of coherent transfer effects in two-dimensional NOE and chemical exchange spectroscopy. *J. Magn. Reson.* **46**, 269–282.
- Makhatadze, G.I. & Privalov, P.L. (1992) Protein interactions with urea and guanidinium chloride. A calorimetric study. *J. Mol. Biol.* **226**, 491–505.
- Makhatadze, G.I. & Privalov, P.L. (1996) On the entropy of protein folding. *Protein Sci.* **5**, 507–510.
- Marion, D., Ikura, K., Tschudin, R. & Bax, A. (1989) Rapid recording of 2D NMR spectra without phase cycling, application to the study of hydrogen exchange in proteins. *J. Magn. Reson.* **85**, 393–399.
- Marqusee, S. & Sauer, R.T. (1994) Contributions of a hydrogen bond/salt bridge network to the stability of secondary and tertiary structure in lambda repressor. *Protein Sci.* **3**, 2217–2225.
- Mayr, L.M. & Schmid, F.X. (1993) Stabilization of a protein by guanidinium chloride. *Biochemistry* **32**, 7994–7998.
- Matthews, B.W., Morton, A.G., Dahlquist, F.W. (1995) Use of NMR to detect water within nonpolar protein cavities. *Science.* **270**, 1847–1848.
- Matthews, B.W. (1996) Structural and genetic analysis of the folding and function of T4 lysozyme. *FASEB J.* **10**, 35–41.
- McIntosh, L.P., Hand, G., Johnson, P.E., Joshi, M.D., Korner, M., Plesniak, L.A., Ziser, L., Wakarchuk, W.W. & Withers, S.G. (1996) The pKa of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a ^{13}C -NMR study of bacillus circulans xylanase. *Biochemistry* **35**, 9958–9966.
- Mondragon, A., Subbiah, S., Almo, S.C., Drott, M. & Harrison, S.C. (1989) Structure of the amino-terminal domain of phage 434 repressor at 2.0 Å resolution. *J. Mol. Biol.* **205**, 189–201.

- Mori, S., Johnson, M.O., Berg, J.M. & Van Zijl, P.C.M. (1994) Water exchange filter (WEX Filter) for nuclear magnetic resonance studies of macromolecules. *J. Am. Chem. Soc.*, **116**, 11982–11984.
- Mori, S., Berg, J.M. & Van Zijl, P.C.M. (1996) Separation of intramolecular NOE and exchange peaks in water exchange spectroscopy using spin-echo filters. *J. Biomol. NMR* **7**, 77–82
- Mrabet, N.T., Van den Broeck, A., Van den Brande, I., Stanssens, P., Laroche, Y., Lambeir, A.M., Matthijssens, G., Jenkins, J., Chiadmi, M., Van Tilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., De Maeyer, M. & Wodak, S.J. (1992) Arginine residues as stabilizing elements in proteins. *Biochemistry* **31**, 2239–2253.
- Murphy, K.P. & Freire, E. (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Protein Chem.* **43**, 313–361.
- Murphy, K.P., Freire, E. & Paterson, Y. (1995) Configurational effects in antibody-antigen interactions studied by microcalorimetry. *Proteins*. **21**, 83–90.
- Myers, J.K., Pace, C. N. & Scholtz, J.M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4** 2138–2148.
- Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. (1989) Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry* **28**, 7510–7516.
- Neri, D., Otting, G. & Wüthrich, K. (1990) ^1H and ^{13}C NMR chemical shifts of the diastereotopic methyl groups of valyl and leucyl residues in peptides and proteins. *Tetrahedron* **46**, 3287–3296.
- Neri, D., Billeter, M. & Wüthrich, K. (1992a) Determination of the NMR solution structure of the DNA-binding domain 1–69 of the 434 repressor and comparison with the X-ray crystal structure. *J. Mol. Biol.* **223**, 743–767.
- Neri, D., Billeter, M., Wider, G. & Wüthrich, K. (1992b) NMR determination of residual structure in a urea-denatured protein, the 434 repressor. *Science* **257**, 1559–1563.
- Novotny, J. & Brucoleri, R.E. (1987) Correlation among sites of limited proteolysis, enzyme accessibility and segmental mobility. *FEBS Lett.* **211**, 185–189.
- Nozaki, Y. & Tanford, C. (1971) The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. *J. Biol. Chem.* **246**, 2211–2217.
- Oda, Y., Yamazaki, T., Nagayama, K., Kanaya, S., Kuroda, Y. & Nakamura, H. (1994) Individual ionization constants of all the carboxyl groups in ribonuclease HI from *Escherichia coli* determined by NMR. *Biochemistry* **33**, 5275–5284.
- Ottiger, M., Zerbe, O., Guentert, P. & Wüthrich, K. (1997) The NMR solution conformation of unligated human cyclophilin A. *J. Mol. Biol.* **272**, 64–81.
- Otting, G. & Wüthrich, K. (1989) Studies of protein hydration in aqueous solution by direct NMR observation of individual protein-bound water molecules. *J. Am.Chem.Soc.* **111**, 1871–1875.
- Otting, G., Liepinsh, E. & Wüthrich, K. (1991a) Protein hydration in aqueous solution. *Science* **271**, 974–980.
- Otting, G., Liepinsh, E., Farmer II, B.T. & Wüthrich, K. (1991b) Protein hydration studied with homonuclear 3D ^1H NMR experiments. *J. Biomol. NMR* **1**, 209–215.
- Otting, G., Liepinsh, E. & Wüthrich, K. (1991c) Proton exchange with internal water molecules in the protein BPTI in aqueous solution. *J. Am. Chem. Soc.* **113**, 4363–4364.
- Otting, G., Liepinsh, E. & Wüthrich, K. (1992) Polypeptide hydration in mixed solvent at low temperature. *J. Am. Chem. Soc.* **114**, 7093–7085.
- Otting, G. & Liepinsh, E. (1995a) Protein hydration viewed by high-resolution NMR spectroscopy: Implications for magnetic resonance image contrast. *Acc. Chem. Res.* **28**, 171–177.
- Otting, G. & Liepinsh, E. (1995b) *J. Biomol. NMR* **5**, 420–426.
- Pace, C.N., K., Shirley, B.A., & Thomson, J.A. (1989) Measuring the conformational stability of a

- protein. In *Protein Structure, a practical approach* (Creighton, T.E. ed.), pp 699–711. IRL press, Oxford.
- Padmanabhan, S., Jimenez, M.A., Gonzalez, C., Sanz, J.M., Gimenez-Gallego, G. & Rico, M. (1997) Three-dimensional solution structure and stability of phage 434 Cro protein. *Biochemistry* **36**, 6424–6436.
- Pellecchia, M., Szyperski, T., Wall, D., Georgopoulos, C. & Wüthrich, K. (1996) NMR structure of the J-domain and the Gly/Phe-rich region of the *Escherichia coli* DnaJ chaperone. *J. Mol. Biol.* **260**, 236–250.
- Pellecchia, M., Iwai, H., Szyperski, T. & Wüthrich, K. (1997a) The 2D NMR experiments H(C)CO₂ and HCCO₂ for assignment and pH titration of carboxylate groups in uniformly ¹⁵N/¹³C-labeled proteins. *J. Magn. Reson.* **124**, 274–278.
- Pellecchia, M., Wider, G., Iwai, H. & Wüthrich, K. (1997b) Arginine side chain assignments in uniformly ¹⁵N-labeled proteins using the novel 2D HE(NE)HGHH experiment. *J. Biomol. NMR* **10**, 193–197.
- Pearson, J.G., Le, H., Sanders, L.K., Godbout, N., Havlin, R.H. & Oldfield, E. (1997) Predicting chemical shifts in proteins: structure refinement of valine residues by using *ab Initio* and empirical geometry optimizations. *J. Am. Chem. Soc.* **119**, 11941–11950.
- Peng, J.W. & Wagner, G. (1994) Investigation of protein motions via relaxation measurements. *Methods in Enzymology* **239**, 563–593.
- Pertsemlidis, A., Saxena, A.M., Soper, A.K., Head-Gordon, T. & Glaeser, R.M. (1996) Direct evidence for modified solvent structure within the hydration shell of a hydrophobic amino acid. *Proc. Natl. Acad. Sci. USA*, **93**, 10769–10774
- Perrin, C.L. (1989) Proton exchange in amides: Surprises from simple systems. *Acc. Chem. Res.* **22**, 268–275.
- Peruz, M.F. (1989) Mechanisms of cooperativity and allosteric regulation in proteins. *Q. Rev. Biophys.* **22**, 139–236.
- Pervushin, K., Billeter, M., Siegal, G. & Wüthrich, K. (1996) Structural role of a buried salt bridge in the 434 repressor DNA-binding domain. *J. Mol. Biol.* **264**, 1002–1012.
- Pitner, T.P., Glickson, J.D., Dadok, J. & Marshall, G.R. (1974) Solvent exposure specific nuclei of angiotensin II determined by NMR solvent saturation method. *Nature* **250**, 582–584.
- Ponder, J.W. & Richards, F.M. (1987) Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* **193**, 775–791.
- Privalov, P.L. (1979) Stability of proteins: small globular proteins. *Adv. Protein Chem.* **33**, 167–241.
- Ptitsyn, O.B. (1992) The molten globule state. In *Protein folding*. Creighton T.E (ed.), Freeman W.H. & company, New York.
- Qi, P.X., Urbauer, J.L., Fuentes, E.J., Leopold, M.F. & Wand, A.J. (1994) Structural water in oxidized and reduced horse heart cytochrome c. *Nat. Struct. Biol.* **1**, 378–382.
- Qian, Y.Q., Otting, G. & Wüthrich, K. (1993) NMR detection of hydration water in the intermolecular interface of a protein-DNA complex. *J. Am. Chem. Soc.* **115**, 1189–1190.
- Qin, J., Clore, G.M. & Gronenborn, A.M. (1994) The high-resolution three-dimensional solution structures of the oxidized and reduced states of human thioredoxin. *Structure* **15**, 503–522.
- Qin, J., Clore, G.M. & Gronenborn, A.M. (1996) Ionization equilibria for side-chain carboxyl groups in oxidized and reduced human thioredoxin and in the complex with its target peptide from the transcription factor NF kappa B. *Biochemistry* **35**, 7–13.
- Radford, S.E., Dobson C.M. & Evans, P.A. (1992) The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **358**, 302–307.
- Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. & Wüthrich, K. (1983) Improved spectral resolution in COSY ¹H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Comm.* **117**, 479–485.
- Richardson, J.S. & Richardson, D.C. (1988) Amino acid preferences for specific locations at the end of helices. *Science* **240**, 1648–1652.

- Richarz, R., Nagayama, K. & Wüthrich, K. (1980) Carbon-13 nuclear magnetic resonance relaxation studies of internal mobility of the polypeptide chain in basic pancreatic trypsin inhibitor and a selectively reduced analogue. *Biochemistry* **19**, 5189–5196.
- Richmond, T.J. (1984) Solvent-accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.* **178**, 63–89.
- Roder, H. & Wüthrich, K. (1986) Protein folding kinetics by combined use of rapid mixing techniques and NMR observation of individual amide protons. *Proteins* **1**, 34–42.
- Rousseau, D.L. (1971) "Polywater" and sweat: similarities between the infrared spectra. *Science* **171**, 170–172.
- Rupley, J.A. & Careri, G. (1991) Protein hydration and function. *Adv. Protein Chem.* **41**, 37–172.
- Sauer, R.T., Pan, J., Hopper, P., Hehir, K., Brown, J. & Poteete, A.R. (1981) Primary structure of the phage P22 repressor and its gene *c2*. *Biochemistry* **20**, 3591–3598.
- Saenger, W. (1987) Structure and dynamics of water surrounding biomolecules. *Ann. Rev. Biophys. Biophys. Chem.* **16**, 93–114.
- Santoro, M.M. & Bolen, D.W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068.
- Savage, H. & Wlodawer, A. (1986) Determination of water structure around biomolecules using X-ray and neutron diffraction methods. *Methods in Enzymology* **127**, 162–183.
- Schellman J.A. (1978) Solvent denaturation. *Biopolymers* **17**, 1305–1322.
- Schirmer, T. & Evans, P.R. (1990) Structural basis of the allosteric behaviour of phosphofructokinase. *Nature* **343**, 140–145.
- Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S.J., Sorensen, O.W. & Griesinger, C. (1994) A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR.* **4**, 301–306.
- Schueler, O. & Margalit, H. (1995) Conservation of salt bridges in protein families. *J. Mol. Biol.* **248**, 125–135.
- Senn, H., Werner, B., Messerle, B.A., Weber, C., Traber, R. & Wüthrich, K. (1989) Stereospecific assignment of the methyl ^1H NMR lines of valine and leucine in polypeptides by nonrandom ^{13}C labelling. *FEBS Lett.* **249**, 113–118.
- Sevilla-Sierra, P., Otting, G. & Wüthrich, K. (1994) Determination of the nuclear magnetic resonance structure of the DNA-binding domain of the P22 *c2* repressor (1 to 76) in solution and comparison with the DNA-binding domain of the 434 repressor. *J. Mol. Biol.* **235**, 1003–1020.
- Shaka, J.A., Lee, C.J. & Pines, A. (1988) Iterative schemes for bilinear operators: application to spin decoupling. *J. Magn. Reson.* **77**, 274.
- Shaka, J.A., Barker, P.B. & Freeman, R. (1985) Computer-optimized decoupling scheme for wide-band applications and low level operation. *J. Magn. Reson.* **64**, 547–552.
- Shang, Z., Isaac, V.E., Li, H., Patel, L., Catron, K.M., Curran, T., Montelione, G.T. & Abate, C. (1994) Design of a "minimAl" homeodomain: the N-terminal arm modulates DNA binding affinity and stabilizes homeodomain structure. *Proc. Natl. Acad. Sci. USA.* **91**, 8373–8377.
- Sobol, A.G., Wider, G., Iwai, H. & Wüthrich, K. Solvent magnetization artifacts in high field NMR studies of macromolecular hydration. (1998) *J. Magn. Reson.* **130**, 262–271.
- Sönnichsen, F.D., Van Eyk, J.E., Hodges, R.S. & Sykes, B.D. (1992) Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochemistry* **31**, 8790–8798.
- Spera, S. & Bax, A. (1991) Empirical correlation between protein backbone conformation and $\text{C}\alpha$ and $\text{C}\beta$ ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**, 5490–5492.
- Spolar, R.S. & Record, M.T. Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA. *Science* **263**, 777–784.
- Sundaralingam, M. & Sekharudu, Y.C. (1989) Water-inserted alpha-helical segments implicate re-

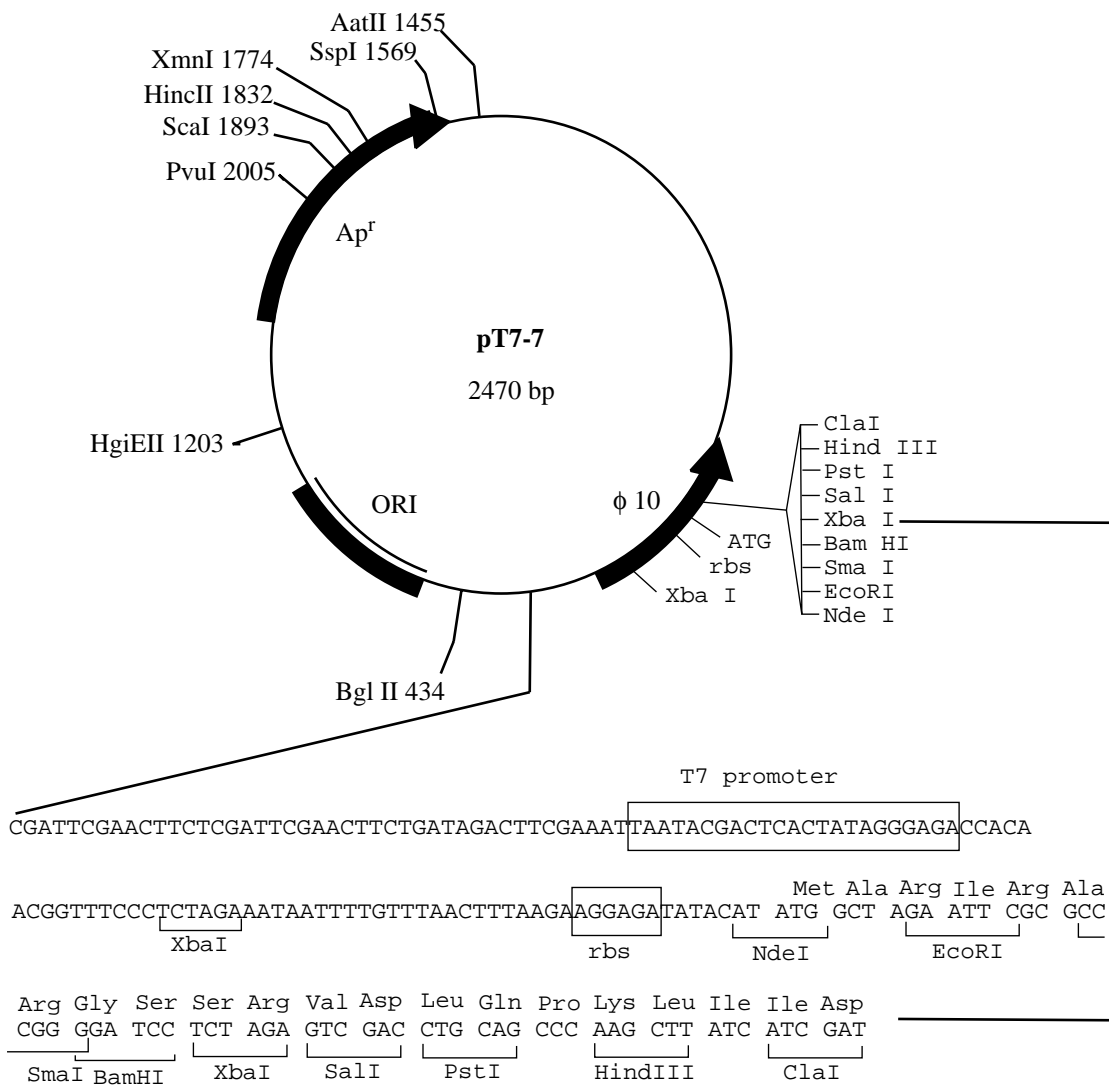
- verse turns as folding intermediates. *Science* **244**, 1333–1337.
- Szyperski, T., Güntert, P., Otting, G. & Wüthrich, K. (1992) Determination of scalar coupling constants by inverse Fourier transformation of in-phase multiplets. *J. Magn. Reson.* **99**, 552–560.
- Szyperski, T., Wider, G., Bushweller, J.H. & Wüthrich, K. (1993a) 3D ^{13}C - ^{15}N -heteronuclear two-spin coherence spectroscopy for polypeptide backbone assignments in ^{13}C - ^{15}N -double-labeled proteins. *J. Biomol. NMR* **3**, 127–132.
- Szyperski, T., Wider, G. Bushweller, J.H. & Wüthrich, K. (1993b) Reduced dimensionality in triple-resonance NMR experiments. *J. Am. Chem. Soc.* **115**, 9307–9308.
- Szyperski, T., Antuch, W., Schick, M., Betz, A., Stone, S.R. & Wüthrich, K. (1994) Transient hydrogen bonds identified on the surface of the NMR solution structure of hirudin. *Biochemistry* **33**, 9303–9310.
- Tabor, S. & Richardson, C.C. (1985) A bacteriophage T7 RNA polymerase/polymerase system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci.* **82**, 1074–1078.
- Tabor, S. (1990) Expression using the T7 RNA polymerase/promotor system. In *Current Protocols in molecular biology* (Ausubel, F.A., Brent, R., Kingston, R.E., Moore, D.D., Seidmann, J.G., Smith, J.A. & Struhl, K.) pp 16.2.1–16.2.11, Greene Publishing and Wiley Interscience.
- Tanford, C. (1968) Protein denaturation. Part A. Characterization of the denatured state. *Adv. Protein Chem.* **23**, 121–217.
- Tanford, C. (1980) *The hydrophobic effect*. Wiley, New York.
- Takano, K., Funahashi, J., Yamagata, Y., Fujii, S. & Yutani, K. (1997) Contribution of water molecules in the interior of a protein to the conformational stability. *J. Mol. Biol.* **274**, 132–142.
- Tanaka, T., Hayashi, M., Kimura, H., Oobatake, M. & Nakamura, H. (1994) De novo design and creation of a stable artificial protein. *Biophys. Chem.* **50**, 47–61.
- Teeter M. M. (1984) Water structure of a hydrophobic protein at atomic resolution: pentagon rings of water molecules in crystals of crambin. *Proc. Natl. Acad. Sci. USA.* **81**, 6014–6018.
- Teeter, M.M., Roe, S.M. & Heo, N.H. (1993) Atomic resolution (0.83 Å) crystal structure of the hydrophobic protein crambin at 130 K. *J. Mol. Biol.* **230**, 292–311.
- Thanki, N., Umrانيا, Y., Thornton, J.M. & Goodfellow, J.M. (1988) Distributions of water around amino acid residues in proteins. *J. Mol. Biol.* **202**, 637–657.
- Thanki, N., Umrانيا, Y., Thornton J.M. & Goodfellow, J.M. (1991) Analysis of protein main-chain solvation as a function of secondary structure. *J. Mol. Biol.* **221**, 669–691.
- Timasheff, S.N. & Arakawa, T. (1989) Stabilization of protein structure by solvents. In *Protein Structure, a practical approach* (Creighton, T.E. ed.) pp331–345. IRL press, Oxford.
- Tissot, A.C., Vuilleumier, S. & Fersht, A.R. (1996) Importance of two buried salt bridges in the stability and folding pathway of barnase. *Biochemistry* **35**, 6786–6794.
- Tjandra, N., Kuboniwa, H., Ren, H. & Bax, A. (1995) Rotational dynamics of calcium-free calmodulin studied by ^{15}N -NMR relaxation measurements. *Eur. J. Biochem.* **230**, 1014–1024.
- Tüchesen, E. & Woodward, C. (1985) Mechanism of surface peptide proton exchange in bovine pancreatic trypsin inhibitor salt effects and O-protonation. *J. Mol. Biol.* **185**, 421–430.
- Waldburger, C.D., Schildbach, J.F. & Sauer, R.T. (1995) Are buried salt bridges important for protein stability and conformational specificity? *Nat. Struct. Biol.* 122–128.
- Wagner, G. & Wüthrich, K. (1978) Dynamic model of globular protein conformations based on NMR studies in solution. *Nature* **275**, 247–248.
- Wagner, G., Stassinopoulou, C.I. & Wüthrich, K. (1984) Amide proton exchange studies by two-dimensional correlated ^1H NMR in two chemically modified analogs of the basic pancreatic trypsin inhibitor. *Eur. J. Biochem.* **145**, 431–436.
- Wang Y.-X., Freedberg, D.I., Yamazaki, T., Wingfield, P.T., Stahl, S.J., Kaufman, J.D., Kiso, Y. & Torchia, D. (1996) Solution NMR evidence that the HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272. *Biochemistry* **35**, 9945–9950.

- Wang, Y.-X., Freedberg, D.I., Wingfield, P.T., Stahl, S.J., Kaufman, J.D., Kiso, Y., Bhat, T.N., Erickson, J.W. & Torchia, D.A. (1996) Bound water molecules at the interface between the HIV-1 protease and a potent inhibitor, KNI-272, determined by NMR. *J. Am. Chem. Soc.* **118**, 12287–12290.
- Warshel, A., Naray-Szabo, G., Sussman, F. & Hwang JK (1989) How do serine proteases really work? *Biochemistry* **28**, 3629–3637.
- Weshhof, E. (ed.) (1993) *Water and Biological macromolecules*. Macmillan press ltd.
- Weiner, S.J., Kollman, P.A., Nguyen, D.T. & Case, D.A. (1986) An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **7**, 230–252.
- Wharton, R.P. & Ptashne, M. (1985) Changing the binding specificity of a repressor by redesigning an alpha-helix. *Nature* **316**, 601–605.
- Wider, G., Riek, R. & Wüthrich, K. (1996) Diffusion filters for separation of solvent-protein and protein-protein nuclear Overhauser effects (HYDRA). *J. Am. Chem. Soc.* **118**, 11629–11634.
- Wishart, D.S., Sykes, B.D. & Richards, F.M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222**, 311–333.
- Wishart, D.S. & Sykes, B.D. (1994) The ^{13}C Chemical-Shift Index: A simple method for the identification of protein secondary structure using ^{13}C chemical shift data. *J. Biomol. NMR* **4**, 171–180.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. & Sykes, B.D. (1995) ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135–140.
- Wittekind, M. & Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonance with the alpha- and beta-carbon resonances in proteins. *J. Mag. Res. B.* **101**, 201–205.
- Wüthrich, K., Billeter, M. & Braun, W. (1984) Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton–proton distances. *J. Mol. Biol.* **180**, 715–740.
- Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley, New York.
- Wüthrich, K., Billeter, M., Güntert, P., Luginbühl, P., Riek, R. & Wider, G. (1996) NMR studies of the hydration of biological macromolecules. *Faraday Discuss.* **103**, 245–253.
- Xu, R.X., Meadows, R.P. & Fesik, S.W. (1993) Heteronuclear 3D NMR studies of water bound to an FK506 binding protein/immunosuppressant complex. *Biochemistry* **32**, 2473–2480.
- Yamazaki, T., Yoshida, M. & Nagayama K. (1993) Complete assignments of magnetic resonances of ribonuclease H from *Escherichia coli* by double- and triple-resonance 2D and 3D NMR spectroscopies. *Biochemistry* **32**, 5656–5669.
- Yamazaki, T., Hinck, A.P., Wang, Y.X., Nicholson, L.K., Torchia, D.A., Wingfield, P., Stahl, S.J., Kaufman, J.D., Chang, C.H., Domaille, P.J. & Lam, P.Y. (1996) Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy. *Protein Sci.* **5**, 495–506.
- Yang, A.S. & Honig, B. (1993) On the pH dependence of protein stability. *J. Mol. Biol.* **231**, 459–474.
- Ye, Q.-Z., Johnson, L.L. & Baragi, V. (1992) Gene synthesis and expression in *E.coli* for Pump, a human matrix metalloproteinase. *Biochem. Biophys. Res. Commun.* **186**, 143–149.

9. APPENDICES

Appendix A

Plasmid map of pT7-7



Appendix B

DNA sequence of P22/434(0–63) from the plasmid pT2434

ATG TTG ATG GGT GAG CGT ATT CGC GCT CGA AGA ATT CAG CTT GGA CTT AAC CAG GCT
 M_0 L_1 M_2 G_3 E_4 R_5 I_6 R_7 A_8 R_9 R10 I11 Q12 L13 G14 L15 N16 Q17 A18

 GAA CTT GCT CAA AAG GTG GGG ACT ACC CAG CAG TCG ATC GAG CAG CTC GAA AAC GGT
 E19 L20 A21 Q22 K23 V24 G25 T26 T27 Q28 Q29 S30 I31 E32 Q33 L34 E35 N36 G37

 AAA ACT AAG CGA CCA CGC TTT TTA CCA GAA CTT GCG TCA GCT CTT GGC GTA AGT GTT
 K38 T39 K40 R41 P42 R43 F44 L45 P46 E47 L48 A49 S50 A51 L52 G53 V54 S55 V56

 GAC TGG CTG CTC AAT GGC ACC TAA
 D57 W58 L59 L60 N61 G62 T63 STP

Appendix C

DNA sequence of P22c2(5-68) from the plasmid pT77/4AF

ATG TTG ATG GGT GAG CGT ATT CGC GCT CGA AGA AAA AAA CTC AAG ATT AGA
 M_0 L_5 M_6 G_7 E_8 R_9 I10 R11 A12 R13 R14 K15 K16 L17 K18 I19 R20

 CAA GCC GCT CTT GGT AAG ATG GTG GGA GTG TCT AAT GTT GCA ATA TCG CAA
 Q21 A22 A23 L24 G25 K26 M27 V28 G29 V30 S31 N32 V33 A34 I35 S36 Q37

 TGG GAG CGC TCG GAG ACT GAG CCA AAT GGG GAG AAC CTG TTG GCA CTT TCG
 W38 E39 R40 S41 E42 T43 E44 P45 N46 G47 E48 N49 L50 L51 A52 L53 S54

 AAG GCT CTT CAG TGC TCC CCT GAC TAT TTG CTG AAA GGA GAT TAA
 K55 A56 L57 Q58 C59 S60 P61 D62 Y63 L64 L65 K66 G67 D68 STP

Appendix D

DNA sequence of the plasmid of pGFC containing the final dh434(0–63) with a replacement of A50R.

```

1/1                               31/11
ATG TTG ATG GGT GAG CGT ATT CGC GCT CGA AGA ATT CAG CTT GGA CTT AAC CAG GCT GAA
M  L  M  G  E  R  I  R  A  R  R  I  Q  L  G  L  N  Q  A  E
61/21                              91/31
CTT GCT CAA AAG GTG GGG GTC GAC CAG CAG GCG ATC GAG CAG CTC GAA AAC GGT AAA GCG
L  A  Q  K  V  G  V  D  Q  Q  A  I  E  Q  L  E  N  G  K  A
121/41                             151/51
AAG CGA CCA CGC TTT TTA CCA GAA CTT GCG CGA GCT CTT GGC GTA GCG GTT GAC TGG CTG
K  R  P  R  F  L  P  E  L  A  R  A  L  G  V  A  V  D  W  L
181/61                              211/71
CTC AAT GGC ACC TAA GCT TAT CAT CGA TGA TAA GCT GTC AAC
L  N  G  T  *  A  Y  H  R  *  *  A  V  N

```

Appendix E

DNA sequence of plasmid pT434HK

```

1/1                               31/11
AGA CCA CAA CGG TTT CCC TGT AGA AAT ATT TTT GTT TAA CTT TAA GAA GGA GAT ATA GAT
R  P  Q  R  F  P  C  R  N  I  F  V  *  L  *  E  G  D  I  D
61/21                              91/31
ATG AGT ATT TCT TCC AGG GTA AAA AGC AAA AGA ATT CAG CTT GGA CTT AAC CAG GCT GAA
M  S  I  S  S  R  V  K  S  K  R  I  Q  L  G  L  N  Q  A  E
121/41                             151/51
CTT GCT CAA AAG GTG GGG ACT ACC CAG CAG TCG ATC GAG CAG CTC GAA AAC GGT AAA ACT
L  A  Q  K  V  G  T  T  Q  Q  S  I  E  Q  L  E  N  G  K  T
181/61                              211/71
AAG CGA CCA CGC TTT TTA CCA GAA CTT GCG TCA GCT CTT GGC GTA AGT GTT GAC TGG CTG
K  R  P  R  F  L  P  E  L  A  S  A  L  G  V  S  V  D  W  L
241/81                              271/91
CTC AAT GGC AAA CAC CAC CAT CAT CAC CAT ATG TAA GCT TAA C
L  N  G  K  H  H  H  H  H  H  M  *  A  *

```


Appendix F

 ^1H chemical shifts table of dh434(0–63)

	H^{N}	H^{α}	H^{β}	Others
Met 0	-	4.13	2.07, 2.16	γCH_2 2.60, 2.60; ϵCH_3 2.09
Leu 1	9.12	4.62	1.73, 1.84	γH 1.75; δCH_3 0.95, 0.97
Met 2	9.10	3.74	2.08, 2.17	γCH_2 2.53, 2.53; ϵCH_3 2.00
Gly 3	9.23	3.60, 4.20		
Glu 4	6.87	4.14	2.09, 2.36	γCH_2 2.36, 2.47
Arg 5	7.82	4.09	1.82, 1.89	γCH_2 1.15, 1.42; δCH_2 1.49, 2.23; ϵNH 6.54; ηNH_2 6.32
Ile 6	8.31	3.47	1.99	γCH_2 0.81, 0.81; γCH_3 1.09; δCH_3 0.88
Arg 7	7.87	4.02	1.73, 1.88	γCH_2 1.63, 1.63; δCH_2 3.25, 3.25; ϵNH 7.35; ηNH_2 -, -
Ala 8	8.26	4.09	1.59	
Arg 9	8.45	4.32	2.30, 2.32	γCH_2 1.75, 1.83; δCH_2 2.90, 3.58; ϵNH 6.96; ηNH_2 6.41
Arg 10	9.05	3.77	1.72, 2.31	γCH_2 1.26, 1.93; δCH_2 2.81, 2.85; ϵNH 11.88; ηNH_2 6.75
Ile 11	8.46	3.84	1.94	γCH_2 1.15, 1.88; γCH_3 0.95; δCH_3 0.88
Gln 12	8.26	4.09	2.32, 2.38	γCH_2 2.58, 2.58; ϵNH_2 6.91, 7.62
Leu 13	7.78	4.27	1.65, 1.91	γH 1.65; δCH_3 0.85, 0.94
Gly 14	8.16	3.91, 4.05		
Leu 15	7.57	4.86	1.46, 1.72	γH 1.59; δCH_3 0.78, 0.81
Asn 16	9.27	4.78	2.89, 3.42	δNH_2 7.03, 7.50
Gln 17	8.99	3.67	1.72, 2.25	γCH_2 2.10, 2.62; ϵNH_2 7.63, 7.75
Ala 18	8.44	3.88	1.49	
Glu 19	8.94	4.07	1.93, 2.25	γCH_2 2.29, 2.63
Leu 20	8.06	4.09	1.23, 1.93	γH 1.48; δCH_3 0.89, 0.94
Ala 21	8.44	3.50	1.35	
Gln 22	7.96	4.10	2.21, 2.45	γCH_2 2.55, 2.55; ϵNH_2 6.88, 7.53
Lys 23	7.63	4.11	1.99, 2.00	γCH_2 1.76, 1.76; δCH_2 1.64, 1.69; ϵCH_2 2.97, 3.10
Val 24	7.66	3.92	2.01	γCH_3 0.80, 0.89
Gly 25	8.00	3.86, 4.11		
Val 26	7.57	4.64	2.02	γCH_3 0.05, 0.37
Asp 27	7.91	4.73	2.73, 2.94	
Gln 28	8.97	3.58	2.06, 2.20	γCH_2 2.31, 2.31; ϵNH_2 6.97, 7.62
Gln 29	8.74	4.01	2.04, 2.12	γCH_2 2.42, 2.42; ϵNH_2 6.88, 7.70
Ala 30	7.80	4.20	1.59	
Ile 31	7.15	3.93	2.33	γCH_2 1.25, 1.72; γCH_3 0.92; δCH_3 0.57

Glu 32	8.35	4.02	2.05, 2.21	γCH_2 2.28, 2.36
Gln 33	8.39	4.28	2.04, 2.09	γCH_2 2.44, 2.68; ϵNH_2 6.95, 7.57
Leu 34	7.53	4.26	1.64, 2.07	γH 1.81; δCH_3 0.82, 0.94
Glu 35	8.54	4.04	1.89, 1.95	γCH_2 2.21, 2.70
Asn 36	8.25	4.74	2.85, 3.01	δNH_2 6.72, 7.57
Gly 37	7.51	4.11, 4.15		
Lys 38	8.52	4.28	1.83, 1.94	γCH_2 1.33, 1.46; δCH_2 1.68, 1.68; ϵCH_2 3.00, 3.00
Ala 39	7.75	4.47	1.28	
Lys 40	8.38	4.31	1.69, 1.88	γCH_2 1.42, 1.47; δCH_2 1.73, 1.73; ϵCH_2 3.02, 3.02
Arg 41	8.52	4.22	1.57, 1.76	γCH_2 1.50, 1.50; δCH_2 3.16, 3.16; ϵNH 7.27; ηNH_2 -, -
Pro 42	-	4.45	0.37, 1.73	γCH_2 1.44, 1.54; δCH_2 3.36, 3.45
Arg 43	9.24	4.22	1.93, 2.00	γCH_2 1.80, 1.80; δCH_2 3.29, 3.29; ϵNH 7.36; ηNH_2 -, -
Phe 44	6.46	5.33	2.86, 3.67	δH 7.27, 7.27; ϵH 7.31, 7.31; ζH 7.22
Leu 45	7.38	4.08	1.75, 1.75	γH 1.65; δCH_3 0.85, 0.92
Pro 46	-	4.16	2.40, 2.40	γCH_2 1.92, 2.31; δCH_2 3.78, 3.82
Glu 47	8.90	4.10	2.14, 2.47	γCH_2 2.39, 2.74
Leu 48	8.66	4.15	1.68, 2.21	γH 1.01; δCH_3 0.97, 1.01
Ala 49	8.77	3.82	1.57	
Arg 50	7.89	4.16	1.97, 1.97	γCH_2 1.64, 1.83; δCH_2 3.30, 3.30; ϵNH 7.43; ηNH_2 -, -
Ala 51	8.13	4.13	1.41	
Leu 52	7.78	4.30	1.61, 1.96	γH 1.68; δCH_3 0.82, 0.94
Gly 53	8.07	3.91, 3.97		
Val 54	8.00	4.61	2.13	γCH_3 0.33, 0.88
Ala 55	8.46	4.59	1.61	
Val 56	9.10	3.47	2.08	γCH_3 0.96, 1.05
Asp 57	8.71	4.38	2.68, 2.70	
Trp 58	7.28	4.19	3.39, 3.57	$\delta^1\text{H}$ 7.22; $\epsilon^1\text{NH}$ 10.43; $\epsilon^3\text{H}$ 7.30; $\zeta^2\text{H}$ 7.55; $\zeta^3\text{H}$ 6.74; $\eta^2\text{H}$ 7.28
Leu 59	7.74	3.57	1.69, 2.22	γH 1.96; δCH_3 0.88, 0.94
Leu 60	7.99	4.08	1.42, 1.69	γH 1.67; δCH_3 0.83, 0.85
Asn 61	8.09	4.99	2.74, 2.79	δNH_2 6.99, 7.74
Gly 62	8.31	3.00, 3.63		
Ala 63	7.71	4.25	1.42	

Appendix G

^{15}N , ^{13}C chemical shifts table of dh434(0–63)

#	N	C $^{\alpha}$	C $^{\beta}$	others
Met 0	-	54.8	33.6	C γ 30.9; ϵ C 16.5
Leu 1	124.0	53.6	43.2	γ C 26.8; δ C 22.1, <u>25.1</u>
Met 2	121.8	59.6	33.5	γ C 30.7; ϵ CH $_3$ 16.6
Gly 3	105.6	48.4		
Glu 4	118.1	58.2	29.7	γ C 36.6
Arg 5	121.5	60.1	30.0	γ C 29.1; δ C 42.4
Ile 6	118.9	66.0	38.1	γ C 30.1; γ CH $_3$ 16.3; δ C 15.4
Arg 7	118.4	59.2	30.3	γ C 27.1; δ C 43.2
Ala 8	117.9	54.9	17.8	
Arg 9	118.9	56.7	28.9	γ C 28.0; δ C 41.8
Arg 10	120.7	60.8	28.6	γ C 27.9; δ C 41.0; η NH $_2$ 9.0
Ile 11	118.4	64.9	37.9	γ C 29.6; γ CH $_3$ 16.5; δ CH $_3$ 13.5
Gln 12	123.6	58.9	28.3	γ C 33.3; N ϵ 2 111.6
Leu 13	117.0	54.9	42.6	γ C 26.9; δ C 21.6, <u>25.5</u>
Gly 14	108.7	45.7		
Leu 15	118.3	53.3	44.5	γ C 26.0; δ C 22.7, <u>25.9</u>
Asn 16	121.8	51.2	38.1	N δ 1 110.9
Gln 17	117.0	60.9	29.0	γ C 34.4; N ϵ 2 112.5
Ala 18	121.9	55.3	17.4	
Glu 19	119.6	58.3	30.2	γ C 36.0
Leu 20	119.1	57.2	40.9	γ C 26.8; δ C <u>22.7</u> , 26.6
Ala 21	120.2	55.2	17.2	
Gln 22	116.3	58.7	27.9	γ C 33.5; N ϵ 2 111.4
Lys 23	117.8	58.5	32.8	γ C 24.9; δ C 29.0; ϵ C 41.6
Val 24	113.8	62.2	32.2	γ C <u>22.3</u> , 23.1
Gly 25	108.3	45.4		
Val 26	111.6	58.1	33.6	γ C 16.9, <u>20.2</u>
Asp 27	117.1	52.2	41.9	
Gln 28	120.9	59.8	27.4	γ C 32.9; N ϵ 2 111.6
Gln 29	118.4	58.8	27.6	γ C 34.0; N ϵ 2 112.1
Ala 30	121.3	54.6	17.7	
Ile 31	115.7	60.6	34.7	γ C 26.4; γ CH $_3$ 18.2; δ C 37.7
Glu 32	121.6	59.3	28.4	γ C 34.5
Gln 33	115.2	57.8	28.3	γ C 34.2; N ϵ 2 111.7
Leu 34	119.2	57.1	42.4	γ C 26.3; δ C 25.2, <u>26.4</u>
Glu 35	118.2	59.5	29.4	γ C 36.9
Asn 36	114.3	52.1	38.2	N δ 1 109.4
Gly 37	106.9	45.7		
Lys 38	120.1	56.5	32.6	γ C 25.1; δ C 28.7; ϵ C 41.9
Ala 39	121.7	51.2	20.4	

Lys 40	120.3	55.4	33.4	γC 24.7; δC 28.6; ϵC 41.8
Arg 41	119.4	53.4	30.0	γC 26.6; δC 43.2
Pro 42	-	60.7	31.3	γC 26.6; δC 49.5
Arg 43	121.4	58.4	29.7	γC 26.8; δC 42.9
Phe 44	110.7	54.3	36.6	δC 132.7; ϵC 131.4; ζC 129.9
Leu 45	121.7	59.0	39.1	γC 25.0; δC <u>24.6</u> , 24.8
Pro 46	-	66.4	30.4	γC 28.6; δC 50.0
Glu 47	119.4	60.2	28.7	γC 37.0
Leu 48	123.0	57.7	42.6	γC 26.7; δC <u>24.9</u> , 26.7
Ala 49	121.3	55.4	17.5	
Arg 50	116.4	59.1	29.8	γC 27.2; δC 43.1
Ala 51	121.8	54.5	18.9	
Leu 52	112.9	53.3	42.8	γC 26.3; δC 22.1, <u>26.7</u>
Gly 53	111.4	46.8		
Val 54	109.9	57.7	35.8	γC 17.9, <u>20.6</u>
Ala 55	122.0	50.7	19.2	
Val 56	123.7	67.4	31.3	γC <u>21.0</u> , 23.4
Asp 57	115.8	57.0	40.2	
Trp 58	119.2	61.4	28.2	$\delta^1\text{C}$ 127.0; $\epsilon^3\text{C}$ 119.4; $\zeta^2\text{C}$ 114.6; $\zeta^3\text{C}$ 120.4; $\eta^1\text{C}$ 124.0; $\text{Ne}1$ 130.0
Leu 59	118.0	58.4	41.4	γC 26.2; δC 23.7, <u>25.8</u>
Leu 60	115.6	57.5	43.0	γC 26.5; δC 22.9, <u>25.2</u>
Asn 61	113.1	53.0	40.9	$\text{N}\delta 1$ 114.7
Gly 62	111.0	44.9		
Ala 63	127.2	53.9	20.3	

For the diastereotopic methyl groups of valines and leucines, the pro-(R) methyl groups ($\gamma 1$ and $\delta 1$) are underlined. The chemical shifts are relative to DSS with the following Ξ ratio: $^{13}\text{C}/^1\text{H} = 0.251449530$, $^{15}\text{N}/^1\text{H} = 0.101329118$ (Wishart et al., 1995).

Appendix H

Comparison of the primary structures of 434(1–63) variants and P22c2.

dh434 (0–63)	MLMGERIRARRIQLGLNQAELAQKVGVDQQAIEQLENGKAK-RPRFLPELAAALGVAVDWLLNGA
434 (1–63)	SISSRVKSKRIQLGLNQAELAQKVGTTQOSIEQLENGKTK-RPRFLPELASALGVSDWLLNGT
P22/434 (0–63)	MLMGERIRARRIQLGLNQAELAQKVGTTQOSIEQLENGKTK-RPRFLPELASALGVSDWLLNGT
P22c2 (5–68)	MLMGERIRARRKCLKIRQAALGKMVGVSNAISQWERSETEPNGENLLALSKALQCSPDYLLKGD

Comparison of the primary structures of the homologous proteins.

	434 (1–63)	SISSRVKSKRIQLGLNQAELAQKVGTTQOSIEQLENGKTK-RPRFLPELASALGVSDWLLNGT
Cro Proteins	RPC2_BPP22	MNTQLMGERIRARRKCLKIRQAALGKMVGVSNAISQWERSETEPNGENLLALSKALQCSPDYLLKGDLSQTNVAYHSRH
	434Cro	MQTLSERLKKRRIALKMTQTELATKAGVKQOSIQLIEAGVTK-RPRFLFEIAMALNCDPVWLQYGT
	Lamda Cro	MEQRITLKDYAMRFGTKTAKDLGVYQSAINKAIHAGRKIFLTINADGSVYAEVKKPFPSNKKTTA
	RCRO_BPHK0	MNEVTFGEVIKSVRVSVVADVCGLTPKAIYKWLERSLPRTEFTGETEYADKIAKASGGKYSAAQIRRIGKQQFVM
	P22Cro	MYKKDVIDHFGTQRAVAKALGISDAAVSQWKEVIPEKDAYRLEIVTAGALKYQENAYRQAA
Cl repressors	RPC1_BPPH8	SSISERIKFLLAREGLKQRDLAEALSTSPQTVNNWIKRDAL-SREAAQQISEKFGYSLDWLLNGEGS
	RPC1_BPD3	MELKDRIKAARKHAKLTQAQLAQRVGLDQTSISNLEQGKSQ-GTSYIAQLASACGVSALWLAAGHGEMNSN
	RPC1_BP186	MRIDSLGWSNVDVLDRICEAYRFSQKIQLANHFDIASSLSNRYTRGAI-SYDFAAHCALETGANLQWLLTGEGEAFVNNRESSDAKRIEG
	RPC1_BPHP1	MKEFIGGKDVISRIMEAYGFANRRLLAEHLGMPHSTFGTWAKRGFF-PAELVIRCVSETGARLDYVAYGN
	RPC1_BPMU	MAADGMPGSVAGVHYRANVQGWTKQKKEGVKGGKAVEYDVMSMP-TKEREQVIAHLGLSTPDTGAQANEKQDSSELINKLTTTLINMIEEL
	RPC1_BPP1	MINYVYGEQLYQEFVSVFRDLFLKKAVARAQHVDAASDGRVPRVVLVLPFKETDSIQAEIDKWTLMARELEQYPLNIPKTIILYVPVNIILRGVRKVTYQT
	RPC1_BPP7	MINYVYGEQLYQEFVSVFRDLFLKKAVARAQHVDAASDGRVPRVVLVLPFKETDSIQAEIDKWTLMARELEQYPLNIPKTIILYVPVNIILRGVRKVTYQT
Ner & Homologues	Ner (Mu)	CSNEKARDWHRADVIAGLKKRKLSSALSRSRQFGYAPTTLANALERHWP-KGEQIIANALETKEPEVIWPSRYQAGE
	DBNE_BPD10	MHMNKRTNRDWHRADIVAE LRKRNMSLAELGRSNHLSSTLKNALDKRYP-KAEKIIADALGMTPPQDIWPSRY
	NLP_ECOLI	MESNFIDWHPADIIAGLRKKGTSMAAESRRNGLSSSTLANALSRPWP-KGEMI IAKALGTDPPVVIWPSRYHDPQTHEFIDRTQLMRSYTK
	NLP_HAEIN	EKPKKTAEQDWHRADILAE LKKNWSLSLAKEGQVSYNTLKTVLDKSYP-KMERLVANAIGVPPPEVIWAGRFAERNKRPTLQHKY

Curriculum vitae

Personal:

Name **Hideo Iwai**
Date of Birth **8th May 1968**
Place of Birth **Kurobe, Japan**
Nationality **Japan**

Education:

1975 – 1981 Mikkaichi Primary School, Kurobe, Japan
1981 – 1984 Sakurai Junior High School, Kurobe, Japan
1984 – 1987 Uozu High School, Toyama, Japan
1988 – 1990 Natural Sciences II, The College of Arts and Sciences, The University of Tokyo, Japan
1990 – 1992 Department of Pharmaceutical Science, The University of Tokyo, Japan.
B.Sc. (Pharmaceutical Science)
1992 – 1994 Division of Pharmaceutical Sciences, Graduate School, The University of Tokyo, Japan
M.Sc. (Pharmaceutical Science)
Supervised by Prof. Dr. Yoji Arata.
Subject: *Structural studies of the antibacterial peptides by NMR spectroscopy.*
1994 – 1996 Research Fellowship from the Japan Society for the Promotion of the Science for Young Scientists
1994 – 1998 Ph.D student in Biology, Swiss Federal Institute of Technology, Zürich (ETH, Zürich)
Supervised by Prof. Dr. Kurt Wüthrich
Subject: *Expression, Stabilities, and NMR Studies of Hydration of Recombinant Variants of the DNA Binding Domain of Phage 434 Repressor.*

Publication

- Iwai, H.**, Nakajima, Y., Natori, S., Arata, Y. & Shimada, I. (1993) Solution conformation of an antibacterial peptide, sarcotoxin IA, as determined by ^1H -NMR. *Eur. J. Biochem.* **217**, 639–644.
- Kim, J.I., **Iwai, H.**, Kurata, S., Takahashi, M., Masuda, K., Shimada, I., Natori, S., Arata, Y. & Sato, K. (1994) Synthesis and characterization of sapecin and sapecin B. *FEBS Letters* **342**, 189–192.
- Kim, J.I., Konishi, S., **Iwai, H.**, Kohno, T., Gouda, H., Sato, K., Shimada, I. & Arata, Y. (1995) Three-dimensional solution structure of the calcium channel antagonist omega-agatoxin IVA: consensus molecular folding of calcium channel blockers. *J. Mol. Biol.* **250**, 659–671.
- Pellecchia, M., **Iwai, H.**, Szyperski, T. & Wüthrich, K. (1997) The 2D NMR experiments H(C)CO_2 and HCCO_2 for assignment and pH-titration of carboxylate groups in uniformly $^{15}\text{N}/^{13}\text{C}$ labelled proteins. *J. Magn. Reson.* **124**, 274–278.
- Luginbühl, P., Pervushin, P.V., **Iwai, H.** & Wüthrich, K. (1997) Anisotropic molecular rotational diffusion in ^{15}N spin relaxation studies of protein mobility. *Biochemistry* **36**, 7305–7312.
- Pellecchia, M., Wider, G., **Iwai, H.** & Wüthrich, K. (1997) Arginine side chain assignments in uniformly ^{15}N labelled proteins using the novel 2D HE(NE)HGHH experiments. *J. Biomol. NMR* **10**, 193–197.
- Szyperski, T., Ono, A., Fernández, C., **Iwai, H.**, Tate, S., Wüthrich, K. & Kainosho, M. (1997) Measurement of $^3\text{J}_{\text{C}_2\text{p}}$ scalar couplings in a 17KDa protein complex with $^{15}\text{N}/^{13}\text{C}$ labelled DNA distinguishes the B_I and B_II phosphate conformations of the DNA. *J. Am. Chem. Soc.* **119**, 9901–9902.
- Sobol, A., Wider, G., **Iwai, H.** & Wüthrich, K. (1998) Solvent magnetization artifacts in high field NMR studies of macromolecular hydration. *J. Magn. Reson.* **130**, 262–271.
- Fernández, C., Szyperski, T., Ono, A., **Iwai, H.**, Tate, S., Kainosho, M. & Wüthrich, K. (1998) NMR with ^{15}N , ^{13}C -doubly-labeled DNA: *Antennapedia* homeodomain complex with a 14mer DNA duplex. *J. Biomol. NMR. in press*