This is to certify that the

dissertation entitled

Biochemical and NMR Studies of Human Cellular Retinoic Acid Binding Proteins

presented by

Lincong Wang

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Biochemistry

Major professor

Date 8/21/98

DATE DUE	DATE DUE	DATE DUE

BIOCHEMICAL AND NMR STUDIES OF HUMAN CELLULAR RETINOIC ACID BINDING PROTEINS

By

Lincong Wang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1998

BIOCHEMICAL AND NMR STUDIES OF HUMAN CELLULAR RETINOIC ACID BINDING PROTEINS

By

Lincong Wang

ABSTRACT

The structure-function relationships of cellular retinoic acid binding proteins (CRABPs) have been investigated by site-directed mutagenesis, biochemical and NMR methods. The major results and conclusions are summarized as follows.

Electrostatic interactions have been suggested to be critical for binding of retinoic acid (RA) by CRABPs. However, the roles of two conserved arginine residues (Arg111 and Arg131 in type-I CRABP; Arg111 and Arg132 in type-II CRABP) that interact with the carboxyl group of RA have not been evaluated. A novel competitive binding assay was developed for measuring the relative dissociation constants of the site-directed mutants including the two mutants R111M and R132M. Binding studies of the two mutants showed that Arg111 and Arg132 contribute to the binding energy by ~2.2 and 1.2 kcal/mol, respectively.

The crystal structure of holo-CRABP suggests that RA cannot enter or exit the deep binding pocket without major conformational changes in the protein. Dimerization has been proposed to be the mechanism by which CRABPs open their binding pockets. The solution structure of apo-CRABPII has been determined by NMR spectroscopy. Although the apo solution structure was similar to the holo crystal structure, the ligand entrance was greatly enlarged in the former and readily accessible to RA. The

enlargement was mainly due to a concerted conformational change in three structural elements, namely the second helix, the βC-βD loop and the βE-βF loop. The ligand-binding pocket of apo-CRABPII was rather dynamic. CRABPII is predominately monomeric in solution. Although the formation of transient dimers could not be ruled out, dimerization apparently is not a prerequisite for entry of retinoic acid into the binding pocket of CRABPII.

Preliminary NMR analysis of holo-CRABPII showed that RA binds to CRABPII in solution in the same manner as determined by crystallography. NMR studies indicated that the ligand entrance of holo-CRABPII is much less flexible than that of apo-CRABPII. The results taken together suggested that RA binding indeed induces major changes in the conformation and dynamics of CRABPII.

The solution structure of R111M has been determined by NMR spectroscopy. Although apo-R111M had a structure similar to that of wild-type apo-CRABPII, there were significant conformational differences between the two proteins, mainly localized to three segments, clustered around the ligand entrance more than 17 Å from the site of point mutation. Furthermore, the ligand-binding pocket of apo-R111M, especially the ligand entrance, was much less flexible than that of apo-CRABPII. The results suggested that Arg111 play a critical role in determining the structure and dynamics of CRABPII.

To my family

ACKNOWLEDGMENTS

I would like to express my gratitude to my thesis advisor, Dr. Honggao Yan, for his financial support and for providing me the unique opportunity to learn both protein biochemistry and NMR. I have learned a lot from him and become a better researcher.

I would like to thank my guidance committee members Dr. David McConnell, Dr. John McCraken, Dr. Robert Hausinger, Dr. John Wilson and Dr. William Wells for their criticism and advice, which truly guided me through the tough journey of graduate school.

I would like to thank Mr. Kermit Johnson for teaching me NMR. His kindness and patience made doing NMR experiments much more enjoyable. Thanks also go to Dr. Yue Li for teaching me molecular biology. Special thanks go to Dr. Robert Cukier of Chemistry Department for his advice on my thesis. His comment, "You have summarized a great deal of (complicated) work", encourages me to go further.

In experimental aspect I would like to thank Dr. Yue Li for establishing the expression systems for CRABPs and for generating the mutants, Dr. Frits Abildgaard at the National Magnetic Resonance Facility at Madison, Wisconsin for recording some of the triple resonance spectra of apo-CRABPII and Dr. Honggao Yan for acquiring the NMR spectra of uniformly ^{15}N-labeled apo-R111M.

Finally, I am grateful to my wife for her understanding and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

LIST OF ABBREVIATIONS

BPP theory:	Bloembergen-Purcell-Pound theory
CD:	circular dichroism
CMPG:	Carr-Purcell-Meiboom-Gill
COSY:	correlated spectroscopy
CRABPI:	type-I cellular retinoic acid binding protein
CRABPII:	type II-cellular retinoic acid binding protein
CRBPI:	type-II cellular retinol binding protein
CRBPII:	type-II cellular retinol binding protein
CSA:	chemical-shielding anisotropy
DEPT:	distortionless enhancement by polarization transfer
DGSA:	distance geometry-simulated annealing
DMSO:	dimethyl sulfoxide
DQF-COSY:	double-quantum-filtered correlated spectroscopy
E.COSY:	exclusive correlated spectroscopy
FABP:	fatty acid binding protein
FID:	free-induction decay
HMQC:	heteronuclear multiple-quantum coherence
HSQC:	heteronuclear single-quantum coherence
iLBP:	intracellular lipid binding protein
INEPT:	insensitive nuclei enhanced by polarization transfer
IPTG:	isopropylthiogalactoside

NMR:	nuclear magnetic resonance
NOE:	nuclear Overhauser effect
NOESY:	nuclear Overhauser enhancement and exchange spectroscopy
P.E.COSY:	primitive exclusive correlated spectroscopy
RA:	retinoic acid
RAR:	retinoic acid receptor
RMSD:	root-mean-square deviation
ROESY:	rotating frame Overhauser effect spectroscopy
RXR:	retinoid X receptor
TOCSY:	total correlation spectroscopy
TPPI:	time proportional phase incrementation
2D:	two-dimensional
3D:	three-dimensional
4D:	four-dimensional

CHAPTER 1

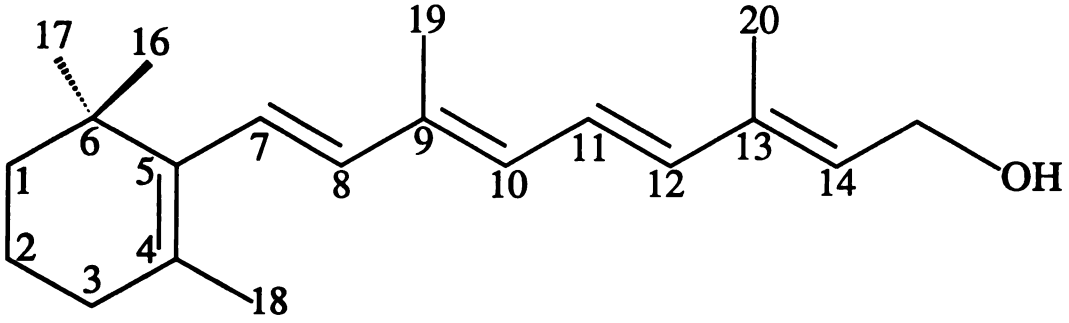
Literature Review

Part 1. Cellular Retinoic Acid Binding Proteins

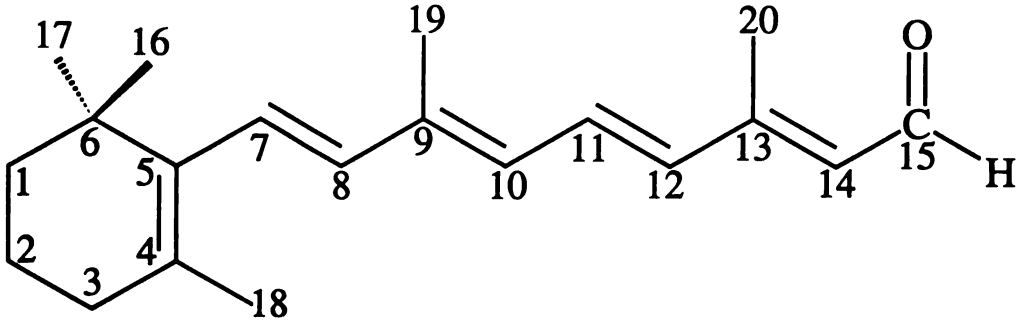
Section 1. Biological Functions of CRABPII

Vitamin A (retinol) and its metabolic derivatives (naturally occurring retinoids) such as retinal and retinoic acid (RA) (Figure 1.1) are essential for vertebrates, with profound effects on a wide variety of biological processes such as vision, cell differentiation, embryonic development and homeostasis (Sporn et al., 1994; Chambon et al., 1996). Both the naturally occurring retinoids and their synthetic analogs have been used to treat several kinds of human diseases including skin disorders (Peck & DiGiovanna, 1994), and certain cancers, such as cervical cancer and acute promyelocytic leukemia (Hong & Itri, 1994). A unified picture of the apparent pleiotropic effects of retinoids is gradually emerging, and some key components of the retinoid signal transduction pathway have been discovered over the past five decades. It has been shown that retinal is a key component of visual signal transduction (Wald, 1968), while RAs,

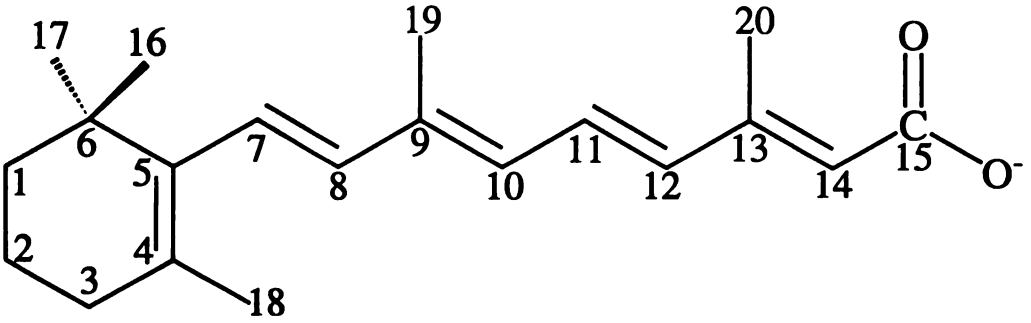
Figure 1.1 Chemical structures and the numbering conventions of all-*trans*-retinol (vitamin A), all-*trans*-retinal and all-*trans*-retinoic acid (RA).



All-*trans*-retinol (Vitamin A)



All-*trans*-retinal



All-*trans*-retinoic acid (RA)

especially all-*trans*-RA and 9-*cis*-RA, are the active forms of vitamin A in most of the other biological processes. RA administration can prevent or reverse the majority of the defects originating from post-natal vitamin A deficiency. Two distinct classes of RA-binding proteins have been found to be involved in retinoid signal transduction: nuclear retinoid receptors and cellular RA binding proteins (CRABPs).

Nuclear retinoid receptors. There exist in cells two retinoid receptor families: the RA receptor (RAR) family with three subtypes RAR-α, -β and -γ, and the retinoid X receptor (RXR) family also with three subtypes RXR-α, -β and -γ (Mangelsdorf et al., 1994; Mangelsdorf & Evans, 1995; Chambon, 1996). The nuclear retinoid receptors play critical roles in retinoid signal transduction. They have been shown to mediate many of the biological effects observed with administration of vitamin A. For example, all abnormalities exhibited by fetuses of vitamin A deficient dams are recapitulated in RAR or RXR single mutants and RAR-RXR double mutants of mouse (Chambon, 1996).

Amino acid sequence comparison suggests that retinoid receptors belong to a protein superfamily of ligand-activated transcriptional factors (Evans, 1988; Tsai & O'Malley, 1994) that includes also the receptors for steroid and thyroid hormones and vitamin D_3. The RAR proteins are activated by both all-*trans*-RA (K_d = ~1-5 nM) and 9-*cis*-RA with similar effectiveness, whereas the RXR proteins bind exclusively 9-*cis*-RA (K_d = ~10 nM) (Heyman et al., 1992; Levin et al., 1992). *In vitro* studies have demonstrated that the functional units of the retinoid receptors are either the RAR-RXR heterodimers consisting of one RXR subunit and one RAR subunit, or the homodimers of two RXR subunits. These two forms of dimer appear to have different functions. The heterodimers bind the DNA response elements with direct repeats of the hexamer (purine-

G(G/T)TCA) separated by one, two, or five base pairs, while the RXR homodimers prefer the DNA sequence with direct repeats of the hexamer separated by only one base pair. For the same response element the heterodimers have higher affinity than the homodimers. Interestingly, RXR proteins also form heterodimers with other nuclear receptors such as thyroid hormone and vitamin D_3 receptors, and are required for the biological functions of these receptors (see, for example, Yu et al., 1991). The diversity of the biological effects of retinoids can be partly explained by the existence of two families of receptors and the interactions between the two families of receptors themselves and with other receptors. It appears that the availability of RA in cells is regulated by other molecules such as the two isoforms of CRABP.

Biological significance of CRABPs. There are two types of cellular retinoic acid binding protein (CRABPI and CRABPII) (Sani & Hill, 1974; Ong & Chytil, 1975; Bailey & Siu, 1988; Kitamoto et al., 1988). They have been found in all vertebrates relying on vitamin A (Ong et al., 1994), and are highly conserved proteins. For example, rat, mouse and cow CRABPI (Nilsson et al., 1988; Stoner & Gudas, 1989; Shubeita et al., 1987) all have identical 136 residues, and human CRABPI has only one residue difference (Astrom et al., 1991). Of 137 residues of mouse CRABPII (Giguere et al., 1990) and human CRABPII (Astrom et al., 1991; Eller et al., 1992) 130 are identical.

Although CRABPs have been suggested to play important roles in mediating the biological functions of RA, their physiological functions have not been well defined. They may be involved in directing the spatial organization of cells during development, especially at the time of the limb generation (Maden, 1991; Hofmannn & Eichele, 1994; Morriss-Kay & Sokolova, 1996). Experimental studies have shown that both CRABPI

and CRABPII are widely expressed at all stages of mouse embryo development. In particular, the mRNAs of CRABPs form an anteroposterior gradient opposite to the gradient of all-*trans*-RA in the budding of a chick limb (Maden et al., 1988). It has also been proposed that CRABPs may protect cells by preventing RA from being incorporated into cell membranes. The solubility of RA in aqueous solution is about 200 nM (Szuts & Harosi, 1991), which is higher than its endogenous concentration (~50 nM). However, RA molecules prefer the membrane to the cytoplasm with ~6 $\times 10^5$ fold higher affinity. Indeed, in tissues that express CRABPs, RA is predominately in the protein-bound forms. Both CRABPI and CRABPII have very high affinity for all-*trans*-RA with an apparent dissociation constant (K_d) less than 2 nM (Norris et al., 1994; Wang et al., 1997). CRABPs, CRABPI in particular, may participate in the catabolism of all-*trans*-RA (Fiorella & Napoli, 1991, 1994; Boylan & Gudas, 1992). The rate of the breakdown of RA into inactive metabolites catalyzed by cytochrome P450 increases more than twenty-fold when it binds CRABPI. CRABPs may also be directly involved in the transport of RA to the nucleus (Takase et al., 1986; Blomhoff et al., 1992). Taken together, these *in vitro* studies lead to the proposal that CRABPs may act as 'buffer' to control the actual level of 'free' intracellular RA available for binding to the nuclear retinoid receptors. However, the physiological functions of CRABPs remain controversial. Genetic experiments have demonstrated that the mutant mice lacking CRABPI and/or CRABPII are essentially indistinguishable from the wild-type (de Bruijn et al., 1994; Gorry et al., 1994; Fawcett et al., 1995; Lampron et al., 1995) with the exception of a minor limb malformation. These results suggest that their actual function may be to control

physiological levels of intracellular RA when the supply of vitamin A from diet is insufficient.

The biological functions of CRABPI and II seem to be different. The homology within either CRABPI or CRABPII sequences across different species is much higher than the homology between CRABPI and CRABPII from the same species. For example, human CRABPI and CRABPII have only 74% identity. Moreover, CRABPI and CRABPII exhibit different expression patterns during development or in adults. In mouse embryo they are expressed in distinct, non-overlapping patterns. In adults CRABPI is expressed in a variety of tissues, while CRABPII is mainly restricted to the epidermis (Astrom et al., 1991; Elder et al., 1992). The CRABPII gene, not CRABPI, has a retinoid response element in its promoter region so that CRABPII can be induced markedly in cultured cells and human skin by 9-*cis*-RA and, to a lesser extent, by all-*trans*-RA (Giguere et al., 1990; Astrom et al., 1991; Durand et al., 1992; Melhus et al., 1994). It has been shown that production of RA is correlated with the expression of CRABPII but not CRABPI (Bucco et al., 1997). Overexpression of CRABPII enhances cellular response to RA in breast cancer cells (Jing et a., 1997) but overexpression of CRABPI decreases the biological potency of RA in F9 teratocarcinoma cells (Boylan & Gudas, 1991). However, no difference in binding specificity has been reported between CRABPI and CRABPII, and their affinities for retinoids are very similar.

Section 2. Biochemical Characterizations of CRABPs

Extensive biochemical studies have been performed on CRABPs. These studies have been expedited by cloning and overexpression of the genes encoding CRABPs in *E. coli* (Giguere et al., 1990; Astrom et al., 1991). The *in vitro* biochemical functions of CRABPs have been established by ligand binding assays, and the residues critical for such functions have been investigated by site-directed mutagenesis.

Binding assay. The ligand-binding properties of CRABPs have been studied by several methods, which could be classified into two categories: non-equilibrium methods and fluorescence titration methods. The non-equilibrium methods remove the unbound ligand either by the use of dextran-coated charcoal (Daly & Redfern, 1988; Fogh et al., 1993) or by chromatography (Siegenthaler et al., 1992). A potential problem with such methods is that the bound ligand can come off the protein during the separation process. The K_ds measured by these methods depend on the off-rate of the ligand, and may be an overestimation of its true value. The fluorescence methods (Cogan et al., 1976) measure the concentration of either the bound ligand or the bound protein at equilibrium. However, reported fluorescence measurements have been plagued by poor experimental designs and data analysis. The problems have been recognized recently (Norris et al., 1994; Wang et al., 1997). Furthermore, because of the low solubility of RA in aqueous solution, the fluorescence methods are unsuitable for measuring the dissociation constants of site-directed mutants with greatly reduced affinity for RA. The reported K_d of CRABPI for all-*trans*-RA ranges from 0.4 nM to 10 nM (Ong & Chytil, 1978; Fiorella & Napoli, 1991; Norris, et al., 1994; Fogh et al., 1993), and the K_d of CRABPII from 1.8 nM to 65 nM (Bailey & Siu, 1988; Fiorella & Napoli, 1993; Fogh et al., 1993; Norris et al., 1994; Wang et al., 1997). These discrepancies may be caused by the different methods used, the

assay conditions and the sources of the proteins. The reported relative K_d of CRABPI *vs.* CRABPII ranges from three-fold (Fiorella, et al., 1993; Wang et al., 1997), as measured by competitive binding assay, to more than ten-fold based on their individual absolute values. The value measured by competitive binding assay appears to be more reliable. The stoichiometry of either CRABPI or II for all-*trans*-RA has been found to be one-to-one.

Binding specificity. No significant differences in binding specificity for retinoids have been reported between CRABPI and CRABPII. Both proteins show no apparent affinities for all-*trans*-retinol, all-*trans*-retinal and esters of RA, and have relatively low affinities for the *cis*-isomers of RA. For example, their affinities for 9-*cis*-RA range from ~50 nM (Fiorella et al., 1993) to ~200 nM (Norris et al., 1994), and their affinities for 13-*cis*-RA are at least 3-4 fold lower than that for 9-*cis*-RA (Fiorella et al., 1993) or too low to be measured (Redfern & Wilson, 1993; Fogh et al., 1993; Norris et al., 1994). However, CRABPs have high affinity for retinoids with a modified β-ionone ring such as 3,4-didehydroretinoic acid (Torma et al., 1991), 4-hydroxyretinoic acid, 4-oxo-retinoic acid, 18-hydroxyretinoic acid, 16-hydroxy-4-oxo-retinoic acid (Fiorella et al., 1993) and acitretin (Norris et al., 1994) (an aromatic retinoid used in the treatment of psoriasis). Both proteins also have high affinity for retinoids with modifications at C-7 and C-8 positions of the isoprene chain (Kleywegt et al., 1994). Although 3,4-didehydroretinoic acid exists in cells and binds CRABPs as tightly as all-*trans*-RA does, the latter has been shown to be the endogenous ligand for CRABPI and possibly CRABPII (Ross et al., 1980; Sarri et al., 1982). Taken together, these experimental data indicate that the

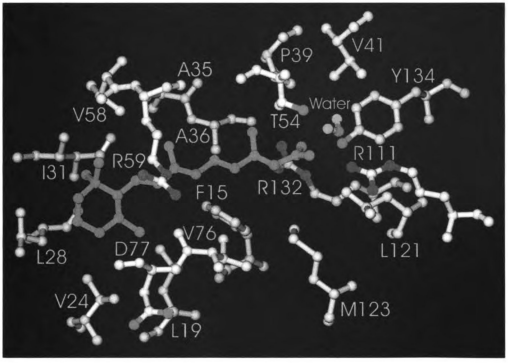
carboxylic group of RA is critical for binding to CRABPs. Modifications at the β-ionone ring and, C-7 and C-8 can be tolerated.

Site-directed mutagenesis studies. The interactions between CRABPs and retinoids have been investigated by site-directed mutagenesis. The single mutants R111Q and R131Q, double mutant R111Q/R131Q and triple mutant R111Q/R131Q/Y133F of bovine/murine CRABPI all have much lower affinity for all-*trans*-RA than the wild-type (Zhang et al., 1992). The two arginines and one tyrosine are conserved residues in CRABPs and interact with the carboxyl group of RA in the crystal structures of holo CRABPs (Figure 1.2A, B) (Kleywegt et al., 1994). Although no quantitative K_d has been reported for these mutants, a CD signal induced by the binding of RA has been observed in the R131Q mutant as in the wild type. No such signal has been detected in the mutant R111Q even with a five-fold excess of all-*trans*-RA. It appears that in CRABPI Arg111 is more critical for binding of RA than Arg131. By contrast, it has been reported that the R111A mutant of mouse CRABPII has a slightly lower affinity for all-*trans*-RA than the wild type, while either R132A or R132Q mutant has much lower affinity (Chen et al., 1995), suggesting that only Arg132 is important for binding of RA by CRABPII. However, no quantitative K_d has been reported for either R132A or R132Q mutant.

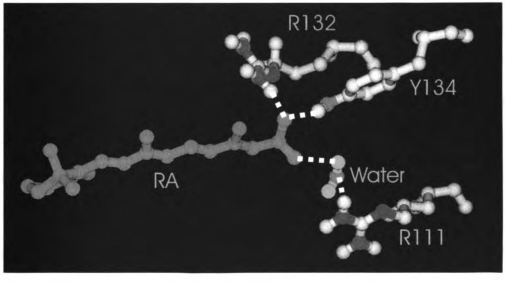
Interestingly, there is no change in the retinoid specificity of CRABPI or CRABPII because of the mutations. In particular, the mutants show no increase in affinity for all-*trans*-retinol and all-*trans*-retinal compared to their wild-types. In other words, neither CRABPI nor CRABPII could be converted into a retinol-binding protein by the single, double or triple mutations. Similarly, type-II cellular retinol binding protein (CRBPII), a close relative of CRABPs, could not be converted into a RA-binding protein

Figure 1.2 (A) The ligand binding pocket of human CRABPII showing the residues (white) interacting with all-*trans*-RA (green). The two hydrogen atoms (yellow) of a water molecule, the nitrogen atoms (blue) of the side chains, and the oxygen atoms (red) of the side chains and the RA molecule and the water are colored. (B) expanded view of the interactions between the carboxyl group of RA and the side chains of Arg111, Arg132 and Tyr134 of the protein. Dotted line indicates the possible hydrogen bond interaction.

A


V41
P39
A35
Y134
V58
Water
T54
A36
R111
I31
R59
F15
R132
L28
D77
V76
L121
V24
M123
L19

B


R132
Y134
RA
Water
R111

by site-directed mutagenesis. The two single mutants Q109R and Q129R (Gln109 and Gln129 of CRBPII are structurally equivalent to Arg111 and Arg132 of CRABPII, respectively) display reduced affinity for all-*trans*-retinol but no apparent affinity for all-*trans*-RA (Cheng et al., 1991). Surprisingly, the mutant protein Q108R of rat CRBPI (Gln108 in rat CRBPI is structurally equivalent to Arg111 of human CRABPII) has similar affinity for all-*trans*-retinol, all-*trans*-RA and 13-*cis*-RA. The mutation reduces the affinity for all-*trans*-retinol by only three-fold compared to the wild-type (Stump et al., 1991).

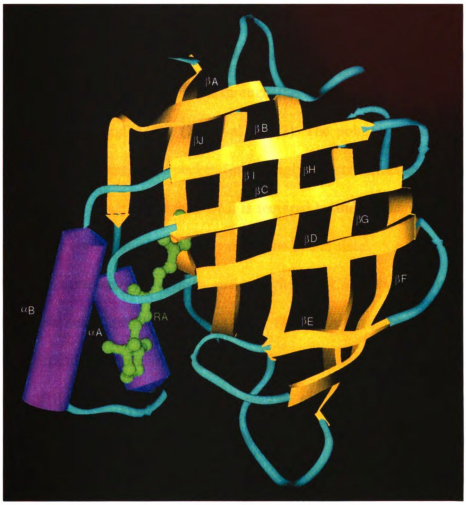
Section 3. Structural and Dynamical Characterizations of CRABPs

The atomic details of the interactions between CRABPs and retinoids, and their internal motions have been analyzed by physical methods, primarily the x-ray diffraction of single crystals and solution NMR spectroscopy. The structural and dynamical information provides a framework for understanding the nature of the interactions and the ligand binding process, and for identifying the factors contributing to the binding affinity and specificity.

The structural features of CRABPs. The crystal structures of apo-CRABPI (Thompson et al., 1995), holo-CRABPI and holo-CRABPII (both in complex with all-*trans*-RA) (Kleywegt et al., 1994) have been determined at 2.7 Å, 2.9 Å and 1.8 Å resolutions, respectively. These studies have demonstrated that the apo- and holo-

CRABPs have very similar secondary and tertiary structures. Moreover, they all have structural features typical of a protein family called intracellular lipid binding proteins (iLBPs) (Banaszak et al., 1994). These proteins all bind hydrophobic ligands such as fatty acid, lipids and retinoids. In addition to CRABPs, the crystal structures of more than twenty-five proteins of the iLBP family have been determined at various resolutions with or without bound ligands. The structures reported include rat CRBPI (Cowan et al., 1993) and CRBPII (Winter et al., 1993), several fatty acid binding proteins (FABPs) including those from chicken liver (Scapin et al., 1990), bovine heart (Muller-Fahrnow et al., 1991), human muscle (Zanotti et al., 1992; Young et al., 1994) and rat intestine (Sacchettini et al., 1989a,b; Sacchettini et al., 1992; Eads et al., 1993), bovine myelin, (Jones et al., 1988; Cowan et al., 1993) and adipocyte lipid binding protein (ALBP) (Xu et al., 1992; 1993). Although the sequence identities among the proteins are rather low (~20%), they all have similar tertiary structures. All the proteins adopt a compact, single domain structure consisting of a helix-turn-helix motif and two antiparallel β-sheets (Figure 1.3). Both sheets are composed of five β-strands, the first βA, βB, βC, βD and βE, and the second βF, βG, βH, βI and βJ. They are twisted into a nearly orthogonal flattened barrel with a simple up-down topology. The regular hydrogen bond ladders between adjacent β-strands are broken between βD and βE. One β-bulge occurs in the middle of the first strand, resulting in a change of its direction so that it forms parts of the two sheets. βF is also shared by the two sheets. A binding cavity is formed between the two β-sheets by the side chains from eight of the ten β-strands. The surface of the cavity could be divided into two regions: a hydrophobic region and a more centrally located polar region (Figure 1.2A). The former is formed by the side chains from the helix-turn-helix, the βC-βD loop

Figure 1.3. The backbone fold and the secondary structural elements of human CRABPII in complex with all-*trans*-RA (green) showing the structural motif of iLBP family: ten antiparallel β–strands (yellow) and two α helices (violet) forming a flattened barrel. The loops or turns are colored in cyan.


βA
βB
βJ
βI
βH
βC
βG
βD
βF
αB
αA
βE

and the βE-βF loop. The latter is formed by the side chains of the residues deep inside the binding pocket such as Arg111, Arg132 and Tyr134 (CRABPII numbering). Generally, only a small opening to the external solvent is observed in the crystal structures, which is supposed to be the ligand entrance. The rear portion of the pocket, opposite to the opening, is completely blocked and inaccessible to solvent. The ligand is either totally buried inside the protein as observed in the majority of the crystal structures or its hydrophobic tail extends to the surface through the ligand entrance as in the crystal structures of CRABPs and ALBP (Xu et al., 1992). The polar group of the ligand is bound in the innermost part of the pocket. Interestingly, in most structures, the volume of the binding pocket is much larger than what is required by the ligand. The extra volume may provide room for the fluctuation of ligand in the bound state.

The structural basis for affinity and specificity of CRABPs. The high-resolution structures of the holo-CRABPs provide a basis for the interpretation of their binding affinity and specificity in terms of the relative positions of atoms. The binding specificity and affinity of iLBPs can be partially explained by variation in the positions and the nature of amino acid residues, and by the tightly bound solvent molecules in the binding pocket. Particularly, the interaction between the polar end of the ligand and the protein, (Banazak et al., 1994), seems to be most critical. In the crystal structure of holo-CRABPII, one of the carboxylate oxygens of RA forms hydrogen bonds with the hydroxyl of Tyr134 and the guanidinium group of Arg132. The other interacts with the guanidinium group of Arg111 mediated by a water molecule (Figure 1.2B). Similar interactions between the carboxyl group and the two arginine and one tyrosine resiudes have been observed in other iLBPs. The importance of such interactions for binding of

the ligands has been verified by site-directed mutagenesis studies. As described earlier, mutation of these residues to other amino acids reduces greatly the binding affinities of iLBPs.

In the structures of holo-CRABPs extensive hydrophobic interactions have been observed between the isoprene chain of RA and CRABPs (Figure 1.2A). Furthermore, a network of water molecules in the binding pocket is often detected in the high-resolution crystal structures. For example, fourteen water molecules have been identified in the crystal structure of holo-CRABPII (Kleywegt et al., 1994). Eight water molecules have been identified in the crystal structure of holo-CRBPI (Cowan et al., 1993) and holo-CRBPII (Winter et al., 1993). They have been suggested to be an integral part of the binding pockets (LaLonde et al., 1994) and to be related to the binding affinity and specificity of these proteins.

Some clues on the binding specificity of CRABPs can be drawn from the comparison of the amino acid sequences and the structures of iLBPs. For example, the two arginines and one tyrosine of CRABPs and FABPs interacting with the carboxyl groups of their ligands are replaced, respectively, by two glutamines and one phenylalanine in CRBPs. And, the arginine-carboxyl interactions in CRABPs and FABPs are substituted by the interactions between the two glutamines and the hydroxyl group of retinol. Indeed, replacement of Gln108 of rat CRBPI with arginine has changed the specificity of the protein so that it binds all-*trans*-RA and 13-*cis*-RA in addition to all-*trans*-retinol (Stump et al., 1991).

Mechanisms of ligand entry or exit. Previous structural studies, mainly by means of x-ray diffraction, have shown that most of the proteins of the iLBP family undergo

only small changes upon ligand binding, and there exists no entrance large enough for the entry or exit of the ligand. In the holo crystal structures of CRBPI, CRBPII, intestinal FABP and ALBP, the ligands are completely buried. The ligand entrances of the unliganded forms are almost identical to those of the holo-forms. The crystal structures of CRABPs show that RA is not completely buried (Kleywegt et al., 1994). The edge of the β-ionone ring of bound RA is exposed to solvent. It appears, however, that significant conformational changes are required for release of the ligand. The ligand entrance of apo-CRABPI is slightly more open than that of holo-CRABPI, and appears to be large enough to admit RA (Thompson et al., 1995). Surprisingly, apo-CRABPI is dimeric in the crystalline state, held together by an intermolecular β-sheet. The slight opening of the ligand entrance has been suggested to be caused by formation of the intermolecular β-sheet.

Biochemical studies, by contrast, have suggested that ligand binding induces conformational and dynamical changes in iLBPs. It has been shown that a microsomal retinol esterifying enzyme, lecithin-retinol acyltransferase, can discriminate between the apo- and the holo-CRBPs (Herr & Ong, 1992). It has also been demonstrated that ligand binding reduces the protease susceptibility of the ligand entrances of CRBPI, CRBPII, CRABPI and heart FABP (Jamison et al., 1994). The results suggest that the apo proteins have more open binding pockets or their entrances are more flexible in solution. These biochemical results are consistent with the physiological functions of these proteins, which require that there exist an entrance large enough for ligand admittance, at least temporarily. Furthermore, if holo-CRABPs and holo-CRBPs are the substrates for metabolic enzymes (Herr & Ong, 1992; Fiorella & Napoli, 1991, 1994; Boylan & Gudas,

1992), there should exist structural features that can be used by the enzyme to distinguish the apo-forms from the holo-forms. Indeed, lecithin-retinol acyltransferase can distinguish the apo and holo forms of CRBPs (Herr & Ong, 1992).

NMR studies of the iLBP family. NMR spectroscopy has the advantage over x-ray diffraction in providing dynamical information. Such information is critical for understanding the ligand binding process. ^{1}H and ^{15}N resonance assignments have been reported for both apo-CRABPI (pH 7.5) and holo-CRABPI (pH 3.8) (Rizo et al., 1994). Secondary structures derived from NOE and chemical shift indices indicate that both the apo- and the holo-CRABPI share the backbone fold of the iLBP family: ten β-strands and a helix-turn-helix. However, significant differences in internal motions between the apo- and holo-forms have been observed in the ligand entrance region. The ligand entrance of apo-CRABPI is rather flexible in solution, and the flexibility is reduced greatly after the binding of RA. Remarkable motional differences in the ligand entrance region have been reported between rat apo and holo intestinal FABP at pH 7.2 (Hodsdon et al., 1995; Hodsdon et al., 1996; Hodsdon & Cistola, 1997a, b). Furthermore, it has been proposed that the decrease in the mobility of the ligand entrance in the holo form is brought about by a series of long-range cooperative interactions that cap and stabilize the C-terminal half of the second helix. In contrast to what was observed in the crystal structure, the ligand entrance of bovine heart FABP is more mobile than the other parts of the molecule, even in the holo form (Lucke et al., 1992; Lassen et al., 1995). Proton resonance heterogeneity has been observed for Ser22-Thr36 in the helix-turn-helix, Thr53-Thr60 in the βC–βD loop and His119-Ala122, indicating that these residues have two or more slowly exchanging conformations in solution.

NMR studies of the bound RA. The conformation of CRABP-bound RA has been studied by NMR (Norris et al., 1995). The NMR results indicate that the bound RA in CRABPII adopts a conformation with the 6-s dihedral angle being -60° skewed from a *cis* conformation, in contrast to the -33° observed in the crystal structure of holo-CRABPII (Kleywegt et al., 1994). However, RA in holo-CRABPI appears to be quite flexible about the 6-s bond. The 8-s and 10-s dihedral angles are 180° ± 30° in both holo-CRABPI and holo-CRABPII in solution as well as in the crystalline state.

References

Astrom, A., Tavakkol, A., Petersson., U., Cromie, M., Elder, J. T., & Voorhees, J. J. (1991) *J. Biol. Chem. 266*, 17662-17666.

Bailey, J. S., & Siu, C. H. (1988) *J. Biol. Chem. 263*, 9326-9332.

Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S. W., & Jones, T. A. (1994) *Adv. Protein Chem. 45*, 89-151.

Blomhoff, R., Green, M. H., Berg, T., & Norum, K. R. (1992) *Science 250*, 399-404.

Boylan, J. F., & Gudas, L. J. (1991) *J. Cell. Biol. 112*, 965-978.

Boylan, J. F., & Gudas, L. J. (1992) *J. Biol. Chem. 267*, 21486-21491.

Bucco, R. A., Zheng, W., Davis, J. T., Sierra-Rivera, E., Osteen, K. G., Chaudhary, A. K., & Ong, D. E. (1997) *Biochemistry 36*, 4009-4014.

Chambon, P. (1996) *FASEB. J 10*, 940-954.

Chambon, P. (1994) *Sem. Cell Biol. 5*, 115-125.

Chambon, P., Olson, J. A., & Ross, A. C. (coordinators) (1996) *The retinoid revolution. FASEB. J 10*, 938-1048.

Chen, L. X., Zhang, Z. -P., Scafonas, A., Cavalli, R. C., Gabriel, J. L., Soprano, K. J., & Soprano, D. R. (1995) *J. Biol. Chem. 270*, 4518-4525.

Cheng, L., Qian, S-J., Rotthschild, C., d'Avignon, A., Lefkowith, J. B., Gordon, J. I., & Li, E. (1991) *J. Biol. Chem. 266*, 24404-24412.

Cogan, U., Kopelman, M., Mokady, S., Shinitzky, M. (1976) *Eur. J. Biochem. 65*, 71-78.

Cowan, S. W., Newcomer, M. E., & Jones, A. (1993) *J. Mol. Biol. 230*, 1225-1240.

Daly, A. K., & Redfern, C. P. F. (1988) *Biochim. Biophys. Acta 985*, 118-126.

de Bruijn, D. R., Oerlemans, F., Hendriks, W., Baats, E., Ploemacher, R., Wieringa, B., & Geurts van Kessel, A. *Differentiation 58*, 141-148.

Durand, B., Saunders, M., Leroy, P., Leid, M., & Chambon, P. (1992) *Cell 71*, 73-85.

Eads, J., Sacchettini, J. C., Kromminga, A., Gordan, J. I. (1993) *J. Biol. Chem. 268*, 26375-26385.

Eller, M. S., Oleksiak, M. F., McQuaid, T. J., McAfee, S. G., & Gilchrest, B. A. (1992) *Exp. Cell Res. 199*, 328-336.

Evans, R. M. (1988) *Science 240*, 889-895.

Fiorella, P. D., Giguere, V., & Napoli, J. L. (1993) *J. Biol. Chem. 268*, 21545-21552.

Fiorella, P. D., & Napoli, J. L. (1991) *J. Biol. Chem. 266*, 16572-16579.

Fiorella, P. D., & Napoli, J. L. (1994) *J. Biol. Chem. 269*, 10538-10544.

Fogh, K., Voorhees, J. J., & Astrom, A. (1993) *Arch. Biochem. Biophys. 300*, 751-755.

Giguere, V., Lyn, S., Yip, P., Siu, C-H., & Amin, S. (1990) *Proc. Natl. Acad. Sci. 87*, 6233-6237.

Giguere, V., Ong, E. S., Segui, P., & Evans, R. M. (1987) *Nature 330*, 624-629.

Herr, F. M., & Ong, D. E. (1992) *Biochemistry 31*, 6748-6755.

Heyman, R. A., Mangelsdorf, D. J., Dyck, J. A., Stein, R. B., Eichele, G., Evans, R. M. & Thaller, C. (1992) *Cell 68*, 397-406.

Hodson, M. E., & Cistola, D. P. (1997a) *Biochemistry 36*, 1450-1460.

Hodson, M. E., & Cistola, D. P. (1997b) *Biochemistry 36*, 2278-2290.

Hodson, M. E., Ponder, J. W., & Cistola, D. P. (1996) *J. Mol. Biol. 264*, 585-602.

Hodson, M. E., Toner, J. J., & Cistola, D. P. (1995) *J. Biomol. NMR 6*, 198-210.

Hofmannn, C., & Eichele, G. (1994) In *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., & Goodman, D. S., eds.), *pp.* 387-441, 2nd ed., Raven Press, New York.

Hong, W. K., & Itri, L. M. (1994) In *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., & Goodman, D. S., eds.), *pp.* 597-630, 2nd ed., Raven Press, New York.

Jamison, R. S., Newcomer, M. E., & Ong, D. E. (1994) *Biochemistry 33*, 2873-2879.

Jing, Y., Waxman, S., & Mira-y-Lopez, R. (1997) *Cancer Res. 57*, 1668-1672.

Jones, T. A., Bergfos, T., Sedzik, J., & Unge, T. (1989) *EMBO J. 7*, 1597-1604.

Kitomoto, T., Momoi, T., & Momoi, M. (1988) *Biochem. Biophys. Res. Commun. 157*, 1302-1308.

Kleywegt, G. J., Bergfors, T., Senn, H., le Motte, P., Gsell, B., Shuao, K., & Jones, T. A. (1994) *Structure 2*, 1241-1258.

LaLonde, J. M., Bernlohr, D. A., & Banaszak, L. J. (1994) *Biochemistry 33*, 4885-4895.

Lassen, D., Lucke, C., Kveder M., Mesgarzadeh, A., Schmidt, J. M., Specht, B., Lezius, A., Spener, F., & Ruterjans, H. (1995) *Eur. J. Biochem.* 230, 266-280.

Levin, A. A., Sturzenbecker, L. J., Kazmer, S., Bosakowski, T., Huselton, C., Allenby, G., Speck, J., Kratzeisen, C., Rosenberger, M., Lovey, A., & Grippo, J. F. (1992) *Nature 355*, 359-361.

Lucke, C., Lassen, D., Kreienkamp, H.-J., Spener, F., & Ruterjans, H. (1992) *Eur. J. Biochem. 210*, 901-910.

Mangelsdorf, D. J., & Evans, R. M. (1995) *Cell 83*, 841-850.

Mangelsdorf, D. J., Ong, E. S., Dyck, J. A., & Evans, R. M. (1990) *Nature 345*, 224-229.

Mangelsdorf, D. J., Umesono, K., & Evans, R. M. (1994) The retinoid receptors. In *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., & Goodman, D. S., eds.), *pp.* 319-349, 2nd ed., Raven Press, New York.

Maden, M. (1991) *Dev. Biol. 2*, 161-170.

Maden, M., Ong, D. E., Summerbell, D., & Chytil, F. (1988) *Nature 335*, 733-735.

Melhus, H., Gobl, A., & Ljunghall, S. (1994) *Biochem. Biophys. Res. Comm. 200*, 1125-1129.

Morriss-Kay, G. M., & Sokolova, N. (1996) *FASEB. J 10*, 961-968.

Muller-Fahrnow, A., Egner, U., Jones, A., Ruedel, H., Spener, F., & Saenger, W. (1991) *Eur. J. Biochem. 199*, 271-276.

Nilsson, M. H., Spurr, N. K., Saksena, P., Busch, C., Norlinder, H., Peterson, P. A., Rask, L., & Sundelin, J. (1988) *Eur. J. Biochem. 173*, 45-51.

Norris, A. W., Cheng, L., Giguere V., Rosenberger, M., & Li, E. (1994) *Biochim. Biophys. Acta 1209*, 10-18.

Norris, A. W., Rong, D., d'Avignon, D. A., Rosenberger, M., Tasaki, K., & Li. E. (1995) *Biochemistry 34*, 15564-15573.

Ong, D. E., & Chytil, F. (1975) *J. Biol. Chem. 250*, 6113-6117.

Ong, D. E., & Chytil, F. (1978) *J. Biol. Chem. 253*, 4551-4554.

Ong, D. E., Newcomber, M. E., & Chytil, F. (1994) In *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., & Goodman, D. S., eds), *pp.* 283-317, 2nd ed., Raven Press, New York.

Peck, G. L., & DiGiovanna, J. J. In *The Retinoids: Biology, Chemistry, and Medicine.*

(Sporn, M. B., Roberts, A. B., & Goodman, D. S., eds.), *pp.* 631-658, 2nd ed., Raven Press, New York.

Petkovich, M., Brand, N. J., Krust, A., & Chambon, P. (1987) *Nature 330*, 444-450.

Rizo, J., Liu, Z. P., & Gierasch, L. M. (1994) *J. Biomol. NMR 4*, 741-760.

Ross, A. C., Adachi, N., & Goodman, D. S. (1980) *J. Lipid Res. 21*, 100-109.

Sacchettini, J. C., Gordan, J. I., & Banaszak, L. J. (1988) *J. Biol. Chem. 263*, 5815-5819.

Sacchettini, J. C., Gordan, J. I., & Banaszak, L. J. (1989a) *Proc. Natl. Acad. Sci. U. S. A. 86*, 7736-7740.

Sacchettini, J. C., Gordan, J. I., & Banaszak, L. J. (1989b) *J. Mol. Biol. 208,* 327-339.

Sacchettini, J. C., Scapin, G., Gopaul, D., & Gordan, J. I. (1992) *J. Biol. Chem. 267*, 23534-23545.

Sani, B. P., & Hill, D. L. (1974) *Biochem. Biophys. Res. Commun. 61*, 1276-1282.

Sarri, J. C. (1982) *Methods Enzymol. 81*, 819-826.

Scapin, G., Spadon, P., Mammi, M., Zanotti, G., & Monaco, H. L. (1990) *Mol. Cell. Biochem. 98*, 95-99.

Shubeita, H. E., Sambrook, J. F., & McCormick, A. M. (1987) *Proc. Natl. Acad. Sci. U. S. A.* 84, 5645-5649.

Siegenthaler, G., Tomatis, I., Chatellard-Gruaz, D., Jaconi, S., Erikson, U., & Saurat, J. H. (1992) *Biochem. J. 287*, 383-389.

Sporn, M. B., Roberts, A. B., & Goodman, D. S. (1994) *The Retinoids: Biology, Chemistry, and Medicine*, 2nd ed., Raven Press, New York.

Stoner, C. M., & Gudas, L. J. (1989) *Cancer Res. 49*, 1497-1504.

Stump, D. G., Lloyd, R. S., & Chytil, F. (1991) *J. Biol. Chem. 266*, 4622-4630.

Szuts, E. Z., & Harosi, F. I. (1991) *Arch. Biochem. Biophys. 287*, 297-304.

Takase, S., Ong, D. E., & Chytil, F. (1986) *Arch. Biochem. Biophys. 247*, 328-334.

Thompson, J. R., Bratt, J. M., & Banaszak, L. J. (1995) *J. Mol. Biol. 252,* 433-446.

Torma, H., Stenstrom, E., Andersson, E., & Vahlquist, A. (1991) *Skin Pharmacol. 4*, 246-253.

Tsai, M- J., & O'Malley, B. W. (1994) *Ann. Rev. Biochem. 63*, 451-486.

Wald, G. (1968) *Science 162*, 230-239.

Wang, L., Li, Y., & Yan, H. (1997) *J. Biol. Chem. 272*, 1541-1547.

Winter, N., Bratt, J., & Banaszak, L. J. (1993) *J. Mol. Biol. 230,* 1247-1259.

Xu, Z., Bernlohr, D. A., & Banaszak, L. J. (1992) *Biochemistry 31*, 3484-3492.

Xu, Z., Bernlohr, D. A., & Banaszak, L. J. (1993) *J. Biol. Chem. 268*, 7874-7884.

Young, A. C. M., Scapin, G., Kromminga, A., Patel, A. B., Veerkamp, J. H., & Sacchettini, J. C. (1994) *Structure 2*, 525-534.

Yu, V. C., Delsert, C., Andersen, B., Holloway, J. M., Devary, O. V., Naar, A. M., Kim, S. Y., Boutin, J. M., Glass, C. K., & Rosenfeld, M. G. (1991) *Cell 67*, 1251-1266.

Zanotti, G., Scapin, G., Spadon, P., Veerkamp, J. H., & Sacchettini, J. C. (1992) *J. Biol. Chem. 267*, 18541-18550.

Zhang, J., Liu, Z., Jones, T. A., Gierasch, L. M., & Sambrook, J. F. (1992) *Proteins: Structure, Function, and Genetics 13*, 87-99.

Part 2. Physical Basis of Protein NMR

Section 1. Theory of Fourier Transform Pulse NMR

Nuclear magnetic resonance (NMR) spectroscopy, the detection of the Zeeman levels of nuclear spins in bulk condensed phase by resonant method through electromagnetic means, was pioneered by two research groups headed by Felix Bloch (Bloch et al., 1946a) and Edward M. Purcell (Purcell et al., 1946), respectively. It has since become an important experimental tool for physicists, chemists and more recently for structural biologists as well as radiologists (medical diagnosis). Its wide applications are made possible by continuous improvement in experimental techniques and advances in theory. The former is stimulated by technological breakthroughs in other fields such as electronics, computer science (e.g. fast Fourier transformation algorithm and the sophisticated controls of experimental parameters), and the manufacturing of stable high field superconducting magnets. The latter takes advantage of the weakness of the interactions involving nuclear spins. The weakness makes it relatively easy to manipulate the interaction Hamiltonian through experimental methods such as radio frequency pulses, and justifies the treatment of the evolution of the states of a spin system by perturbation methods. The simplicity of the derived mathematical expressions facilitates

study biological macromolecules. In this review I will summarize the underlying physical principles of NMR, and will show how they can be applied to practical problems especially those in structural biology.

Basic equations. Nuclear spin, a quantum mechanical phenomenon without classical analogs, can only be treated by quantum mechanics. However, for an isolated system consisting of non-interacting spins, the equation for the expectation value of the total magnetization $\langle \boldsymbol{M} \rangle$ is the same as that for the classical magnetization under the influences of a static field $\boldsymbol{H}_0$ and a rotating external radio-frequency field $\boldsymbol{H}_1$ (r. f. field) (Rabi et al., 1954).

$$d\langle \boldsymbol{M} \rangle / dt = \gamma \langle \boldsymbol{M} \rangle \times \boldsymbol{H}_{\text{eff}} \qquad (1)$$

where $\boldsymbol{H}_{\text{eff}}$ is the total magnetic field in a frame rotating with the frequency of the r. f. field, γ is the gyromagnetic ratio. Therefore, it is legitimate to treat such a system classically. In reality the spin system is not isolated but interacts with the molecular environment (bath) including all degrees of freedom other than the spins. The interaction causes a decay of the observed magnetization. With the assumptions that the decay is exponential, and the interactions with the external fields and with the bath are additive, Bloch derived a phenomenological equation for a non-interacting spin system (Bloch, 1946b).

$$d\boldsymbol{M}/dt = \gamma\, \boldsymbol{M} \times \boldsymbol{H} - \boldsymbol{M}^{+}/T_2 - \mathbf{k}\,(\boldsymbol{M}_z - \boldsymbol{M}_0)/T_1 \qquad (2)$$

where $\boldsymbol{H}$ is the sum of a static field $\boldsymbol{H}_0 = \mathbf{k}H_0$ and a r. f. field $\boldsymbol{H}_1$, $\mathbf{k}$ is the unit vector on the z-axis, $\boldsymbol{M}_0 = \chi_0\boldsymbol{H}_0$ is the equilibrium magnetization with χ_0 being the static susceptibility, $\boldsymbol{M}$ is the total magnetization, $\boldsymbol{M}^+$ is the transverse magnetization. T_1 and T_2 are, respectively, the longitudinal and transverse relaxation times representing the interactions between the spin system and the bath. The nuclear spins in a real spin system interact with each other. The descriptions of such a spin system must resort to quantum mechanism. For convenience, the Hamiltonian H of the conserved whole system is divided into three parts: the Hamiltonian $H_s(s, t)$ of the spin system including the interactions with the external static and r. f. fields, the Hamiltonian $F(f)$ of the bath and the interaction Hamiltonian $G(f, s, t)$ between them. The state of a spin system can be represented by a reduced density operator $\sigma(t)$ (von-Neumann, 1927, 1955; Fano, 1957), which could be defined as an incoherent superposition of pure states. The dynamical evolution of $\sigma(t)$ under the influence of the spin Hamiltonian $H_s(t)$ follows the Liouville-von Neumann equation.

$$\partial\sigma / \partial t = -i\,[H_s(t), \sigma(t)] \tag{3}$$

$\sigma(t)$ is simply related to the density operator $\rho(t)$ of the whole system: $\sigma(t) = \mathrm{tr}_f\{\rho(t)\}$, here tr_f means the partial trace over the variables of the bath. The expectation value $\langle Q(t)\rangle$ of any spin operator $Q(t)$ can be calculated according to $\langle Q(t)\rangle = \mathrm{tr}_s\{Q(t)\sigma\}$, here tr_s means the partial trace over the variables of the spin system. The separation of the whole system into two subsystems (the spin system and bath) is very helpful in the

development of NMR theory because of the weakness of the interaction between the spin system and the bath. For example, in the operator product formalism discussed later, an exponential term multiplied into a spin operator can be used to represent the effect of relaxation on the spin system caused by the bath. In relaxation theory, irreversible processes are assumed to occur exclusively in the spin system though only reversible processes could happen in the conserved whole system.

Pulse NMR. For most pulse NMR experiments, the exact analytic solutions of the Liouville-von Neumann equation (3) are very difficult to find because of the complexities of the interactions among the spin system itself, and between the spin system and its environment. To make the problem tractable and provide useful qualitative insights into NMR experiments, various approximations have been proposed to simplify the Hamiltonians and/or the reduced density operator depending on the nature of the practical system under investigation. One of the simplified versions of the density operator formalism called operator product formalism (Banwell & Primas, 1962; Sorensen et al., 1983; Van de Ven & Hibers, 1983) has been frequently employed to describe the experiments performed on a weakly coupled spin system in the context of high resolution NMR. Following a suggestion by Fano (Fano, 1957), the main idea behind this formalism is to expand the $\sigma(t)$ in a tensor base constructed from the products of spin operators (E, σ_x, σ_y, σ_z of the Pauli matrices) (Sorensen et al., 1983; Ernst et al., 1987).

$$\sigma(t) = \Sigma_s \, b_s(t) \, B_s \quad and \quad B_s = 2^{(q-1)} \prod_{k=1}^{N} (I_{kv})^{a_{sk}} \tag{4}$$

where B_s is a base operator in the Liouville space consisting of all spin operators, N = the total number of I = ½ nuclei in the spin system, k = index of nucleus, v = x, y, or z, q = number of single-spin operators in the product, $a_{sk} = 1$ for q nuclei and $a_{sk} = 0$ for the N-q remaining nuclei. The spin Hamiltonian H_s of a weakly coupled spin system in the context of high resolution NMR in liquid phase could be approximated by $\Omega_k I_{kz}$, $\pi J_{kl} 2I_{kz} I_{lz}$ and βI_{kv} (here β is the flip angle of per unit time about v axis), respectively. The first two are time-independent, representing the interaction with the static field (the chemical shifts) and the interaction between nuclear spins (spin-spin coupling), respectively. The other is a time-dependent term, representing the interaction with external r. f. fields (pulses), which could be made time-independent for a finite time segment by selecting a suitable rotating frame (Rabi et al., 1954). The solution of the equation (3) for a time-independent propagator U is very simple

$$\sigma(t) = \exp(-iUt)\, \sigma(0)\, \exp(+iUt) \tag{5}$$

Most of the NMR experiments could be calculated by means of this equation and commutation relations. With the simplified Hamiltonians mentioned above the we could derive all the results from operator product formalism. However, the physical picture was more clear in the latter, whose advantage originates from the simplification of the spin Hamiltonian H_s such that each of the three terms is itself a base operator. Therefore, the evolution of any product operator under the effect of any of the three terms could be

calculated according to the commutation rules among them. If there exists the following commutation relation among operators *A*, *B* and *C* (and their cyclic permutations)

$$[A, B] = iC \tag{6}$$

the operator A under the propagator $U = H_s = C$ will evolve as

$$\exp(-i\theta C)\, A \exp(+i\theta C) = A\cos\theta + B\sin\theta \tag{7}$$

Therefore, the space spanned by the three independent operators *A*, *B* and *C* is a subspace of the Liouville space. The majority of pulse NMR experiments applied to the structural biological problems, where the H_s could be approximated by the three terms above, could be easily treated in this manner. However, this formalism is not adequate for a complete mathematical description of NMR relaxation in liquid.

Fourier Transform NMR. The method of recording NMR signals free of pulse was widely used in the T_1 and T_2 measurements by the spin-echo method (Hahn, 1950; Carr & Purcell, 1954; Meiboom & Gill, 1958). A non-echo method has also been developed to record NMR signals free of pulse by means of *Fourier* transform. It has been shown that after applying a strong external r. f. field $\boldsymbol{H}_1$ for a very short time τ (r. f. pulse) the transverse magnetization $\boldsymbol{M}^+(t) = \boldsymbol{M}_x + i\,\boldsymbol{M}_y$ evolves in the absence of a pulse according to (Lowe & Norberg, 1957)

$$\boldsymbol{M}^+(t) = \boldsymbol{M}_0 \sin\omega_1\tau \exp(i\omega_0 t) \underbrace{\int f(\omega_0 + \omega)\exp(-i\omega t)\, d\omega}_{G(t)} \tag{8}$$

where M_0 is the equilibrium magnetization, ω_0 is the central Larmor frequency, ω_1 is the frequency of the applied r. f. field, $f(\omega_0 + \omega)$ is the shape function related to the imaginary part of the r. f. susceptibility χ''. The integral is from $-\infty$ to $+\infty$. $M^+(t)$ processes with a time-dependent amplitude $G(t)$ called correlation function, which is the Fourier transform of the shape function $f(\omega_0 + \omega)$. Therefore, the recording of the decay of a NMR signal (called free induction decay) represented by $G(t)$ gives the same information about the shape function as the observation of the resonance with a vanishing small r. f. field (the continuous wave NMR). The superiority of the Fourier transform NMR over the continuous wave NMR becomes obvious when it was applied to complex systems (Ernst & Anderson, 1966). One of the advantages is the sensitivity improvement, which becomes more apparent for complicated high resolution spectra in liquid. Furthermore, Fourier transformation consists in the basis for extending the dimensionality of NMR experiments.

Multidimensional NMR. Two-dimensional (2D) NMR experiments, first proposed by Jeener (Jeener, 1971), are natural extensions of double resonance experiments (Bloch, 1954, 1958; Royden, 1954; Bloom & Schoolery, 1955). Both techniques are designed to observe the properties of a system, which are related to its nonlinear responses to external perturbations such as the connectivity of transitions in various energy levels. In contrast to the normal double resonance techniques, the parameter t_1 (time in the evolution period) in 2D techniques is changed systematically as a variable, and a Fourier transform is applied to both time variables t_1 and t_2 (Aue et al., 1976).

$$S(\omega_1, \omega_2) = \int dt_1 \exp(-i\omega_1 t_1) \int dt_2 [\exp(-i\omega_2 t_2)] s^+(t_1, t_2) \quad (9)$$

$$s^+(t_1, t_2) = \mathrm{tr}_s\{F^+\sigma(t_1, t_2)\} = \sum_{rs}\sum_{tu}\exp\{(-i\omega_{rs}^{(d)} - \lambda_{rs}^{(d)})t_2\} Z_{rs\,tu} \exp\{(-i\omega_{tu}^{(e)} - \lambda_{tu}^{(e)})t_1\} \quad (10)$$

$$\sigma(t_1, t_2) = \exp\{-(i\mathcal{H}_s^{(d)} + \Gamma_s^{(d)})t_2\} \mathcal{R} \exp\{-(i\mathcal{H}_s^{(e)} + \Gamma_s^{(e)})t_1\} \mathcal{P}_s\sigma_0 \quad (11)$$

where $S(\omega_1, \omega_2)$ represents the cross-peaks of the 2D spectrum, $s^+(t_1, t_2)$ is the observed complex magnetization signal (FID) and $F^+ = \sum I_x + iI_y$ is the observable spin operator. $\mathcal{H}_s^{(d)}$, $\Gamma_s^{(d)}$ and $\mathcal{H}_s^{(e)}$, $\Gamma_s^{(e)}$ are the Hamiltonian and relaxation superoperators in the detection and evolution periods, respectively. $Z_{rs\,tu} = F^+_{sr} \mathcal{R}_{rs\,tu} (\mathcal{P}_s\sigma_0)_{tu}$, here $\mathcal{P}_s\sigma_0$ is the density operator after the preparation period, $\mathcal{P}_s$ is the preparation superoperator and σ_0 is the density operator at $t = 0$. $\mathcal{R}$ is a rotation superoperator representing the r. f. pulses, which couples the coherences in the evolution period with those in the detection period. In other words, $\mathcal{R}$ induces coherence transfers which are representative of the molecular system under investigation, and is the characteristic feature distinguishing the various forms of 2D experiments. In three- and four-dimensional experiments, there may be two or three evolution periods, and two or three rotational superoperator $\mathcal{R}$'s transferring the coherences between the different evolution periods and between the last evolution period and the final detection period.

Selection of pathway. For the success of a 2D experiment, certain desired coherence transfer pathways must be selected while others are suppressed by either phase-cycling (Wokaun & Ernst, 1977) or pulsed field gradients (Maudsley et al, 1978; Bax et al., 1980). Both methods rely on the different transformational behaviors of the

coherences with different orders under phase-shifting or gradients. In NMR experiments coherences are transferred between the different orders by a propagator $U(\varphi)$ (5), which changes with its phase φ in the following manner.

$$U(\varphi) = \exp(-i\varphi F_z)\, U(0)\, \exp(+i\varphi F_z) \tag{12}$$

where φ is defined as the angle of the displacement from the x-axis towards the y-axis, $U(0)$ is the propagator with phase zero and $F_z = \sum I_z$. A coherence (σ^p) with order p transforms under a rotation about the z-axis in a manner characteristic of its order: the change in phase is proportional to the order.

$$\exp(-i\varphi F_z)\sigma^p \exp(+i\varphi F_z) = \sigma^p \exp(-ip\varphi) \tag{13}$$

Therefore, different coherence order can be separated from each other by means of shifting the phase of the propagator U according to the equation (12) and (13). More importantly, according to (13) a certain desired coherence order such as p can be selected by means of a discrete Fourier transformation of the combined data acquired with distinct phases φ.

The selection of the desired coherence transfer pathways by pulsed field gradients is essentially based on the same mechanism: the proportionality between the phase angle ϕ_r rotated by the application of a shaped gradient and the coherence order p (Maudsley et al, 1978; Bax et al., 1980; John et al, 1991; Davis et al., 1992)

$$\phi_r = \gamma(sG\tau)rp \tag{14}$$

where s is the shape factor, G is the gradient strength, τ is the gradient duration, $sG\tau$ is the area of the gradient pulse and r is the distance from the gradient isocenter.

Pure phase spectrum. In 2D NMR, a single coherence transfer pathway invariably leads to mixed peaks with unwanted features, each of them is a superposition of a pure absorptive peak and a pure dispersive peak. A 2D spectrum with pure phase can be obtained, for example, by either adding or subtracting the signals from the two mirror-image coherence pathways followed by a 2D Fourier transformation, which is real with respect to t_1 and complex with respect to t_2 (other factors in equations (9) and (10) not relevant to the discussion here are dropped unless stated otherwise).

$$s(t_1, t_2) = 2\cos(\omega_{tu}^{(e)} t_1) \exp(-i\omega_{rs}^{(d)} t_2) \tag{15}$$

$$S(\omega_1, \omega_2) = 2\int dt_1 \cos(\omega_1 t_1) \int dt_2 [\exp(-i\omega_2 t_2)]\, s(t_1, t_2)$$

$$= \{a_{tu}(\omega_1) + a_{tu}(-\omega_1)\}\{a_{rs}(\omega_2) - i d_{rs}(\omega_s)\} = a_{tu}(\omega_1)a_{rs}(\omega_2) + \text{additional terms} \tag{16}$$

where $a(\omega)$ and $d(\omega)$ are the Lorentzian absorption and dispersion component, respectively. It is assumed that the two pathways have the same magnitude. Here, only the added signal is retained and the subtracted one is disregarded. Therefore, without regard to the noise, the sensitivity decreases two-fold in order to obtain a pure phase spectrum.

Quadrature detection. In the direct detection dimension (t_2), the sign of a resonance frequency $\omega_{rs}^{(d)}$ can be discriminated by recording complex signals $s^+(t_1, t_2) =$

$s_x(t_1, t_2) + i\, s_y(t_1, t_2)$ (Redfield & Gupta, 1971). Following the same principle the sign of a resonance frequency in the indirect (t_1) dimension could be distinguished (States method) by recording the two signals $s_A(t_1, t_2)$ and $s_B(t_1, t_2)$ resulting, respectively, from the two distinct phases $\varphi_A = 0$ and $\varphi_B = \pi / 2|p|$ of the pulse at the beginning of the evolution period (t_1) (States et al., 1982). Either of them may be understood as a superposition of the two equally weighted mirror-image coherence transfer pathways with frequencies $\omega_{tu}^{(e)}$ and $-\omega_{tu}^{(e)}$. After a complex Fourier transformation with respect to t_2 we have

$$s_A(t_1, t_2)=2\cos(\omega_{tu}^{(e)}t_1)\exp(-i\omega_{rs}^{(d)}t_2) \text{ and } S_A(t_1, \omega_2)=2[a_{rs}(\omega_2)- id_{rs}(\omega_2)]\cos(\omega_{tu}^{(e)}t_1) \quad (17)$$

and

$$s_B(t_1, t_2)=2\sin(\omega_{tu}^{(e)}t_1)\exp(-i\omega_{rs}^{(d)}t_2) \text{ and } S_B(t_1, \omega_2)=2[a_{rs}(\omega_2)-id_{rs}(\omega_2)]\sin(\omega_{tu}^{(e)}t_1) \quad (18)$$

With the assumption that the magnitudes of the two signals are the same irrespective of the possible differences of their transfer pathways due to the relaxation, a pure phase spectrum with sign discrimination can be obtained by first constructing a complex signal $s_C(t_1, \omega_2) = \mathrm{Re}S_A(t_1, \omega_2) + i\,\mathrm{Re}S_A(t_1, \omega_2)$, and subsequently transforming it by a complex Fourier transformation with respect to t_1 (keeping only the pure absorption term).

$$S_C(\omega_1, \omega_2) = 2\, a_{rs}(\omega_2)\, a_{tu}(\omega_1) \quad (19)$$

Please note that the signal from one of the two mirror-image coherence pathways is discarded with two-fold decrease in sensitivity without regard to the noise. The sign discrimination in the indirect detection dimension can also be achieved by increasing the

phase of the pulse at the beginning of the evolution period (t_1) in proportional to the evolution time t_1 (TPPI method) (Bodenhausen et al., 1980b; Marion & Wuthrich, 1983)

$$\varphi = \pi/(2|p|t_1\Delta t_1) = \omega^{(TPPI)} / t_1 \text{ with } \omega^{(TPPI)} = \pi/(2|p|\Delta t_1) \qquad (20)$$

where $1/\Delta t_1$ is the sampling rate in t_1 dimension. Therefore, the signals in t_1 dimension appear to be shifted in frequency $\omega_1 = \omega_{nu}^{(e)} + p_{nu}\,\omega^{(TPPI)}$, here p_{nu} is the coherence order in the evolution period. The signals from the coherences with opposite orders such as ± 1 are shifted in opposite directions. A pure phase spectrum could be obtained by superimposing the different signals calculated by real Fourier transformations with respect to t_1. The third method (States-TPPI) to discriminate the sign is essentially a States method with the exception that the axial peaks are moved to the edge of the spectrum by the TPPI method (Marion et al, 1989a).

Sensitivity enhancement. As shown before, the signal from one of the two mirror-image pathways is usually discarded to obtain a pure phase spectrum. However, under certain circumstances, it can be retained to increase the sensitivity, for example, by shifting the phase of a pulse in the mixing period (Cavanagh et al., 1991, Palmer et al., 1991). This can be best illustrated with a simple pulse sequence on a system consisting of non-interacting spin 1/2.

$$90^\circ_\beta - t_1 - 90^\circ_x - 90^\circ_\phi - t_2 \text{ (detection)}$$

where $\phi = \pm x$ and $\beta = y$ or $-x$. The two signals correspond to $\beta = y$ and $\phi = \pm x$ are

$$s_y^{-x}(t_1, t_2) = C_y^{-x} \exp(-i\omega_{rs}^{(d)} t_2) \, \mathrm{Re}[I^+ \exp(-i\omega_{tu}^{(e)} t_1)] \text{ and}$$

$$s_y^{x}(t_1, t_2) = C_y^{x} \exp(-i\omega_{rs}^{(d)} t_2) \, \mathrm{Re}[I^+ \exp(i\omega_{tu}^{(e)} t_1)] \qquad (21)$$

where *C's* represent other factors in equations (9) and (10). These two signals look like they originate from two mirror-image coherence transfer pathways, and are related with each other by equation (12) with $\varphi = \pi$. Therefore, amplitude-modulated signals (pure phase spectrum) can be obtained by either adding (s_y^a) or subtracting (s_y^s) the two signals assuming $C_y^{-x} = C_y^{x} = C_y$.

$$s_y^a = C_y \, 2I_x \cos(\omega_{tu}^{(e)} t_1) \exp(-i\omega_{rs}^{(d)} t_2) \text{ and } s_y^s = C_y \, 2iI_y \sin(\omega_{tu}^{(e)} t_1) \exp(-i\omega_{rs}^{(d)} t_2) \qquad (22)$$

The two signals with $\beta = -x$ can be obtained from (20) by equation (12).

$$s_{-x}^{-x}(t_1, t_2) = C_{-x}^{-x} \exp(-i\omega_{rs}^{(d)} t_2) \, \mathrm{Re}[\exp(-i\pi/2 \, F_z) I^+ \exp(+i\pi/2 \, F_z) \exp(-i\omega_{tu}^{(e)} t_1)]$$

$$= C_{-x}^{-x} \exp(-i\omega_{rs}^{(d)} t_2) \, \mathrm{Re}[\exp(-i\pi/2) I^+ \exp(-i\omega_{tu}^{(e)} t_1)]$$

$$s_{-x}^{x}(t_1, t_2) = C_{-x}^{x} \exp(-i\omega_{rs}^{(d)} t_2) \, \mathrm{Re}[\exp(-i\pi/2) I^+ \exp(i\omega_{tu}^{(e)} t_1)] \qquad (23)$$

The signals similar to (21) could be obtained by either adding (s^a_{-x}) or subtracting (s^s_{-x}) the two signals assuming $C_{-x}^{-x} = C_{-x}^{x} = C_{-x}$.

$$s^a_{-x} = -C_{-x} \, 2I_x \sin(\omega_{tu}^{(e)} t_1) \exp(-i\omega_{rs}^{(d)} t_2) \text{ and } s^s_{-x} = C_{-x} \, 2iI_y \cos(\omega_{tu}^{(e)} t_1) \exp(-i\omega_{rs}^{(d)} t_2) \qquad (24)$$

With the assumption that $C_{-x} = C_y = C$ new complex time domain data can be constructed by combining the two signals $s^r(t_1, t_2)$ and $s^i(t_1, t_2)$ with respect to t_1 as in the usual States method prior to the complex Fourier transformation.

$$s^r(t_1, t_2) = s_y^a(t_1, t_2) - \mathrm{i}\, s^s{}_{-x} = 4C\cos(\omega_{nu}^{(e)} t_1)\exp(-\mathrm{i}\omega_{rs}^{(d)} t_2) \text{ and}$$

$$s^i(t_1, t_2) = s^a{}_{-x}(t_1, t_2) + \mathrm{i}\, s^s{}_y = -4C\sin(\omega_{nu}^{(e)} t_1)\exp(-\mathrm{i}\omega_{rs}^{(d)} t_2) \qquad (25)$$

The sensitivity is enhanced by two-fold compared to the usual States method (19). It originates from the preservation and combination of the two orthogonal transverse magnetization components I_x and I_y with the condition that I_x and I_y must be equivalent in the coherence transfer processes. However, the enhancement is only $2^{1/2}$ with regard to the noise, and in reality it is still lower because the four C's above are not necessarily equal to each other due to the different relaxation rates of different coherence transfer pathways. This method has been successfully incorporated into many pulse sequences for studying the structure and dynamics of proteins.

Composite sequence. Composite pulse sequences have been widely used to suppress off-resonance effects, and to accomplish broadband decoupling and homonuclear isotropic mixing. They are designed to have the same effect as a single pulse but can tolerate the imperfections of r. f. pulses. In particular, cyclic composite sequences designed to make their propagators as close as possible to the unit operators have been employed most frequently. Of them, WALTZ-16 (Shaka et al., 1983), GARP-1 (Shaka et al., 1985), DIPSI-1, 2 or 3 (Shaka et al., 1988), MELEV-16 (Levitt et al., 1983) and SEDUCE-1 (McCoy & Muëller, 1992a, b) have been used for broadband decoupling.

While WALTZ-16, MELEV-17 (Bax & Davis, 1985) and DIPSI-2 or 3 have been employed for homonuclear isotropic mixing.

Water suppression. The huge solvent signal of an H_2O sample can be suppressed by presaturation (Anil-Kumar et al., 1980b; Wider et al., 1983), 1-1 hard pulse combined with phase alternated 1-1 hard pulse (Sklenar & Bax, 1987), spin-lock (Messerle et al., 1989), WATERGATE (Piotto et al., 1992) and water flip-back techniques (Grzesiek & Bax, 1993a). The last technique has been incorporated into many pulse sequences of multidimensional heteronuclear NMR experiments to minimize the dephasing and saturation of the solvent signal, and consequently to minimize the exchange of amide protons with the solvent.

Section 2. NMR Experiments for Protein Studies

NMR experiments for sequential resonance assignments. Homonuclear 2D ^{1}H-^{1}H through-bond correlation experiments (Wüthrich, 1986) such as COSY (Aue et al., 1976), DQF-COSY (Piantini et al., 1982) and TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985; Griesinger et al., 1988) can be used to identify amino acid spin systems. Homonuclear 2D ^{1}H-^{1}H NOESY, a through-space correlation experiment, (Jeener et al., 1979; Macura & Ernst, 1980; Anil-Kumar et al., 1980a) can be used to establish the sequential connectivities along the polypeptide chain. These homonuclear experiments can only be used for resonance assignments of proteins with ~100 residues

because the through-bond experiments require that the coupling constant $^3J_{HH}$ be not much smaller than the ^{1}H resonance line width. The first sequential resonance assignment of a protein called BPTI was obtained by this method (Wagner & Wüthrich, 1982). Furthermore, it has been extended for larger proteins with M_r less than 15 kDa by adding another dimension (^{15}N) to these 2D experiments to alleviate spectral overlap.

The assignments of proteins with M_r larger than 15 kDa, but less than 25 kDa, can be achieved by a series of triple resonance experiments (Ikura et al., 1990a; Kay et al., 1990a) using a ^{13}C and ^{15}N doubly labeled protein. These 3D heteronuclear correlation experiments are designed to take advantage of the relatively large heteronuclear coupling constants to alleviate the problems associated with large line widths. INEPT (Morris & Freeman, 1979), or refocused INEPT (Burum & Ernst, 1980), or DEPT (Doddrell et al., 1983) sequences have been incorporated into these experiments to accomplish the coherence transfers between the heteronuclei ^{1}H and ^{15}N and ^{13}C. For relatively rigid proteins with M_r about 20 kDa, these transfer steps are quite efficient since the line widths of ^{1}H, ^{15}N and ^{13}C resonances are normally significantly smaller than the relevant one-bond ($^1J_{HN}$, $^1J_{HC}$, $^1J_{NC\alpha}$, $^1J_{NCO}$, $^1J_{C\alpha CO}$, $^1J_{CC}$) (Figure 1.4) and two-bond ($^2J_{NC\alpha}$) coupling constants. The sequential resonance assignment is obtained in a conformation-independent or through-bond manner. These experiments have been further optimized by techniques such as constant-time (Grzesiek & Bax, 1992a; Zhu & Bax, 1990), gradient coherence selection (Kay et al., 1992a; Kay, 1995) and gradient artifact suppression (Bax & Pochapsky, 1992; Wider et al., 1993), composite pulse, and sensitivity enhancement (Kay et al., 1992a). The experiments most frequently performed for sequential assignments are HNCO, HNCA, HCACO, HN(CO)CA and their extensions such as

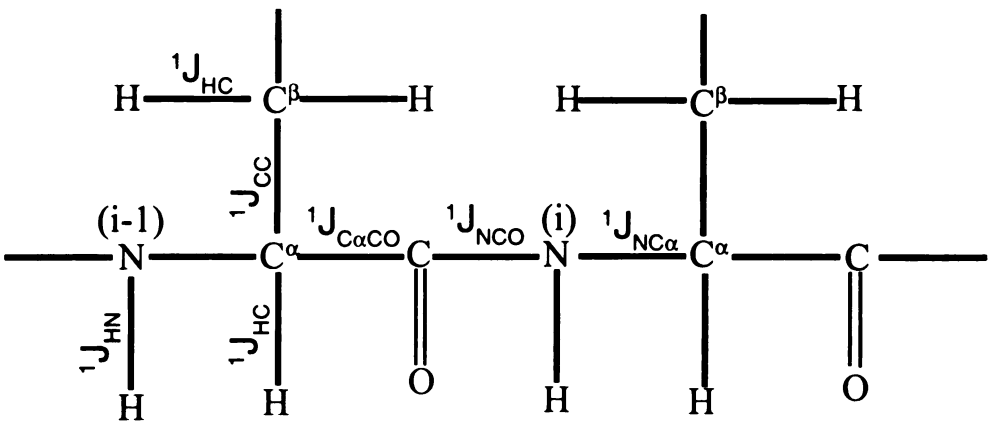


Figure 1.4. The polypeptide backbone and side chains showing the one–bond coupling constants through which the coherences are transferred in double and triple resonance NMR experiments.

HNCACB (Wittekind & Müeller, 1993), CBCA(CO)NH (Grzesiek & Bax, 1992b) and HBHA(CO)NH (Grzesiek & Bax, 1993b). The side-chain resonances are mainly assigned by HCCH-COSY (Kay et al., 1990b) and HCCH-TOCSY experiments (Bax et al., 1991; Kay et al., 1993), and in favorable cases by C(CO)NH and H(CCO)NH experiments (Logan et al., 1993; Grzesiek et al., 1993) (Figure 1.4). For proteins with M_r larger than 25 kDa, the recently developed deuterium labeling is normally needed to simplify the spectrum, to reduce the line width, and to enhance the efficiency of coherence transfer. For example, the HCCH-COSY and HCCH-TOCSY experiments which depend on $^1J_{CC}$ (~30-35 Hz) become less efficient for proteins with M_r larger than 25 kDa. With the help of a series of experiments specifically designed for ^{2}H, ^{13}C and ^{15}N triply labeled proteins, the solution structure of a protein with 259 residues has been reported recently (Garret et al., 1997).

NMR experiments for extracting structural restraints. The distance restraints used for structural calculation are derived from NOEs. The experiments most frequently performed include 2D homonuclear NOESY, 3D ^{15}N-edited NOESY-HMQC (Fesik & Zuiderweg, 1988; Zuiderweg & Fesik, 1989; Marion et al., 1989b) or NOESY-HSQC (Marion et al., 1989b), 3D ^{13}C-edited NOESY-HMQC (Ikura et al., 1990b; Zuiderweg et al., 1990) as well as 4D ^{15}N, ^{13}C-edited NOESY (Kay et al., 1990c; Muhandiram et al., 1993) and 4D ^{13}C, ^{13}C-edited NOESY (Zuiderweg et al., 1991). These experiments are built on two highly sensitive 2D heteronuclear correlation experiments HMQC (Müeller, 1979; Bax et al., 1983) and HSQC (Bodenhausen & Ruben, 1980) ($^1J_{NH}$ = ~92 Hz and $^1J_{CH}$ = ~120-160 Hz). The torsion angle restraints are obtained from coupling constants

according to the Karplus equation (Karplus, 1959). The experiments most frequently performed to extract such coupling constants are homonuclear 2D COSY, E. COSY (Griesinger et al., 1985), P. E. COSY (Mueller, 1987) and heteronuclear 2D ^{1}H-^{15}N HMQC-J (Kay & Bax, 1990) and 3D HNHA (Vuister & Bax, 1993). The last two experiments need ^{15}N-labeled samples and are used to measure $^{3}J_{HNH\alpha}$. Furthermore, ^{13}C secondary chemical shifts have also been included in the structural refinements (Kuszewski et al., 1995). Stereospecific assignment at a chiral center can be obtained from the three-bond scalar coupling, where appropriate, with intraresidue and sequential interresidue NOE data (Clore & Gronenborn, 1991a, b). The first protein NMR solution structure (Williamson et al., 1985) was determined by distance geometry (Havel & Wüthrich, 1984). However, most of the recent NMR structures are calculated by the hybrid distance geometry-simulated annealing (DGSA) protocol (Nilges et al., 1988).

NMR experiments for studying other structural and biochemical problems. In addition to structure determination and dynamical studies, solution NMR techniques can be used to study other structural and biochemical problems. Bound waters can be identified by employing the different translational diffusive behaviors of water and protein molecules (Wider et al., 1996) in terms of intermolecular translational Brownian motions (Torrey, 1953, 1954; Resing & Torrey, 1963; Stejskal & Tanner, 1965). Binding sites for ligands or substrates can be identified by measuring intermolecular NOEs (Gronenborn & Clore, 1995), or quickly but approximately by monitoring the changes in the chemical shifts of the residues upon ligand binding. A protocol for screening the ligands and subsequently optimize them has been developed for drug discovery (Shukert et al., 1996). The ionization states of the ionizable groups in proteins can also be

measured by NMR. For example, two aspartyl groups in the active site of HIV-1 protease have different ionization states. The carboxyl of Asp 25 is protonated while that of Asp125 is not over a pD range extending from 2.5 to 6.2 (Wang et al., 1996). This critical information has been used in the design of its inhibitor.

Section 3. Theories of NMR Relaxation in Liquid

Internal motions are critical for many functions of proteins in solution such as ligand binding and enzymatic catalysis (Karplus & McCammon, 1983; Karplus & Petsko, 1990), although their relevance in many systems has yet to be proven experimentally. Solution NMR spectroscopy, especially NMR relaxation measurement, has become one of the most powerful techniques to quantitatively characterize internal motions. For example, the rotations of aromatic side chains were discovered by NMR (Snyder et al., 1975; Wüthrich & Wagner, 1975). More recently, NMR relaxation study has shown that the long helix connecting the two globular domains of the protein calmodulin observed in x-ray diffraction is not stable in solution (Ikura et al., 1992). Internal motions that occur in partially folded proteins are important for understanding protein folding pathways. The residual structures of unfolded proteins have been identified by heteronuclear ^{1}H-^{15}N NOE experiment (Farrow et al., 1997). Relaxation study can be beneficial to NMR spectroscopy itself. The quantitative characterization of the relaxation behaviors helps to optimize the design of NMR experiments. It is well known that the efficiency of

coherence transfer, no matter whether by means of spin-spin coupling or cross-relaxation, depends on both the global and internal motions. The fast decay of the signals of large proteins on the NMR time scale is the key factor limiting the applicability of NMR techniques to such systems. The quantitative relaxation information may also be used to improve the precision of NMR-derived structures because the distance restraints used for structural calculation are derived from cross-relaxation rates, which in turn depend on internal motions. In the following I will review two of the most frequently used methods for extracting quantitative information about the internal motion from NMR relaxation measurement. Although NMR relaxation measurements have been reported for quite a few proteins, the interpretation of such data is itself far from an easy task.

Fourier series method. A general theory of relaxation, an irreversible dissipative process by which a disturbed closed system returns to its equilibrium state in the absence of external perturbation (see, e. g. Landau & Lifshitz, 1969), is a difficult quantum-statistical problem starting from the time reversal *Schrödinger* equation (see, e. g. Prigogine, 1962). It has not been completely solved. In contrast to the thermodynamic fluctuations at equilibrium, irreversibility means a direction in time. To make the problem tractable, it is usually assumed that relaxation is caused by the same underlying random process (Brownian motion) as that giving rise to the thermodynamic fluctuations at equilibrium. Irreversibility is introduced as an ad hoc assumption. The random process of Brownian motion is assumed to be a stationary *Markoff* process to be described to the accuracy of a joint probability or conditional probability (Wang & Uhlenbeck, 1945). There exist at least two methods to describe Brownian motion, the *Fourier* series method (spectrum method) and the diffusion equation method. The latter will be discussed in

later sections. The former analyzes the Brownian motion by *Fourier* series or integral (Rice, 1944). A random function $F_a(t)$ of time can be expanded as a Fourier integral, the spectrum of the random process.

$$F_a(t) = \int_{-\infty}^{+\infty} d\omega \, A_a(\omega) e^{i\omega t} \tag{26}$$

where the random variable $A_a(\omega)$ is the amplitude with frequency ω of the *Fourier* expansion. A random process with two random functions $F_a(t)$ and $F_b(t)$ can be characterized by the spectral density $J_{ab}(\omega)$ defined as Lim $_{T\to\infty} \langle A_a(\omega) A_b^*(\omega) \rangle / T$, a time average. The same random process can also be described in terms of a time correlation function $C_{ab}(t, t+\tau)$ defined as

$$C_{ab}(t, t+\tau) = \mathrm{Lim}_{T\to\infty} \left(\int_{-\infty}^{+\infty} F_a(t) \, F_b^*(t+\tau) \, dt \right) / T \tag{27}$$

$C_{ab}(t, t+\tau)$ depends only on the time interval τ. The equivalence of the two descriptions is established by the *Wiener-Khintchine* theorem, which states that the time correlation function of a random function is a Fourier transform of its spectral density (Wang & Uhlenbeck, 1945).

$$C_{ab}(t, t+\tau) = 2 \int_0^{+\infty} d\omega \, J_{ab}(\omega) \cos\omega\tau \tag{28}$$

Furthermore, the ergodic hypothesis is employed to justify the replacement of time average with ensemble average

$$C_{ab}(t, t+\tau) = \mathrm{Lim}_{T\to\infty}\, 1/T \int_{-\infty}^{+\infty} F_a(t)\, F_b^*(t+\tau)\, dt = \langle F_a(t)\, F_b^*(t+\tau)\rangle_{\mathrm{Av}} \qquad (29)$$

where angular bracket represents ensemble average.

From the exact *Liouville-von Neumann* equation (3), several authors have derived relaxation theories in terms of the correlation function or spectral density of the molecular environment (bath) appropriate for a spin system with interactions such as scalar, dipolar and quadrupolar couplings (Wangsness & Bloch, 1953; Fano, 1953, Bloch, 1956, 1957; Redfield, 1958, 1965; Abragam, 1961; Hubbard, 1961; Argyres & Kelley, 1964). These theories share features similar to the general Onsager's thermodynamic theory on irreversible processes (Zwanzig, 1961), and are constructed by means of various simplifying assumptions on the bath and the magnitude of the coupling between the bath and the spin system. Their final results are a set of equations relating the NMR-measurable relaxation parameters such as relaxation time and NOE enhancement to the microscopic properties characteristic of the bath. The key steps of the derivations are recapitulated as follows.

Quantum relaxation theory in operator form. An isolated system with total conserved Hamiltonian H is divided into two subsystems, a bath with time-independent Hamiltonian $\hbar F(f)$ and a spin system with time-dependent Hamiltonian $\hbar H_s(s, t)$ including

the interactions with external r. f. fields (Hubbard, 1961). These two subsystems interact with each other and the interaction Hamiltonian $\hbar G(f, s, t)$ is a random function of time.

$$H = \hbar H_s(s, t) + \hbar F(f) + \hbar G(f, s, t) \tag{30}$$

The separation is necessary since the two subsystems possess quite different properties and play distinct roles in NMR relaxation theory.

Several simplifying assumptions are introduced for the bath at the different stages of the derivation. The bath is assumed to possess quasi-continuous energy levels, and to be in thermal equilibrium at all times and independent of the spin system.

$$\rho(f, s, t) = \sigma(s, t)\, \rho^T(f) = \sigma(s, t) \exp(-\hbar F / kT) \,/\, \mathrm{tr}_f\{\exp(-\hbar F / kT)\} \tag{31}$$

where $\rho(f, s, t)$, $\rho^T(f)$ and $\sigma(s, t)$ are density operators for the whole system, the bath in thermal equilibrium and the spin system, respectively. T is the absolute temperature, k is the Boltzmann constant and tr_f means taking the trace over an ensemble of the bath subsystems. These assumptions characterize the basic feature of a relaxation process: irreversibility. With these assumptions, a symmetrized quantum-mechanical correlation function $C_{kl}(\tau)$ can be defined as

$$C_{kl}(\tau) = \langle F^l(t)\, F^{k\bullet}(t+\tau)\rangle_{\mathrm{Av}} = \tfrac{1}{2}\,[\mathrm{tr}_f\{\rho^T F^{k\bullet}(t+\tau)\, F^l(t)\} + \mathrm{tr}_f\{F^{k\bullet}(t+\tau)\, \rho^T F^l(t)\}] \tag{32}$$

where F's are the variables of the bath whose dissipative behavior is rather loosely characterized by a temperature-dependent correlation time τ_c with the property that $|C_{kl}(\tau)| \approx 0$ if $\tau >> \tau_c >> \hbar / kT$. It is further assumed that the bath correlation time τ_c is much shorter than the relaxation time of the spin system, that is, $|R|^{-1}$ or $|N|^{-1} >> \tau_c$, where $|R|^{-1}$ or $|N|^{-1}$ is a function of both the spin operators and the spectral densities of the bath, the latter is defined as $J_{kl}(\omega) = \int C_{kl}(\tau) \exp(i\omega\tau)\, d\tau$ {see equation (75) and (76) of Hubbard, 1961}. The magnitude of $|R|^{-1}$ or $|N|^{-1}$ represents a measure of the relaxation times of the spin system.

Two assumptions are also made for the interaction Hamiltonian $\hbar G(s, f, t)$. It is assumed to be so weak that the relaxation of a spin system caused by it can be treated by time-dependent perturbation up to the second order.

$$\partial\sigma / \partial t = - \mathrm{tr}_f \{ i [G(t), \rho(0)] + \int_0^t d\tau [G(t), [G(t-\tau), \rho(0)]] \} \qquad (33)$$

Furthermore, $G(t)$ is expanded as a sum of products of a bath operator $F^k(t)$ and a spin operator $V^k(t)$ with the latter itself being expanded as a Fourier series or integral

$$G(t) = \Sigma_k F^k(t)\, V^k(t) \text{ and } V^k(t) = \Sigma_k V_r^k \exp(i\omega_r^k t) \qquad (34)$$

where ω_r^k is the difference of the energy levels in frequency of the spin system. With these assumptions the change in $\sigma(s, t)$ for a time interval Δt satisfying the condition $|R|^{-1}$ or $|N|^{-1} >> \Delta t >> \tau_c$ could be expressed as

$$\sigma(t+\Delta t) = \sigma(t) + \int_t^{t+\Delta t} \{R(\sigma(t), t') + i[\sigma(t), N(t')]\}\, dt' \qquad (35)$$

Finally, with the assumption that $\sigma(t)$ in (34) can be replaced by $\sigma(t')$ (this is equal to the introduction of the irreversibility into the problem according to Fano), the integrodifferential equation (35) becomes a linear differential equation.

$$\partial\sigma(t)/\partial t = i[\sigma(t), N(t)] + R(\sigma(t), t) \qquad (36)$$

This approximation implies that the response of the spin system to an external perturbation is simultaneous without any memory effect (Zwanzig, 1961). Hubbard argued that this assumption was the same as the assumption $|R|^{-1}, |N|^{-1} >> \Delta t >> \tau_c$.

In order to take advantage of the symmetrical properties of the bilinear dipolar-dipolar, quadrupolar and chemical shift anisotropy (CSA) interactions $G(t)$ could be expanded as a scalar contraction of second-rank spherical irreducible tensor operators T^{μ}_{q} of the spin system and $F^{\mu}_{q}(t)$ of the molecular degrees of the bath (see, e. g. Brink & Satchler, 1993; Goldman, 1984)

$$G(t) = \Sigma_{\mu q} (-1)^q T^{\mu}_{q} F^{\mu}_{-q}(t) \qquad (37)$$

where μ specifies different interactions and the summation q runs from $-l$ to $+l$, here l represents the rank of the tensor, in this case, $l = 2$. The corresponding correlation function involving only $F^{\mu}_{q}(t)$ can be defined similarly. This expansion makes it much

easier to treat the cross-correlation between different relaxation mechanisms but with the same rank (Werbelow & Grant, 1977; Vold & Vold, 1978). The cross correlation between different ranks vanishes irrespective of the value of q.

Semiclassical theory. In semiclassical approximation the bath is further simplified to act as a randomly fluctuating magnetic field with infinite temperature. The classical bath correlation function is identified with the symmetrized quantum-mechanical correlation function (31) without regard to the commutation relation and the density operator of the bath $\rho^T(f)$.

$$C_{kl}(\tau) = \mathrm{tr}_f\{F^{k\bullet}(t+\tau)\, F^l(t)\} = \mathrm{tr}_f\{F^{k\bullet}(\tau)\, F^l(0)\} = \langle F^l(t)\, F^{k\bullet}(t+\tau)\rangle_{\mathrm{Av}} \qquad (38)$$

The semiclassical formalism widely used in the interpretation of protein NMR relaxation data is due to *Abragam* (Abragam, 1961). In an interaction representation with $\sigma^*(t) = \mathrm{tr}_f\{\rho^*(t)\} = \mathrm{tr}_f\{\exp(i H_s+F)\, \rho\, \exp(-iH_s-F)\}$ equation (32) and (34) can be simplified as

$$G^*(t) = \exp(iH_s+F)\, G(t)\, \exp(-iH_s-F) = \sum_k F^k(t)\, A^k(t) = \sum_{k,r} F^k(t)\, V_r^k \exp(i\omega_r^k t) \qquad (39)$$

$$\partial\sigma^*/\partial t = -1/2 \sum_{k,r} J_k(\omega_r^k)\, [V_r^{-k}, [V_r^k, \sigma^* - \sigma^T]] \qquad (40)$$

where spectral density $J_k(\omega_r^k) = \int \langle F^k(t)\, F^{k\bullet}(t+\tau)\rangle \exp(i\omega_r^k t)\, d\tau$ is defined as before, and $k^q(t)$ is a spin operator. σ^T represents the reduced density operator in a thermal equilibrium state. There is a difference between the $H_s(s, t)$ in *Hubbard*'s notation and the $H_s(s, t)$ in *Abragam*'s notation, namely, the former includes the average over an ensemble of baths in thermal equilibrium while the latter does not, so no σ^T appears in (36). The cross-

correlation between different dipolar-dipolar interactions, and between the dipolar and chemical shift anisotropy (CSA) is neglected in (40). They are supposed to be suppressed experimentally (Kay et al., 1992b).

Usually T_1 and T_2 relaxation and heteronuclear NOE enhancement $\sigma(H_z \rightarrow N_z)$ such as that between ^1H and ^{15}N are dominated by the dipolar-dipolar interaction between ^{15}N and the attached proton, as well as the CSA interaction with an external magnetic field. These three relaxation parameters can be related to spectral densities $J(\omega)$'s at four frequencies (Abragam, 1961)

$$1/T_1 = \gamma^2_H \gamma^2_N \hbar^2 /4r^6_{NH} \{J^{(0)}(\omega_H-\omega_N) + 3J^{(1)}(\omega_N) + 6J^{(2)}(\omega_H+\omega_N) \} + 1/3\Delta^2 \omega^2_N J^{cc}(\omega_N) \quad \text{.....(41)}$$

$$1/T_2 = \gamma^2_H \gamma^2_N \hbar^2 /8r^6_{NH} \{4J^{(0)}(0)+J^{(0)}(\omega_H-\omega_N) + 3J^{(1)}(\omega_N) + 6J^{(1)}(\omega_H) +6J^{(2)}(\omega_H+\omega_N)\}$$
$$+ 1/3\Delta^2\omega^2_N \{2/3\, J^{cc}(0) + 1/2\, J^{cc}(\omega_N)\} \quad (42)$$

$$\sigma(H_z \rightarrow N_z) = \gamma^2_H \gamma^2_N \hbar^2 /4r^6_{NH} \{6J^{(2)}(\omega_H+\omega_N) - J^{(0)}(\omega_H-\omega_N)\} \quad (43)$$

where $J^{(q)}(\omega)$'s are the spectral densities for the different components of the dipolar-dipolar interaction, and $J^{cc}(\omega)$ is the spectral density for CSA interaction, respectively. The CSA is assumed to have an axially symmetrical axis z with $\Delta = \sigma_{zz} - \frac{1}{2}(\sigma_{xx}+\sigma_{yy})$. r_{NH} is the distance between the two nuclei. It is often further assumed that all the spectral densities $J(\omega)$'s are the same

$$J^{(0)}(\omega) = J^{(1)}(\omega) = J^{(2)}(\omega) = J^{cc}(\omega) = J^{cd}(\omega) / P_2(\cos\theta) \quad (44)$$

where the $J^{cd}(\omega)$ is the spectral density for the cross-correlation between the dipolar-dipolar and the chemical shift anisotropy interactions if it is not suppressed, $P_2(\cos\theta)$ is the second *Legendre* polynomial, and θ is the angle between the internuclear vector and the symmetrical axis of CSA. This assumption is valid for isotropic rotational diffusions, and usually a good approximation for most practical systems in liquid. According to equations (41)-(43) it is obvious that only certain frequencies of the bath are directly responsible for NMR relaxation. Those are *0*, ω_H, ω_N, $\omega_H-\omega_N$ and $\omega_H+\omega_N$ causing zero, single and double quantum transitions of the spin system. Similar equations for the relaxation of antiphase and two spin order coherences have also been derived (see, e. g. Peng & Wagner, 1992), and the equations keeping the cross-correlation term between the dipolar-dipolar and CSA interactions have been published (Goldman, 1984).

Relaxation theory in matrix form. The formalisms of the relaxation theories above are expressed in terms of operators. The main result of the semiclassical formalism equation (39) could also be cast in a matrix form. The result is a set of linear equations for the evolution of the matrix elements $\sigma_{\alpha\alpha'}$'s in a representation of the eigenstates α, α', β, β' etc. of the unperturbed spin Hamiltonian $H_s(s, t)$ (Redfield, 1965).

$$d\sigma_{\alpha\alpha'} / dt = \mathrm{i}\,(\alpha'-\alpha)\sigma_{\alpha\alpha'} + \Sigma_{\beta\beta'}\, R_{\alpha\alpha'\beta\beta'}\,[\sigma_{\beta\beta'} - \delta_{\beta\beta'}\sigma^{T}_{\beta\beta'}] \qquad (45)$$

where $\delta_{\beta\beta'}$ is the Kronecker delta. $R_{\alpha\alpha'\beta\beta'}$ is a time-independent element of the relaxation matrix, and is related to the spectral density $J_{\alpha\alpha'\beta\beta'}$ of the bath.

$$R_{\alpha\alpha'\beta\beta'}=J_{\alpha\beta\alpha'\beta'}(\beta'-\alpha') + J_{\alpha\beta\alpha'\beta'}(\alpha-\beta) - \delta_{\alpha'\beta'}\Sigma_\gamma J_{\gamma\beta\gamma\alpha}(\gamma-\alpha) - \delta_{\alpha\beta}\Sigma_\gamma J_{\gamma\alpha'\gamma\beta'}(\alpha'-\gamma) \quad (46)$$

The spectral density $J_{\alpha\beta\alpha'\beta'}$ is the Fourier transform of the time correlation function $C_{\alpha\beta\alpha'\beta'}(\tau)$.

$$J_{\alpha\beta\alpha'\beta'}(\alpha-\beta) = \int d\tau\, e^{i(\alpha-\beta)\tau} C_{\alpha\beta\alpha'\beta'}(\tau) \text{ and } C_{\alpha\beta\alpha'\beta'}(\tau) = \langle G_{\alpha\beta}(t)\, G^{\dagger}_{\alpha'\beta'}(t-\tau)\rangle \quad (47)$$

The interaction Hamiltonian $G(t)$ responsible for the relaxation of the spin system is a random function of time, and $G_{\alpha\alpha'}(t)$ could be further expanded as before

$$G_{\alpha\alpha'}(t) = \Sigma_k F^{k}(t)\, V^{k}_{\alpha\alpha'} \quad (48)$$

where $V^{k}_{\alpha\alpha'}$'s are the spin operators and $F^{k}(t)$'s are the random functions representing local fluctuating fields. A spectral density involving only bath operators could be defined as

$$k_{kk'}(\alpha-\beta) = \int d\tau\, e^{i(\alpha-\beta)\tau} \langle F^{k}(t)\, F^{k'*}(t-\tau)\rangle \quad (49)$$

The semiclassical theory in matrix form presented above can be easily generalized into a quantum-mechanical theory. One result of such extensions is that the temperature appears in the definition of the quantum spectral density $j_{k\,k'}(\alpha-\beta) = k_{kk'}(\alpha-\beta)\exp[\hbar(\beta-\alpha)/2kT]$, here k is the Boltzman constant. Therefore, we have

$$J_{k'k}(\alpha-\beta) = j_{k'k}(\beta-\alpha)\exp[\hbar(\beta-\alpha)/\kappa T] \tag{50}$$

This equation represents mathematically the fact that the bath induced transition where the bath gains the energy $\hbar(\beta-\alpha)$ is more probable than the opposite one by a factor $\exp[\hbar(\beta-\alpha)/\kappa T]$. Therefore, in quantum-mechanical relaxation theory $R_{\alpha\alpha,\beta\beta} \neq R_{\beta\beta,\alpha\alpha}$, and the disturbed spin system will approach its equilibrium state as it should be. This is in contrast to the semiclassical theory where $R_{\alpha\alpha,\beta\beta} = R_{\beta\beta,\alpha\alpha}$, and the approach to an equilibrium state is introduced in (40) and (45) as an ad hoc assumption.

BPP theory. The semiclassical theory becomes identical with BPP theory for an isolated system consisting of non-interacting nuclear spins (Bloembergen et al., 1948) by an additional assumption that there exists no phase correlation between the eigenstates of the spin Hamiltonian. Consequently, all the off-diagonal elements of the reduced density operator $\sigma(t)$ can be neglected. The underlying mechanism is clear in this theory. Random fluctuations of the bath cause not only single quantum transitions between the nuclear spin energy levels with energy transfer between the bath and the spin system, but also zero and double quantum transitions with energy conservation within the spin system alone. A master equation about the probability of a transition per unit time $W_{\alpha\beta}$ $(= R_{\alpha\alpha\beta\beta})$ could be derived

$$dp_\alpha / dt = \Sigma_\beta W_{\alpha\beta}[(p_\beta - p_\beta^0) - (p_\alpha - p_\alpha^0)] \tag{51}$$

where p_α and p_β are, respectively, the populations of α and β energy states and p_α^0, p_β^0 their corresponding equilibrium values. The macroscopic magnetization such as M_z, which commutes with the spin Hamiltonian of a system of nuclear spins I could be calculated according to

$$M_z(t) = \Sigma'_{-I} m\, p_m(t) \tag{52}$$

where $p_m(t)$ is the population of an energy level with magnetic number m.

Cross-Relaxation. The BPP theory was extended to a two spin system coupled by dipolar-dipolar interaction with four eigenstates $|1\rangle = |\alpha\alpha\rangle$, $|2\rangle = |\alpha\beta\rangle$, $|3\rangle = |\beta\alpha\rangle$ and $|4\rangle = |\beta\beta\rangle$. The following equations for longitudinal components I_z, S_z, and similar ones for transverse components could be derived (Solomon, 1955).

$$dI_z / dt = -\rho\,(I_z - I_0) - \sigma\,(S_z - S_0) \text{ and } dS_z / dt = -\rho'\,(S_z - S_0) - \sigma\,(I_z - I_0) \tag{53}$$

$$\rho = U_0 + 2U_1 + U_2 \text{ and } \rho' = U_0 + 2U_1' + U_2 \tag{54}$$

$$\sigma = U_2 - U_0 \tag{55}$$

and

$$U_0 = W_{32} \cong W_{23},\ U_1 = W_{13} = W_{24},\ U_2 = W_{14} \text{ and } U_1' = W_{12} = W_{34} \tag{56}$$

where W_{ij}, $i, j = 1, 2, 3, 4$, is the transition probability per unit time between the eigenstates $|i\rangle$ and $|j\rangle$ of the longitudinal component of the combined system of spin I and S. ρ and ρ′ are auto relaxation rates of spin I and S, respectively. σ is the cross-relaxation

rate between the two spins *I* and *S*. One manifestation of the cross-relaxation is nuclear Overhauser effect (NOE) predicted by Overhauser (Overhauser, 1953) and extended to the case of nuclear dipolar-dipolar interaction by Bloch (Bloch, 1954). The distance restraint used in NMR structure determination is extracted from the measurement of the cross-relaxation rate σ between protons by a NOESY type experiment. For a system consisting only of two protons $^1H_1$ and $^1H_2$ their cross-relaxation rate σ depends on the internuclear distance r, and the spectral densities $J^{(q)}(\omega)$ at two frequencies.

$$\sigma\,(^1H_1 \leftrightarrow {}^1H_2) = \gamma^4_H \hbar^2/4r^6\,\{6\,J^{(2)}(2\omega_H) - J^{(0)}(0)\} \qquad (57)$$

In NMR structural determination it is assumed that $\sigma \propto 1/r^6$ for any pair of protons in a protein. This essentially is equivalent to the assumption of an isotropic global rotation without internal motion. However, in reality the situation is much more complicated. The spectral densities are related to both the global and internal motion. Moreover, global motion is anisotropic for non-spherical molecules. For any system with three protons sharing a common partner there exists cross-correlation (Hubbard, 1970). Other mechanisms such as random-field may also contribute to the relaxation (Werbelow & Grant, 1977). The compromise of extracting only distance range rather than one accurate number from the NOESY experiment may be the best choice. However, the structures derived in such a manner may be biased, especially for proteins with high internal flexibility. This offers an excellent example that a clear understanding of the underlying

physical principles is a prerequisite for the assessments of the results obtained through NMR techniques.

Bloch phenomenological equation. The simplest classical relaxation theory (1) was proposed by Felix Bloch in 1946 (Bloch, 1946b) based on phenomenological arguments. It assumes that relaxation is solely determined by two distinct relaxation times T_1 and T_2.

Spectral density mapping. The theories outlined above correspond to the treatment of the Brownian motion by the spectrum method. Their final results are similar: a set of equations such as (41)–(43) relating the measurable NMR relaxation parameters to the spectral densities of the bath at several frequencies. The equations (41)–(43) and two similar ones for two spin relaxation are the starting point of the so-called spectral density mapping method. The original version (Peng & Wagner, 1992) tried to determine the spectral density at the five frequencies 0, ω_X, ω_{H+X}, ω_H, ω_{H-X} (here X represents heteronucleus) through the measurements of both the one- and two-spin relaxation. Modified versions dealing with the spectral density at only three frequencies have been developed to compensate for the experimental errors at high frequencies by assuming that $J_{XH}(\omega_{H-X}) \approx J_{XH}(\omega_H) \approx J_{XH}(\omega_{H+X})$ (Farrow et al., 1995; Peng & Wagner, 1995; Lefevre et al., 1996), here $J_{XH}(\omega)$ is the spectral density of the X-H vector. It is true that spectral densities at certain frequencies are the only parameters that can be directly determined by NMR relaxation experiments at least nowadays. However, the information context about internal motion obtained by the mapping method is very limited. Furthermore, without resort to a certain model the result can not be correlated or compared with the microscopic properties of the bath obtained by means of theoretical or other experimental

methods. In fact, employing a model to describe relaxation is related to the diffusion equation method (Wang & Uhlenbeck, 1945) for analyzing the Brownian motion. The two methods are not completely equivalent under all circumstances but they give identical results for the processes much longer than the correlation times characteristic of the Brownian motion of the bath (Chandrasekhar, 1943). In contrast to the mapping method the macroscopic diffusion constant obtained from the diffusion equation method can be related to other microscopic quantities of the bath.

Diffusion equation method. This method assumes that NMR relaxation in liquid is caused by the discrete random walk processes through short distances or over small angular orientations. Passing over to the continuous case the probability distribution satisfies a particular partial differential equation of diffusional type with certain boundary conditions (Chandrasekhar, 1943). For a random *Markoff* process there exists a general solution $P(f, t)$, and a *Green* function $P_2(f_1, t; f_2, t+\tau)$, here $P(f, t)\, df$ is *a priori* probability (at equilibrium) of finding f in the range of $(f, f+df)$ at a time t, and $P_2(f_1, t; f_2, t+\tau)\, df_2$ is the conditional probability of given f_1 at time t one finds f_2 in the range of $(f_2, f_2 + df_2)$ a time τ later. These two probabilities are simply related

$$P(f_2, t+\tau) = \int P(f_1, t) P_2(f_1, t; f_2, t+\tau)\, df_1 \qquad (58)$$

Furthermore, $\lim_{\tau\to\infty} P_2(f_1, t; f_2, t+\tau)$ is equal to $P(f_2, t)$ for the thermodynamic fluctuations at equilibrium but for an irreversibility process it becomes $P(f_2, \infty)$ of the newly established equilibrium state. For NMR relaxation we have $P(f_2, \infty) = P(f_2, t)$ according to

the assumptions about the bath (31). The equivalence of the spectrum and the diffusion equation methods is established by the following basic relationship.

$$C_{ab}(t, t+\tau) = \langle F_a(t)\, F_b^*(t+\tau)\rangle_{Av} = \iint P(f_1, t)P_2(f_1, t; f_2, t+\tau)\, F_a(f_1)F_b^*(f_2)df_1\, df_2 \quad (59)$$

Motional models. As is well known both the overall and internal rotations contribute to NMR relaxation in liquid. The overall rotation can be either isotropic or anisotropic depending on the shape of the molecule under investigation. The internal rotation can be either free or restricted, about one bond or multiple bonds, and the bond can be either fixed or wobbling. A model is distinguished by the particular diffusion equation, and by its boundary condition suggested by physical instinct and the nature of the system. A general model appropriate for an anisotropic overall rotation, and a restricted internal rotation about multiple wobbling bonds will be quite complicated. Some simplifications are usually necessary for extracting the dynamical information from NMR relaxation data.

Sphere without internal motions. The simplest model considers a protein in solution as an isotropically rotating sphere without internal motions (Abragam, 1961). Its correlation function $C(\tau)$ is

$$C(\tau) = \exp(-|\tau| / \tau_c) \quad (60)$$

where the parameter τ_c is the correlation time characteristic of the bath. The irreversibility is encoded in the absolute value of $|\tau|$. The conditional probability $P(\Omega_0, \Omega, \tau)$, where Ω_0

and Ω are the *Euler* angles specifying the orientation of the molecule in a laboratory frame, is the solution of a diffusion equation for the free Brownian rotational motion with the initial condition $\psi(\Omega, 0) = \delta(\Omega - \Omega_0)$

$$\partial\psi/\partial\tau = D_s \Delta_s\psi \tag{61}$$

where D_s is the rotational diffusional coefficient and Δ_s the spherical *Laplacian* operator. This diffusion equation is the special case of the more general ***Smoluchowski*** equation with the potential V

$$\partial\psi/\partial t = \mathrm{div}[D\ \mathrm{grad}\psi + \psi f^{-1}\ \mathrm{grad}\ V] \tag{62}$$

where D and f are the diffusion constant and the frictional coefficient of the rotational motion, respectively (Chandrasekhar, 1943). The solution $P(\Omega_0, \Omega, \tau)$ can be found by first expanding ψ in spherical harmonics Y_m^l, and then substituting the result into (61).

$$P(\Omega_0, \Omega, \tau) = \Sigma_{l,m}\, Y_m^{l*}(\Omega_0)\, Y_m^l(\Omega) \exp[-\tau l(l+1)D_s] \tag{63}$$

With *a priori* probability $P(\Omega_0) = 1/(8\pi^2)$ the corresponding correlation function can be obtained from $P(\Omega_0, \Omega, \tau)$ by (59).

Arbitrary shape without internal motion. The distribution function $P(\Omega_0, \Omega, t)$ for the overall motion of an arbitrary shaped molecule is a solution of the following

diffusion equation for the free Brownian rotational motion with the initial condition $\psi(\Omega, 0) = \delta(\Omega - \Omega_0)$ (Favro, 1960),

$$\partial\psi/\partial t = \Sigma_k D_{kk} L_k^2 \psi \tag{64}$$

where $k = x, y, z,$ and L_x^2, L_y^2, L_z^2 are the components of angular momentum $\boldsymbol{L}$. The molecule-fixed frame with the axes x, y and z diagonalizes the diffusion tensor D, and D_{xx}, D_{yy} and D_{zz} are the three principal diffusional coefficients. The solution could be obtained similarly as before except that ψ is now expanded in terms of functions $\boldsymbol{D}^l_{mn}(\Omega)$, which are the elements of *Wigner* rotation matrix $\boldsymbol{D}^l(\Omega)$ (Hubbard, 1970). They are the eigenfunctions of the total angular momentum of a rigid body (Brink & Satchler, 1993). These eigenfunctions form a complete set in the space of *Euler* angles.

$$P(\Omega_0, \Omega, t) = \Sigma^{\infty}_{l=0} (2l+1)/(8\pi^2)\, \mathrm{tr}\{[\exp(-\boldsymbol{Q}^l t)]\, \boldsymbol{D}^{l\dagger}(\Omega_0)\, \boldsymbol{D}^l(\Omega)\} \tag{65}$$

The $(2l+1) \times (2l+1)$ real, symmetric matrix $\boldsymbol{Q}^l$ has matrix elements

$$\boldsymbol{Q}^l_{mn} = \langle lm|\, D_+ L^2 + (D_{zz} - D_+) L_z^2 + D_- (L_x^2 - L_y^2)\, |ln\rangle \tag{66}$$

where $|lm\rangle$ is the eigenket of L^2 and L_z^2 with eigenvalues $l\,(l+1)$ and m, respectively, and $D_{\pm} \equiv 1/2\,(D_{xx} \pm D_{yy})$. If one introduces the quantities

$$b_{lm} \equiv l(l+1)\, D_{+} + m^2(D_{zz} - D_{+}) = b_{l,-m} \tag{67}$$

then in the special case with axially symmetry ($D_{-} = 0$) the matrix $\boldsymbol{Q}^l$ becomes diagonal,

$$[\exp(-\boldsymbol{Q}^l(t)]_{mn} = \delta_{mn} \exp(-b_{lm}t) \tag{68}$$

and

$$P(\Omega_0, \Omega, t) = \Sigma_{l,m,n}\, (2l+1)\, /(8\pi^2)\, [\exp(-b_{lm}t)]\, \boldsymbol{D}^l_{nm}(\Omega_0)\, \boldsymbol{D}^l_{nm}(\Omega) \tag{69}$$

where summation l runs from 0 to $+\infty$ and m, n from $-l$ to $+l$. For relaxation dominated by the dipolar-dipolar and CSA interactions l is equal to 2. Isotropic rotation of a sphere corresponds to $D_{xx} = D_{yy} = D_{zz}$ and (69) becomes (63). With *a priori* probability $P(\Omega_0) = 1/(8\pi^2)$ the corresponding correlation functions can be obtained from (59).

Model with internal motion. With internal motion the solutions to the diffusion equation become much more complicated. As the first step of simplification it is usually assumed that the internal motion is independent of the overall rotation. This is valid for the isotropic overall rotations of spherical molecules, and normally a good approximation for arbitrary shaped molecules. The correlation function in the laboratory frame can be related to that in a molecule-fixed frame diagonalizing the diffusion tensor by *Wigner* rotation matrices D^2_{mn}'s, since the elements of the correlation functions of most practical NMR systems are second-rank spherical tensors (Brink & Satchler, 1993).

$$C_{mm'}(t) = \Sigma_{n,n'} \langle D^2_{mn}{}^{*} [\Omega(0)]\, D^2_{m'n'} [\Omega(t)]\rangle \langle D^2_{n0}{}^{*} [\theta(0), \phi(0)]\, D^2_{n'0} [\theta(t), \phi(t)]\rangle \tag{70}$$

The summation n, n' runs from -2 to $+2$. Where Ω is referred to the laboratory frame and θ, ϕ to the molecule-fixed frame, respectively. The overall rotation is characterized by the first average while the internal motion by the second average over a bath ensemble.

Model-free approach. One of the simplest ways to include internal motion is to assume that the correlation function $C_I(t)$ for the internal motion can be divided into two parts (Lipari & Szabo, 1982a, b): the average value S^2 of the components of $C_I(t)$ representing the restrictions of the internal motion, and a weighted normal correlation function $Q(t)$ with the properties that $\lim_{t\to\infty} Q(t) = 0$ and $Q(0) = 1$ characterizing the randomness of the motion. In other words, the average value of $Q(t)$ is zero.

$$C_I(t) = S^2 + (1 - S^2)Q(t) \tag{71}$$

S^2 is called order parameter and equal to $\lim_{t\to\infty} C_I(t)$. It does not depend on time and is not a random function. With the assumption that the overall rotation is isotropic, and is independent of the internal motion the total correlation function $C(t)$ becomes

$$C(t) = C_0(t)\, C_I(t) = [1/5 \exp(-t/\tau_M)]\, C_I(t) \tag{72}$$

where $C_0(t)$ is the correlation function for the overall motion. The corresponding spectral density $J(\omega)$ is

$$J(\omega) = 2S^2\tau_M / 5\, [1+(\omega\tau_M)^2] + 2/5\, (1 - S^2) \int_0^{+\infty} dt\, \cos\omega t\, Q(t) \exp(-t/\tau_M) \tag{73}$$

If we assume that $Q(t) = \exp(-t/\tau_e)$ we have

$$J(\omega) = 2/5\{S^2\tau_M / [1+(\omega\tau_M)^2] + (1-S^2)\,\tau / [1+(\omega\tau)^2]\} \tag{74}$$

where $\tau^{-1} = \tau_M^{-1} + \tau_e^{-1}$, and τ_e is defined according to $\tau_e\,(1-S^2) = \int_0^{+\infty} (C_l(t) - S^2)\,dt$. If the internal motion is much faster than the overall motion, and lies in the extreme narrowing range, that is, the correlation time for internal motion τ_e satisfies the condition $\tau_e << \tau_M$ and $\tau_e << \omega$, then (74) becomes

$$J(\omega) = 2/5\{S^2\tau_M / [1+(\omega\tau_M)^2]+(1-S^2)\tau_e\} \tag{75}$$

For internal motion much slower than overall motion we instead have $J(\omega) = 2\tau_M/5\,[1+(\omega\tau_M)^2]$.

Order parameter. The physical meaning of the order parameter S^2 is quite clear according to the definition of the correlation function (59). With the modified spherical harmonics C_{2m} (Brink & Satchler, 1993) substituting for the general random function in (59) we have

$$\begin{aligned} S^2 &= \lim_{\tau\to\infty} C(\Omega_1, t; \Omega_2, t+\tau) = \lim_{\tau\to\infty} \Sigma_m \langle C_{2m}(\Omega_1, t)\, C_{2m}^*(\Omega_2, t+\tau)\rangle_{Av} \\ &= \Sigma_m \langle C_{2m}(\Omega_1, t)\rangle_{Av} \langle C_{2m}^*(\Omega_2, \infty)\rangle_{Av} = \Sigma_m |\langle C_{2m}(\Omega, t)\rangle_{Av}|^2 \end{aligned} \tag{76}$$

where the summation runs from -2 to 2, and $C_{2m}^{*}(\Omega_2, \infty)$ is the spherical harmonics at the final equilibrium state reached by relaxation. We have $<C_{2m}(\Omega_1, t)>_{Av} = <C_{2m}(\Omega_2, \infty)>_{Av} = <C_{2m}(\Omega_2, t)>_{Av}$ based on the assumptions about the bath introduced before (31). In terms of the probability distribution function the order parameter S^2 is

$$
\begin{aligned}
S^2 &= \lim_{\tau\to\infty} \Sigma_m \iint P(\Omega_1, t) P_2(\Omega_1, t; \Omega_2, t+\tau)\, C_{2m}(\Omega_1, t)\, C_{2m}^{*}(\Omega_2, t+\tau)\, d\Omega_1\, d\Omega_2 \\
&= \Sigma_m \iint P(\Omega_1, t) P_2(\Omega_1, t; \Omega_2, \infty)\, C_{2m}(\Omega_1, t)\, C_{2m}^{*}(\Omega_2, \infty)\, d\Omega_1\, d\Omega_2 \\
&= \Sigma_m \{\int P(\Omega_1, t)\, C_{2m}(\Omega_1, t)\, d\Omega_1\} \{\int C_{2m}^{*}(\Omega_2, t)\, P_2(\Omega_2, t)\, d\Omega_2\} \\
&= \Sigma_m |\{\int P(\Omega, t)\, C_{2m}(\Omega, t)\, d\Omega\}|^2 = \Sigma_m |<C_{2m}(\Omega, t)>_{Av}|^2
\end{aligned}
\tag{77}
$$

where the relations $P_2(\Omega_1, t; \Omega_2, \infty) = P_2(\Omega_2, \infty) = P_2(\Omega_2, t) = P_1(\Omega_1, t) = P(\Omega, t)$ of the equilibrium distribution have been used. Therefore, S^2 is nothing more than the average value over a bath ensemble of the spherical harmonics of the *Euler* angles specifying the orientation of a vector relative to a molecule-fixed frame. It is also the first nontrivial term in the expansion of $P(\Omega, t)$ in a series of *Legendre* polynomials (Lipari & Szabo, 1980). However, how well it represents $P(\Omega, t)$ depends on how fast this series converges. Therefore, a model is needed to establish the relationship between S^2 and the distribution of the orientation. The information context of S^2 corresponds to the description of a random process by only the first set of the probability distribution $P(\Omega, t)$ (Wang & Uhlenbeck, 1945). The conformation with high energy or rare probability is not being sampled properly by S^2. The claim that S^2 is insensitive to motions slower than nanosecond is misleading. The insensitivity is not due to the time scale but due to the energies of these conformations. Though $P(\Omega, t)$ can not be easily obtained for a complex

system such as a protein even a random *Markoff* process must be specified by a joint probability or conditional probability $P_2(\Omega_1, t; \Omega_2, t+\tau)$ together with the equilibrium distribution $P(\Omega, t)$.

Extension of model-free approach. For some proteins it was found that two exponential terms instead of one were needed to fit their relaxation data by the *Lipari & Szabo* approach (Clore et al., 1990). Moreover, an exchanging term R_{ex} has been added to the transverse relaxation rate constant $1/T_2$ to account for the contribution to it from the possible slow motions (Kay et al., 1989; Clore et al., 1990) other than the dipolar-dipolar and CSA interaction. The term R_{ex} is usually explained as the contribution from conformational exchanges. However, the physical meanings of these added terms are not well defined. They are related to the diffusion tensor in some complicated ways. Furthermore, for proteins with arbitrary shapes the total correlation function can not be separated rigorously into an overall and an internal part. Therefore, one should always bear in mind the range of validity of the original formalism or its extension when using it to interpret NMR relaxation data. Recently the *Lipari & Szabo* approach has been extended to the cross-correlation by several authors (Kay & Torchia, 1991; Zhu et al., 1995; Daragan & Mayo, 1995; Daragan & Mayo; 1997). The cross-correlation is responsible for the difference of relaxation behaviors in the two outer lines from the two inner ones, and manifests itself through the nonexponential behavior of the relaxation. It has been treated in detail in two reviews (Werbelow & Grant, 1977; Vold & Vold, 1978).

Restricted rotational diffusion model. Several more sophisticated models are in vogue. The first class may be referred to as restricted rotational diffusion model (Wittebort & Szabo, 1978, London & Avitabile, 1978). The conditional probability $P(\gamma', t$

; γ, *0)* of the internal rotation is the solution of one dimensional diffusion equation $\partial\psi/\partial t = D\,\partial^2\psi/\partial^2\gamma$ with the initial condition $\psi(\Omega, 0) = \delta(\Omega - \Omega_0)$ and the boundary condition $\partial\psi(\gamma)/\partial\gamma \mid_{\pm\gamma_0} = 0$, where γ is azimuthal angle. The boundary condition is the mathematical statement that the internal rotational diffusion is restricted between $\pm\gamma_0$.

$$P(\gamma', t; \gamma, 0)=1/(2\gamma_0)\{1 + \Sigma_{n=1} \cos[n\pi(\gamma-\gamma_0)/2\gamma_0] \cos[n\pi(\gamma'-\gamma_0)/2\gamma_0] \exp(-t/\tau_n)\} \quad (78)$$

where $\tau_n = 4\gamma_0^2/(Dn^2\pi^2)$. With *a priori* probability $P(\Omega_0) = 1/(2\gamma_0)$ the corresponding correlation function can be obtained from (59). With this model we can include excluded volume effects accounting for the restricted space an internuclear vector could have to diffuse. The *Woessner* model (Woessner, 1962) of free internal rotation is included as a special case of this general model with $\pm\gamma_0 = \pm 180°$. The model above is good for the internal rotation about one bond. It can be easily generalized into a model for internal rotations about N bonds (Wallach, 1967). The transformation from a frame F attached to a nucleus whose relaxation is of interest to the laboratory frame L can be achieved by a series of Wigner's rotation matrix D^2's with *Euler* angles $\Omega_{D1}, \Omega_{12}, \cdots, \Omega_{N-1\,N}$ specifying the orientations of the successive coordinate systems.

$$D^2_{q0}(\Omega_{LF}) = \Sigma_{a,b_1\cdots b_n} D^2_{qa}(\Omega_{LD}) D^2_{ab_1}(\Omega_{D1}) D^2_{b_1b_2}(\Omega_{12}) \cdots D^2_{b_{n-1}b_n}(\Omega_{N,N-1})\, D^2_{b_n 0}(\Omega_{NF}) \quad (79)$$

The internal rotations are usually assumed to be independent of each other in order to simplify the problem.

Discrete-jump model. The discrete-jump model with M configurations (Wittebort & Szabo, 1978) tries to give a more realistic description of concerted internal rotations with mutual dependence. It is also an extension of the *Woessner* jump model with only three configurations (Woessner, 1962). The underlying mathematics is the same: the conditional probability $P(k, t; l, 0)$ of the side chain of a macromolecule has configuration k at time t if it has configuration l at time 0 is the solution of the mater equation (81) with the initial conditions (80):

$$P(k, t; l, 0) = \delta_{kl}\,,\ t = 0 \text{ and } lim_{t\to\infty} P(k, t; l, 0) = P_{eq}(k) \text{ for all } l. \tag{80}$$

$$\partial\psi/\partial t = \Sigma^{M}_{j=1} R_{ij}\psi \tag{81}$$

where R_{ij} is the rate constant for the transition from configuration j to i with M total available configurations. $P_{eq}(k)$ is *a priori* probability at equilibrium satisfying the equation $\Sigma^{M}_{j=1} R_{ij}\ P_{eq}(j) = 0$. The solution can be obtained by matrix method, and is expressed in terms of the eigenvectors $X_k' = (X_k^0, \cdots X_k^M)$ and eigenvalues λ_n's of the matrix $\boldsymbol{Q}$ with elements $Q_{ij} = (R_{ij}\ R_{ji})^{1/2}$.

$$P(k, t; l, 0) = X_k^0\ (X_l^0)^{-1}\ \Sigma_n\ X_k^n\ X_l^n \exp(-\lambda_n t) \tag{82}$$

where $X_k^0 = [P_{eq}(k)]^{1/2}$.

Wobbling in a cone model. This model assumes that an internuclear vector is wobbling uniformly in a cone (Kinoshita et al., 1977; Lipari & Szabo, 1980). The

conditional probability $P(\Omega_0; \Omega, t)$ is the solution of a diffusion equation for the free Brownian rotational motion with the initial condition $\psi(\Omega, 0) = \delta(\Omega - \Omega_0)$, and the boundary condition $\partial\psi(\theta)/\partial\theta\,|_{\theta_0} = 0$, here θ is the polar angle.

$$\partial\psi/\partial t = D_w \Delta_s \psi \tag{83}$$

where D_w is the diffusion coefficient of the wobbling motion. The boundary condition is the mathematical statement that the internal rotational diffusion is restricted within a polar angle θ_0 but does not depend on the azimuthal angle. The solution could be obtained as an expansion of associated *Legendre* functions. However, the resulting expressions for $P(\Omega_0; \Omega, t)$, and consequently for the correlation function $C_l(t)$ are not closed analytic functions.

$$C_l(t) = \sum_i A_i \exp(-D_w t/\sigma_i) \tag{84}$$

where A_i and σ_i are functions of θ_0. Instead an approximate expression for $C_l(t)$ is derived which is exact at time $t = 0$ and $t = \infty$ (order parameter S^2), and has the property that the area under the approximate correlation function is the same as that under the exact correlation function.

$$C_l(t) = S^2 + (1 - S^2) \exp(-D_w t/\langle\sigma\rangle) \tag{85}$$

where $S^2 = A_\infty$ and $\langle\sigma\rangle = \sum_{i\neq\infty} A_i\sigma_i / (1- S^2)$. The *a priori* probability $P(\theta_0) = [2\pi(1-\cos\theta_0)]^{-1}$. The wobbling motion has been combined with a free rotation to give a more realistic description of the internal motion (Richarz et al., 1980).

Intermediate and slow motions. The NMR relaxation measurement discussed above can provide dynamical information about the fast motion on the time scale ranging from 10^{-12} to 10^{-8} second. The motions in the intermediate time scale from microsecond to millisecond are much more difficult to be studied by NMR. The measurement of the exchange term R_{ex}, sensitive to motions on the microsecond to millisecond time scale, will be discussed later. The slow motion in the time scale ranging approximately from milliseconds to seconds gives cross-peaks between separate resonances which could be observed in homonuclear NOESY, ROESY and in heteronuclear longitudinal magnetization or two-spin order exchange experiments. Most of NMR techniques for measuring intermediate or slow motions have been discussed in a recent review (Palmer et al., 1996).

Pulse sequences for measuring NMR relaxation parameters. The 2D pulse sequences for NMR relaxation experiments have been reviewed in a recent paper (Peng & Wagner, 1994), and the versions with pulsed field gradients are also available (Farrow et al., 1994). Frequently the inversion-recovery (Carr & Purcell, 1954; Vold et al., 1968) and the Carr-Purcell-Meiboom-Gill (CPMG) (Carr & Purcell, 1954; Meiboom & Gill, 1958) techniques are used to measure T_1 and T_2, respectively.

$$M_z(t) = M_0 [(1-2 \exp(-t/T_1)] \tag{86}$$

$$M_{x,y}(2nt) = M_x(0) \exp(-2nt/T_2) \exp(-2\gamma^2 DG^2 nt^3/3) \tag{87}$$

where G is the gradient, D diffusion coefficient and n the number of echo. The heteronuclear NOE enhancement is measured by steady-state NOE technique. The experimental techniques for measuring two-spin order or antiphase coherence have also been developed (Peng & Wagner, 1992).

These pulse sequences normally include pulses to suppress cross-correlation for minimizing its contribution to the relaxation (Boyd et al., 1990; Kay et al., 1992; Palmer et la., 1992). The interference from the chemical exchange or slow motions on the microsecond time scale can be measured by $T_1\rho$ measurement (Szypersky et al., 1993) or spectral density mapping method. The exchange term R_{ex} manifests itself by an anomalous decrease in T_2 relative to that predicted for dipolar, CSA or quadrupolar relaxation in a free-precession NMR spectrum (the line width), or in a CPMG (Allerhand & Thiele, 1968), or $T_{1\rho}$ experiment. In the $T_{1\rho}$ experiment R_{ex} can be identified by directly measuring τ_{ex} according to the equation (Davis et al., 1990, where a number of general cases are also discussed)

$$R_{ex} = (\delta\omega)^2 p_A p_B \tau_{ex} / (1+\tau_{ex}^2\omega_e^2) \qquad (88)$$

by varying the effective spinlock field ω_e (Akke & Palmer, 1996). Where $\delta\omega$ is the difference in chemical shift, and p_A, p_B are the time fractions spent in the conformation A and B, respectively.

The future of NMR relaxation study. A more realistic model needs to be derived from a general diffusion equation with external forces (Chandrasekhar, 1943) with the

considerations of more relevant factors such as solvent effect and cross-correlation effect. A unified relaxation theory is also very helpful for distinguishing and quantitating the contributions to the total relaxation of the various internal motions on different time scales. However, even with a quite realistic model the concerted movements of large fragments, which may be the most important internal motions for protein functions, could not be obtained because the dynamical information obtained from NMR experiments is local. Certain calculation procedures need to be developed to identify the collective motion. Most importantly, experimental techniques need to be improved or developed to measure accurately the relaxation parameters, or to measure new relaxation parameters. Finally, dynamical information should be combined with biochemical data to establish the roles of internal motions in the biochemical functions of proteins.

References

Abragam, A. (1961) *Principles of Nuclear Magnetism*, Clarendon Press, Oxford. Chapter VIII, XXI.

Akke, M., & Palmer A. G. (1996) *J. Am. Chem. Soc. 118*, 911-912.

Allerhand, A., & Thiele, E. (1968) *J. Chem. Phys. 45*, 902-916.

Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1980a) *Biochem. Biophys. Res. Commun. 95*, 1-6.

Anil-Kumar, Wagner, G., Ernst, R. R., & Wüthrich, K. (1980b) *Biochem. Biophys. Res. Commun. 96*, 1156-1163.

Argyres, P. N., & Kelley, P. L. (1964) *Phys. Rev. 134A*, 98-111.

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229-2246.

Banwell, C. N., & Primas, H. (1962) *Molec. Phys. 6*, 225-256.

Basus, V. J. (1989) *Methods Emzymol. 177*, 132-149.

Bax, A., Clore, G. M., & Gronenborn, A. M. (1991) *J. Magn. Reson. 88*, 425-431.

Bax,A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355-360.

Bax, A., DeJong, P. G., Mehlkopf, A. F. , & Smidt, J. (1980) *Chem. Phys. Lett. 69*, 567-570.

Bax, A., Driscoll, P. C., Clore, G. M., Gronenborn, A. M., Ikura, M., Kay, L. E. (1990) *J. Magn. Reson. 87*, 620-627.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Magn. Reson. 55*, 301-315.

Bax, A., & Pochapsky, S. S. (1992) *J. Magn. Reson. 99*, 638-643.

Bloch, F. (1946b) *Phys. Rev. 70*, 460-474.

Bloch, F. (1954) *Phys. Rev. 93*, 944-944.

Bloch, F. (1956) *Phys. Rev. 102,* 104-135.

Bloch, F. (1957) *Phys. Rev. 105,* 1206-1222.

Bloch, F. (1958) *Phys. Rev. 111*, 841-853.

Bloch, F., Hansen, W. W., & Packard, M. (1946a) *Phys. Rev. 69*, 127-127.

Bloembergen, N., Purcell, E. M., & Pound, R. V. (1948) *Phys. Rev. 73,* 679-712.

Bloom, A. L., & Schoolery, J., N. (1955) *Phys. Rev. 97,* 1261-1265.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185-188.

Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93-106.

Boyd, J., Hommel, U., Campbell, I. D. (1990) *Chem. Phys. Lett. 175*, 477-482.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53,* 521-528.

Brink, D. M., & Satchler, G. R. (1993) *Angular Momentum*, Oxford University Press, New York.

Burum, D. R., & Ernst R. R. (1980) *J. Magn. Reson. 39,* 163-168.

Carr, H.Y., & Purcell, E. M. (1954) *Phys. Rev. 94,* 630-638.

Cavanagh, J., Palmer, A. G., Wright, P. E., & Rance, M. (1991) *J. Magn. Reson. 91*, 429-436.

Chandrasekhar, S. (1943) *Rev. Mod. Phys. 15*, 1-89.

Clore, G. M., Bax, A., Driscoll, P. C., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 8172-8184.

Clore, G. M., Driscoll, P. C., Wingfield, P. T., Gronenborn, A. M. (1990) *Biochemistry*

29, 7387-7401.

Clore, G. M., & Gronenborn, A. M. (1991a) *Science 252*, 1390-1399.

Clore, G. M., & Gronenborn, A. M. (1991b*) Annu. Rev. Biophys. Chem. 21*, 29-63.

Clore, G. M., Szabo, A., Bax, a., Kay, L. E., Driscoll, P. C., & Gronenborn, A. M. (1990) *J. Am. Chem. Soc. 112*, 4989-4991.

Daragan, V. A., & Mayo, K. H. (1995) *J. Magn. Reson. B107*, 274-278.

Daragan, V. A., & Mayo, K. H. (1997) *Prog. NMR Spectrosc. 31*, 63-105.

Davis, D. G., Perlman, M. E., London, R. E. (1990) *J. Magn. Reson. B104*, 266-275.

Davis, A. L., Boelens, R., & Kaptein, R. (1992) *J. Biomol NMR 2*, 395-400.

Dayie, K. T., Wager, G., & Lefevre, J. F. (1996) *Annu. Rev. Phys. Chem.* 47, 243-282.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995) *J. Biomol.NMR 6*, 277-293.

Doddrell, D. M., Pegg, D. T., & Bendall, M. R. (1982) *J. Magn. Reson. 48*, 323-327.

Ernst, R. R., & Anderson, W. A. (1966) *Rev. Sci. Instrum. 37,* 93-102.

Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.

Fano, U. (1953) *Phys. Rev. 96,* 869-873.

Fano, U. (1957) *Rev. Mod. Phys. 29*, 74-93.

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Kay, J. D. F., & Kay, L. E. (1994) *Biochemistry 33*, 5984-6003.

Farrow, N. A., Zhang, O., Forman-Kay, J. D., & Kay, L. E. (1997) *Biochemistry 36*, 2390-2402.

Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. & Kay, L. E. (1995) *J. Biomol. NMR 6*, 153-162.

Farvo, L. D. (1960) *Phys. Rev. 119,* 53-62.

Fesik, S. W., & Zuiderweg, F. R. P. (1988) *J. Magn. Reson. 78*, 588-593.

Garret, D. S., Seok, Y., Liao d., Peterkofsky, A., Gronenborn, A. M., & Clore, G. M. (1997) *Biochemistry 36*, 2517-2530.

Goldman, M. (1984) *J. Magn. Reson. 60*, 437-452.

Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. P. (1988) *J. Am. Chem. Soc. 110*, 7870-7872.

Griesinger, C., Sørensen, O. W., & Ernst, R. P. (1985) *J. Am. Chem. Soc. 107*, 6394-6396.

Gronenborn, A. M., & Clore, G. M. (1995) *CRC Crit. Rev. Biochem. Mol. Biol. 30*, 351-385.

Grzesiek, S., Anglister, J., & Bax, A. (1993) *J. Magn. Reson. B101*, 114-119.

Grzesiek, S., & Bax, A. (1992a) *J. Magn. Reson. 96*, 432-440.

Grzesiek, S., & Bax, A. (1992b) *J. Am. Chem. Soc. 114*, 6291-6293.

Grzesiek, S., & Bax, A. (1993a) *J. Am. Chem. Soc. 115*, 12593-12594.

Grzesiek, S., & Bax, A. (1993b) *J. Biomol. NMR 3*, 185-204.

Hahn, E. L. (1950) *Phys. Rev. 80,* 580-594.

Havel, T. F., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673-698.

Hubbard, P. S. (1961) *Rev. Mod. Phys. 33*, 249-164.

Hubbard, P. S. (1970) *J. Chem. Phys. 52*, 563-568.

Huntress, W. T. (1970) *Adv. Magn. Reson. 4*, 1-37.

Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., & Bax, A. (1992) *Science 256,* 632-638.

Ikura, M., Kay, L. E., & Bax, A. (1990a) *Biochemistry 29,* 4659-4667.

Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990b) *J. Magn. Reson. 86*, 204-209.

Jeener, J. (1971) *Ampere International Summer School*, Basko Polje, Yugoslovia

Jeener, J., Meier, B. H., & Ernst, R. R. (1978) *J. Chem. Phys.* 71, 4546-4553.

John, B. K., Plant, D., Heald, S. H., & Hurd, R. D. (1991) *J. Magn. Reson. 94*, 664-669.

Karplus, M. (1959) *J. Phys. Chem. 30*, 11-15.

Karplus, M., & McCammon, J. A. (1983) *Ann. Rev. Biochem. 53*, 253-300.

Karplus, M., & Petsko, G. A. (1990) *Nature 347*, 631-639.

Kay, L. E. (1995) *Prog. Biophys. Molec. Biol. 63*, 277-299.

Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110-126.

Kay, L. E., Clore, G. M., Bax, A., & Gronenborn, A. M. (1990c) *Science 249*, 411-414.

Kay, L. E., Ikura, M., & Bax, A. (1990b) *J. Am. Chem. Soc. 112*, 888-889.

Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (1990a) *J. Magn. Reson. 89*, 495-514.

Kay, L. E., Keiffer, P., & Saarinen, T. (1992a) *J. Am. Chem. Soc. 114*, 10663-10665.

Kay, L. E., Nicholson, L. D., Delaglio, F., Bax, A., & Torchia, D. (1992b) *J. Magn. Reson. 97*, 359-375.

Kay, L. E., & Torchia, D. (1991) *J. Magn. Reson. 95*, 536-547.

Kay, L. E., Torchia, D., & Bax, A. (1989) *Biochemistry 28*, 8972-8979.

Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., & Forman-Kay, J. D. (1993) *J. Magn. Reson. B101*, 333-337.

Kinoshita, K. Jr., Kawato, S., & Ikegami, A. (1977) *Biophys. J. 20*, 289-305.

Kuszewski, L., Qin, J., Clore, G. M., & Gronenborn, A. M. (1995) *J. Magn. Reson. B106*, 92-96.

Landau, L. D., & Lifshitz, E. M. (1969) *Statistical Physics, pp. 6*, Addison-Wesley, Reading, Massachusetts.

Lefevre, J. F., Dayie, K. T., Peng, J. W., & Wagner, G. (1996) *Biochemistry 35*, 2674-2686.

Levitt, M. H., Freeman, R., & Frenkiel, T. (1983) *Adv. Magn. Reson. 11*, 47-110.

Lipari, G., & Szabo, A. (1980) *Biophys. J. 30*, 489-506.

Lipari, G., & Szabo, A. (1982a) *J. Am. Chem. Soc. 104*, 4546-4559.

Lipari, G., & Szabo, A. (1982b) *J. Am. Chem. Soc. 104*, 4559-4570.

Logan, T. M., Olejniczak, E. T., Xu, R. X., & Fesik, S. W. (1993) *J. Biomol. NMR 3*, 225-231.

London, R. E., & Avitabile, J. (1978) *J. Am. Chem. Soc. 99*, 7765-7776.

Lowe, I. J., & Norberg, R. E. (1957) *Phys. Rev. 107*, 46-61.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95-117.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfeld, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989d) *Biochemistry 28*, 6150-6156.

Marion, D., Ikura, M., & Bax, A. (1989c) *J. Magn. Reson. 84*, 425-430.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989a) *J. Magn. Reson. 85*, 393-399.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., Bax, A. (1989b) *J. Am. Chem. Soc. 111*, 1515-1517.

Marion, D., & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967-974.

Maudsley, A. A., Wokaun, A., & Ernst, R. R. (1978) *Chem. Phys. Lett. 55*, 9-14.

McCoy, M. A., & Müeller, L. (1992a) *J. Am. Chem. Soc. 114*, 2108-2112.

McCoy, M. A., & Müeller, L. (1992b) *J. Magn. Reson. B98*, 674-679.

Meiboom, S., & Gill, G. (1958) *Rev. Sci. Instrum. 29,* 668-691.

Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989) *J. Magn. Reson. 85*, 608-613.

Morris, G. A., & Freeman, R. (1979) *J. Am. Chem. Soc. 101*, 760-762.

Müeller, L. (1979) *J. Am. Chem. Soc. 101*, 4481-4484.

Müeller, L. (1987) *J. Magn. Reson. 72*, 191-196.

Muhandiram, D. R., & Kay, L. E. (1994) *J. Magn. Reson. B 103*, 203-216.

Muhandiram, D. R., Xu, G., & Kay, L. E. (1993) *J. Biomol. NMR 3*, 463-470.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 129-136.

Overhauser, A. W. (1953) *Phys. Rev. 89*, 689-700; *ibid*, *92*, 411-415.

Palmer, A. G., Cavanagh, J., Wright, P. E., & Rance, M. (1991) *J. Magn. Reson. 93,* 151-170.

Palmer, A. G., Skelton ,N. J., Chazin, W. J., Wright, P. E., & Rance, M. (1992) *Mol. Phys. 75*, 699-711.

Palmer, A. G., Williams, J., & McDermott, A. (1996) *J. Phys. Chem. 100,* 13293-13310.

Peng, J. W., & Wagner, G. (1992) *J. Magn. Reson. 98,* 308-332.

Peng, J. W., & Wagner, G. (1994) *Methods in Enzymology 239*, 563-596.

Peng, J. W., & Wagner, G. (1995) *Biochemistry 34,* 16733-16752.

Piantini, U., Sørensen, O. W., & Ernst, R. P. (1982) *J. Am. Chem. Soc. 104*, 6800-6801.

Piotto, M., Saudek, V., & Sklenar, V. (1992) *J. Biomol. NMR 2*, 661-665.

Prigogine, I. (1962) *Non-Equilibrium Statistical Mechanics*, John Wiley & Sons, New

York, New York.

Purcell, E. M., Torrey, H. C., & Pound, R. V. (1946) *Phys. Rev. 69,* 37-38.

Rabi, I. I., Schwinger, J., & Ramsey, N. F. (1954) *Rev. Mod. Phys. 26*, 167-171.

Redfield, A. G. (1958) *IBM J. Res. Dev. 1*, 19-31.

Redfield, A. G. (1965) *Adv. Magn. Reson. 1*, 1-32.

Redfield, A. G., & Gupta, R. K. (1971) *Adv. Magn. Reson. 5,* 81-115.

Resing, H. A., & Torrey, H. C. (1963) *Phys. Rev. 131,* 1102-1104.

Rice, S. O. (1944) in *Selected Papers on Noise and Stochastic Processes* edited by Nelson Max *pp. 133-294*, Dover, New York, New York.

Richarz, R., Nagayama, K., & Wuthrich, K. (1980) *Biochemistry 19*, 5189-5196.

Royden, V. (1954) *Phys. Rev. 96*, 543-544.

Shaka, A. J., Barker, P. B., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547-552.

Shaka, A. J., Keeler, J., Frenkiel, T., & Freeman, R. (1983) *J. Magn. Reson. 52*, 335-338.

Shaka, A. J., Lee, C. J., & Pine, A. (1988) *J. Magn. Reson. 77*, 274-293.

Shukert, S. B., Hajduk, P. J., Meadows, R. P., & Fesik, S. W. (1996) *Science 274*, 1531-1534.

Sklenar, V., & Bax, A. (1987) *J. Magn. Reson. 74*, 469-479.

Snyder, G. H., Rowan, R., Karplus, S., & Sykes, B. D. (1975) *Biochemistry 14*, 3765-3777.

States, D. J., Haberkon, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.

Stejskal, E. O., & Tanner, J. E. (1965) *J. Chem. Phys. 42,* 288-292.

Solomon, I. (1955) *Phys. Rev. 99,* 559-565.

Sorensen, O. W., Eich, G. W., Levitt, M. H., Bodenhausen, G., & Ernst, R., R. (1983)

Prog. NMR Spectrosc. 16, 163-192.

Szypersky, T., Luginbuhl, P., Otting, G., Guntert, P., & Wüthrich, K. (1993) *J. Biomol. NMR 3*, 151-164.

Torrey, H. C. (1953) *Phys. Rev. 92*, 962-969.

Torrey, H. C. (1954) *Phys. Rev. 96*, 690-690.

Van de Ven, F. J. M., & Hibers, C. W. (1983) *J. Magn. Reson. 54*, 512-520.

Vold, R. L., & Vold, R. R. (1978) *Prog. NMR Spectrosc. 12,* 79-133.

Vold, R. L., Waugh, J. S., Klein, M. P., Phelps, D. E. (1968) *J. Chem. Phys. 48*, 3831-3832.

Von-Neumann, J. (1927) *J. Gottinger Nachr.* 245 and 273.

Von-Neumann, J. (1955) *Mathematical Foundations of Quantum Mechanics,* Princeton University Press, Princeton. New Jersey.

Vüister, G. W., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772-7777.

Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347-366.

Wallach, D. (1967) *J. Chem. Phys. 47*, 5258-5268.

Wang, Y-X., Freedberg, I. D., Yamazaki, T., Wingfield, P. T., Stahl, S. J., Kaufman, J. D., Kiso, Y., & Torchia, D. A. (1996) *Biochemistry 35*, 9945-9950.

Wang, M., & Uhlenbeck, G. E. (1945) *Rev. Mod. Phys. 17*, 323-342.

Wangsness, R. K., & Bloch, F. (1953) *Phys. Rev. 89,* 728-739.

Werbelow, L. G., & Grant, D. M. (1977) *Adv. Magn. Reson. 9*, 189-299.

Wider, G., Hosur, R. V., & Wüthrich, K. (1983) *J. Magn. Reson. 52*, 130-135.

Wider, G., Riek, R., & Wüthrich, K. (1996) *J. Am. Chem. Soc. 118*, 11629-11634.

Wider, G., & Wüthrich, K. (1993) *J. Magn. Reson. A102*, 239-241.

Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295-315.

Wittebort, R. J., & Szabo, A. (1978) *J. Chem. Phys. 69,* 1722-1737.

Wittekind, M., & Müeller, L. (1993) *J. Magn. Reson. 101*, 201-205.

Woessner, D. E. (1962) *J. Chem. Phys. 36*, 1-4.

Wokaun, A., & Ernst, R. R. (1977) *Chem. Phys. Lett. 52*, 407-412.

Wüthrich, K. (1986) *NMR of Proteins and Nuclear Acids*, Wiley, New York.

Wüthrich, K., & Wagner, G. (1975) *FEBS Lett. 50,* 265-268.

Zhang, O., Kay, L. E., Olivier, J. P., & Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845-858.

Zhu, G., & Bax, A. (1990) *J. Magn. Reson. 90*, 405-410.

Zhu, L., Kemple, M. D., Landy, S. B., & Buckley, P. (1995) *J. Magn. Reson. B109*, 19-30.

Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387-2391.

Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., & Fesik, S. W. (1990) *J. Magn. Reson. 86*, 210-216.

Zuiderweg, E. P. R., Petros, A. M., Fesik, S. W., Olejniczak, E. T. (1991) *J. Am. Chem. Soc. 113,* 370-372.

Zwanzig, R. (1961) *Phys. Rev. 124,* 983-992.

CHAPTER 2

Structure-Function Relationships of Cellular Retinoic Acid Binding Proteins

(Reprinted with the permission of the American Society of Biochemistry and Molecular Biology)

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 3, Issue of January 17, pp. 1541–1547, 1997
Printed in U.S.A.

Structure-Function Relationships of Cellular Retinoic Acid-binding Proteins

QUANTITATIVE ANALYSIS OF THE LIGAND BINDING PROPERTIES OF THE WILD-TYPE PROTEINS AND SITE-DIRECTED MUTANTS*

(Received for publication, August 1, 1996, and in revised form, September 27, 1996)

Lincong Wang, Yue Li, and Honggao Yan‡

From the Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824

It has been suggested that electrostatic interactions are critical for binding of retinoic acid by cellular retinoic acid-binding proteins (CRABP-I and CRABP-II). However, the roles of two conserved arginine residues (Arg-111 and Arg-131 in CRABP-I; Arg-111 and Arg-132 in CRABP-II) that interact with the carboxyl group of retinoic acid have not been evaluated. A novel competitive binding assay has been developed for measuring the relative dissociation constants of the site-directed mutants of CRABPs. Arg-111 and Arg-132 of CRABP-II were replaced with methionine by site-directed mutagenesis. The relative dissociation constants of R111M and R132M ($K_{d\,(R111M)}/K_{d\,(CRABP\text{-}II)}$ and $K_{d\,(R132M)}/K_{d\,(CRABP\text{-}II)}$) were determined to be 40–45 and 6–8, respectively. The ring protons of the aromatic residues of the wild-type CRABP-II and the two mutants were sequentially assigned by two-dimensional homonuclear NMR in conjunction with three-dimensional heteronuclear NMR. Detailed analysis of the nuclear Overhauser effect spectroscopy spectra of the proteins indicated that the conformations of the two mutants are highly similar to that of the wild-type CRABP-II. These results taken together showed that Arg-111 and Arg-132 are important for binding retinoic acid but contribute to the binding energy only by ~2.2 and 1.2 kcal/mol, respectively. In addition, the relative dissociation constant of CRABP-II and CRABP-I ($K_{d\,(CRABP\text{-}II)}/K_{d\,(CRABP\text{-}I)}$) was determined to be 2–3, in close agreement with that calculated using the apparent K_d values determined under the same conditions by fluorometric titrations.

Retinoic acid (RA),[1] a hormonally active metabolite of vitamin A, has profound effects on cell growth, differentiation, and morphogenesis. Two types of proteins have been found to bind RA: nuclear retinoic acid receptors (RARs and RXRs) and cellular retinoic acid-binding proteins (CRABPs). RARs and RXRs are RA-activated transcriptional factors that regulate expression of target genes (1). Although the physiological roles of CRABPs are not clear at present, they are thought to be involved in cellular transport and metabolism of RA (2). Two isoforms (CRABP-I and CRABP-II) have been characterized. Both CRABP-I and CRABP-II bind specifically to all-*trans*-retinoic acid, but they differ in several respects: (i) CRABP-I has higher affinity for RA than CRABP-II (3–6); (ii) CRABP-I is expressed in many adult mouse and human tissues, but the expression of CRABP-II is limited to skin (7, 8); and (iii) RA stimulates expression of CRABP-II but not that of CRABP-I (8). It appears that the two isoforms may have distinct functions. The idea is supported by the fact that the sequence identity of human and mouse CRABP-I (99.3%) or human and mouse CRABP-II (93.5%) is much higher than the sequence identity (73.7%) between the two isoforms from the same source.

CRABPs are members of a family of intracellular lipid-binding proteins that bind small hydrophobic molecules such as retinoids and fatty acids (9). The family of proteins includes CRABPs, cellular retinol-binding proteins, fatty acid-binding proteins, P2 myelin protein, a mammary gland protein, and gastropin. The structures of 11 different intracellular lipid-binding proteins, including CRABP-I and CRABP-II, have been determined by x-ray crystallography (9–11). Although the sequence identity among intracellular lipid-binding proteins is rather low (~20%), these crystal structures are remarkably similar with respect to backbone folding. They are composed of two nearly orthogonally packed five-stranded β-sheets and two short α-helices. The ligand binding pocket in each protein is deep inside the interior of the β-barrel formed by the two β-sheets. The helix-turn-helix motif is located at the ligand entrance. These proteins apparently lack a true hydrophobic core that is important for folding and stability of many other proteins (12).

Electrostatic interactions are thought to play major roles in binding of RA by CRABPs because the two proteins bind only retinoic acid but not retinol and retinal. Two arginine residues (Arg-111 and Arg-131 in CRABP-I; Arg-111 and Arg-132 in CRABP-II) have been identified by crystallography to interact with the carboxyl group of the bound RA (10). However, previous site-directed mutagenesis studies suggest that both Arg-111 and Arg-131 are critical for CRABP-I to bind RA (13) but that only Arg-132 is important for CRABP-II to bind RA (14). Because of the limitations of the current methods for measuring RA binding, the dissociation constants of the critical mutants have not been measured. Furthermore, the conformations of the CRABP-I mutants have been characterized only by CD, and the conformations of the CRABP-II mutants have not been characterized. Thus, the contributions of the electrostatic interactions between the arginine residues and RA to binding are still uncertain.

We are interested in studying the quantitative structure-function relationships of CRABPs with respect to binding of

* This work was supported by funds from the REF Center of Protein Structure and Design and the Cancer Center at Michigan State University. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 517-353-8786; Fax: 517-353-9334; E-mail: yan@nmr1.bch.msu.edu.

[1] The abbreviations used are: RA, all-*trans*-retinoic acid; CRABP, cellular retinoic acid-binding protein; DQF-COSY, double quantum-filtered correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; RAR, retinoic acid receptor; RXR, retinoid X receptor; TOCSY, total correlation spectroscopy; WT, wild-type CRABP-II; MT, CRABP-II mutant.

RA. In this paper, we report a novel competitive binding assay developed for measuring the dissociation constants of the site-directed mutants of CRABPs. We have used the novel method to evaluate the contributions of both Arg-111 and Arg-132 of CRABP-II to binding of RA in conjunction with site-directed mutagenesis and NMR. The results show that like CRABP-I, both Arg-111 and Arg-132 in CRABP-II are important for binding of RA, contrary to the results of Chen *et al.* (14). However, Arg-111 and Arg-132 contribute to the overall binding energy only by ~2.2 and 1.2 kcal/mol, respectively.

EXPERIMENTAL PROCEDURES

Materials—Nonradioactive RA was purchased from Sigma. [11,12-^{3}H]RA was purchased from DuPont NEN. A DNA sequencing kit was obtained from U.S. Biochemical Corp. Enzymes for recombinant DNA experiments were purchased from Life Technologies, Inc., or New England Biolabs. Other chemicals were analytical or reagent grade from commercial sources.

Cloning and Expression—The cDNAs encoding human CRABP-I and CRABP-II were kindly provided by Dr. Anders Åström (8). The genes were subcloned into the expression vector pET-17b (Novagen) by polymerase chain reaction. The primers for polymerase chain reaction were: 5'-GGAATTCCATATGCCCAACTTCGCCGGC-3' (forward, CRABP-I), 5'-CGGGATCCTCATTCCCGGACATAAATTC-3' (reverse, CRABP-I), 5'-CGGGATCCATATGCCAAACTTCTCTGGCAACTG-3' (forward, CRABP-II), and 5'-GGAATTCTCACTCTCGGACGTAGAC-3' (reverse, CRABP-II). To ensure that there were no undesired mutations in the amplified genes, they were sequenced by double-stranded DNA sequencing from both orientations. The expression constructs were then transformed into the bacterial strain BL21(DE3)pLysS (15). The transformed bacterial cells were grown at 37 °C in a LB agar plate containing both ampicillin (100 μg/ml) and chloramphenicol (25 μg/ml). Expression of CRABPs was verified by growing the colonies in 5 ml of LB medium followed by induction with isopropyl-1-thio-β-D-galactopyranoside and SDS-polyacrylamide gel electrophoresis of the harvested cells.

Site-directed Mutagenesis—The oligonucleotides for making human CRABP-II mutants were 5'-TCGTGGACCATGGAACTGACCA-3' (R111M) and 5'-GTGTGCACCATGGTCTACGTC-3' (R132M). The mutants were generated by the method of Kunkel (16) and screened by DNA sequencing. In order to ensure that there were no unintended mutations in the mutants, the entire sequences of the mutated genes were determined.

Protein Purification—All proteins were purified by the same procedure as described below. 100 ml of LB medium containing 100 μg/ml of ampicillin and 25 μg/ml of chloramphenicol was inoculated by a small piece of a frozen seed culture and incubated at 37 °C with vigorous shaking (200 rpm) overnight. It was then used to inoculate 4 liters of LB medium containing both antibiotics and grown at 37 °C. When A_{600} of the culture reached 0.6–0.8, isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 0.4 mM. The culture was allowed to grow further at 30 °C for 3–5 h. The bacterial cells were then harvested by centrifugation and suspended in 100 ml of buffer A (10 mM Tris-HCl, 1 mM dithiothreitol, pH 8.2). The suspension was sonicated on ice and centrifuged (27,000 × *g*) at 4 °C for 30 min. The supernatant was applied to a DEAE-cellulose column equilibrated with the same buffer. The column was washed with buffer A until A_{280} of the eluent was less than 0.05. Elution of the column was achieved by a linear NaCl gradient (0–200 mM in buffer A) and monitored by A_{280} and 15% SDS-polyacrylamide gel electrophoresis. The fractions containing CRABP were pooled and concentrated by an Amicon ultrafiltration cell using a YM 10 membrane. The concentrated protein solution was centrifuged (27,000 × *g*) for 20 min. The supernatant was loaded onto a Sephadex G-50 column equilibrated with phosphate-buffered saline (4 mM NaH_2PO_4, 16 mM Na_2HPO_4, 150 mM NaCl, pH 7.3) and eluted with the same buffer. The fractions containing CRABP were located by A_{280} and 15% SDS-polyacrylamide gel electrophoresis. The fractions of greater than 99% purity were pooled and concentrated. The protein solution was dialyzed against double-distilled water and lyophilized.

Fluorometric Titration—Fluorescence binding assays were carried out by a procedure modified from that of Cogan *et al.* (17) using a Hitachi 4500 fluorometer. Briefly, CRABPs were dissolved in phosphate-buffered saline. The concentrations of the protein stock solutions were measured by A_{280}. The absorption coefficients used were 19,868 M^{-1} cm^{-1} for CRABP-I and 19,480 M^{-1} cm^{-1} for CRABP-II determined by the method of Gill and Hippel (18). RA stock solutions were prepared in absolute ethanol. The concentrations of the RA stock solutions were determined by A_{336} using the absorption coefficient of 45,000 M^{-1} cm^{-1}. The final ethanol content for each titration was kept less than 2%. The samples were excited at wavelength of 283 nm with a slit width of 2.5 nm. The emission wavelengths were 330 nm for CRABP-I and 338 nm for CRABP-II with a slit width of 20 nm. The excitation shutter was closed between measurements. The inner filter effects were not corrected because they were negligible at the protein and RA concentrations we used in the experiments. The data were analyzed by nonlinear least square fit to Equation 1.

$$\frac{F}{F_f} = 1 + \left(\frac{F_b}{F_f} - 1\right) \times \left(\frac{P_t + R_t + K_d - \sqrt{(P_t + R_t + K_d)^2 - 4P_tR_t}}{2P_t}\right) \quad \text{(Eq. 1)}$$

where F is the observed fluorescence, F_b is the fluorescence of the bound CRABP, F_f is the fluorescence of the free CRABP, P_t is the total protein concentration, and R_t is the total RA concentration.

Competitive Binding Assay—The assay was carried out in a Spectra/Por equilibrium dialyzer at room temperature. The two compartments of each dialysis cell were separated by a semipermeable membrane with a molecular mass cut-off of 6–8 kDa. One compartment was filled with the wild-type CRABP-II (1 ml), and the other was filled with a CRABP-II mutant or the wild-type CRABP-I. The buffer used was phosphate-buffered saline plus 5 mM dithiothreitol. An equal amount of [^{3}H]RA (100 nM) was added to each compartment. The protein concentrations in both compartments were much greater (at least 20-fold) than the RA concentration. 100 μl of samples was taken from the two compartments after various times of incubation at room temperature, mixed with 5 ml of scintillation fluids, and counted by a liquid scintillation counter. The equilibria in the two compartments that contained the wild-type CRABP-II and mutant proteins can be described by Equations 2 and 3 (similar equations can be written for CRABP-I and CRABP-II),

$$K_{d\,(WT)} = \frac{[WT][RA]}{[WT \cdot RA]} \quad \text{(Eq. 2)}$$

$$K_{d\,(MT)} = \frac{[MT][RA]}{[MT \cdot RA]} \quad \text{(Eq. 3)}$$

where WT, MT, WT·RA, and MT·RA represent the wild-type CRABP-II, a CRABP-II mutant, and their RA complexes, respectively. Therefore,

$$\frac{K_{d\,(MT)}}{K_{d\,(WT)}} = \frac{[MT][WT \cdot RA]}{[WT][MT \cdot RA]} \quad \text{(Eq. 4)}$$

Since the concentrations of the proteins were much greater than their respective dissociation constants and the concentration of RA, $[WT] \sim [WT]_{total}$, $[MT] \sim [MT]_{total}$, $[WT{\cdot}RA] \gg [RA]$, and $[MT{\cdot}RA] \gg [RA]$. Then the relative dissociation constant can be calculated by Equation 5,

$$\frac{K_{d\,(MT)}}{K_{d\,(WT)}} = \frac{[MT]_{total}C_{WT}}{[WT]_{total}C_{MT}} \quad \text{(Eq. 5)}$$

where C_{WT} and C_{MT} are the measured radioactivities of the two compartments containing the wild-type CRABP-II and the mutant, respectively. It turned out that the system could not reach equilibrium in 2 days, presumably because of few free RAs in solution to diffuse across the membrane. Since RA is not stable even in the dark, the assay was redesigned to match the equilibrium conditions by varying the ratio of the protein concentrations of the wild-type and the mutant ($[MT]_{total}/[WT]_{total}$). Thus, the concentration of the mutant was varied while keeping the concentration of the wild type at ~2 μM. Initially, the concentration of the mutant was increased in an exponential manner (*e.g.* 2, 20, 200 μM). Then it was varied in a small range. Since an equal amount of RA was added to the two compartments of the dialysis cell, the two compartments should have the same RA concentration and radioactivity at the beginning of the assay. If $[MT]_{total}/[WT]_{total} \neq K_{d\,(MT)}/K_{d\,(WT)}$, there would have a net transfer of RA across the semipermeable membrane separating the two compartments. Thus, the radioactivity counts of the two compartments (C_{WT} and C_{MT}) would differ after incubation for a certain period. When $[MT]_{total}/[WT]_{total} < K_{d\,(MT)}/K_{d\,(WT)}$, then $C_{WT} - C_{MT} > 0$. When $[MT]_{total}/[WT]_{total} > K_{d\,(MT)}/K_{d\,(WT)}$, then $C_{WT} - C_{MT} < 0$. When $[MT]_{total}/[WT]_{total} = K_{d\,(MT)}/K_{d\,(WT)}$, then $C_{WT} - C_{MT} = 0$.

NMR Spectroscopy—All NMR measurements were performed at 32 °C on a Varian VXR-500 spectrometer operating at a proton frequency of 500 MHz. The proteins were dissolved in 20 mM sodium phosphate, pH 7.5 (direct pH meter reading), 100 mM NaCl, 5 mM dithiothreitol in D_2O. The protein concentration was ~2 mM. The spec-

tral width was 7200 Hz. For wild-type CRABP-II, one phase-sensitive DQF-COSY spectrum (19), two MLEV-17 clean TOCSY spectra (20–22) with mixing times of 20 and 40 ms and one NOESY spectrum (23, 24) with a mixing time of 150 ms were acquired. For R111M, one DQF-COSY spectrum, three TOCSY spectra with mixing times of 20, 40, and 75 ms and one NOESY spectrum with a mixing time of 150 ms were acquired. For R132M, one TOCSY spectrum with a mixing time of 40 ms and one NOESY spectrum with a mixing time of 150 ms were acquired. All of the spectra were acquired in the hypercomplex mode with standard phase cycling schemes. The data were usually acquired with 2048 complex points in the t2 dimension and 256 complex points in the t1 dimension. 96 transients were collected for each FID. Data processing was performed on a Sun Sparc 10 station using VNMR software from Varian. The time domain data were zero-filled once and multiplied by shifted sine bell or Gaussian functions before Fourier transformation in both dimensions. Base lines were corrected in t2 dimensions using a 5-order polynomial. Chemical shifts were referenced to internal sodium 3-(trimethyl silyl)-propionate-2,2,3,3-d_4.

RESULTS

Fluorometric Titration—Probably because of the simplicity of the method, fluorometric titration has been frequently used for measuring the affinities of CRABPs for RA. However, there are large discrepancies in the K_d values obtained by the method. These could be due to variations in the assay methodology and/or conditions. Some problems with the method have been recently discussed (2, 4, 6). In order to compare the ligand binding properties of CRABPs under the same conditions, we carried out fluorometric titration experiments. The results are shown in Fig. 1. The apparent K_d values obtained by nonlinear fitting were 0.63 nM for CRABP-I and 1.9 nM for CRABP-II. While our work was in progress, Norris *et al.* (6) reported the apparent K_d values of mouse CRABPs measured by an improved fluorescence titration method. The major changes were the lowering of protein and RA concentrations and the use of nonlinear least square fitting for data analysis. These modifications were in line with our approach. Our fluorometric titration data were analyzed by nonlinear least square fitting. The protein concentrations we used in the assays (~50 nM) were the lowest among those used in previous studies except those of Norris *et al.* (6). We did not include the additive gelatin used by Norris *et al.* (6) in the assays. The apparent K_d values determined by us are in close agreement with those (0.4 nM for CRABP-I and 2 nM for CRABP-II) reported by Norris *et al.* (6).

Competitive Binding Assay—In order to measure the affinities of the mutants of CRABPs for RA, we developed a novel competitive binding assay. We first used the method to measure the relative affinity of CRABP-I and CRABP-II for RA. The result is shown in Fig. 2*A*. When the ratio of the concentrations of the two proteins ([CRABP-II]/[CRABP-I]) was ≤2, there was a net transfer of RA from the compartment containing CRABP-II to the compartment containing CRABP-I. When the ratio of the concentrations of the two proteins was ≥3, there was a net transfer of RA from the compartment containing CRABP-I to the compartment containing CRABP-II. As described under "Experimental Procedures," the relative K_d of the two proteins lies between the points with opposite net transfers. Thus the K_d of CRABP-II relative to that of CRABP-I (K_d(CRABP-II)/K_d(CRABP-I)) is 2–3. It is in close agreement with the results of the fluorometric titration studies. We then used the competitive binding assay to measure the K_d values of R111M and R132M relative to that of the wild-type CRABP-II. The results are shown in Fig. 2 (*B* and *C*). In the case of R111M (Fig. 2*B*), there was a net transfer of RA from the compartment containing R111M to the compartment containing the wild-type CRABP-II when [R111M]/[WT] ≤ 40. When [R111M]/[WT] ≥ 45, there was a net transfer of RA from the compartment containing the wild-type protein to the compartment containing R111M. Therefore, the relative affinity of R111M for RA ($K_{d\,(R111M)}/K_{d\,(WT)}$) is 40–45. As shown in Fig. 2*C*, the relative affinity of R132M for RA ($K_{d\,(R132M)}/K_{d\,(WT)}$) is 6–8.

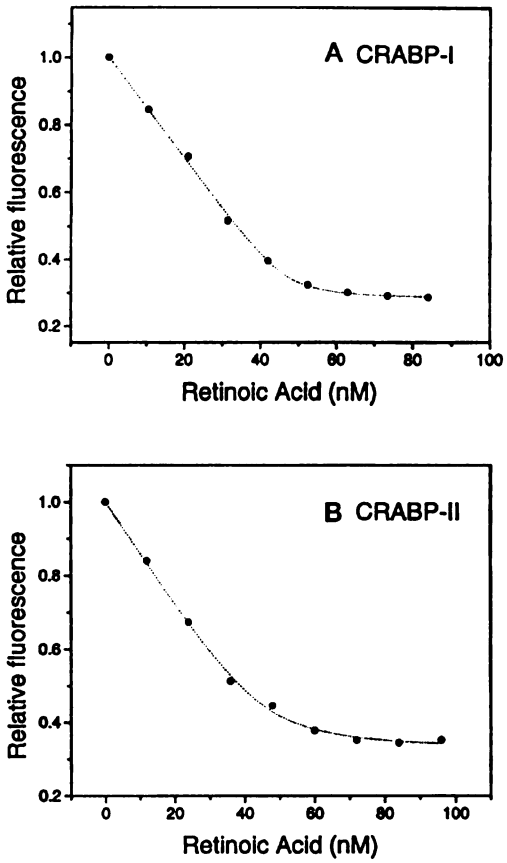


FIG. 1. Fluorometric titration of wild-type CRABP-I (*A*) and CRABP-II (*B*). The concentrations of CRABP-I and CRABP-II were 40 and 50 nM, respectively.

Sequential Resonance Assignment of the Ring Protons of the Aromatic Residues of the Wild-type CRABP-II and Site-directed Mutants—Since perturbations in the ligand binding property of a mutant may be due to conformational changes caused by the mutation (25), we characterized the conformations of the mutants by NMR. A prerequisite for NMR structural analysis is sequential resonance assignment. Currently, we are in the process of making total sequential resonance assignments of the wild-type CRABP-II and R111M by multidimensional multinuclear NMR. We have sequentially assigned the ring protons of aromatic residues of the wild-type and mutants based on the homonuclear two-dimensional NMR (DQF-COSY, TOCSY, and NOESY) and heteronuclear three-dimensional NMR.[2] The chemical shifts of the ring protons are listed in Table I. The sequential resonance assignment is labeled in Fig. 3 for R132M. In comparison with the chemical shifts of the wild-type ring protons, there are quite a few chemical shift changes in the ring protons of the mutants. Some chemical shift changes can be easily rationalized, while the causes for other chemical shift changes are not obvious. For example, the guanidino group of Arg-132 stacks on the aromatic ring of

[2] L. Wan, Y. Li, and H. Yan, unpublished results.

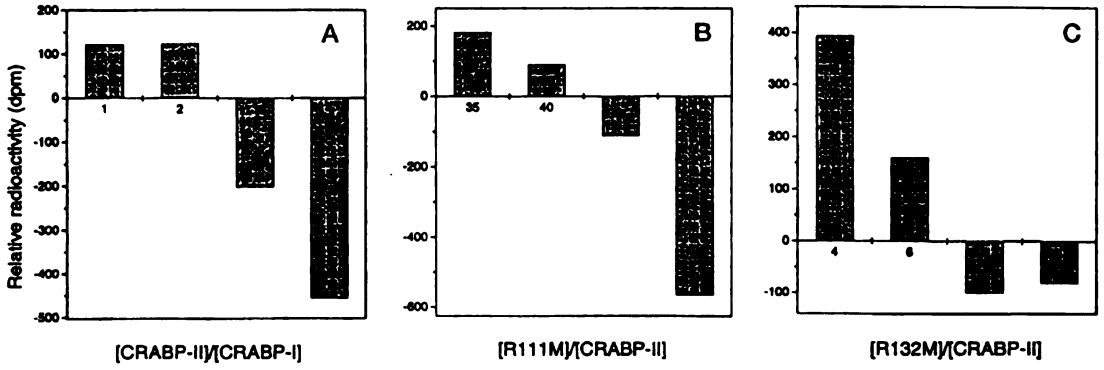


FIG. 2. **Competitive binding assays for measuring the relative dissociation constants of CRABPs.** The relative radioactivity in *A* is the radioactivity count of the compartment containing CRABP-I minus that of the compartment containing CRABP-II. The relative radioactivity in *B* is the radioactivity count of the compartment containing CRABP-II minus that of the compartment containing R111M. The relative radioactivity in *C* is the radioactivity count of the compartment containing CRABP-II minus that of the compartment containing R132M.

TABLE I

Chemical shifts of the ring protons of the aromatic residues of wild-type CRABP II, R111M mutant, and R132M mutant

The underlined values are the mutant resonances that differ by >0.02 ppm from the corresponding resonances of the wild-type CRABP-II.

Residue	Wild-type CRABP-II				R111M				R132M			
Phe-3	7.60[a]	7.34[b]	7.52[c]		7.61[a]	7.38[b]	7.45[c]		7.60[a]	7.34[b]	7.52[c]	
Phe-15	7.42[a]	7.02[b]	7.11[c]		7.39[a]	7.12[b]	7.20[c]		7.40[a]	7.11[b]	7.11[c]	
Phe-50	6.66[a]	7.26[b]	6.99[c]		6.68[a]	7.24[b]	6.99[c]		6.64[a]	7.24[b]	6.99[c]	
Phe-65	6.42[a]	6.70[b]	6.48[c]		6.38[a]	6.70[b]	6.44[c]		6.42[a]	6.69[b]	6.42[c]	
Phe-71	7.48[a]	7.00[b]	6.70[c]		7.44[a]	6.99[b]	6.73[c]		7.48[a]	7.00[b]	6.69[c]	
Tyr-51	7.12[d]	6.79[e]			7.12[d]	6.74[e]			7.13[d]	6.78[e]		
Tyr-134	6.98[d]	6.65[e]			7.00[d]	6.86[e]			7.03[d]	6.75[e]		
Trp-7	7.25[f]	7.00[g]	7.41[h]	7.30[i]	7.22[f]	6.99[g]	7.39[h]	7.29[i]	7.25[f]	6.99[g]	7.43[h]	7.30[i]
Trp-87	8.16[f]	7.13[g]	7.39[h]	7.80[i]	8.16[f]	7.04[g]	7.39[h]	7.76[i]	8.17[f]	7.05[g]	7.39[h]	7.80[i]
Trp-87[j]									8.07[f]	7.05[g]	7.38[h]	7.76[i]
Trp-109	7.08[f]	6.53[g]	6.61[h]	6.87[i]	7.01[f]	6.59[g]	6.59[h]	6.89[i]	7.09[f]	6.52[g]	6.61[h]	6.87[i]

[a-c] 2,6H, 3,5H, and 4H of phenylalanine, respectively.
[d,e] 2,6H and 3,5H of tyrosine, respectively.
[f-i] 4H, 5H, 6H, and 7H of tryptophan, respectively.
[j] Trp-87 has two sets of resonances because of conformational heterogeneity.

Phe-15 and is hydrogen-bonded to the hydroxyl group of Tyr-134 in the crystal structure of the wild-type CRABP-II (10). Removal of the guanidino group is likely to be the cause for the chemical shift changes in Phe-15 and Tyr-134 of R132M. The crystal structure also reveals that the guanidino group of Arg-111 interacts with Trp-109 through an amido/aromatic-ring hydrogen bond (10). Thus, removal of the guanidino group of Arg-111 may cause the chemical shift changes in Trp-109 of R111M. Some small chemical shift differences may result from changes in pH, temperature, and other conditions. They are not necessarily indicative of conformational changes. Conformational changes are best characterized by analysis of NOESY spectra as described below. Interestingly, 4H and 7H of the Trp-87 ring protons in R132M have two sets of NMR signals (Fig. 3), indicating that there is localized conformational heterogeneity and that the Trp ring is in two distinct conformational states. The lifetimes of the two conformational states are at least 50 ms. The ratio of the populations in the two states (major and minor forms) is about 4:1. 1H of the Trp-87 ring protons in the wild-type protein and R111M also show two distinct resonance signals (data not shown), indicating that the two proteins also have localized conformational heterogeneities. The lifetimes of the two conformational states are at least 40 ms in the wild-type protein and at least 33 ms in R111M. The ratio of the major and minor forms is about 3:1 in the wild-type protein and about 4:1 in R111M. Conformational dynamics may be important for RA to get into the deep binding pocket of the protein. However, Trp-87 is not at the RA-binding pocket and is far away from the ligand entrance. The functional significance of the conformational heterogeneity is not clear.

Conformational Comparison of the Wild-type CRABP-II, R111M, and R132M—The aromatic-aromatic NOEs are shown in the *lower panels* of Fig. 4 (wild-type CRABP-II), Fig. 5 (R111M), and Fig. 6 (R132M). Some aromatic-aliphatic NOEs are shown in the *upper panels* of the figures. Interresidue NOEs are labeled, because they are indicative of the conformation of a protein. The interresidue NOE peaks of the three proteins are the same except for some minor variations in peak intensity. The observed NOEs are also consistent with the atomic distances measured from the coordinate of the crystal structure of CRABP-II complexed with RA. The identities of the aliphatic protons shown in the *upper panels* of these figures are not known at present. However, qualitatively, the three proteins have very similar aromatic-aliphatic NOE patterns. The results of the NOESY analysis suggest that the conformations of R111M and R132M are highly similar to that of the wild-type protein. Thus, the decreases in affinity for RA of the two mutants are most likely due to the disruption of the interactions of the arginine residues with RA by mutagenesis. Arg-111 and Arg-132 contribute to the binding energy by ~2.2 and 1.2 kcal/mol, respectively.

DISCUSSION

Competitive Binding Assay—Two types of methods have been in general use for measuring binding of RA to CRABPs: fluorometry and radiometry. The radiometric method involves

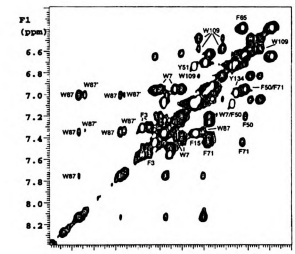


FIG. 3. TOCSY spectrum of the aromatic protons of R132M at 32 °C showing two sets of resonances for Trp-87. The mixing time was 40 ms. All of the cross-peaks have been sequentially assigned. The chemical shifts of the resonances are listed in Table I.

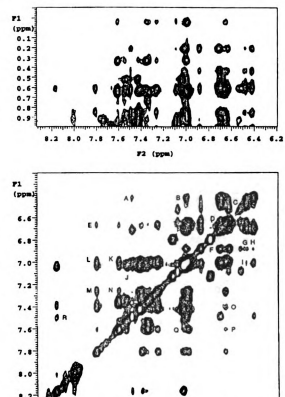


FIG. 4. Parts of the 500-MHz NOESY spectrum of the wild-type CRABP-II at 32 °C. The mixing time was 150 ms. Only the interresidue NOEs are labeled. The identities of the NOEs are as follows. *A*, Phe-65 2,6H to Phe-71 2,6H; *B*, Phe-65 2,6H to Phe-71 3,5H; *C*, Phe-65 4H to Trp-109 6H; *D*, Phe-65 3,5H to Trp-109 6H; *E*, Phe-50 2,6H to Trp-87 7H; *F*, Phe-71 4H to Trp-109 7H; *G*, Phe-65 4H to Trp-109 7H; *H*, Phe-65 2,6H to Trp-109 7H; *I*, Phe-50 3,5H to Phe-65 4H; *J*, Trp-87 5H to Phe-3, 4H; *K*, Phe-50 3,5H to Phe-3 2,6H; *L*, Phe-50 3,5H to Trp-87 7H; *M*, Phe-50 4H to Trp-87 7H; *N*, Phe-50 4H to Phe-3 2,6H; *O*, Phe-50 2,6H to Trp-87 6H; *P*, Phe-50 2,6H to F3–2,6H; *Q*, Trp-87 5H to Phe-3 2,6H; *R*, Phe-3 4H to Trp-87 4H.

separation of bound from free RA by dextran-coated charcoal, gel filtration, and other means. Substantial loss of bound ligand during the separation process makes the method unsuitable for measuring the dissociation constants of site-directed mutants with greatly decreased affinity for RA. The very limited solubility of RA in water (~200 nM) (26) also makes the fluorometric method inapplicable for determining the dissociation constants of these mutants. Studies of the quantitative structure-function relationships of CRABPs have been hampered by the lack of methods for measuring the affinities of site-directed mutants for RA (13, 14). We have developed a novel competitive binding assay for measuring the affinities of site-directed mutants for RA relative to that of the wild-type CRABP. The essence of the method is to monitor the competition between a mutant and the wild-type protein for binding of limited RA. The method uses an equilibrium dialyzer. The two compartments of each dialysis cell are filled with the wild-type and mutant proteins, respectively. The absolute concentration of RA is not important as long as the concentration of free RA is much smaller than that of bound RA. There is no need to separate bound from free RA by dextran-coated charcoal, gel filtration, and other means. The transfer of RA from one compartment to the other is determined by measuring the radioactivities of the samples taken from the two compartments. The direction of the net transfer is dependent on the relative affinity of the proteins and the ratio of the protein concentrations of the two compartments. If the solubility of a mutant is the same as that of the wild-type CRABP-II (~2 mM), then the ratio of the concentrations of the wild-type and mutant proteins can be varied by more than 1,000-fold, and a relative K_d of as large as 1000 ($K_{d\,(MT)}/K_{d\,(WT)} = 1000$) can still be measured by the method. More importantly, determination of the relative dissociation constant of a point mutant is sufficient for estimating the energetic contribution of the amino acid residue to ligand binding ($\Delta\Delta G = RT\ln(K_{d\,(MT)}/K_{d\,(WT)})$).

In addition to vitamin A, many other hydrophobic molecules such as vitamin D and steroid hormones play vital roles in a variety of cellular processes. Studies of the structure-function relationships of the receptor proteins for these bioactive hydrophobic molecules are of fundamental interest and biomedical significance. The method can also be used for measuring the relative dissociation constants of the site-directed mutants of these important proteins and may facilitate these studies.

Relative Affinity of CRABP-I and CRABP-II for RA—CRABP-I and CRABP-II differ in affinity for RA, expression pattern, and regulation. Although it is generally accepted that CRABP-I has a higher affinity for RA than CRABP-II, it has been difficult to determine precisely the relative affinity of the two proteins. A wide range of apparent K_d values have been reported. The apparent K_d values for purified CRABP-I are in the range of 0.4–39 nM (4–6, 13, 27–29). The apparent K_d values for purified CRABP-II are in the range of 2–64 nM (3–6, 14). The best estimation of the relative affinity of the two proteins for RA was made by Fiorella *et al.* (5). In their study, RA was first incubated with CRABP-II and then mixed with CRABP-I. The protein·RA complexes (CRABP-I·RA and CRABP-II·RA) were separated by chromatofocusing. The result indicated that CRABP-II has about 3-fold lower affinity for RA than CRABP-I. More recently, Norris *et al.* (6) compared the affinities of CRABP-I and CRABP-II for RA by fluorometric titration under the same conditions. The results showed that

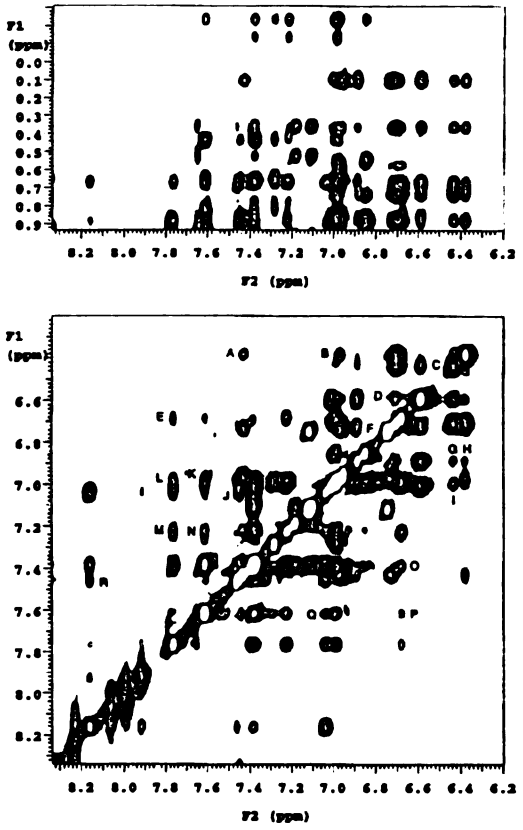


FIG. 5. **Parts of the 500-MHz NOESY spectrum of R111M at 32 °C.** The mixing time was 150 ms. Only the interresidue NOEs are labeled. The identities of the NOEs are given in the legend to Fig. 4.

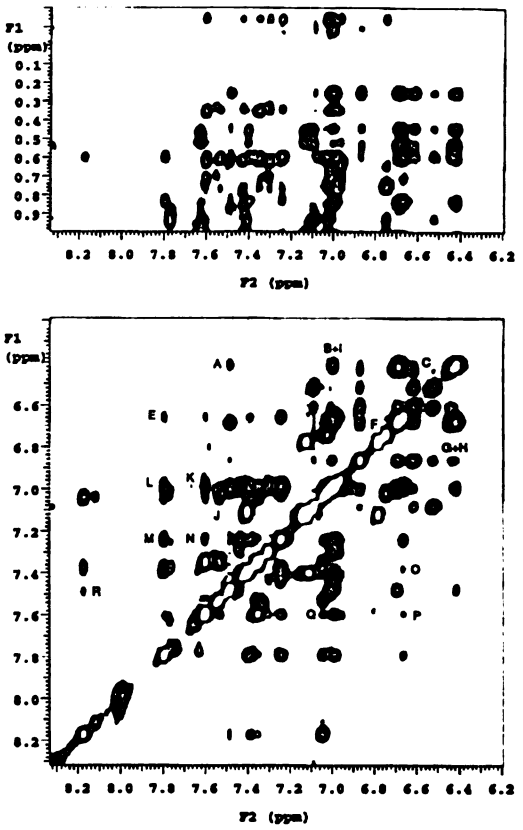


FIG. 6. **Parts of the 500-MHz NOESY spectrum of R132M at 32 °C.** The mixing time was 150 ms. Only the interresidue NOEs are labeled. The identities of the NOEs are given in the legend to Fig. 4.

CRABP-I has 4-fold higher affinity for RA than CRABP-II (although a 10-fold stronger affinity of CRABP-I for RA, compared with CRABP-II, was stated in the paper, apparently due to an error in calculation). The apparent dissociation constant of CRABP-II relative to that of CRABP-I ($K_{d\,(\text{CRABP-II})}/K_{d(\text{CRABP-I})}$) is 3, based on our fluorometric measurements. The results of our competitive binding assays indicated that the relative apparent dissociation constant of CRABP-II ($K_{d\,(\text{CRABP-II})}/K_{d(\text{CRABP-I})}$) is 2–3. The results of our fluorometric measurements and competitive binding studies are in close agreement. They are also in line with the results of Fiorella *et al.* (5). The relative affinity of CRABP-I and CRABP-II measured by Fiorella *et al.* (5) may be the upper limit, because some RA bound to CRABP-II could be preferentially lost during separation of the protein·RA complexes by chromatofocusing. The difference in affinity for RA between CRABP-I and CRABP-II is rather small (1–2-fold).

Both Arg-111 and Arg-132 of CRABP-II Are Important for Binding of RA—Arg-111, Arg-132, and Tyr-134 of CRABP-II have been predicted to be involved in binding of RA on the basis of homologous modeling (13, 30, 31). The crystal structure of CRABP-II in complex with RA reveals that one of the carboxylate oxygens of RA is hydrogen-bonded to the guanidino group Arg-132 and the hydroxyl group of Tyr-134, and the other carboxylate oxygen of RA interacts with Arg-111 via a water molecule (10). We have substituted Arg-111 and Arg-132 with methionine. Our biochemical and structural characterizations show that guanidino groups of Arg-111 and Arg-132 contribute to the binding energy by ~2.2 and 1.2 kcal/mol, respectively. The results indicate that both Arg-111 and Arg-132 are indeed involved in binding of RA and that Arg-111 contributes more to the binding energy than Arg-132. Fluorometric titration also suggests that R132M has higher affinity for RA than R111M (data not shown). Zhang *et al.* (13) have replaced the corresponding residues in CRABP-I with glutamine (R111Q and R131Q). Quantitative comparison of our results with those of Zhang *et al.* (13) is not possible, because the dissociation constants of their mutants have not been determined. Qualitatively, the ligand binding properties of our mutants are similar to those of the corresponding CRABP-I mutants. Zhang *et al.* (13) observed that both the wild-type CRABP-I and R131Q can induce CD of RA in the wavelength range between 320 and 360 nm, but R111Q cannot induce the CD signal. The results of the CD experiments with CRABP-I are consistent with our measurements of the energetic contributions of Arg-111 and Arg-132 to the binding of RA. Chen *et al.* (14) have mutated Arg-111 of CRABP-II to alanine (R111A) and Arg-132 of CRABP-II to alanine and glutamine (R132A and R132Q). In contrast to our R111M mutant, the apparent dissociation constant of their R111A mutant is very similar to that of the wild-type CRABP-II (only 50% higher than that of the wild type). The causes of the discrepancy are not clear, but several possibilities can be ruled out. (i) It is unlikely that there are mutations other than substitution of Arg-111 with a methionine in our R111M mutant. We have sequenced the entire gene after mutagenesis and subcloning into the expression vector. (ii) Steric hindrance is unlikely to be the cause, because the side chain of methionine is smaller than that of arginine. (iii) Although small conformational changes cannot be ruled out, NMR characterizations suggest that R111M is properly folded, and its conformation is

very similar to that of the wild-type protein. Furthermore, R111M has been crystallized. The preliminary crystal structure at 2.2-Å resolution shows that the structure of R111M is indeed very similar to that of the wild-type protein.[3] The ligand binding property of our R132M mutant is qualitatively similar to those of the R132A and R132Q generated by Chen *et al.* (14). However, quantitative comparison is again not possible, because the apparent dissociation constants of the two mutants have not been determined.

The results of our mutagenesis studies appear to be partly inconsistent with the crystallographic data. The crystal structure of holo-CRABP-II suggests that Arg-132 interacts directly with the carboxyl group of RA, whereas the interaction between Arg-111 and the carboxyl group of RA is mediated by a water molecule (10). Our results agree with the crystallographic observation that both Arg-111 and Arg-132 are involved in binding of RA. However, in contrast to the crystallographic data, our mutagenesis results indicate that Arg-111 is more important than Arg-132 for RA binding. But this is not necessarily inconsistent with the crystallographic studies. Since Arg-132 and Tyr-134 are hydrogen-bonded to the same carboxylate oxygen of RA, the loss of the hydrogen bond between Arg-132 and RA in R132M may be partially compensated by the interaction between Tyr-134 and RA. On the other hand, the other carboxylate oxygen of RA only interacts with Arg-111 (albeit via a water molecule), the loss of the interaction in R111M may not be compensated. Alternatively, in solution, the position of the carboxyl group of the bound RA may be slightly different from that observed in the crystals. It is noted that the RA binding cavity is much larger than necessary to accommodate the ligand, and a large portion of the deep binding cavity is not occupied by RA. It appears that the β-ionone ring of the bound RA is well fixed, but the isoprene tail and carboxyl group have room to move. The guanidino group of Arg-111 could be closer to the carboxyl group of RA than to that of Arg-132 in solution. Solution NMR study of holo-CRABP-II is in progress.

CRABPs have very stringent retinoid specificity (2). The proteins bind only RA and reject both retinol and retinal. On the other hand, cellular retinol-binding proteins bind both retinol and retinal and reject RA. Amino acid sequence analysis reveals that the two conserved arginine residues in CRABPs are replaced with glutamine in cellular retinol-binding proteins. It has long been thought that the electrostatic interactions between the guanidino groups of the two conserved arginine residues and the carboxyl group of RA play major roles in binding of RA (31). Our mutagenesis studies indicate that Arg-111 and Arg-132 are indeed involved in RA binding, but the electrostatic interactions between the guanidino groups of Arg-111 and Arg-132 and the carboxyl group of RA contribute to the overall binding energy only by ~2.2 and 1.2 kcal/mol. Thus, amino acid residues that interact with the hydrophobic moiety of RA may be more important for binding of RA.

[3] Ji, X., L. Wang, Y. Li, and H. Yan, unpublished results.

Acknowledgment—We are indebted to Dr. Anders Åström for providing the cDNA clones of human CRABP-I and CRABP-II. We thank Drs. David G. McConnell and Yanling Zhang for critical reading of the manuscript and Dr. Xinhua Ji for useful discussion.

REFERENCES

1. Mangelsdorf, D. J., Umesono, K., and Evans, R. M. (1994) in *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Rorberts, A. B., and Goodman, D. S., eds) pp 319–349, Raven Press, New York
2. Ong, D. E., Newcomer, M. E., and Chytil, F. (1994) in *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Rorberts, A. B., and Goodman, D. S., eds) pp 283–317, Raven Press, New York
3. Bailey, J. S., and Siu, C.-H. (1988) *J. Biol. Chem.* **263**, 9326–9332
4. Fogh, K., Voorhees, J. J., and Åström (1993) *Arch. Biochem. Biophys.* **300**, 751–755
5. Fiorella, P. D., Giguère, V., and Napoli, J. L. (1993) *J. Biol. Chem.* **268**, 21545–21552
6. Norris, A. W., Cheng, L., Giguère, V., Rosenberger, M., and Li, E. (1994) *Biochim. Biophys. Acta* **1209**, 10–18
7. Giguère, V., Lyn, S., Yip, P., Siu, C.-H., and Amin, S. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 6233–6237
8. Åström, A., Tavakkol, A., Pettersson, U., Cromie, M., Elder, J. T., and Voorhees, J. J. (1991) *J. Biol. Chem.* **266**, 17662–17666
9. Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S. W., and Jones, T. A. (1994) *Adv. Protein Chem.* **45**, 89–151
10. Kleywegt, G. J., Bergfors, T., Senn, H., Le Motte, P., Gsell, B., Shudo, K., and Jones, T. A. (1994) *Structure* **2**, 1241–1258
11. Thompson, J. R., Bratt, J. M., and Banaszak, J. (1995) *J. Mol. Biol.* **252**, 433–446
12. Scapin, G., Gordon, J. I., and Sacchettini, J. C. (1992) *J. Biol. Chem.* **267**, 4253–4269
13. Zhang, J., Liu, Z.-P., Jones, T. A., Gierasch, L. M., and Sambrook, J. F. (1992) *Proteins Struct. Funct. Genet.* **13**, 87–89
14. Chen, L. X., Zhang, Z.-P., Scafonas, A., Cavalli, R. C., Gabriel, J. L., Soprano, K. J., and Soprano, D. R. (1995) *J. Biol. Chem.* **270**, 4518–4525
15. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1987) *Methods Enzymol.* **1985**, 60–89
16. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 488–492
17. Cogan, U., Kopelman, M., Mokady, S., and Shinitzky, M. (1976) *Eur. J. Biochem.* **65**, 71–78
18. Gill, S. C., and Hippel, P. T. (1989) *Anal. Biochem.* **182**, 319–326
19. Piantini, U., Sorensen, O. W., and Ernst, R. P. (1982) *J. Am. Chem. Soc.* **104**, 6800–6801
20. Braunschweiler, L., and Ernst. R. P. (1983) *J. Magn. Reson.* **53**, 5521–5528
21. Bax, A., and Davis, D. V. (1985) *J. Magn. Reson.* **65**, 355–360
22. Griesinger, C., Otting, G., Wuthrich, K., and Ernst, R. P. (1988) *J. Am. Chem. Soc.* 110, 7870–7872
23. Jenneer, J., and Ernst, R. P. (1979) *J. Chem. Phys.* **71**, 4546–4553
24. Macura, S., and Ernst, R. P. (1980) *Mol. Phys.* **41**, 95–117
25. Tsai., M.-D., and Yan, H. (1991) *Biochemistry* **30**, 6806–6818
26. Szuts, E. Z., and Harosi, F. I. (1991) *Arch. Biochem. Biophys.* **287**, 297–304
27. Daly, A. K., and Redfern, C. P. F. (1988) *Biochem. Biophys. Acta* **965**, 118–126
28. Fiorella, P. D., and Napoli, J. L. (1991) *J. Biol. Chem.* **266**, 16572–16579
29. Ong, D. E., and Chytil, F. (1978) *J. Biol. Chem.* **253**, 4551–4554
30. Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1993) *J. Mol. Biol.* **230**, 1225–1246
31. Jones, T. A., Bergfors, T., Sedzik, J., and Unge, T. (1988) *EMBO J.* **7**, 1597–1604

CHAPTER 3

Solution Structure of Type-II Human Cellular Retinoic Acid Binding Protein by NMR

Introduction

One of the key questions concerning the family of iLBPs including CRABPs is how the ligands get in and out of their respective binding pockets, since they are encapsulated by the β-sheets and helix-turn-helix motif in most cases. The crystal structures of holo-CRABPI and holo-CRABPII (both in complex with all-*trans*-RA) have been determined (Kleywegt et al., 1994). In each structure, RA is buried deeply in the RA-binding pocket, with its carboxyl group interacting with the side-chains of two arginines and one tyrosine at the bottom of the pocket. The β-ionone ring which is fixed snugly at the entrance of the RA-binding pocket with only one edge of the ring accessible to the solvent, is forced to adopt an unusual *cis*-like conformation. It appears that RA cannot enter or exit the deep binding pocket in the absence of major conformational changes in the protein.

More recently, the crystal structure of apo-CRABPI has been solved (Thompson

More recently, the crystal structure of apo-CRABPI has been solved (Thompson et al., 1995). The ligand entrance of apo-CRABPI is slightly more open than that of holo-CRABPI and thus a bit more accessible to RA. Surprisingly, apo-CRABPI is a dimer in the crystalline state, held together by an intermolecular β-sheet. Dimerization has been suggested to be the mechanism that allows RA to enter or exit the RA-binding pocket and by which CRABPI interacts with RA-metabolizing enzymes such as cytochrome P450s. However, it is unclear whether apo-CRABPI is dimeric in solution. CRABPI has been studied by NMR (Rizo et al., 1994), but no solution NMR structure has been reported.

In this chapter, the solution structure of human apo-CRABPII determined by NMR spectroscopy is described. The sequential assignments of the ^{1}H, ^{13}C and ^{15}N resonances of apo-CRABPII were established by multinuclear multidimensional NMR experiments. The solution structure of apo-CRABPII was derived from 2,382 experimental NMR restraints. The solution structure of apo-CRABPII is similar to the crystal structure of holo-CRABPII, but significant conformational differences were observed, especially in the ligand entrance region. In comparison to holo-CRABPII, apo-CRABPII exhibits a concerted movement of the second helix, the βC-βD loop and the βE-βF loop, so that the ligand entrance of apo-CRABPII is greatly enlarged and readily accessible to RA. Furthermore, the ligand-binding pocket of apo-CRABPII is rather disordered. The results suggest that binding of RA induces significant changes in the conformation and dynamics of CRABPII.

Experimental Procedures

Sample preparation. [^{13}C] D_6-glucose, $^{15}NH_4Cl$, [^{15}N]-leucine and [^{15}N]-valine were purchased from Cambridge Isotopic Laboratories. Unlabeled amino acids and other compounds were purchased from Sigma (St. Louis, Missouri).

Human CRABPII gene was over-expressed and the unlabeled protein was purified as detailed in Chapter 2. The same procedure was used for isotopic labeling of CRABPII except for some minor modifications as described below. M9 medium was used for the uniform isotopic labeling of CRABPII with $^{15}NH_4Cl$ and [^{13}C] D_6-glucose as the sole nitrogen and carbon sources for ^{15}N- and ^{13}C-labeling, respectively. For ^{15}N-labeling of leucine and valine residues, a rich medium containing [^{15}N]-valine or [^{15}N]-leucine was used (Muchmore et al., 1989). The *E. coli* strain DL49PS pLysS (kindly provided by Dr. David M. LeMaster) was used for the selective isotopic labeling. D_2O samples were prepared by dissolving lyophilized unlabeled CRABPII in PBS buffer (20 mM sodium phosphate, 150 mM sodium chloride in 99.6% D_2O, pD 7.5, uncorrected). H_2O samples were prepared by directly concentrating CRABPII fractions eluted with the PBS buffer (pH 7.3) from a gel-filtration column at the final step of the purification procedure (Chapter 2). About 10% D_2O was added to the concentrated protein solutions. The protein concentrations of the NMR samples were ~2 mM.

NMR spectroscopy. All NMR spectra were acquired at 25°C. Typical carrier frequencies for various experiments were as follows: ^{1}H, 4.70 ppm; ^{15}N, 119.7 ppm; ^{13}CO, 175 ppm; $^{13}C^{\alpha}$, 54 ppm and $^{13}C^{\alpha/\beta}$, 43 ppm. Quadrature detection in the

nonacquisition dimensions was achieved by the hypercomplex method (States et al., 1982) for homonuclear experiments and by the States-TPPI method (Marion et al., 1989a) for heteronuclear experiments.

Homonuclear 2D spectra were recorded on a Varian INOVA-600 spectrometer. The acquisition times and numbers of complex points were as follows: DQF-COSY (Piantini et al., 1982) F1 61.0 ms, 512 and F2 243.8 ms, 2048 (128 transients); NOESY (Jeener et al., 1978; Macura & Ernst, 1980; Kumar et al., 1980) F1 38.1 ms, 320 and F2 243.8 ms, 2048 (64 transients); SSNOESY (Smallcombe, 1993) and a WATERGATE (Piotto et al., 1992) NOESY, F1 61.0 ms, 512; F2 243.8 ms, 2048 (64 transients and 150 ms mixing time); clean-TOCSY (Griesinger et al., 1988) F1 38.1 ms, 320 and F2 243.8 ms, 2048 (64 transients). Two TOCSY spectra were acquired, one with a mixing time of 30 ms and one of 50 ms. Two NOESY spectra were recorded, one with a mixing time of 100 ms and one of 150 ms. The data were processed with the program VNMR v. 5.2F (Varian Associates). A gaussian and minus LB combination was applied in F2 dimension and a ~75°-shifted sine-bell in F1 dimension in processing NOESY and TOCSY spectra. A ~30°-shifted sine-bell was applied in both F1 and F2 dimensions in processing of DQF-COSY data. The data sets were zero-filled to 4096 × 2048 real points for TOCSY and NOESY and 8192 × 4096 real points for DQF-COSY. In general, a five-order polynomial was applied for baseline correction in F2 dimension after Fourier transformation.

Heteronuclear double resonance NMR spectra of uniformly ^{15}N labeled protein were recorded with the following acquisition times and complex data points: 2D ^{1}H-^{15}N HSQC (Bodenhausen & Ruben, 1980; Kay et al., 1992; Zhang et al., 1994), ^{1}H (F2) 121.9

ms, 1024, ^{15}N (F1) 53.3 ms, 128 (32 transients); 3D ^{1}H-^{15}N NOESY-HSQC (Zhang et al., 1994), ^{1}H (F3) 141.5 ms, 1024, ^{15}N (F2) 34.0 ms, 68, ^{1}H (F1) 33.2 ms, 200 (16 transients, 150 ms mixing time). 2D ^{1}H-^{15}N HMQC spectra (Müeller, 1979; Bax et al., 1983) of the selectively labeled proteins ([^{15}N]-valine or [^{15}N]-leucine) were recorded with the following acquisition times and complex data points: ^{1}H (F2) 141.5 ms, 1024 and (F1) ^{15}N 32.0 ms, 64 (32 transients). Heteronuclear triple resonance NMR spectra (Kay et al., 1990; Grzesiek & Bax, 1992a; Muhandiram & Kay, 1994; Kay, 1995) of $^{15}N/^{13}C$-doubly labeled protein were recorded with the following acquisition times and complex data points: CT-HNCO (Grzesiek & Bax, 1992a), ^{15}N (F1) 16.0 ms, 32, ^{13}CO (F2) 34.0 ms, 64, ^{1}H (F3) 127.80 ms, 1024 (16 transients); CT-HNCA and CT-HN(CO)CA (Grzesiek & Bax, 1992a), ^{15}N (F1) 16.0 ms, 32, $^{13}C^{\alpha}$ (F2) 10.9 ms, 64, ^{1}H (F3) 127.80 ms, 1024 (16 transients); HNCACB (Wittekind & Müeller, 1993; Muhandiram & Kay, 1994), $^{13}C^{\alpha/\beta}$ (F1) 7.1 ms, 64, ^{15}N (F2) 15.2 ms, 32, ^{1}H (F3) 128.0 ms, 1024 (32 transients); CBCA(CO)NH (Grzesiek & Bax, 1992b; Muhandiram & Kay, 1994), $^{13}C^{\alpha/\beta}$ (F1) 6.4 ms, 64, ^{15}N (F2) 16.0 ms, 32, ^{1}H (F3) 127.80 ms, 1024 (32 transients); C(CO)NH (Grzesiek et al., 1993; Muhandiram & Kay, 1994), ^{13}C (F1) 9.7 ms, 96, ^{15}N (F2) 15.2 ms, 32, ^{1}H (F3) 128.0 ms, 1024 (32 transients); H(CCO)NH (Grzesiek et al., 1993; Muhandiram & Kay, 1994), ^{1}H (F1) 15.9 ms, 128, ^{15}N (F2) 25.0 ms, 50, ^{1}H (F3) 127.80 ms, 1024 (32 transients); gradient HCCH-TOCSY (Bax et al., 1991; Kay et al., 1993), ^{1}H (F1) 35.8 ms, 128, ^{13}C (F2), 10.5 ms, 75, ^{1}H (F3) 127.80 ms, 1024 (16 transients, 16 ms mixing time). A low power GARP-I sequence (Shaka et al., 1985) was applied for ^{15}N broad-band decoupling during the data acquisition. ^{1}H broadband decoupling in the CT-HNCO, CT-HNCA, CT-HN(CO)CA, C(CO)NH and H(CCO)NH experiments were achieved by a

DIPSI-2 sequence (Shaka et al., 1988), and in the CBCA(CO)NH experiment by a WALTZ-16 sequence (Shaka et al., 1983). The CO decoupling during the $C^{\alpha/\beta}$ chemical shift and $C^{\alpha/\beta}$-^{15}N scalar coupling evolution periods in the HNCACB experiment, and C^{α} decoupling during the constant time t2 (^{15}N) in the CBCA(CO)NH experiment were accomplished by applying a SEDUCE-1 sequence (McCoy & Müeller, 1992a, b). ^{13}C isotropic mixings in the 3D HCCH-TOCSY (Bax et al., 1991; Kay et al., 1993), H(CCO)NH and C(CO)NH experiments were accomplished by a DIPSI-3 sequence (Shaka et al., 1988). The HSQC, HMQC, HNCACB and C(CO)NH spectra were acquired on the INOVA-600 spectrometer and processed with the program VNMR v. 5.3 (Varian Associates). All other spectra were recorded on a Brüker DMX-500 spectrometer at the National Magnetic Resonance Facility at Madison and processed by Felix 95 (Biosym Technologies inc.). A ~45°-shifted sine-bell was applied in both dimensions of the 2D HSQC and HMQC data. The HSQC data was zero-filled to 2048 × 512 real points and the HMQC data to 2048 × 256 real points. A ~75°-shifted sine-bell was applied in both F1 and F3 dimensions, and a cosine-bell in the F2 dimension of the 3D NOESY-HSQC and HCCH-TOCSY data. For the other 3D data, a 75°-shifted sine-bell was applied in the F3 dimension and a cosine-bell in the other two dimensions. The first points of both F1 and F2 dimensions were calculated by linear predictions. The data size in indirect detection dimension was extended by backward-forward linear predication (Zhu & Bax, 1992). Solvent suppression was improved by convolution of time domain data (Marion et al., 1989b). The data size was doubled in each dimension by zero-filling once. The empty regions of the 3D data (the aliphatic parts in the F3 dimension of the ^{15}N-edited data and the amide part of the HCCH-TOCSY data) were removed to reduce the data size.

Derivation of structural restraints. Two types of structural restraints were derived from experimental NMR data: interproton distance restraints and hydrogen bond restraints. Approximate interproton distance restraints were derived from the NOE data. NOE cross peaks between aliphatic protons were picked from the homonuclear 2D NOESY spectrum with a mixing time of 100 ms using the program VNMR v. 5.2. NOE cross peaks involving amide protons were picked from the 3D ^{1}H-^{15}N NOESY-HSQC with a mixing time of 150 ms using the program Felix 95, except for those of Ile9, Ile10, Ile63, Arg132, Val133 and Tyr134, which were picked from the 2D homonuclear NOESY data. The integrated peak volumes were converted into approximate interproton distances by normalizing them against the calibrated volumes of NOE cross-peaks between backbone protons within the identified β-sheet regions. The upper limits of the interproton distances were calibrated according to the equation $V_a = V_b (r_b/r_a)^6$, here V_a, V_b were the volumes and r_a, r_b the internuclear distances. The distance bounds were set to 1.8–2.7 Å (1.8–2.9 Å for NOE cross peaks involving amide protons), 1.8–3.3 Å (1.8–3.5 Å for NOE cross peaks involving amide protons) and 1.8–5.0 Å corresponding to strong, medium and weak NOEs, respectively. Pseudoatom corrections were made for non-stereospecifically assigned methylene and methyl resonances (Wüthrich et al., 1983). An additional 0.5 Å was added to the upper bounds for methyl protons.

Protein backbone hydrogen bonds were derived from patterns of cross peaks remaining in the fingerprint regions of the homonuclear 2D TOCSY and NOESY spectra recorded in D_2O. Hydrogen bond restraints were set to 2.0 ± 0.5 Å and 3.0 ± 0.3 Å for H^N···O and N···O, respectively.

Structure calculation. The structures were calculated by the hybrid distance geometry-simulated annealing (DGSA) protocol (Nilges et al., 1988) in the program X-PLOR v. 3.1 (Brünger, 1992) on an SGI IndigoII workstation. A square-well potential function with a force constant of 50 $kcal.mol^{-1}.Å^{-2}$ was applied for the distance restraints. The X-PLOR f_{repel} function was used to simulate van der Waals interaction with atomic radii set to 0.80 times their CHARMM values (Brooks et al., 1983). Hydrogen bond restraints within the regions of regular secondary structures were introduced at a later stage of structural refinement. Seventy structures were generated using this protocol at the beginning. The structures were inspected with the programs QUANTA96 (Molecular Simulations. inc.) and InsightII (Biosym Technologies inc.), and analyzed by PROCHECK–NMR (version 3.4.4) (Laskowski et al., 1993; 1997). An iterative strategy was used for the structural refinement. The first round calculation employed mainly NOEs between backbone protons. Then the newly computed NMR structures were employed to assign more NOE restraints, to correct wrong assignments and to loosen NOE distance bounds if spectral overlapping was deduced. Then another round of structural refinement was carried out with the modified NMR restraints. After several rounds of such refinement, an ensemble of 25 structures was selected according to their best-fits to the experimental NMR restraints and the low values of their total energies.

Results

Sequential resonance assignment. HNCACB and CBCA(CO)NH spectra were first analyzed to obtain sequential connectivities. Representative results are shown in Figure 3.1. About one hundred residues with relatively strong peaks in the 2D ^{1}H-^{15}N HSQC (Figure 3.2) spectrum could be easily linked on the basis of these two sets of data. These residues were then sequentially assigned by the combined analyses of the 3D C(CO)NH, H(CCO)NH, HCCH-TOCSY, and homonuclear 2D TOCSY spectra.

The majority of the remaining residues had weak or no detectable peaks in the 2D ^{1}H-^{15}N HSQC spectrum, even with minimum water perturbation using the water flip-back technique (Grzesiek & Bax, 1993). Because these residues also had weak or no detectable peaks in the triple resonance experiments, their sequential assignment posed a big challenge. The following residues had weak cross-peaks in the 2D ^{1}H-^{15}N HSQC spectrum: Lys8–Glu17, Leu28, Lys30, Ala32–Ala35, Ala40, Val41, Glu62, Ile63, Leu119, Leu121 and Thr131–Val135. They were assigned by spin system identifications combined with sequential and long range (inter-strand) NOEs. Because of the excellent chemical shift dispersion of the protein, 2D homonuclear TOCSY and NOESY data were especially useful for assigning these residues. In particular, most of the inter-strand NOE cross-peaks between H^{α}-H^{α} protons of the adjacent β-strands were well resolved in the 2D NOESY spectra. The sequential assignments of many of these residues were also facilitated by the comparisons with the sequential assignments of the wild type holo-CRABPII in complex with all-*trans*-RA (Chapter 5) and the apo-form of the site-directed

Figure 3.1. Strips of the HNCACB (a) and CBCA(CO)NH (b) spectra showing sequential connectivities for residues Gly68–Gln74. Both inter-residue and intra-residue peaks in the HNCACB strips (a) are labeled. The dotted lines indicate negative peaks.

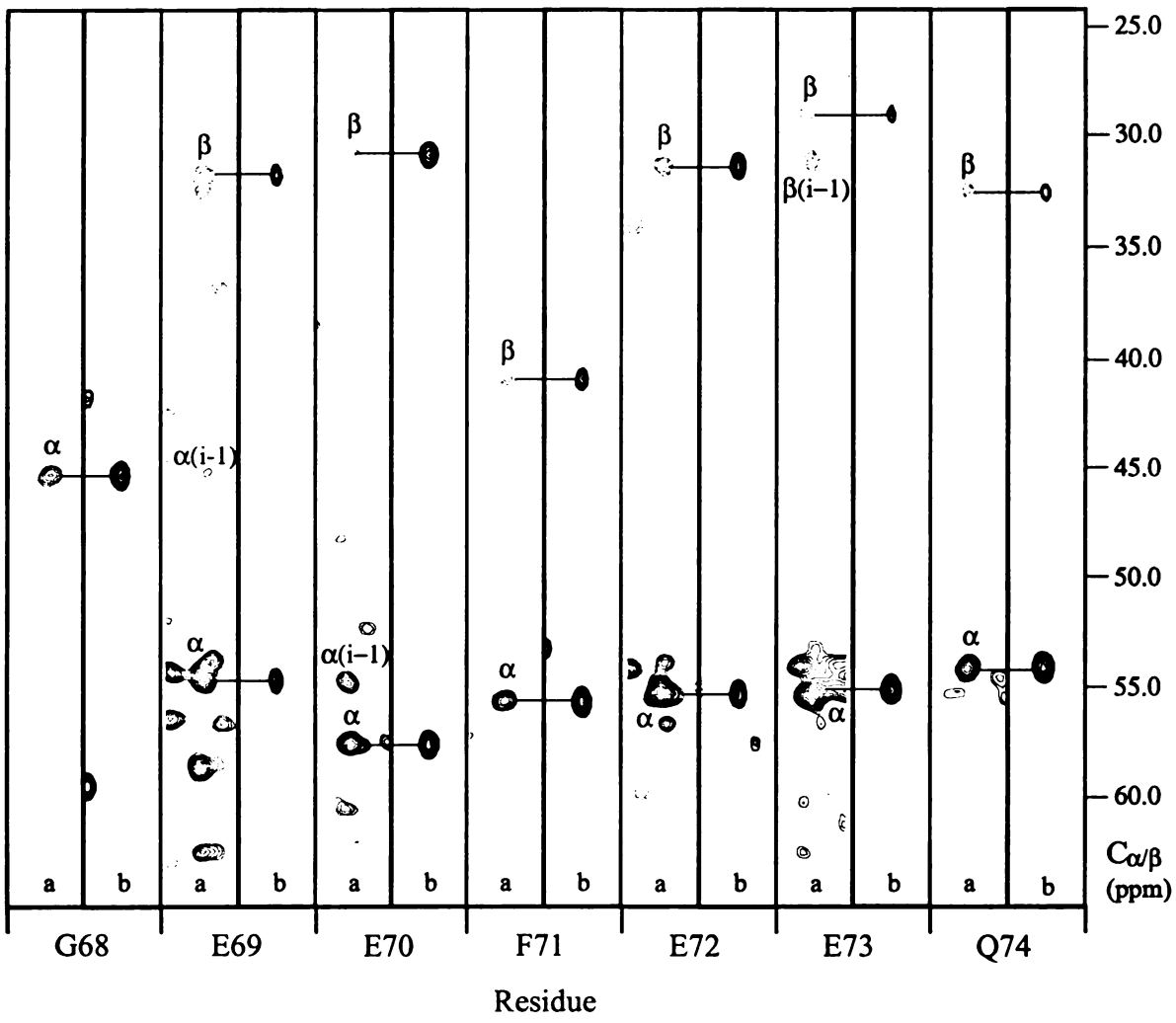
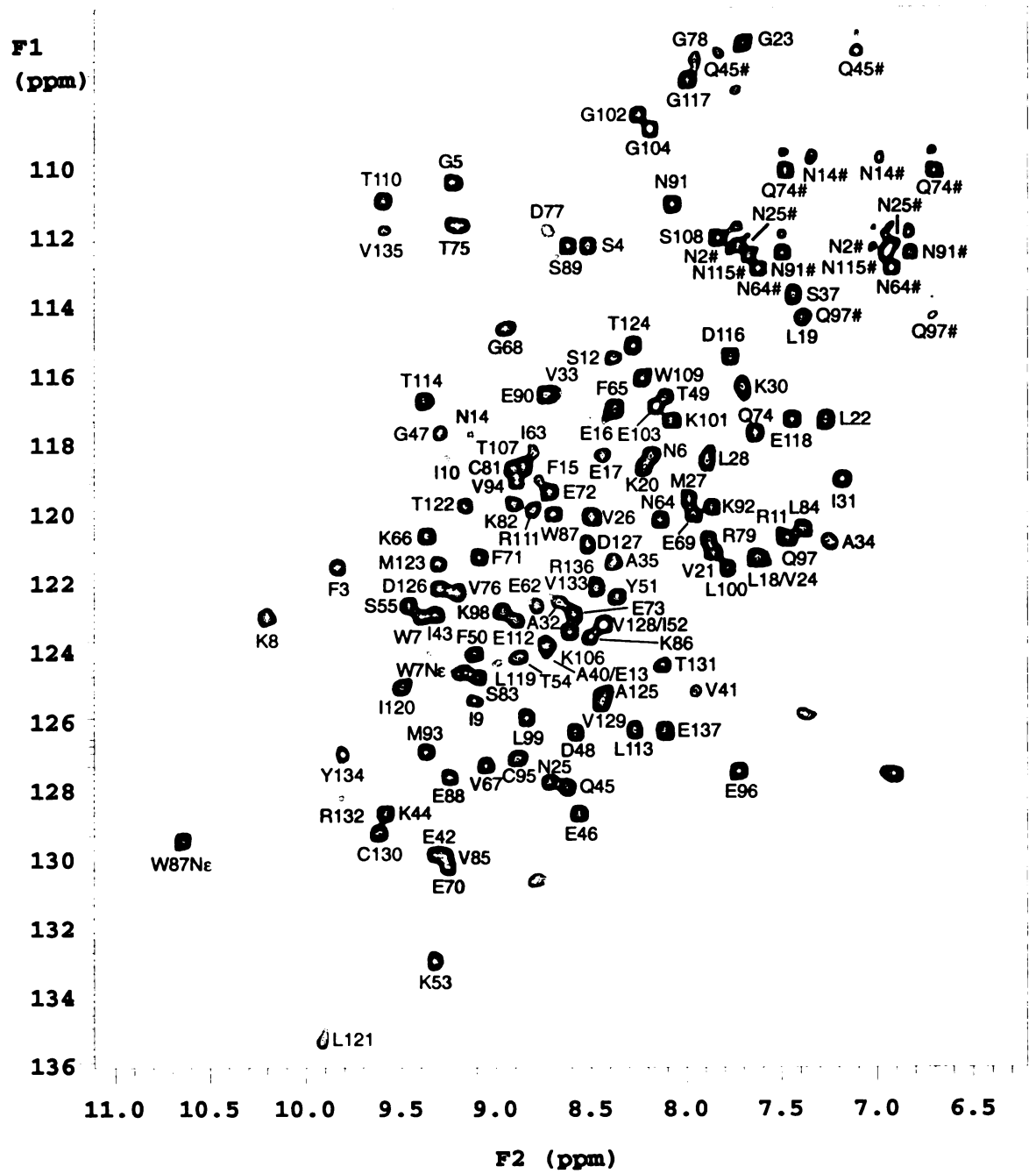

25.0
30.0
35.0
40.0
45.0
50.0
55.0
60.0
Cα/β (ppm)
β
α
α(i-1)
β(i−1)
a
b
G68
E69
E70
F71
E72
E73
Q74
Residue

Figure 3.2. Gradient- and sensitivity-enhanced 2D ^{1}H–^{15}N HSQC spectrum of uniformly ^{15}N-enriched apo-CRABPII. Sequential assignments are indicated with one-letter amino acid code and residue number. The side-chain amides of Asn and Gln are indicated by #. The unlabeled peaks are most probably from the side-chains of arginines.


F1
(ppm)
110
112
114
116
118
120
122
124
126
128
130
132
134
136
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
F2 (ppm)
G78
G23
Q45#
Q45#
G117
G102
G104
G5
T110
N91
Q74#
N14#
N14#
Q74#
D77
N25#
N25#
V135
T75
S108
S4
N2#
N2#
N91#
S89
N115#
N91#
N115#
N64#
S37
N64#
Q97#
Q97#
T124
D116
G68
L19
S12
V33
W109
T114
F65
K30
E90
T49
N14
K101
Q74
L22
G47
I63
E16
E103
E118
T107
N6
L28
I10
C81
F15
E17
K20
M27
V94
E72
I31
T122
N64
K92
L84
K82
V26
R11
K66
R111
W87
D127
E69
R79
A34
M123
F71
R136
A35
Q97
V21
F3
D126
V76
E62
V133
Y51
L100
L18/V24
S55
E73
K98
A32
V128/I52
K8
W7
I43
F50
E112
K86
K106
T131
W7Nε
L119
T54
A40/E13
S83
A125
V41
I120
I9
V129
E137
M93
L99
D48
L113
Y134
N25
C95
V67
Q45
E96
E88
R132
K44
E46
E42
C130
V85
W87Nε
E70
K53
L121

mutant R111M (Chapter 4). Most weak peaks in the 2D ^{1}H-^{15}N HSQC spectrum of wild type apo-CRABPII became much stronger upon the binding of RA or with the point mutation, and thus could be assigned unambiguously for holo-CRABPII and apo-R111M. Furthermore, neither the binding of RA nor the mutation changed the chemical shifts of the majority of residues. Such comparisons were especially helpful for the assignments of the selectively ^{15}N-labeled residues such as Leu119, Val33, Val41, Val76, Val133 and Leu121. Such comparisons also expedited greatly the assignments of the forty-six residues with amide protons remaining in the fingerprint regions of the homonuclear 2D spectra in D_2O (Figure 3.3). Almost exactly the same residues remained in the 2D homonuclear D_2O spectra of apo-CRABPII and apo-R111M (Figure 4.8 of Chapter 4). In particular, such comparisons helped the assignments of residues Lys8, Ile9, Ile10, Ser12, Ile63, Thr131, Arg132 and Tyr134. Nine residues, namely Glu13, Asn14, Phe15, Lys30, Ile31, Ala40, Asp77, Gly78 and Cys130, were assigned based on NOEs and with the help of the comparisons. The H^N and ^{15}N resonance assignments of Val33 were based on selective labeling, spin system identification and several NOEs between the ring H^{δ}s of Phe15 and H^{γ}s of Val33 after the first round of structural calculation. The H^N and ^{15}N resonances of Ala32, Ala34 and Ala35 were assigned based on the comparisons and a few weak, sequential NOEs. The H^N or ^{15}N resonance of Ser37 was assigned tentatively by spin system identification and the comparisons of the three sets of 2D ^{1}H-^{15}N HSQC spectra. The peaks corresponding to the two adjacent residues, Ala36 and Lys38, could not be observed in the 2D ^{1}H-^{15}N HSQC spectrum of apo-CRABPII. The aliphatic protons of the residues with weak peaks in the 2D ^{1}H-^{15}N HSQC spectrum were assigned by the inter-strand NOEs, by spin system identifications obtained from the homonuclear

2D spectra and 3D HCCH-TOCSY spectrum, and by the comparisons with the assignments of holo-CRABPII and apo-R111M. For examples, the spin systems of Ser12, Ala40, Val41, Thr54, Glu62 and Val76 could be easily identified. The H^{α} of Arg132 was assigned by its inter-strand NOE to Ser12 observed in the homonuclear 2D NOESY spectrum.

The following eleven residues were not observed in the 2D ^{1}H-^{15}N HSQC spectrum: Asn2, Arg29, Ala36, Lys38, Thr56–Thr61 and Asn115. They were first grouped into several spin systems based on the homonuclear 2D DQF-COSY, TOCSY and 3D HCCH-TOCSY experiments. The aliphatic protons of Asn2, Thr56, Thr57, Val58, Thr60, Thr61 and Asn115 were assigned by spin system identifications, the inter-strand NOEs, and by the comparisons with the assignments of holo-CRABPII and apo-R111M. The assignments of the aliphatic protons of Thr56 and Thr57 were further helped by the similar but unique chemical shifts of their H^{α} and H^{β} resonances in the spectra of apo- and holo-CRABPII and apo-R111M: H^{α} 4.83 ppm and H^{β} 4.61 ppm for Thr56, and H^{α} 4.00 ppm and H^{β} 4.23 ppm for Thr57 (in apo-R111M). Two possible alanine peaks from Ala32 and Ala36 could be identified in the 2D homonuclear and 3D HCCH-TOCSY spectra, but their sequential assignments could not be obtained due to the lack of strong, well-resolved sequential NOEs. None of the ^{1}H, ^{13}C and ^{15}N resonances of Arg29, Lys38, Pro39 and Arg59 were assigned (Table 3.1).

The ring protons of aromatic residues, except for Phe15 and Tyr134, could be easily assigned by analysis of the 2D DQFCOSY and TOCSY data, and by the comparisons with the assignments of holo-CRABPII and apo-R111M (Chapters 4, 5). Their identities were confirmed later by sequential assignment and the NOEs between H^{β}-

Figure 3.3. The fingerprint region of the homonuclear 2D TOCSY spectrum with 40 ms mixing time. The TOCSY experiment was initiated more than 24 h after dissolving a lyophilized protein sample in PBS, pD 7.5, in D_2O.

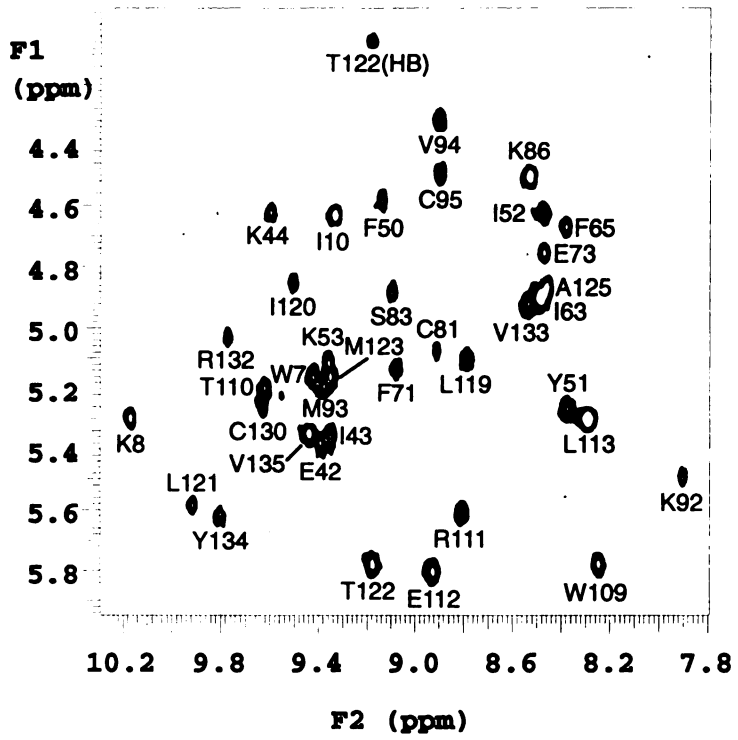

F1
(ppm)
4.4
4.6
4.8
5.0
5.2
5.4
5.6
5.8
T122(HB)
V94
K86
C95
I52
F65
K44
I10
F50
E73
A125
I120
S83
I63
C81
V133
K53
M123
R132
T110
W7
F71
L119
Y51
M93
K8
C130
I43
L113
V135
E42
L121
K92
Y134
R111
T122
E112
W109
10.2
9.8
9.4
9.0
8.6
8.2
7.8
F2 (ppm)

Table 3.1. ^{1}H, ^{15}N and ^{13}C assignments of apo-CRABPII in PBS at 25°C.

Residue	^{15}N	CO	C^{α}	C^{β}	Others
P1			63.0 (4.32)	33.2 (2.37, 1.91)	H^{γ} 2.10, 2.08; H^{δ} 3.32, 3.25
N2		171.9	52.7 (5.03)	39.3 (2.95, 2.89)	
F3	121.9 (9.82)	173.5	60.4 (4.32)	32.3 (3.28, 2.70)	2, 6H 7.42; 3, 5H 7.16; 4H 7.34
S4	112.5 (8.49)	171.8	60.7 (4.70)	64.5 (4.30, 3.99)	
G5	110.6 (9.22)	167.9	44.9 (3.96, 3.76)		
N6	118.6 (8.15)	171.5	53.0 (5.52)	40.2 (2.69, 2.52)	
W7	123.3 (9.35)	172.5	56.7 (5.14)	32.2 (3.06, 2.66)	2H 6.78; 4H 7.07; 5H 6.82; 6H 7.23; 7H 7.12 H^{ε} 9.20; N^{ε} 124.6
K8	123.2 (10.17)	172.4	54.1 (5.28)	36.1 (1.90, 1.85)	H^{γ} 1.63, 1.52; H^{δ} 1.37, 1.33; H^{ε} 2.85
I9	125.6 (9.05)	173.3	55.5 (3.92)	33.6 (1.76)	H^{γ} 1.65, 1.56; H^{δ} 0.48; $H^{\gamma m}$ 1.39
I10	118.8 (9.24)	173.3	61.3 (4.61)	41.8 (1.83)	$H^{\gamma m}$ 0.87
R11	120.8 (7.57)				
S12	115.8 (8.33)	171.0	62.8 (5.12)	(3.61, 3.43)	
E13	124.0 (8.71)		61.7 (4.80)	40.1 (2.19, 2.08)	H^{γ} 1.93
N14	117.8 (9.25)		55.7 (4.95)	41.4 (3.31, 3.22)	H^{δ} 7.35, 7.00; N^{δ} 109.8
F15	119.3 (8.72)		55.3 (4.59)	33.7 (2.98, 2.85)	2, 6H 7.14; 3, 5H 6.84; 4H 6.93
E16	117.4 (8.36)	172.1	63.2 (3.30)	32.6 (1.96)	H^{γ} 2.16
E17	118.5 (8.36)	175.9	59.6 (3.68)	28.1 (1.84, 1.73)	H^{γ} 2.38, 2.19
L18	121.4 (7.59)	174.9	58.5 (3.73)	42.4 (2.17, 1.34)	H^{γ} 0.81; H^{δ} 0.74, 0.70
L19	114.5 (7.36)	176.0	57.6 (3.62)	41.1 (1.62, 1.56)	H^{γ} 1.12; H^{δ} 0.36, 0.32
K20	118.9 (8.18)	178.3	60.5 (3.83)	32.6 (1.93, 1.61)	H^{γ} 1.35, 1.14; H^{ε} 2.91
V21	121.3 (7.83)	174.9	66.0 (3.78)	31.7 (2.15)	H^{γ} 1.10, 0.95
L22	117.5 (7.24)	173.6	55.4 (3.91)	41.5 (1.39)	H^{δ} 0.49
G23	106.5 (7.69)	172.0	45.9 (3.96, 3.58)		
V24	121.4 (7.62)	173.6	63.3 (3.72)	31.7 (1.45)	H^{γ} 0.83
N25	127.9 (8.68)	172.8	54.2 (4.38)	39.3 (3.09, 2.74)	H^{δ} 7.73, 6.91; N^{δ} 112.3
V26	120.3 (8.47)	173.4	66.7 (3.46)	32.1 (2.03)	H^{γ} 0.96, 0.91
M27	119.7 (7.95)		58.5 (4.14)	32.0 (2.07, 2.01)	H^{γ} 2.59, 2.46
L28	118.6 (7.82)		(4.08)	(1.63, 1.32)	H^{γ} 1.09; H^{δ} 0.82, 0.76
R29			(4.32)	(2.20, 1.97)	H^{γ} 1.74; H^{δ} 2.70
K30	116.5 (7.66)	177.7	59.6 (3.87)	32.2 (1.89, 1.81)	H^{γ} 1.68, 1.59
I31	119.2 (7.16)		64.3 (3.72)	39.2 (1.80)	$H^{\gamma m}$ 0.88; H^{δ} 0.80
A32	122.6 (8.54)				
V33	116.7 (8.66)		66.6 (3.62)	(2.07)	H^{γ} 0.94
A34	121.0 (7.31)		(4.08)	(1.38)	
A35	121.6 (8.44)		(4.21)	(1.38)	
A36		174.8	52.2	19.1	
S37	113.9 (7.43)		60.4 (4.70)	(4.23, 4.02)	
K38					
P39					
A40	123.9 (8.71)	174.4	51.9 (4.96)	20.9 (1.45)	
V41	125.2 (7.93)	171.5	61.2 (5.13)	35.3 (1.73)	H^{γ} 0.63, 0.46
E42	130.0 (9.30)	172.1	54.6 (5.36)	33.4 (2.06, 1.99)	H^{γ} 2.25
I43	123.1 (9.30)	173.9	60.0 (5.34)	40.8 (2.27)	H^{γ} 1.66; $H^{\gamma m}$ 0.85; H^{δ} 0.51
K44	128.9 (9.56)	170.8	56.4 (4.60)	35.3 (1.87, 1.76)	H^{γ} 1.54, 1.41; H^{δ} 1.33, 1.24
Q45	128.0 (8.58)	172.2	53.9 (4.45)	30.6 (1.82, 1.71)	H^{γ} 2.03; H^{ε} 7.83, 7.11; N^{ε} 106.8
E46	128.9 (8.53)	173.5	55.1 (4.37)	31.0 (1.62, 1.26)	H^{γ} 2.02
G47	117.9 (9.10)	170.6	48.0 (3.94, 3.58)		
D48	126.6 (8.54)	172.5	54.5 (5.18)	42.1 (3.15, 2.68)	
T49	116.7 (8.08)	170.1	63.0 (4.70)	69.6 (4.14)	H^{γ} 0.99
F50	124.2 (9.08)	173.1	57.6 (4.56)	44.6 (1.65)	2, 6H 6.48; 3, 5H 7.08; 4H 6.81
Y51	122.5 (8.35)	172.8	56.3 (5.24)	41.0 (3.01, 2.70)	2, 6H 6.94; 3, 5H 6.61
I52	123.3 (8.41)	171.6	61.0 (4.60)	40.4 (1.84)	H^{γ} 1.26; $H^{\gamma m}$ 0.80; H^{δ} 0.31
K53	133.1 (9.30)	172.0	54.5 (5.08)	34.9 (1.76, 1.58)	H^{γ} 1.17; H^{δ} 0.44, 0.29; H^{ε} 2.84
T54	124.4 (8.85)		62.1 (4.88)	70.0 (3.86)	H^{γ} 0.94
S55	122.7 (9.43)		58.0 (5.30)	65.0 (3.83, 3.74)	
T56	124.0 (8.71)		(4.85)	(4.61)	H^{γ} 1.07
T57		173.5	64.9 (4.00)	66.7 (4.23)	H^{γ} 1.27
V58			62.2 (4.16)	(1.96)	H^{γ} 0.79, 0.74

Residue	^{15}N	CO	C^{α}	C^{β}	Others
R59					
T60	123.5 (8.58)		60.3 (5.20)	71.7 (3.90)	H^{γ} 1.06
T61			60.1 (4.61)	71.7 (3.97)	H^{γ} 1.06
E62	122.9 (8.73)		(4.98)	(2.02, 1.91)	
I63		170.5	59.5 (4.88)	42.2 (1.85)	H^{γ} 1.16, 1.01; H^{m} 1.09; H^{δ} 0.80
N64	120.4 (8.10)	169.7	52.3 (5.56)	42.8 (2.54, 2.50)	H^{δ} 7.62, 6.92; N^{δ} 113.1
F65	117.2 (8.34)	169.1	55.9 (4.65)	39.8 (2.15)	2, 6H 6.24; 3, 5H 6.52; 4H 6.32
K66	120.8 (9.32)	175.2	53.9 (4.65)	35.4 (1.63, 1.49)	H^{γ} 1.45, 1.28; H^{ϵ} 2.92
V67	127.5 (9.01)	174.1	67.1 (3.23)	31.9 (1.85)	H_{γ} 0.64, 0.43
G68	114.9 (8.91)	170.7	45.7 (4.38, 3.61)		
E69	120.2 (7.94)	171.5	54.9 (4.77)	35.3 (2.16, 1.96)	H^{γ} 2.19
E70	130.3 (9.21)	172.6	57.6 (4.88)	31.4 (1.99, 1.91)	H^{γ} 2.20, 2.13
F71	121.4 (9.06)	169.5	55.8 (5.11)	41.4 (3.53, 3.20)	2, 6H 7.32; 3, 5H 6.82; 4H 6.52
E72	119.6 (8.67)	173.1	55.3 (5.05)	31.9 (2.02,1.91)	H^{γ} 2.30, 2.13
E73	122.4 (8.44)	171.6	55.3 (4.73)	29.7 (2.02,1.90)	H^{γ} 2.14, 2.14
Q74	117.9 (7.59)	174.3	54.2 (5.30)	33.1 (1.73, 1.70)	H^{γ} 2.17, 1.99; H^{ϵ} 7.49, 6.70; N^{ϵ} 110.3
T75	111.8 (9.18)	175.4	62.4 (4.55)	71.2 (4.16)	H^{γ} 1.03
V76	122.4 (9.18)		65.6 (3.82)	32.1(2.00)	H^{γ} 1.01, 0.81
D77	112.0 (8.54)				
G78	107.2 (7.93)		45.9 (4,10, 3.55)		
R79	121.0 (7.86)	171.4	54.0	29. 8	
P80		173.6	63.8 (4.66)	33.0 (2.38, 1.86)	H^{γ} 2.15, 1.99
C81	118.8 (8.87)	168.5	55.9 (5.05)	31.6 (2.58)	
K82	120.0 (8.87)	173.0	54.5 (4.73)	34.3 (1.54, 1.47)	H^{γ} 1.26, 1.22; H^{δ} 1.03
S83	124.9 (9.06)	168.7	58.7 (4.86)	68.0 (2.78, 1.54)	
L84	120.7 (7.36)	173.7	54.9 (4.23)	45.2 (2.00, 1.58)	H^{γ} 1.17; H^{δ} 0.68, 0.51
V85	130.1 (9.20)	171.5	60.6 (4.35)	32.6 (1.35)	H^{γ} 0.23, 0.02
K86	123.8 (8.47)	173.3	54.2 (4.48)	36.9 (1.61, 1.43)	H^{γ} 1.26, 1.12
W87	120.2 (8.65)	174.9	57.0 (5.11)	29.5 (3.22, 2.99)	2H 7.30; 4H 7.98; 5H 6.95; 6H 7.21; 7H 7.62 H^{ϵ} 10.66; N^{ϵ} 129.6
E88	127.7 (9.21)	173.3	58.0 (4.35)	32.3 (1.80, 1.66)	H^{γ} 1.99
S89	112.5 (8.59)	171.2	57.0 (4.55)	65.0 (4.11, 3.98)	
E90	116.7 (8.71)	173.1	59.1 (3.66)	29.5 (1.84, 1.78)	H^{γ} 2.08
N91	111.2 (8.05)	171.3	52.9 (4.96)	40.4 (3.15, 2.85)	H^{δ} 7.50, 6.83; N^{δ} 112.6
K92	120.1 (7.84)	171.0	56.7 (5.46)	37.2 (1.98, 1.47)	H^{γ} 1.22; H^{δ} 1.10; H^{ϵ} 2.68
M93	127.0 (9.35)	171.0	53.4 (5.18)	35.1 (1.99)	H^{γ} 2.19, 2.08; H^{ϵ} 0.41
V94	119.3 (8.86)	170.0	60.3 (4.30)	34.9 (1.77)	H^{γ} 0.73, 0.70
C95	127.3 (8.85)	171.4	55.6 (4.48)	27.9 (0.36, –0.28)	
E96	127.7 (7.70)	172.6	55.2 (4.31)	31.6 (1.81, 1.73)	H^{γ} 2.14, 1.95
Q97	121.0 (7.43)	173.1	54.7 (4.91)	30.8 (1.75)	H^{γ} 2.43, 2.02; H^{ϵ} 7.38, 6.71; N^{ϵ} 114.4
K98	123.1 (8.94)	172.9	54.7 (4.60)	35.0 (2.10)	H^{γ} 1.96, 1.78; H^{δ} 1.24; H^{ϵ} 3.16, 2.84
L99	126.1 (8.78)	175.4	56.0 (4.00)	42.0 (1.98, 1.63)	H^{γ} 1.36; H^{δ} 0.86, 0.65
L100	121.9 (7.77)	174.8	57.0 (4.07)	42.5 (1.64, 1.47)	H^{γ} 1.35; H^{δ} 0.80, 0.76
K101	117.4 (8.05)	172.7	55.4 (4.37)	34.9 (1.72, 1.61)	H^{γ} 1.37, 1.27
G102	108.6 (8.18)	169.7	45.1 (4.02, 3.71)		
E103	117.2 (8.10)	173.1	54.9 (4.60)	32.9 (1.96, 1.78)	H^{γ} 2.10
G104	109.1 (8.14)		45.5 (4.10, 3.90)		
P105		172.8	62.7 (4.48)	32.7 (2.03, 1.33)	H^{γ} 1.84, 1.76; H^{δ} 3.60, 3.29
K106	123.5 (8.58)	175.1	56.6 (4.38)	32.6 (1.81, 1.51)	H^{γ} 1.35, 1.30;
T107	118.7 (8.79)	170.8	60.1 (5.36)	72.5 (4.46)	H^{γ} 1.13
S108	112.3 (7.80)	171.2	57.3 (4.63)	66.8 (3.86, 3.78)	
W109	116.3 (8.21)	171.8	56.1 (5.76)	33.1 (3.48, 3.29)	2H 6.89; 4H 6.90; 5H 6.35; 6H 6.43; 7H 6.59 H^{ϵ} 11.37; N^{ϵ} 137.8
T110	111.2 (9.56)	171.7	61.0 (5.24)	72.2 (4.16)	H^{γ} 1.09
R111	120.1 (8.79)	170.9	56.1 (5.57)	35.6 (1.85, 1.50)	H^{γ} 1.35, 1.23; H^{δ} 3.00
E112	123.3 (8.87)	170.5	53.9 (5.78)	34.9 (1.94)	H^{γ} 2.03
L113	126.4 (8.25)	175.1	52.8 (5.28)	43.9 (1.25)	H^{γ} 0.94; H^{δ} 0.51, 0.44
T114	116.9 (9.35)		(4.76)	(4.70)	H_{γ} 1.25
N115		173.1	55.5 (4.44)	38.1 (2.79, 2.76)	H^{δ} 7.67, 6.95; N^{δ} 112.6
D116	115.7 (7.75)	173.5	54.3 (4.74)	41.2 (2.79, 2.39)	
G117	107.6 (7.96)	171.7	46.1 (4.17, 3.69)		
E118	117.5 (7.43)	172.0	(4.98)	(2.02, 1.98)	

Residue	^{15}N	CO	C^{α}	C^{β}	Others
L119	124.6 (8.73)	172.6	53.4 (5.08)	43.3 (0.89)	H^{γ} 0.44; H^{δ} 0.18, –0.36
I120	125.2 (9.47)		60.8 (4.83)	39.4 (1.77)	H^{γ} 1.35; $H^{\gamma m}$ 0.75; H^{δ} 0.99
L121	135.5 (9.88)	172.8	53.3 (5.56)	45.3 (2.15)	H^{γ} 1.22; H^{δ} 0.85, 0.74
T122	120.1 (9.13)	171.4	60.4 (5.76)	71.0 (4.03)	H^{γ} 1.06
M123	121.7 (9.29)	171.5	55.1 (5.14)	38.3 (1.84, 1.81)	H^{γ} 2.55, 2.28
T124	115.2 (8.26)	172.1	60.7 (5.46)	71.2 (3.93)	H^{γ} 0.96
A125	125.3 (8.39)	172.7	51.6 (4.84)	21.3 (1.03)	
D126	122.4 (9.28)	172.7	57.3 (4.14)	39.6 (2.72)	
D127	121.1 (8.48)	172.8	54.7 (4.50)	41.4 (2.66)	
V128	123.4 (8.41)	171.2	63.6 (3.95)	32.9 (2.14)	H^{γ} 0.92, 0.69
V129	125.7 (8.38)	172.9	61.4 (4.50)	34.4 (1.83)	H^{γ} 0.77, 0.69
C130	129.4 (9.56)		56.4 (5.16)	35.4 (3.00, 2.66)	
T131	124.6 (8.11)		(5.13)	(3.74)	H^{γ} 0.98
R132	128.3 (9.77)		(5.01)	(1.78, 1.67)	H^{γ} 1.50
V133	122.2 (8.50)	171.2	62.2 (4.90)	(1.66)	H^{γ} 0.85, 0.80
Y134	127.2 (9.79)	172.5	56.5 (5.59)	43.6 (3.23, 3.01)	2, 6H 6.82; 3, 5H 6.37
V135	112.1 (9.55)	172.8	59.7 (5.33)	36.2 (2.36)	H^{γ} 1.04, 0.85
R136	123.2 (8.55)	174.5	57.3 (3.89)	30.3 (1.29, 1.11)	H^{γ} 0.59, 0.06; H^{δ} 2.70, 2.79
E137	125.6 (8.08)	178.7	58.9 (3.86)	31.6 (1.80, 1.68)	H^{γ} 2.05, 1.90

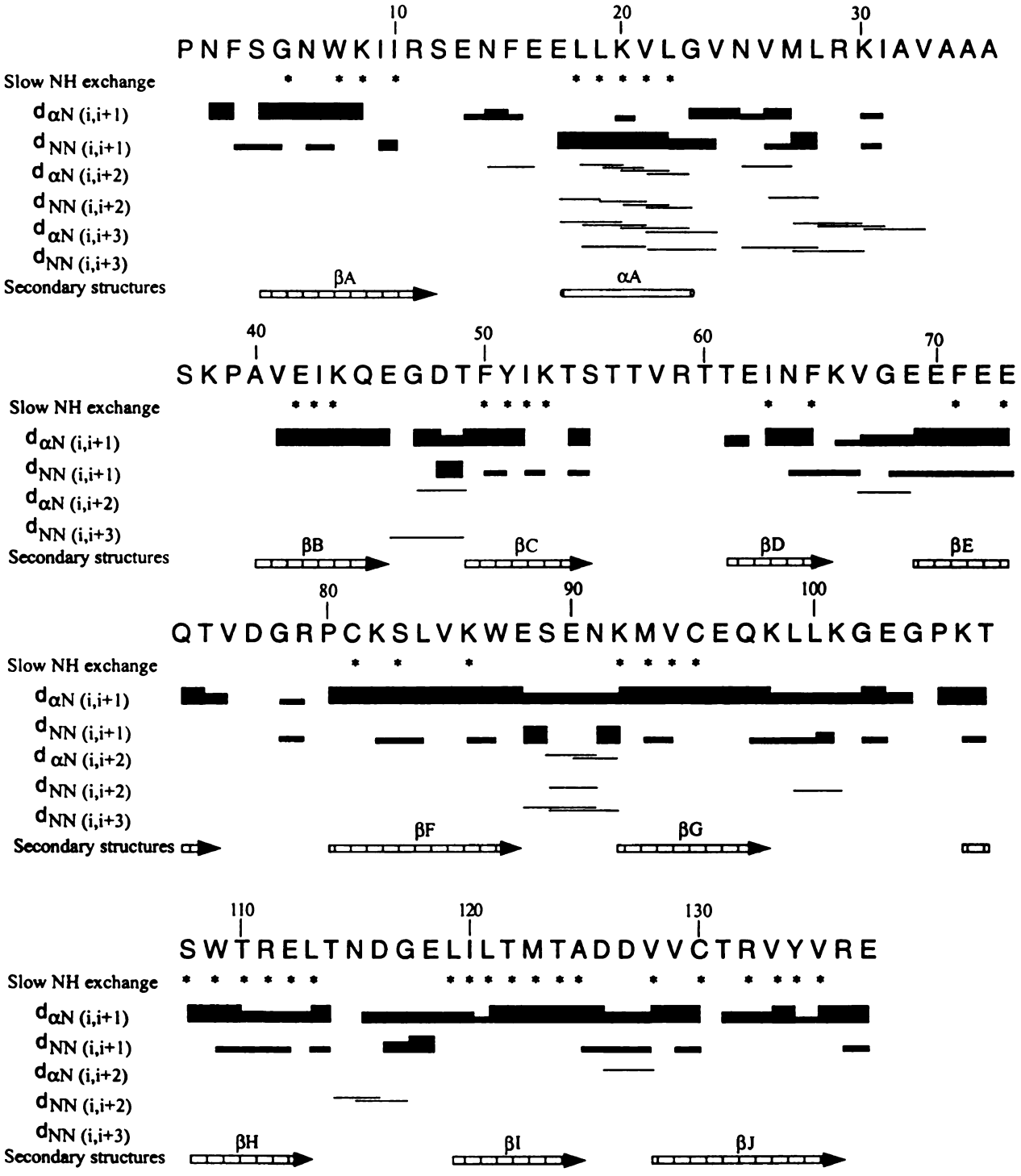
H^{δ} and between H^{α}-H^{δ} for tyrosines and phenylalanines, between H^{β}-$H^{\delta 1}$ and between H^{α}-$H^{\varepsilon 3}$ for tryptophans. The ring protons of Phe15 and Tyr134, which were overlapped with other peaks in the aromatic regions, were assigned based on the NOEs between H^{β}-H^{δ} and between H^{α}-H^{δ} and the comparisons. In the corresponding spectra of holo-CRABPII and apo-R111M they were separated from the other peaks.

Stereospecific resonance assignment. In order to better determine the solution structure of CRABPII, stereospecific assignments were made whenever possible. The methylene H^{β} protons were stereospecifically assigned on the basis of the qualitative estimations of $^{3}J_{\alpha\beta}$ coupling constants from DQF-COSY spectrum in conjunction with the relative volumes of the NOE cross-peaks between H^{α}–H^{β} picked from the 2D homonuclear NOESY with 100 ms mixing time and the NOE cross-peaks between H^{N}–H^{β} from 3D ^{1}H-^{15}N NOESY-HSQC (Basus, 1984). The methyl protons of valine residues were stereospecifically assigned in a similar manner except that the relative volumes of the NOE cross-peaks of H^{α}–H^{γ} and H^{N}–H^{γ} were used. Methyl protons of leucine residues and the ring protons of Phe and Tyr residues were stereospecifically assigned after two rounds of structure refinement. Stereospecific assignments were obtained for H^{β} methylene protons of 43 residues, for methyl protons of six valine residues and four leucine residues.

Secondary structures. Secondary structural elements were obtained from NOE data. The NOEs used to characterize secondary structures were $d_{\alpha\alpha}$ and $d_{\alpha\beta}$ from the homonuclear 2D NOESY and $d_{\alpha N}$, d_{NN} and $d_{\beta N}$ from the 3D ^{1}H-^{15}N NOESY-HSQC and/or homonuclear 2D NOESY data. Regular α-helix was characterized by strong short-range NOEs d_{NN}, $d_{\alpha\beta}(i, i+3)$ and $d_{\beta N}(i, i+3)$ and weak $d_{\alpha N}$, and regular β-sheet by strong short-

range NOEs $d_{\alpha N}$ and $d_{\beta N}$ and strong or medium long-range NOEs $d_{\alpha\alpha}$ and $d_{\alpha N}$ between adjacent strands. The following ten β-strands forming two anti-parallel β-sheets were identified (Figure 3.4): Ser4–Ser12 (βA), Ala40–Glu46 (βB), Thr49–Ser55 (βC), Thr61–Lys66 (βD), Glu69–Thr75 (βE), Pro80–Glu88 (βF), Lys92–Lys98 (βG), K106–L113 (βH), L119–A125 (βI) and V128–R136 (βJ). Long-range NOEs, $d_{\alpha\alpha}$ and $d_{\alpha N}$, were observed between adjacent βA, βB, βC and βD, between adjacent βE, βF, βG, βH, βI and βJ, and also between the N-terminal half of βA and the C-terminal half of βJ. No such NOEs were observed between the adjacent βD and βE. Lys8-Ser12 of βA, Lys53 of βC, Ile63 of βD and Thr131–Val135 of βJ had weak peaks in the 2D ^{1}H-^{15}N HSQC spectrum. However, the amide protons of Lys8, Ile10, Lys53, Ile63 and Arg132–Val135 were observed in the homonuclear 2D spectra recorded in D_2O (Figure 3.3). In particular, most of their inter-strand NOE distances $d_{\alpha\alpha}$ and $d_{\alpha N}$ was observed in the 2D NOESY spectra in D_2O. Therefore, Lys8–Ser12 and Thr131–Val135 were considered to adopt β-sheeted structures. They formed part of the so-called β-bulge structure. Ala40–Val41 of βB and Thr60–Glu62 of βD showed weak or no observable peaks in the 2D ^{1}H-^{15}N HSQC and 3D ^{1}H-^{15}N NOESY-HSQC spectra. However, strong NOEs ($d_{\alpha\alpha}$) were observed for these residues. The first two β-strands were linked by a long segment (Glu13-Pro39). Along this segment a short α-helix αA, Glu17–Gly23, could be recognized. Glu17 was weak in the 2D ^{1}H-^{15}N HSQC spectrum. It was assigned as the N-terminus of the helix according to the NOE $d_{\alpha\beta}(i, i+3)$ observed in the D_2O spectra.

Figure 3.4. Summary of the sequential and medium-range NOEs involving backbone H^N and H^α atoms, slow-exchange backbone amide protons and the deduced secondary structures of apo-CRABPII. Line thickness for $d_{\alpha N}$ and d_{NN} sequential NOE distances reflects the intensities of the cross peaks. Asterisks indicate the residues whose amide protons remained after 24 h of exchange with D_2O in PBS buffer at room temperature and pD 7.5 as determined by 2D homonuclear NMR experiments.

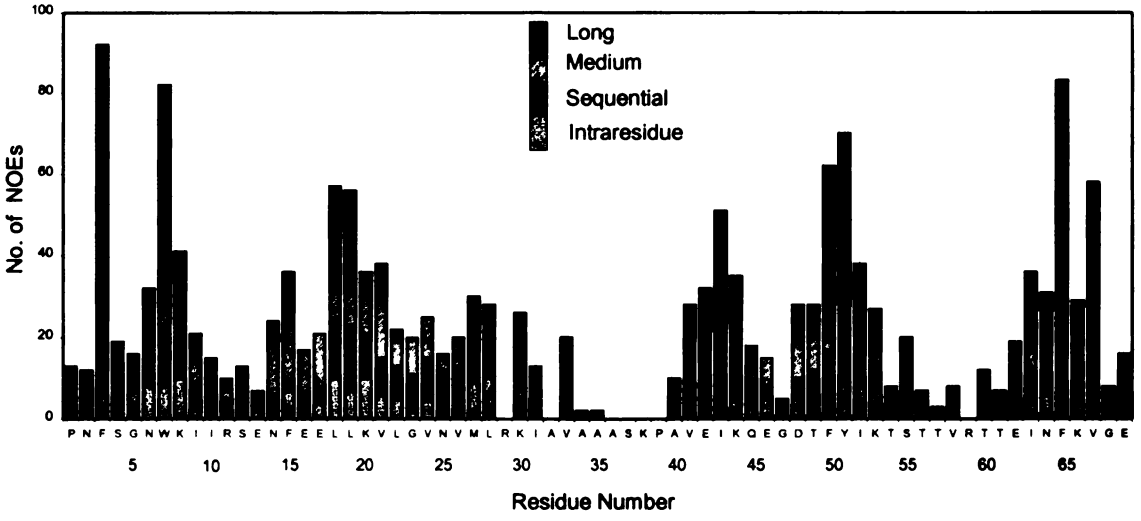

10
20
30
PNFSGNWKIIRSENFEELLKVLGVNVMLRKIAVAAA
Slow NH exchange
dαN (i,i+1)
dNN (i,i+1)
dαN (i,i+2)
dNN (i,i+2)
dαN (i,i+3)
dNN (i,i+3)
Secondary structures
βA
αA
40
50
60
70
SKPAVEIKQEGDTFYIKTSTTVRTTEINFKVGEEFEE
Slow NH exchange
dαN (i,i+1)
dNN (i,i+1)
dαN (i,i+2)
dNN (i,i+3)
Secondary structures
βB
βC
βD
βE
80
90
100
QTVDGRPCKSLVKWESENKMVCEQKLLKGEGPKT
Slow NH exchange
dαN (i,i+1)
dNN (i,i+1)
dαN (i,i+2)
dNN (i,i+2)
dNN (i,i+3)
Secondary structures
βF
βG
110
120
130
SWTRELTNDGELILTMTADDVVCTRVYVRE
Slow NH exchange
dαN (i,i+1)
dNN (i,i+1)
dαN (i,i+2)
dNN (i,i+2)
dNN (i,i+3)
Secondary structures
βH
βI
βJ

Glu13–Glu16 had very weak peaks in the 2D ^{1}H-^{15}N HSQC, homonuclear 2D SSNOESY and WATERGATE NOESY spectra.

NOE distance restraints. A total of 2290 structurally useful NOE distance restraints were obtained from the analyses of the homonuclear 2D NOESY and 3D ^{1}H-^{15}N NOESY-HSQC spectra. Of the NOE restraints, 647 were intra-residue, 560 were sequential, 248 were medium-range and 835 were long-range NOEs (Table 3.2). On average, each residue had ~17 NOE restrains, but the numbers of NOE restraints per residue varied greatly along the amino acid sequence (Figure 3.5). The residues situated in several turns or loops and the entrance region had NOE restraints well below the average. A total of 92 hydrogen bond restraints obtained from 46 hydrogen bonds were included in the structural refinement.

Quality of the structure. Figure 3. 6A shows the C^{α} traces of the superimposed 25 computed structures with no NOE violations exceeding 0.3 Å. The statistics of the structures are summarized in Table 3.2. Except for the entrance region, the structures were well defined. Excluding residues Ala32-Pro39 and Thr57-Glu62, the RMSD of the 25 coordinates was 0.54 Å for the backbone (N, C^{α}, C', O) and 0.92 Å for all heavy atoms. The stereochemical qualities of the structures were checked by the program PROCHECK–NMR (Table 3. 2). The restrained minimized mean structure (Clore & Gronenborn, 1989) had 55.6% residues in the most favorable regions of the Ramachandran plot, and 37.1% in the additional allowed regions. Two residues, Phe3 and Asn115, and eleven others were at the disallowed and generously allowed regions, respectively. The relative low percentage of the residues in the most favorable regions

Figure 3.5. The distribution of NOEs along the amino acid sequence of CRABPII. For intra-residue NOEs only those that are structurally useful are included.


Long
Medium
Sequential
Intraresidue
No. of NOEs
Residue Number

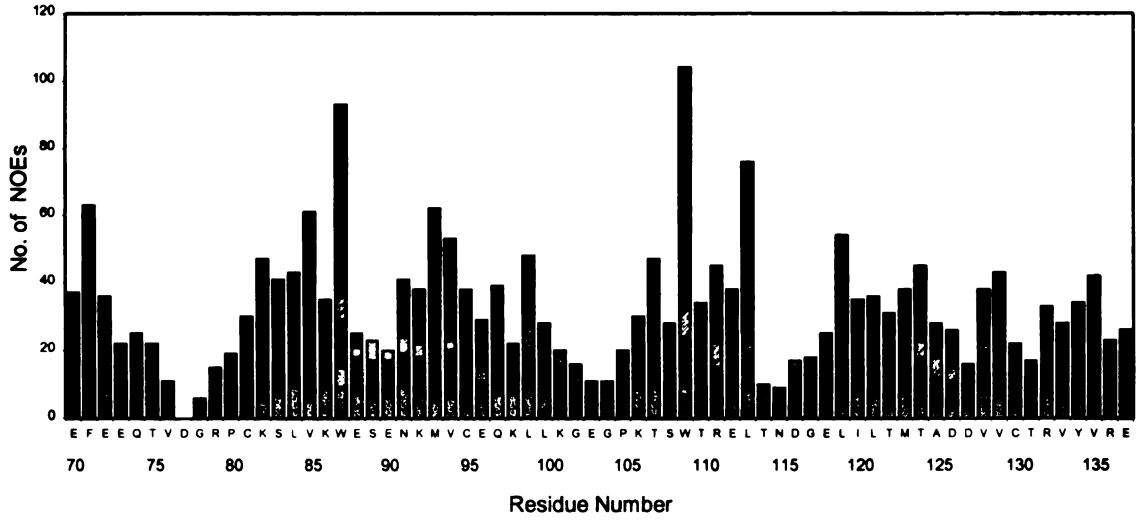

No. of NOEs
Residue Number

Figure 3.6. (A) Stereoview of the C^{α} traces of the superimposed 25 refined structures of apo-CRABPII. (B) Stereoview of the C^{α} traces of the restrained minimized mean structure of apo-CRABPII (thin line) superimposed with the C^{α} traces of the crystal structure of holo-CRABPII (thick line). (C) Stereoview of the C^{α} traces of the restrained minimized mean structure of apo-CRABPII (thin line) superimposed with the C^{α} traces of the crystal structure of apo-CRABPI (molecule A) (thick line).

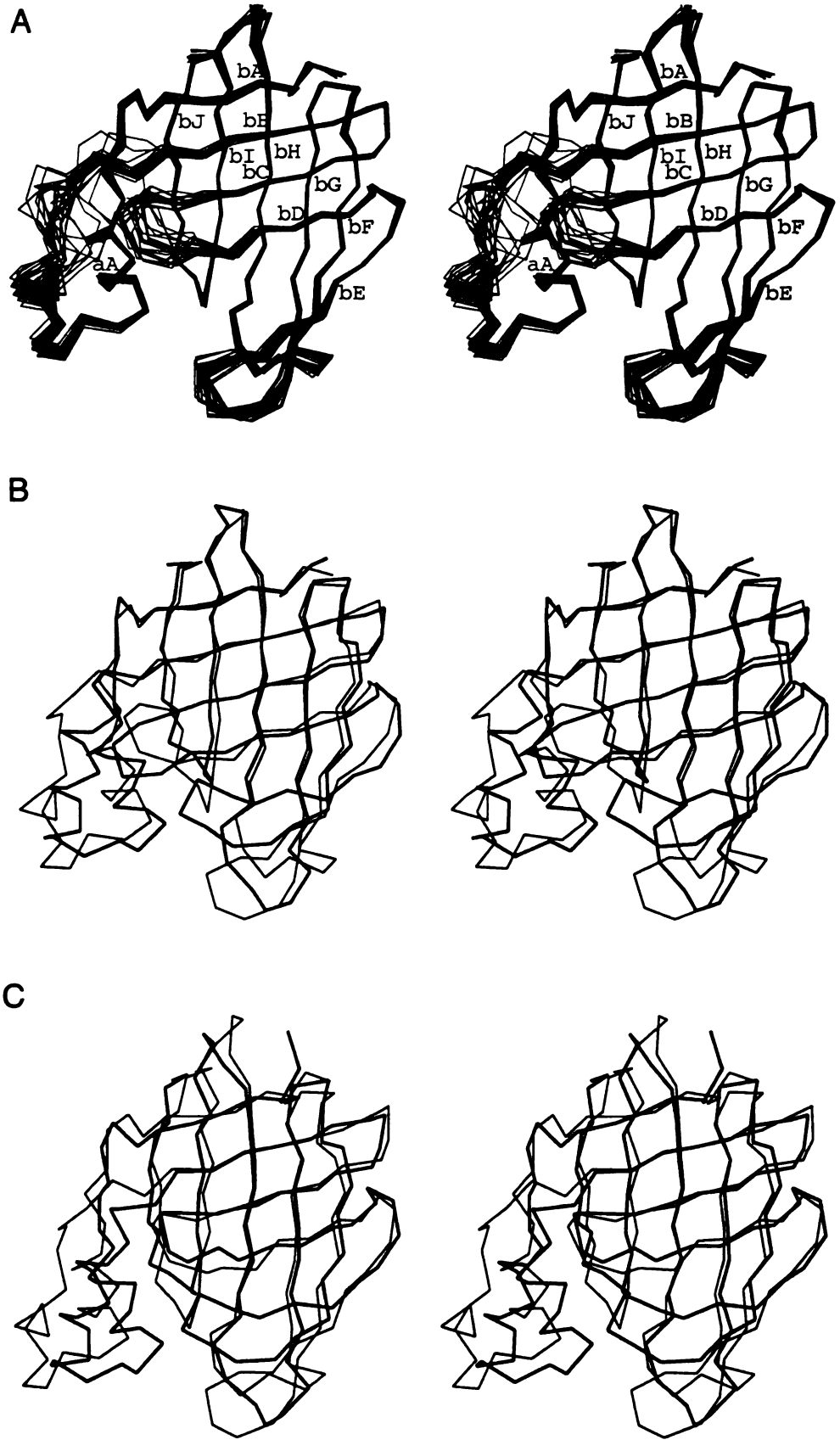

A
bA
bJ
bB
bI
bH
bC
bG
bD
bF
aA
bE
bA
bJ
bB
bI
bH
bC
bG
bD
bF
aA
bE
B
C

Table 3.2. Restraint and structural statistics of apo-CRABPII

Restraint Statistics		
Numbers of experimental NOE restraints		
Intra-residue	647	
Sequential	560	
Medium	248	
Long	835	
Total	2290	
Number of hydrogen bonds	92	
Structural Statistics		
	<SA>[a]	<SA>$_r$[b]
RMSD from experimental distance restraints (Å) (2290)	0.036 ± 0.001	0.028
Deviation from idealized covalent geometry		
Bonds (Å)	0.004 ± 0.001	0.003
Angles (deg)	0.742 ± 0.028	0.583
Impropers (deg)	0.604 ± 0.006	0.482
Measures of Structural Quality (By Procheck)		
% residues in most favorable regions of the Ramachandran plot	55.6 ± 2.4	54.8
% residues in additional allowed regions of the Ramachandran plot	37.1 ± 1.6	37.1
No. of bad contacts	8 ± 2	8
H-bond energy	0.40 ± 0.10	0.50
Overall G-factor	-0.30 ± 0.02	-0.30
Coordinate Precision		
RMSD for C_α trace (Å)	0.84 ± 0.21	0.47 ± 0.15[c]
RMSD for backbone atoms (Å)	0.89 ± 0.25	0.54 ± 0.18[c]
RMSD for all heavy atoms (Å)	1.12 ± 0.28	0.92 ± 0.20[c]

[a] <SA> are the final 25 simulated annealing structures.
[b] <SA>$_r$ is the restrained minimized mean structure obtained by restrained regularization of the mean structure, which is obtained by averaging the coordinates of the individual SA structures best fitted to each other.
[c] Residues from Ala32-Pro39 and Thr57-Glu62 were excluded from the RMSD calculations.

was due to the looseness or absence of the backbone NOE restraints for more than one-fourth of the total 137 residues.

Description of the solution structure of apo-CRABPII. The solution structure of apo-CRABPII mainly consists of two nearly orthogonal anti-parallel β-sheets twisted to form a flattened barrel, a structural feature shared by all the published structures of iLBPs (Banaszak et al., 1994). As in other iLBPs, owing to a large gap, there are no main-chain hydrogen bonds between βD and βE. The gap gradually widens from the N-terminus to the C-terminus of βD (Figure 3. 6A). The C-terminal half of βC and the N-terminal half of βD together with the loop connecting them move away from the main protein body and extend into the solvent. The C-terminal half of βE together with the loop connecting βE and βF also moves away the main protein body. The entrance to the binding pocket is formed by the hydrophobic side-chains of Leu19, Val24, Leu28–Pro39 and the residues in the βC–βD loop and βE–βF loop. The distance from the C^{α} atom of Val58 in the βC–βD loop to the C^{α} atom of Val76 in the βE–βF loop is about 20.7 Å. The distance between the C^{α} atom of Val24 of the loop connecting αA to αB and the C^{α} atom of Val76 is 17.0 Å. The ligand binding pocket is wide open and the side chains of the trio (Arg111, Arg132 and Tyr134) are readily accessible to the external ligand.

Discussion

Assignment and secondary structures. The amide resonances of eleven residues were not observed, and about twice the number of residues had weak peaks in the 2D ^{1}H-^{15}N HSQC spectrum. This posed a challenge for the sequential assignment method based on a series of ^{15}N-edited heteronuclear NMR experiments because these experiments depend on magnetization transfer of amide resonances. Their assignments could not be accomplished without the help of the comparisons with the sequential assignments of both the wild type holo-CRABPII and the site-directed mutant apo-R111M. However, the intensities of the amide proton signals from Arg29, Ala36, Lys38 and Arg59 were so low in both the 2D ^{1}H-^{15}N HSQC and the homonuclear 2D SSNOESY and WATERGATE NOESY spectra that none of their ^{1}H and ^{15}N resonances has been assigned. Secondary structure elements were identified according to the NOE patterns. For residues with weak or no observable amide peaks in NOESY spectra only the $d_{\alpha\alpha}$ and $d_{\alpha\beta}(i, i+3)$ NOEs were used. Several residues had strong peaks in the fingerprint regions of the 2D homonuclear spectra recorded in D_2O but had only very weak peaks in the ^{15}N-edited heteronuclear spectra. Their secondary structural identities were established by the backbone hydrogen bond, and the NOEs observed in the homonuclear spectra. All the available data, in particular, the homonuclear 2D SSNOESY, WATERGATE NOESY spectra were analyzed carefully to assign the resonances of Met27-Ala36, which assume an α-helical structure (αB) according to the crystal structure of holo-CRABPII. Some weak sequential

NOEs typical of the regular helical structure were observed for Met27–Ile31 but their intensities and numbers were well below the average (Figure 3.5). No structurally useful sequential NOEs could be identified for residues Ala32-Ala35. However, the chemical shift indices suggested that Lys30-Ala35 are helical. No NOEs could be identified unambiguously for Ser37 except its intra-residue NOEs. These experimental data suggested that the residues Met27-Ala35 adopt an unstable α-helical structure. Residues Ala32-Pro39 are disordered in the NMR structures because no chemical shift indices were included in the structure refinement (Figure 3.6A).

Comparisons with the crystal structures of holo-CRABPII and apo-CRABPI. Figure 3.6B shows the overlay of the C^{α} traces of the restrained minimized mean structure of the apo-CRABPII and the crystal structure of holo-CRABPII. The structure can be superimposed with C^{α} deviations of 3.52, 2.04, and 1.65 Å for all residues, the regular secondary structures, and all residues excluding Val24-Ser37, Glu74-Pro80 and Leu100-Gly104, respectively. The main differences between the two structures are located in the regions around the ligand entrance. The second helix is well defined in the holo crystal structure but appears flexible in solution, especially the C-terminal half. The βC-βD loop moves away from the main body of the protein by ~5 Å in the apo solution structure. The C-terminal half of βE and the N-terminal half of βF in the apo solution structure do not twist toward the center of the protein to constrict the ligand entrance as they do in the holo crystal structure. Consequently, the long loop between βG and βH also moves away from the main body of the apo protein. The distance between the C^{α} atom of Val76 in the βE–βF loop and the C^{α} atom of Val58 in the βC–βD loop is ~8.0 Å longer in the apo solution structure than in the holo crystal structure. The distance between the C^{α} atom of

Val76 and the C^{α} atom of Val24 in the αA-αB loop is ~6.7 Å longer in the apo solution structure than in the holo crystal structure. As a result, the ligand entrance in the apo solution structure is greatly enlarged with a much more exposed binding pocket. In particular, residues Val24, Leu28, Ile31, Val58, Arg59, and Val76 move away from each other so that the side-chains of Arg111, Arg132 and Tyr134 that interact with the carboxyl group of RA are easily accessible to the ligand. However, the relative positions of Arg111, Arg132, and Tyr134 are quite similar to those observed in the holo crystal structure: their backbone atoms can be superimposed rather well.

Compared with other parts of the protein, the ligand entrance region of apo-CRABPII is not well defined in the solution structure. Most residues in the region have weak or no observable peaks in all ^{15}N-edited spectra. The number of the experimental distance restraints per residue for the region is well below the average (Figure 3.5). However, except for Ala32–Pro39, the differences between the apo solution structure and the holo crystal structure (Figure 3.6B) are larger than between any two conformations from the ensemble of 25 refined NMR solution conformations (Figure 3.6A). Thus, the differences between the apo solution structure and the holo crystal structure are real.

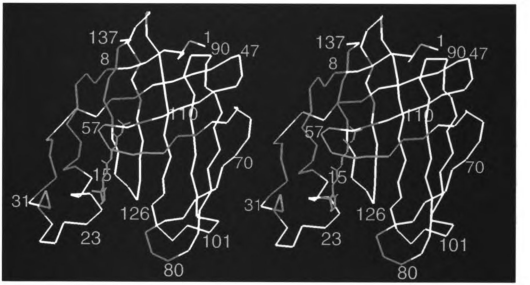
Most of the differences between the apo solution structure and the holo crystal structure of CRABPII are also observed between the solution structure of apo-CRABPII and the crystal structure of apo-CRABPI (Thompson et al., 1995) (Figure 3.6C).

Dynamical properties. As detailed earlier, about a quarter of the residues showed weak or no signals in the 2D ^{1}H-^{15}N HSQC spectrum (Figure 3.2). Some of them were situated in the turns or loops between elements of regular secondary structure. Surprisingly, many others assumed regular α-helical or β-sheet structures. Only the

residues in the βC-βD loop and the βE-βF loop and Asn115 in the βH-βI loop showed weak or no cross-peaks in the ^{1}H-^{15}N HSQC spectrum. The residues in the longest loop between βG and βH showed the highest intensities in the ^{1}H-^{15}N HSQC spectrum. By contrast, residues in the long loop of holo-CRABPII had high temperature factors in the X-ray structure. Interestingly, most of these residues that showed weak or no cross-peaks in the ^{1}H-^{15}N HSQC spectrum are located in the RA-binding pocket (Figure 3.7). Among the 21 residues that constitute the RA-binding pocket (Kleywegt et al., 1994), 16 residues exhibited weak or no detectable cross-peaks in the ^{1}H-^{15}N HSQC spectrum.

The cluster of the weak or missed ^{1}H-^{15}N correlations suggested that the RA-binding pocket of CRABPII is rather dynamic in the absence of RA. It is unlikely that these peaks were affected by water saturation, because the ^{1}H-^{15}N HSQC spectrum was recorded with a "water flip-back" pulse sequence that minimized water saturation and dephasing (Grzesiek & Bax, 1993). Furthermore, for several weak or missing ^{1}H-^{15}N correlations, the amide protons remained two days after the lyophilized protein was dissolved in D_2O and showed cross-peaks in the fingerprint regions of the homonuclear 2D spectra recorded in D_2O (Figure 3.3), indicating that these amide protons exchange slowly with water and should not be perturbed by water saturation. Rather, the ^{1}H-^{15}N correlations were most likely weak or missing as the result of broadening of the amide proton and/or nitrogen resonances due to conformational exchange at rates intermediate on the NMR time scale. For Lys8, Ile10, Leu19, Leu121 and Arg132-Val135, clearly, line broadening of their nitrogen resonances was responsible for the weak peaks in the ^{1}H-^{15}N HSQC spectrum because their amide protons exhibited sharp signals and showed cross-peaks with α-protons in 2D homonulcear spectra recorded in D_2O (Figure 3.3). A

Figure 3.7. The relative peak intensity of the 2D ^{1}H–^{15}N HSQC spectrum of apo-CRABPII color-coded along the C^{α} trace of the solution structure. The strong peaks are in white, the weak peaks in pink and the missing peaks in cyan. RA (green) is positioned on the basis of the superposition of the solution structure of apo-CRABPII and the crystal structure of holo-CRABPII.


137
8
1
90
47
110
57
70
15
31
126
23
101
80
137
8
1
90
47
110
57
70
15
31
126
23
101
80

slow conformational exchange with a lifetime of at least 40 ms was observed as described in Chapter 2, resulting in two sets of cross-peaks for Trp87. It was noted that Trp87 is located at the N-terminus of βF far away from the RA-binding pocket. The residues in the RA-binding pocket likely undergo intermediate conformational exchange, probably on the microsecond to millisecond time scale, because only one set of NMR signals was observed for these residues. The clustering of the mobile residues and the relatively low frequency of conformational exchange suggest that the RA-binding pocket may undergo collective motions in the absence of the RA ligand.

Apart from the RA-binding pocket, the structure of apo-CRABPII appears well ordered and less dynamic. Most residues, including those at the turns and loops, showed intense cross-peaks in the 2D ^{1}H-^{15}N HSQC spectrum. Many of them also showed cross-peaks in the fingerprint regions of the 2D homonuclear spectra recorded in D_2O (Figure 3. 3). Remarkably, the hydroxyl protons of Ser83 and Thr107 were observed in 2D homonuclear spectra recorded in H_2O, and the indole ε^{1}-proton of Trp7 was detected in 2D homonuclear spectra recorded in D_2O, suggesting that these side-chains are not only buried but also immobile.

The dynamic nature of the RA-binding pocket is consistent with biochemical and NMR studies of CRABPI. Limited proteolysis showed that helix α2 of CRABPI is significantly more susceptible to proteolysis in the apo-form than in the holo-form (Jamison et al., 1994), suggesting that RA binding induces a conformational change or that helix α2 is more mobile in the apo-form. Hydrogen exchange measurements revealed that helix α2 and the βC-βD loop of CRABPI have much higher exchange rates in the apo-form than in the holo-form (Rizo, et al., 1994), indicating that these parts of the

molecule are more dynamic in the apo-form. Other iLBPs also have been found to exhibit similar dynamic properties. For example, it was shown by NMR that the second helix of rat intestinal FABP is disordered in solution (Hodsdon and Cristola, 1997a, b).

Is CRABPII dimeric in solution? Unlike other iLBPs, apo-CRABPI was crystallized in a dimeric form (Thompson et al., 1995). The crystalline dimer is held together by an intermolecular β-sheet formed by the βD strands of two CRABPI molecules, resulting in a 20-stranded double β-barrel with slightly more open RA-binding pockets. It was suggested that dimerization may be the mechanism by which CRABPI opens the ligand entrance so that RA can enter or exit the binding pocket without steric hindrance. In light of the potential significance of the dimerization to the ligand entrance problem, we carefully examined the possibility of the formation of the intermolecular β-sheet. We concluded that apo-CRABPII is predominately monomeric in solution for the following reasons. (i) On a Sephadex G-50 gel filtration column, apo-CRABPII eluted with a volume characteristic of monomeric CRABPII. (ii) For the most part, the line widths of the NMR signals of CRABPII were consistent with a monomeric structure. (iii) None of the predicted NOEs indicative of formation of the intermolecular β-sheet was observed. For example, according to the dimeric crystal structure of apo-CRABPI, the distance between H^{α} of Thr61 of molecule A and H^{β} of Thr60 of molecule β is 2.5 Å but the distance between H^{α} of Thr61 and H^{β} of Thr60 of the same molecule is 5.9 Å. Therefore, if there is a stable dimer in solution, we should see a NOE cross-peak between the two protons. On the other hand, if there is no stable dimer, there would be no NOE between the two protons. The two protons were sequentially assigned, but we did not see any NOE cross-peak between them. Although the formation of transient dimers could not

be ruled out, the transient dimers, if they exist, can only be a minor population under the conditions of the NMR or gel filtration experiments. What is the likelihood of the dimerization *in vivo*? Since the cellular concentration of CRABPII is much lower than the concentrations of the NMR samples, the concentration of dimeric species would be even lower *in vivo*.

Implications for RA binding. The crystal structures of holo-CRABPI and CRABPII suggested that RA cannot enter or exit the ligand binding pocket of the proteins in the absence of significant conformational changes. On the basis of limited proteolysis of several intracellular lipid binding proteins, it was proposed that a rigid body movement of the helix-turn-helix motif may serve as the mechanism by which the family of proteins opens up their ligand binding pocket (Jamison et al., 1994). In light of the solution structure of apo-CRABPII, it is unlikely that the two helices move as rigid rods as previously proposed. Rather, the second helix of the helix-turn-helix motif is flexible and undergoes conformational exchange. The dynamical nature of the second helix may be the cause of its high susceptibility to proteolysis.

Based on the comparisons of the crystal structures of apo- and holo-CRABPI, it was suggested that movement of the βC–βD loop is responsible for the opening of the ligand entrance (Thompson et al., 1995). Furthermore, it was also proposed that the movement of the βC–βD loop is dependent on the formation of an intermolecular β-sheet as mentioned above. Comparison of the solution structure of the apo-CRABPII and the crystal structure of holo-CRABPII confirms the hypothesis that βC–βD loop moves when RA is bound. However, the movement of the βC–βD loop that opens the binding pocket

is shown by NMR not to involve dimerization of the protein because CRABPII is predominately monomeric in solution.

The solution structure of apo-CRABPII suggests that the ligand entrance is widely open and readily accessible to RA. The enlargement of the ligand entrance of apo-CRABPII compared with holo-CRABPII is mainly due to a concerted conformational change in three structural elements, namely the second helix, the βC–βD loop and the βE–βF loop. A widely opened entrance may be essential for binding of RA, because RA is a long, relatively rigid, negatively charged molecule. A narrow entrance lined with the hydrophobic side-chains would result in unfavorable interactions with RA. The entry of RA into the deep binding pocket may be further facilitated by the positively charged potentials generated by Arg29, Arg59, Arg111 and Arg132 (Chen et al., 1998). Binding of RA induces significant conformational changes in CRABPII. Furthermore, the interactions between RA and CRABPII may stabilize the structural elements constituting the RA-binding pocket, especially the second helix, the βC–βD loop and the βE–βF loop.

Conclusion

Using multidimensional NMR spectroscopy, we have determined the first solution structure of a CRABP. Comparison of the solution structure of apo-CRABPII with the crystal structure of holo-CRABPII indicates that the largest conformational differences between the two structures are localized at the ligand entrance. The ligand entrance of apo-CRABPII is greatly enlarged and readily accessible to RA, mainly due to a concerted movement of the second helix and the βC-βD and βE-βD loops. Furthermore, the ligand binding pocket of apo-CRABPII is rather dynamic as indicated by analysis of the cross-peaks intensities of the ^{1}H-^{15}N HSQC spectrum. CRABPII is predominately monomeric in solution as determined from NMR and biochemical evidence. Although the formation of transient dimers could not be ruled out, dimerization apparently is not a prerequisite for entry of RA into the ligand binding pocket of CRABPII.

References

Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1981) *Biochem. Biophys. Res. Commun. 95*, 1-6.

Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S. W., & Jones, T. A. (1994) *Adv. Protein Chem. 45*, 89-151.

Basus, V. J. (1989) *Methods Enzymol. 177*, 132-149.

Bax, A., Clore, G. M., & Gronenborn, A. G. (1991) *J. Magn. Reson. 88*, 425-431.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Magn. Reson. 55,* 301-315.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185-188.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., & Karplus, M. (1983) *J. Comput. Chem. 4,* 187-217.

Brünger, A. T. (1992) *X-PlOR: A System for Crystallography and NMR.* Yale University, New Haven, CT.

Chambon, P., Olson, J. A., & Ross, A. C. (coordinators) (1996) *The retinoid revolution. FASEB. J 10*, 938-1048.

Chen, X., Tordova, M., Gilliland, G. L., Wang, L., Li, Y., Yan, H., and Ji, X. (1998) *J. Mol. Biol.*, in press.

Clore, G. M., & Gronenborn, A. G. (1989*) CRC Crit. Rev. Biochem. Mol. Biol. 24*, 479-564.

Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. P. (1988) *J. Am. Chem. Soc. 110*,

7870-7872.

Grzesiek, S., & Bax, A. (1992a) *J. Magn. Reson. 96*, 432-440.

Grzesiek, S., & Bax, A. (1992b) *J. Am. Chem. Soc. 114*, 6291-6293.

Grzesiek, S., Anglister, J., & Bax, A. (1993) *J. Magn. Reson. Ser. B 101*, 114-119.

Grzesiek, S., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593-12594.

Hodson, M. E., & Cistola, D. P. (1997a) *Biochemistry 36*, 1450-1460.

Hodson, M. E., & Cistola, D. P. (1997b) *Biochemistry 36*, 2278-2290.

Jamison, R. S., Newcomer, M. E., & Ong, D. E. (1994) *Biochemistry 33*, 2873-2879.

Jeener, J., Meier, B. H., & Ernst, R. R. (1978) *J. Chem. Phys.* 71, 4546-4553.

Kay, L. E. (1995) *Prog. Biophys. Molec. Biol. 63*, 277-299.

Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 120-126.

Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 89*, 495-514.

Kay, L. E., Keiffer, P., & Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.

Kay, L. E., Xu. G. Y., Singer, A. U., Muhandiram, D. R., & Forman-Kay, J. D. (1993) *J. Magn. Reson. B101*, 333-337.

Kleywegt, G. J., Bergfors, T., Senn, H., le Motte, P., Gsell, B., Shuao, K., & Jones, T. A. (1994) *Structure 2*, 1241-1258.

Laskowski, R., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283-291.

Laskowski, R., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1997) *J. Biomol. NMR 8*, 477-486.

Logan, T. M., Olejniczak, E. T., Xu, R. X., & Fesik, S. W. (1993) *J. Biomol. NMR 3*, 225-231.

Mangelsdorf, D. J., Umesono, K., & Evans, R. M. (1994) In *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., & Goodman, D. S., eds.), *pp.* 319-349, 2nd ed., Raven Press, New York.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989a) *J. Magn. Reson. 85*, 393-399.

Marion, D., Ikura, M., & Bax, A. (1989b) *J. Magn. Reson. 84*, 425-430.

McCoy, M. A., & Müeller, L (1992a) *J. Am. Chem. Soc. 114*, 2108-2112.

McCoy, M. A., & Müeller, L (1992b) *J. Magn. Reson. B98*, 674-679.

Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., & Dahluist, F. W. (1989) *Methods Enzymol. 177*, 44-73.

Muhandiram, D. R., & Kay, L. E. (1994) *J. Magn. Reson. B103*, 203-216.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 129-136.

Ong, D. E., Newcomber, M. E., & Chytil, F. (1994) In *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., & Goodman, D. S., eds.), *pp.* 283-317, 2nd ed., Raven Press, New York.

Piantini, U., Sørensen, O. W., & Ernst, R. P. (1982) *J. Am. Chem. Soc. 104*, 6800-6801.

Rizo, J., Liu, Z. P., & Gierasch, L. M. (1994) *J. Biomol. NMR 4*, 741-760.

Sacchettini, J. C., Gordan, J. I., & Banaszak, L. J. (1988) *J. Biol. Chem. 263*, 5815-5819.

Shaka, A. J., Keeler, J., Frenkiel, T., & Freeman, R. (1983) *J. Magn. Reson. 52*, 335-338.

Shaka, A. J., Lee, C. J., & Pine, A. (1988) *J. Magn. Reson. 77*, 274-293.

Shaka, A. J., Barker, P. B., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547-552.

Smallcombe, S. H. (1993) *J. Am. Chem. Soc. 115*, 4776-4785.

Sporn, M. B., Roberts, A. B. & Goodman, D. S. (1994) *The Retinoids: Biology, Chemistry, and Medicine*, 2nd ed., Raven Press, New York.

States, D. J., Haberkon, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.

Thompson, J. R., Bratt, J. M., & Banaszak, L. J. (1995) *J. Mol. Biol. 252,* 433-446.

Wittekind, M., & Müeller, L. (1993) *J. Magn. Reson. 101*, 201-205.

Wüthrich, K., Billletter, M., & Braün, W. (1983) *J. Mol. Biol. 169*, 949-961.

Zhang, J., Liu, Z., Jones, T. A., Gierasch, L. M., & Sambrook, J. F. (1992) *Proteins: Structure, Function, and Genetics 13*, 87-99.

Zhang, O., Kay, L. E., Olivier, J. P., & Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845-858.

Zhu, G., & Bax, A. (1992) *J. Magn. Reson. 90*, 405-410.

Zuiderweg, F. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387-2391.

CHAPTER 4

Solution Structure of a Site-Directed Mutant (R111M) of Human CRABPII

Introduction

The crystal structures of holo-CRABPs reveal that RA is buried in the binding pocket with its carboxyl group interacting with Arg111, Arg132 and Tyr134 (CRABPII numbering) at the bottom of the pocket (Kleywegt et al., 1994). Arg132 and Tyr134 interact directly with the carboxyl group of RA, whereas the interaction between Arg111 and the carboxyl group of RA is mediated by a water molecule. Electrostatic calculations based on a modeled structure of CRABPI have suggested that the electrostatic repulsion between the guanidinium groups of the above two conserved arginine residues provides the conformational flexibility to CRABPs for RA to enter the binding pockets (Zhang et al., 1992).

Site-directed mutagenesis and competitive binding studies have shown that both Arg111 and Arg132 are important for CRABPII to bind RA but Arg111 contributes more to the binding energy than Arg132 (Chapter 2). This chapter describes the structure determination of the CRABPII mutant protein R111M in solution by NMR spectroscopy.

The results suggest that Arg111 play a critical role in determining the structure and dynamical properties of CRABPII.

Experimental Procedures

Sample preparation. Unlabeled amino acids, nucleotides, vitamins and D-glucose were purchased from Sigma (St. Louis, Missouri). [^{15}N]-labeled amino acids and $^{15}NH_4Cl$ were purchased from Cambridge Isotopic Laboratories.

Site-directed mutagenesis and expression, and the purification of the recombinant mutant protein R111M have been described in Chapter 2. Uniform ^{15}N-labeling and ^{15}N-labeling of leucine were performed as detailed in Chapter 3. The procedure for ^{15}N-labeling of isoleucine and phenylalanine was the same as that for isoleucine and leucine. For ^{15}N-labeling of lysine, the expression strain BL21(DE3)/pLysS containing the plasmid pET-17b/R111M was used. The strain was grown in M9 medium supplemented with 110 mg/L [^{15}N]-lysine added ten minutes before the induction with IPTG. D_2O sample was prepared by dissolving the lyophilized protein in PBS buffer (20 mM sodium phosphate, 100 mM sodium chloride in 99.9% D_2O, pD 7.5, uncorrected). The protein solution was incubated for six hours at 37°C and lyophilized again. The lyophilized protein was then dissolved in 99.96% D_2O. H_2O samples were prepared by directly concentrating the protein fractions eluted with the PBS buffer (pH 7.3) from a gel-filtration column (G-50) at the final step of the purification procedure (Chapter 2). About 10% D_2O was added into the concentrated protein solutions. The protein concentrations of the NMR samples were ~ 2 mM.

NMR spectroscopy. All NMR spectra were recorded at 27°C. Homonuclear 2D spectra of unlabeled R111M in D_2O and heteronuclear 2D ^1H-^{15}N HMQC (Müeller, L. 1979; Bax et al., 1983) spectra of selectively ^{15}N-labeled proteins in H_2O were acquired on a Varian VXR-500S spectrometer. All 2D experiments were recorded in the phase-sensitive mode with F1 quadrature detection achieved by the hypercomplex method (States et al., 1982). The relaxation delay was 1.5 second, and the residual water resonance was suppressed by low-power presaturation. The spectral width was 7238 Hz for ^1H and 2000 Hz for ^{15}N. The following 2D homonuclear spectra were recorded: one DQF-COSY (Piantini et al., 1982) spectrum, one E.COSY (Griesinger et al., 1985) spectrum, three NOESY (Jeener et al., 1978; Kumar et al., 1981) spectra with mixing times of 50, 100 and 150 ms, and three clean-TOCSY (Griesinger et al., 1988) spectra with mixing times of 20, 45 and 75 ms. The time domain data were composed of 2048 × 320 complex points (128 transients) for the E.COSY and DQF-COSY experiments and 2048 × 256 complex points (80 transients) for all other homonuclear 2D experiments. The 2D ^1H-^{15}N HMQC spectra were recorded with 1024 (^1H) × 64 (^{15}N) complex points (32 transients). The data were processed with the program VNMR v. 5.1 (Varian Associates). A ~30°-shifted sine-bell was applied in both dimensions for processing the DQF-COSY and E.COSY data, and a ~45°-shifted sine-bell for the ^1H-^{15}N HMQC data. A gaussian and minus LB combination was applied in F2 dimension and a ~75°-shifted sine-bell in F1 dimension for processing the NOESY and TOCSY data. The time domain data were then zero-filled to 2048 × 512, 4096 × 2048, 4096 × 2048, and 8192 × 4096 real points for the 2D HMQC, TOCSY, NOESY, DQFCOSY and E.COSY, respectively.

A five-order polynomial was applied for baseline correction in the F2 dimension after Fourier transformation.

Heteronuclear NMR experiments of the uniformly ^{15}N labeled protein were carried out on a Brüker DMX-600 spectrometer. The carriers for ^{1}H and ^{15}N were set at 4.70 ppm and at 115.7 ppm, respectively. A 2D ^{1}H-^{15}N HSQC spectrum (Bodenhausen & Ruben 1980; Cavanagh et al., 1991; Palmer et al., 1991; Kay et al., 1992) was acquired with coherent selection by pulsed field gradients and sensitivity enhancement. The acquisition times and complex data points were ^{1}H (F2) 113.9 ms, 1024 and ^{15}N (F1) 56.9 ms, 256 (32 transients). Both 3D ^{1}H-^{15}N TOCSY-HMQC (Marion et al., 1989b) and NOESY-HMQC (Zuiderweg and Fesik, 1989; Marion et al., 1989c) spectra were recorded with the following acquisition times and complex data points: ^{1}H (F3) 113.9 ms, 1024; ^{15}N (F2) 11.9 ms, 32; ^{1}H (F1) 14.3 ms, 128 (8 transients and 45 ms mixing time for TOCSY-HMQC, and 16 transients and 150 ms mixing time for NOESY-HMQC). In both 3D experiments the solvent resonance was suppressed by low-power presaturation. Quadrature detection in indirectly detected dimensions were accomplished by the States-TPPI method (Marion et al., 1989a). Homonuclear isotropic mixing in TOCSY-HMQC experiment was achieved by a DIPSI-3 sequence (Shaka et al., 1987). Two dimensional ^{1}H-^{15}N HMQC-J (Key and Bax, 1990) data was recorded on a Brüker DMX-750 spectrometer at the National Magnetic Resonance Facility at Madison with the following acquisition times and complex data points: ^{1}H 136.4 ms, 2048 and ^{15}N 232.7 ms, 512 (16 transients). A low power GARP-I sequence (Shaka et al., 1985) was applied for ^{15}N broad-band decoupling during the data acquisition periods of these experiments. The spectra were processed with the program NMRPipe (Delaglio et al., 1995). A ~45°-

shifted sine-bell was applied in both F2 (^{1}H) and F1 (^{15}N) dimensions for both 2D ^{1}H-^{15}N HSQC and HMQC-J spectra. The ^{1}H-^{15}N HSQC and HMQC-J data were zero-filled to 2048 (F2) × 512 (F1) and 2048 (F2) × 1024 (F1) real points, respectively. A ~75°-shifted sine-bell was applied in both F1 (^{1}H) and F3 (^{1}H) dimensions and a cosine-bell in F2 (^{15}N) dimension for processing of the 3D data. The first points of F1 and F2 dimensions were calculated by linear prediction. Both 3D data were zero-filled to 2048 (F3) × 128 (F2) × 512 (F1) real points. The aliphatic part in the F3 dimension was removed after Fourier transformation.

Resonance assignment. The sequential assignments were carried out following the standard procedure developed by Wüthrich and others (Wüthrich, 1986; Clore & Gronenborn, 1989 and 1991). Briefly, the spin systems were identified by the analysis of the homonuclear 2D TOCSY and DQF-COSY spectra and the 3D ^{1}H-^{15}N TOCSY-HMQC spectrum or by selective labeling. Sequential linking was obtained from NOE connectivities observed in 3D ^{1}H-^{15}N NOESY-HMQC experiment. The stereospecific assignments of β-mythylene protons were obtained by the joint analyses of the qualitative $^3J_{\alpha\beta}$ coupling constants derived from the E.COSY and DQF-COSY spectra and the relative volumes of the NOE cross peaks between H^{α}–H^{β} and H^{N}–H^{β} from the 2D homonuclear NOESY and 3D ^{1}H-^{15}N NOESY-HMQC spectra (Basus, 1984). The methyl protons of valine and leucine residues and the ring protons of aromatic residues were stereospecifically assigned as detailed in Chapter 3. Secondary structural elements were identified according to their characteristic NOE patterns.

Derivation of structural restraints. Three types of structural restraints were derived from the experimental NMR data: interproton distance restraints, torsion angle

restraints and hydrogen bond restraints. Approximate interproton distance restraints were derived from the homonuclear 2D NOESY with 100 and 150 ms mixing times, and the 3D ^{15}N-edited NOESY-HMQC with 150 ms mixing time. The same procedure for establishing the upper limits of interproton distances was followed as described in Chapter 3. The only difference was that the NOE cross peaks from the 3D ^{1}H-^{15}N NOESY-HMQC were picked by the programs CAPP and PIPP (Garret et al., 1991). The NOEs were classified as strong, medium, weak and very weak corresponding to interproton distance 1.8–2.7 Å (1.8–2.9 Å for NOEs involving NH protons), 1.8–3.3 Å (1.8–3.5 Å for NOEs involving NH protons), 1.8–5.0 Å and 2.5–5.0 Å, respectively. Pseudoatom corrections were made for non-stereospecifically assigned methylene protons and methyl protons (Wüthrich et al., 1983). Additional 0.5 Å was added to the upper bounds for methyl protons. The backbone torsion angle ϕ was restrained according to the value of $^{3}J_{NH\alpha}$ obtained from 2D HMQC-J experiment. It was restrained $-120° \pm 40°$ for $^{3}J_{NH\alpha} \geq 7.5$ Hz, and $-60° \pm 20°$ for $^{3}J_{NH\alpha} \leq 5.0$ Hz and when NOE data indicated the presence of a helix. Backbone hydrogen bond restraints were derived from the cross peaks remaining in the fingerprint regions of the homonuclear 2D TOCSY and NOESY spectra recorded in D_2O. The HN···O and N···O were restrained to be 2.0 ± 0.5 Å and 3.0 ± 0.3 Å, respectively.

Structure calculation. The structures were calculated as detailed in Chapter 3 except that an additional square-well potential function with force constant of 200 kcal.mol^{-1}.rad^{-2} was applied for the dihedral angle restraints.

Results

Isotopic enrichment. The level of ^{15}N enrichment in the uniform ^{15}N-labeling was high as judged by the signal to noise ratio of the 2D ^{1}H-^{15}N HSQC spectrum (Figure 4.1). No ^{15}N isotopic scrambling to other residues was observed for selective labeling of isoleucine, leucine, lysine and phenylalanine. For example, the number of the cross peaks in the 2D ^{1}H-^{15}N HMQC spectrum of the [^{15}N]-leucine-labeled protein was exactly equal to the total number of leucine residues in R111M (Figure 4.2).

Spin system identification. Aliphatic spin systems were first identified by the analysis of the homonuclear 2D DQF-COSY and TOCSY spectra in D_2O and then extended to amide resonances by the 3D ^{1}H-^{15}N NOESY-HMQC experiment. Most of the threonine, alanine, valine and glycine residues could be identified easily by their unique cross-peak patterns. The majority of the residues belonging to either AMX or AM(PT)X spin system could also be identified. However, except for a few, the residues with long side-chains (lysine, leucine, isoleucine, arginine and proline) could not be distinguished from each other because of severe spectral overlapping in the upfield region (1.00 ppm-2.00 ppm) and the inefficiency of the TOCSY experiments for long-relayed coherence transfer. Their spin system identifications were achieved or assisted by selective ^{15}N-labeling. All the leucine, isoleucine and lysine residues could be identified by selective ^{15}N-labeling. Furthermore, proline and arginine could also be distinguished from each other. In general, the H^{δ} protons of prolines have larger down field chemical shifts than

Figure 4.1. Gradient- and sensitivity-enhanced 2D ^{1}H–^{15}N HSQC spectrum of uniformly ^{15}N enriched of apo-R111M. Sequential assignments are indicated with one-letter amino acid code and residue number. The side-chain amides of Asn and Gln are indicated by #.

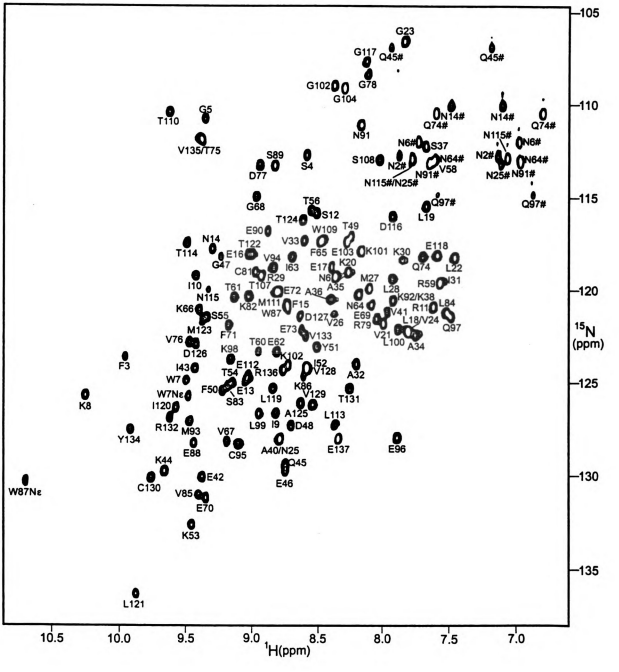
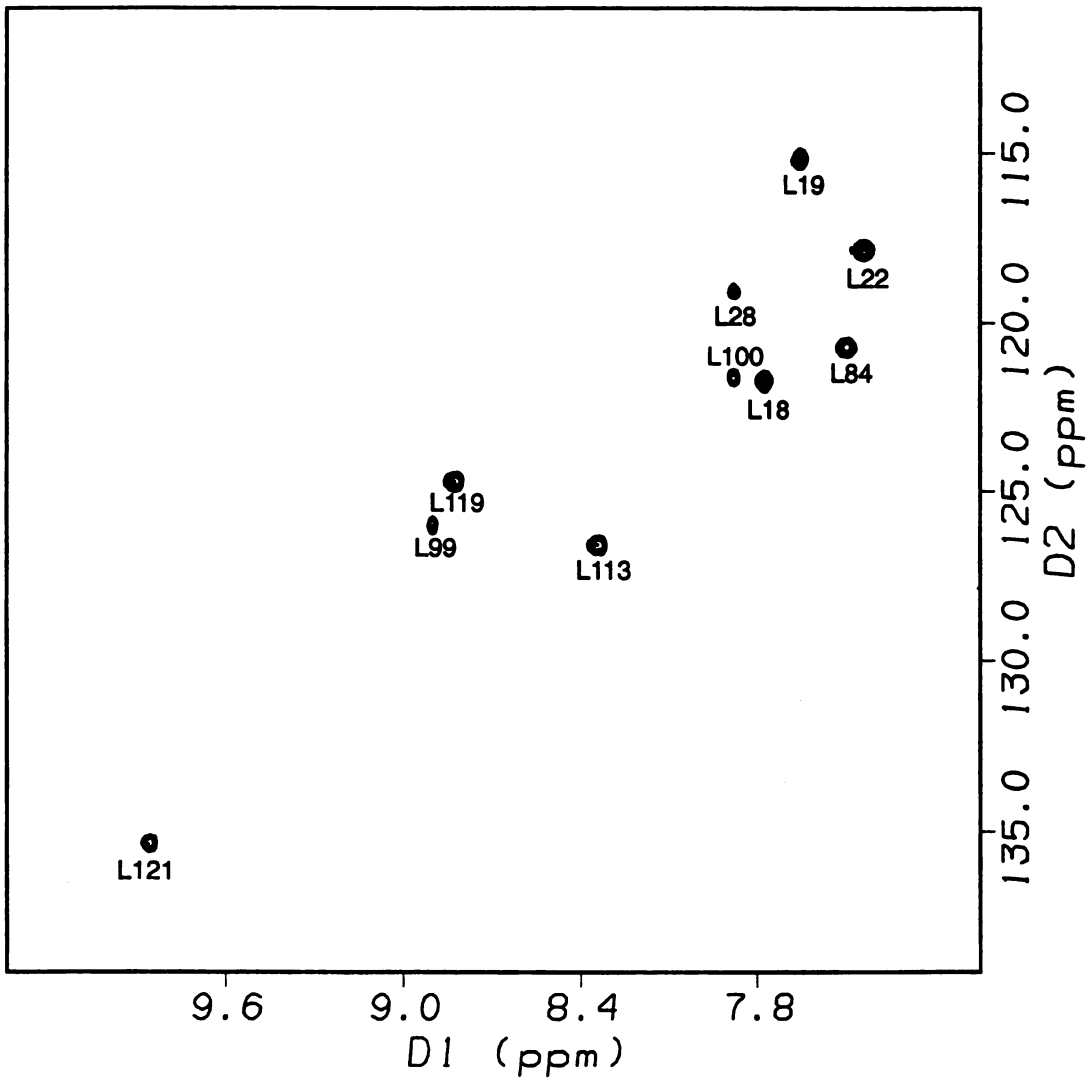

105
110
115
120
125
130
135
^{15}N (ppm)
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
^{1}H(ppm)
G23
G117
Q45#
Q45#
G102
G78
G104
T110
G5
N14#
Q74#
N14#
Q74#
N91
N115#
N6#
V135/T75
S89
S37
N6#
S108
N2#
N64#
N2#
N64#
D77
S4
V58
N91#
N25#
N91#
N115#/N25#
T56
G97#
G97#
G68
S12
L19
T124
E90
W109
T49
D116
N14
T122
V33
E118
T114
E16
V94
F65
E103
K101
K30
Q74
L22
G47
C81
I63
K20
E17
R29
N6
M27
L26
R59
I31
I10
T61
T107
E72
A36
A35
K92/K38
N115
K82
M111
F15
N64
L84
K66
S55
W87
D127
E65
R11
V41
L18/V24
M123
F71
V26
R79
Q97
V76
T60
E62
V133
V21
L100
A34
D126
K98
K102
Y51
F3
I43
E112
I52
T54
R136
V128
A32
W7
K86
W7Nε
E13
V129
T131
K8
S83
L119
I120
A125
L113
R132
M93
V67
L99
I9
D48
Y134
E88
C95
A40/N25
E137
E96
K44
Q45
E46
W87Nε
C130
E42
V85
E70
K53
L121

Figure 4.2. 2D $^{1}H-^{15}N$ HMQC spectrum of the [^{15}N]-leucine selectively labeled apo-R111M. Sequential assignments are indicated with one-letter amino acid code and residue number.


L19
L22
L28
L100
L84
L18
L119
L99
L113
L121
9.6
9.0
8.4
7.8
D1 (ppm)
115.0
120.0
125.0
130.0
135.0
D2 (ppm)

those of arginines. For example, Pro1 was identified in this way. The cross peaks corresponding to H^{ε} protons of four arginines could be identified in the 2D ^{1}H-^{15}N HSQC spectra. Consequently, the H^{ε} was linked to the H^{γ} and H^{β} protons, and eventually to H^{α} protons by the 3D ^{1}H-^{15}N TOCSY-HMQC experiment. The residues belonging to AMX spin system were further distinguished from each other based on distinct chemical shifts and NOEs. Four of the seven serines could be identified by the down field positions of their H^{β} protons. All four asparagine residues could be identified by the NOE cross-peaks between H^{β} and H^{δ} protons observed in the 3D ^{1}H-^{15}N NOESY-HMQC spectrum. All five phenyalanines were identified by selective labeling, and confirmed by the homonuclear 2D NOESY experiments. One tyrosine and two tryptophan residues could be identified after the assignments of their aromatic ring protons. Two glutamines could also be differentiated from the other residues belonging to the same AM(PT)X spin system by the NOE cross-peaks between H^{γ} and H^{ε} protons. All remaining residues were classified into several groups, and identified later along with sequential assignment. The positional ^{1}H resonance assignments of the side-chains were obtained mainly through the 2D DQF-COSY spectrum, and three TOCSY spectra with different mixing times. Remarkably, the hydroxyl protons of both Ser83 and Thr107 were observed in both homonuclear and heteronuclear NMR experiments performed in H_2O (data not shown), and were assigned by the homonuclear TOCSY and NOESY experiments. The indole ε^{1} proton of Trp7 was observed in the NMR spectra recorded on D_2O sample.

The ring protons of aromatic residues. The assignments of aromatic ring protons were derived from the 2D DQF-COSY spectrum, and three TOCSY spectra with different mixing times. The ring protons of phenylalanines and tyrosines were linked to their H^{β},

H^{α} and H^{N} protons by NOE cross peaks between H^{β} and H^{δ} protons and between H^{α} and H^{δ} protons observed in the 2D homonuclear NOESY. The $H^{\varepsilon 1}$ and $H^{\delta 1}$ protons of tryptophans could be assigned by the 3D ^{1}H-^{15}N NOESY-HMQC experiment, and $H^{\varepsilon 3}$ protons by NOEs between H^{α} and $H^{\varepsilon 3}$ protons and between H^{β} and $H^{\varepsilon 3}$ protons observed in the homonuclear 2D NOESY experiments.

Sequential assignment. The sequential assignment (Table 4.1) was primarily based on sequential NOE connectivities such as $d_{\alpha N}$, d_{NN} and $d_{\beta N}$. The assignment was initiated after over sixty percent of the total 137 residues had been identified. Most residues from the ten β-strands and the first α-helix could be easily assigned (Figure 4.3). The sequential assignments of Val26–Lys38 posed a challenge. Val26, Arg29, Ile31 and Lys38 were weak or very weak in the 2D ^{1}H-^{15}N HSQC spectrum. No correlations between their amide and aliphatic protons could be observed in the 3D ^{1}H-^{15}N TOCSY-HMQC spectrum. Only a few NOEs useful for sequential assignments were observed in the 3D ^{1}H-^{15}N NOESY-HMQC spectrum. The ^{1}H chemical shifts of the four alanines (Ala32, Ala34, Ala35 and Ala36) were quite similar. The ^{1}H and ^{15}N resonances of both Val26 and Arg29 were assigned tentatively by comparison with the assignments of wild-type holo-CRABPII (Chapter 5), and a few weak sequential NOE peaks. The amide resonances of both residues became stronger upon RA binding. The aliphatic protons of Val26 were assigned by spin system identification. The assignments of Ile31 and Lys38 were helped greatly by selective labeling, and by comparison with the assignments of wild-type holo-CRABPII. The spin system of Asn115 could be easily identified by the homonuclear NMR experiments but its sequential assignment was achieved only at the later stage because of the weakness of its peaks in all ^{15}N-edited spectra. The H^{α} and H^{β}

Table 4.1. ^{1}H and ^{15}N resonance assignments of apo-R111M in PBS at 27°C

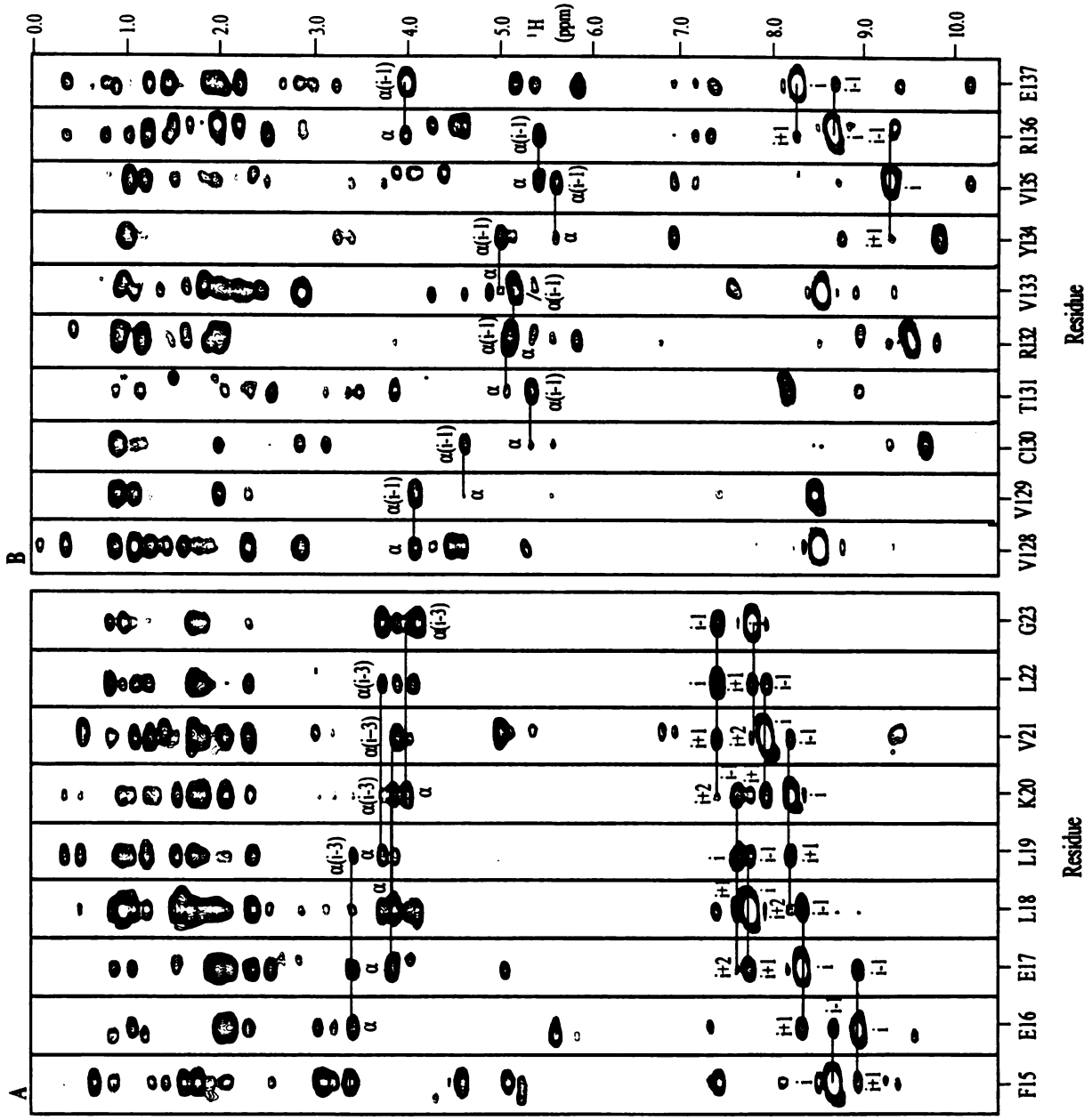
Residue	^{15}N	H^N	H^α	H^β	Others
P1			4.14	2.21, 1.75	H^γ 1.92, 1.94; H^δ 3.13, 3.11
N2			5.11	2.99, 2.83	H^γ 7.64, 6.89; N^γ 112.5
F3	123.2	9.76	4.38	3.40, 3.22	2, 6H 7.43; 3, 5H 7.20; 4H 7.27
S4	112.6	8.39	4.31	3.99, 3.99	
G5	110.7	9.36	4.13, 3.97		
N6	119.0	8.17	5.24	2.74, 2.56	
W7	124.4	9.31	5.08	3.08, 2.91	2H 6.80; 4H 7.04; 5H 6.81; 6H 7.21; 7H 7.11 H^ε 9.28; N^ε 125.2
K8	125.2	10.06	5.27	1.85, 1.85	H^γ 1.65, 1.41; H^δ 0.86, 0.86; H^ε 2.86, 2.86
I9	126.1	8.62	3.67	1.39	H^γ 0.92, 0.92; H^δ 0.60; $H^{\gamma m}$ 0.69
I10	118.9	9.23	4.59	1.87	H^γ 1.21 1.21; H^δ 1.00; $H^{\gamma m}$ 0.81
R11	120.6	7.43	4.60	1.72, 1.72	H^γ 1.59; H^δ 3.07; H^ε 7.29; N^ε 83.8
S12	115.6	8.32	5.28	3.70, 3.61	
E13	124.1	8.83	4.83	1.93, 1.93	H^γ 2.14, 2.21
N14	117.5	9.10	4.96	3.38, 3.27	H^γ 7.29, 6.91; N^γ 109.8
F15	120.5	8.54	4.60	3.14, 2.97	2, 6H 7.21; 3,5H 6.94; 4H 7.02
E16	117.8	8.80	3.32	1.88, 1.88	H^γ 1.98, 2.18
E17	118.5	8.20	3.74	1.97, 1.78	H^γ 2.42, 2.24
L18	121.8	7.62	3.75	1.56, 1.56	H^γ 0.84; H^δ 0.79, 0.79
L19	115.3	7.48	3.64	1.61, 0.81	H^γ 1.09; H^δ 0.38, 0.22
K20	118.7	8.07	3.89	1.96, 1.69	H^γ 1.41, 1.41; H^δ 1.20, 0.95; H^ε 3.22, 2.96
V21	121.4	7.81	3.79	2.19	H^γ 1.10, 0.94
L22	118.0	7.27	3.97	1.70, 1.59	H^γ 0.96; H^δ 0.86, 0.69
G23	106.7	7.65	4.00, 3.62		
V24	121.8	7.62	4.60	1.88	H^γ 0.92, 0.82
N25	127.4	8.59	4.46	3.15, 2.80	
V26	120.7	8.20	3.51	2.05	H^γ 0.95, 0.90
M27	119.6	7.91	4.15	2.09, 2.03	H^γ 2.60, 2.45
L28	119.1	7.74	4.11	1.82, 1.66	H^γ 1.35; H^δ 0.90, 0.79
R29	120.9	8.66			
K30	118.1	7.66	3.92	1.93, 1.93	H^γ 1.74, 1.55; H^δ 1.47, 1.47; H^ε 2.92, 2.92
I31	119.3	7.56	4.07	2.02	H^γ 1.85, 1.85; $H^{\gamma m}$ 1.01
A32	123.5	8.01	4.10	1.37	
V33	117.1	8.41	3.72	2.10	H^γ 0.95, 0.95
A34	122.0	7.58	4.20	1.47	
A35	119.0	8.19	4.35	1.40	
A36	120.2	8.20	4.54	1.40	
S37	112.2	7.48	4.60	3.92, 3.82	
K38	120.1	7.68			
P39			4.90	1.82, 1.40	H^γ 1.07, 0.93; H^δ 3.47, 3.47
A40	127.4	8.60	4.90	1.59	
V41	120.8	7.78	4.98	1.29	H^γ 0.43, 0.41
E42	129.4	9.19	5.40	2.06, 2.10	H^γ 2.30, 2.30
I43	123.7	9.20	5.30	2.25	H^γ 1.66, 1.28; H^δ 0.69; $H^{\gamma m}$ 0.91
K44	129.1	9.48	4.63	1.84, 1.84	H^γ 1.54, 1.35; H^ε 2.86, 2.78

Residue	^{15}N	H^N	H^α	H^β	Others
Q45	127.4	8.59	4.50	1.86, 1.86	H^γ 1.03, 1.03; H^δ 7.71, 6.98; N^δ 106.9
E46	128.9	8.55	4.40	1.84, 1.68	H^γ 2.03, 2.03
G47	117.9	9.04	3.88, 3.62		
D48	126.7	8.51	5.19	3.16, 2.74	
T49	116.9	8.06	4.75	4.18	H^γ 1.02
F50	124.7	9.00	4.60	1.74, 1.74	2, 6H 6.50; 3, 5H 7.06; 4H 6.81
Y51	122.6	8.32	5.20	3.03, 2.77	2, 6H 6.94; 3, 5H 6.56
I52	123.8	8.42	4.72	1.59	H^γ 1.38, 1.28; H^δ 0.59; $H^{\gamma m}$ 0.43
K53	131.9	9.26	5.17	1.77, 1.69	H^γ 1.24, 1.16; H^ε 2.86, 2.86
T54	124.6	8.96	5.10	3.89	H^γ 0.96
S55	121.2	9.16	5.34	3.75, 3.75	
T56	115.5	8.36	4.85	4.72	H^γ 1.16
T57	110.9	8.34	4.05	4.27	H^γ 1.31
V58	113.0	7.43	4.19	2.02	H^γ 0.82, 0.82
R59	119.3	7.38	4.60	1.72, 1.72	H^γ 1.40, 1.40
T60	122.9	8.76	5.23	3.94	H^γ 1.11
T61	120.0	8.95	4.65	4.01	H^γ 1.07
E62	122.9	8.62	5.13	1.93, 1.93	H^γ 2.10, 2.04
I63	117.9	8.50	4.85	1.86	H^γ 1.36, 0.79; H^δ 0.91; $H^{\gamma m}$ 1.11
N64	119.9	8.00	5.59	2.51, 2.51	H^γ 7.51, 6.81; N^γ 113.1
F65	117.1	8.29	4.67	2.28, 2.28	2, 6H 6.20; 3, 5H 6.52; 4H 6.26
K66	120.7	9.20	5.31	1.67, 1.54	H^γ 1.32, 1.32; H^ε 2.96, 2.96
V67	127.5	8.83	3.05	1.67	H^γ 0.46, 0.25
G68	114.8	8.77	4.40, 3.63		
E69	120.4	7.90	4.81	2.00, 2.00	H^γ 2.22, 2.22
E70	130.5	9.18	4.91	2.02, 1.94	H^γ 2.24, 2.15
F71	121.5	8.99	5.15	3.51, 3.22	2, 6H 7.26; 3, 5H 7.81; 4H 6.55
E72	119.7	8.61	5.09	2.06, 1.97	H^γ 2.34, 2.16
E73	121.8	8.42	4.63	1.98, 1.90	H^γ 2.14, 2.14
Q74	117.9	7.78	5.35	1.79, 1.79	H^γ 2.20, 2.03; H^δ 7.35, 6.58; N^δ 110.3
T75	111.8	9.18	4.62	4.18	H^γ 1.06
V76	122.4	9.26	3.79	1.98	H^γ 0.86, 0.86
D77	113.1	8.75	4.40	2.74, 1.98	
G78	108.4	7.93	4.23, 3.58		
R79	121.8	6.86	4.63	1.82, 1.58	H^γ 1.36, 0.69; H^δ 3.21, 3.12; H^ε 8.59; N^ε 87.9
P80			4.70	2.42, 1.90	H^γ 2.05, 2.17; H^δ 4.03, 3.66;
C81	118.7	8.78	5.08	2.64, 2.13	
K82	120.0	8.83	4.79	1.58, 1.51	H^γ 1.28; H^ε 2.82;
S83	124.9	9.02	4.93	2.80, 1.55	H^γ 6.26
L84	120.9	7.33	4.24	1.61, 1.29	H^γ 1.19; H^δ 0.71, 0.71
V85	130.3	9.21	4.40	1.32	H^γ 0.22, –0.03
K86	123.7	8.39	4.52	1.64, 1.51	H^γ 1.31, 1.22
W87	120.5	8.54	5.02	3.28, 2.99	2H 7.38; 4H 7.98; 5H 6.86; 6H 7.21; 7H 7.58 H^ε 10.50 (10.57)*; N^ε 129.6
E88	127.7	9.25	4.38	1.95, 1.76	H^γ 2.22, 2.12
S89	113.2	8.63	4.61	4.18, 4.01	
E90	116.6	8.69	3.77	1.87, 1.87	H^γ 2.12, 2.12
N91	111.1	7.98	4.01	3.12, 2.92	H^γ 7.37, 6.70; N^γ 112.6
K92	120.2	7.74	5.71	1.96, 1.96	H^γ 1.50, 1.50; H^δ 1.23, 0.93; H^ε 2.71, 2.56
M93	126.5	9.27	5.16	1.22, 1.22	H^γ 1.75, 1.75; H^ε 0.31

Residue	^{15}N	H^N	H^α	H^β	Others
V94	118.5	8.64	4.24	1.77	H^γ 0.75, 0.69
C95	127.6	8.91	4.46	0.51, –0.10	
E96	127.4	7.70	4.32	1.86, 1.80	H^γ 2.17, 1.97
Q97	121.0	7.31	4.92	2.03, 1.85	H^γ 2.50, 2.15; H^δ 7.51, 6.68; N^δ 114.8
K98	123.6	8.97	4.60	1.57, 1.39	H^γ 1.29; H^ε 2.85
L99	126.1	8.75	4.00	1.67, 1.51	H^γ 1.39; H^δ 0.69, 0.69
L100	121.7	7.70	4.10	1.47, 1.47	H^γ 0.78
K101	117.6	7.97	4.40	1.82, 1.82	H^γ 1.66, 1.66; H^δ 1.29, 1.29
G102	109.0	8.18	4.06, 3.75		
E103	117.1	8.08	4.63	1.79, 1.79	H^γ 2.12, 1.99
G104	109.1	8.10	4.12, 3.95		
P105			4.54	2.08, 2.08	H^γ 1.87, 1.80; H^δ 3.63, 3.35
K106	123.6	8.54	4.42	1.85, 1.85	H^γ 1.56, 1.40; H^δ 1.81, 1.81; H^ε 2.98, 2.98
T107	118.9	8.75	5.38	4.50	H^γ 1.17, $H^{\gamma 1}$ 6.04
S108	112.9	7.84	4.75	3.83, 3.83	
W109	117.1	8.29	5.72	3.58, 3.33	2H 6.87; 4H 6.83; 5H 6.41; 6H 6.41; 7H 6.71 H^ε 11.17; N^ε 136.4
T110	110.4	9.44	5.15	4.19	H^γ 1.07
M111	119.7	8.62	5.46	1.86, 1.83	H^γ 2.41, 2.25
E112	124.3	8.85	5.78	1.93, 1.93	H^γ 2.09, 2.09
L113	126.7	8.17	5.29	1.31, 0.77	H^γ 1.14; H^δ 0.51, 0.48
T114	117.2	9.29	4.62	4.52	H^γ 1.08
N115	119.6	9.14	4.46	2.82	H^γ 7.54, 6.83; N^γ 112.7
D116	115.8	7.75	4.77	2.82, 2.46	
G117	107.7	7.94	4.21, 3.73		
E118	117.8	7.40	5.03	2.04, 1.92	H^γ 2.35, 2.20
L119	124.8	8.64	5.04	0.95, 0.84	H^γ 0.30; H^δ –0.31, –0.38
I120	125.8	9.38	4.85	1.84	H^γ 1.10, 1.10; H^δ 1.04; $H^{\gamma m}$ 0.79
L121	135.4	9.69	5.54	1.92, 1.36	H^γ 1.60; H^δ 0.77, 0.71
T122	117.8	8.85	5.75	4.08	H^γ 1.08
M123	121.2	9.20	5.20	1.87, 1.84	H^γ 2.45, 2.37
T124	116.0	8.42	5.50	3.97	H^γ 0.99
A125	125.6	8.44	4.88	1.06	
D126	122.5	9.23	4.17	2.77, 2.73	
D127	121.0	8.44	4.54	2.70, 2.70	
V128	123.7	8.39	3.99	2.18	H^γ 0.96, 0.74
V129	125.7	8.35	4.55	1.87	H^γ 0.81, 0.74
C130	129.5	9.57	5.25	3.03, 2.74	
T131	124.8	8.06	5.00	3.77	H^γ 1.02
R132	126.3	9.43	5.04	1.87, 1.84	H^γ 1.75, 1.75; H^δ 3.02, 1.96; H^ε 7.17; N^ε 84.4
V133	122.0	8.42	4.92	1.70	H^γ 0.89, 0.84
Y134	126.9	9.73	5.54	3.30, 3.17	2, 6H 6.82; 3, 5H 6.68
V135	111.8	9.19	5.34	2.38	H^γ 1.08, 0.91
R136	123.8	8.56	3.89	1.31, 1.09	H^γ 0.62, 0.23; H^δ 2.82, 2.70; H^ε 6.86; N^ε 85.2
E137	127.4	8.15	3.93	1.85, 1.75	H^γ 2.09, 2.09

* Two sets of resonances.

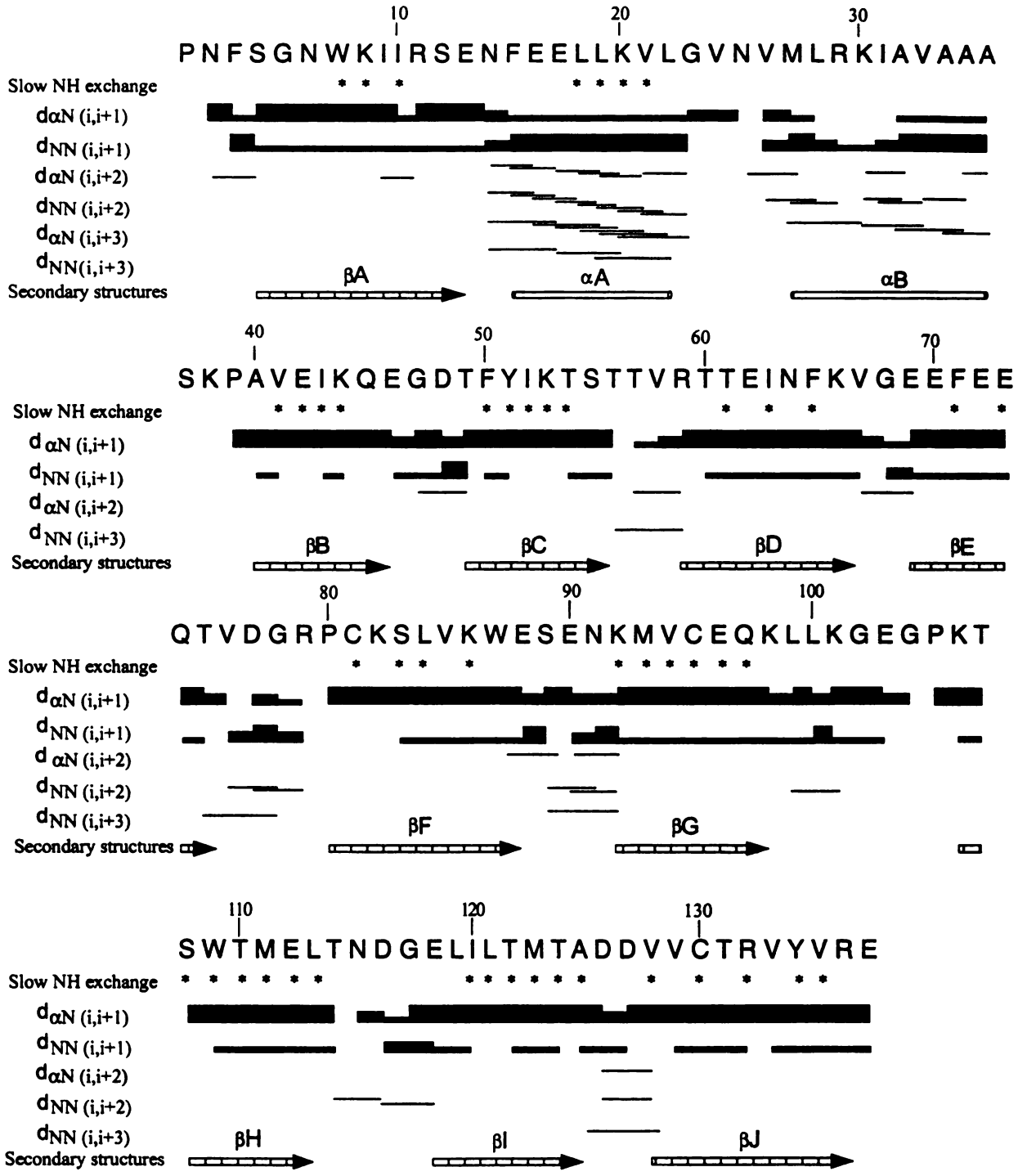
Figure 4.3. Strip plots extracted from the 3D ^{15}N-edited NOESY-HMQC spectrum of apo-R111M showing the sequential and medium-range NOE connectivies of α-helix (αA) (A) and β-strand (βJ) (B).


A
B
^{1}H (ppm)
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
F15
E16
E17
L18
L19
K20
V21
L22
G23
V128
V129
C130
T131
R132
V133
Y134
V135
R136
E137
Residue
Residue
α
α(i-1)
α(i-3)
i
i+1
i+2
i-1

protons of Thr114, whose chemical shifts were close to the water signal, were not observed in the two sets of the 3D spectra due to water presaturation. They were assigned by the analysis of the 2D homonuclear spectra recorded in D_2O helped by the comparison with their assignments in wild-type holo-CRABPII. Asn2 and Thr57 were the only two residues not observed in the 2D ^{1}H-^{15}N HSQC spectrum. Their aliphatic protons were assigned by spin system identification after the sequential assignments. The side-chain assignments were complete except for Arg11, Leu22, Arg29, Ile31, Arg59 and eight lysine residues. Stereospecific assignments were made for the β-methylene protons of 51 residues and the methyl protons of ten valine and six leucine residues.

Secondary structure determination. Regular secondary structure elements were characterized by their distinct NOE patterns. The following ten strands forming two anti-parallel β-sheets could be identified (Kleywegt et al., 1994) (Figure 4. 4): Ser4–Glu13 (βA), Pro39–Glu46 (βB), Thr49–Thr56 (βC), Thr60–Lys66 (βD), Glu69–Thr75 (βE), Pro80–Glu88 (βF), Lys92–Lys98 (βG), K106–L113 (βH), E118–A125 (βI) and V128–R136 (βJ). The second strand βB began at Pro39, one residue shorter than the corresponding βB in the crystal structure of holo-CRABPII, where it began at Lys38. Lys38 was very weak in the 2D ^{1}H-^{15}N HSQC spectrum. Long-range NOEs were observed between the backbone protons of all adjacent β-strands except between βD and βE. There were also NOEs between the backbone protons of the N-terminal half of βA and the C-terminal half of βJ. The first two strands βA and βB were connected by two anti-parallel α-helices. The first (αA), Phe15–Leu22, showed NOE connectivity and hydrogen bonding patterns typical of a regular α-helical structure. The N-terminal half

Figure 4.4. Summary of the sequential and medium-range NOEs involving backbone H^N and H^α atoms, slow-exchange backbone amide protons and the deduced secondary structures of apo-R111M. Line thickness for $d_{\alpha N}$ and d_{NN} sequential NOE distances reflects the intensities of the cross peaks. Asterisks indicate the residues whose amide protons remained after 24 h of exchange with D_2O in PBS buffer at room temperature and pD 7.5 as determined by 2D homonuclear NMR experiments.


10
20
30
PNFSGNWKIIRSENFEELLKVLGVNVMLRKIAVAAA
Slow NH exchange
dαN (i,i+1)
dNN (i,i+1)
dαN (i,i+2)
dNN (i,i+2)
dαN (i,i+3)
dNN(i,i+3)
Secondary structures
βA
αA
αB
40
50
60
70
SKPAVEIKQEGDTFYIKTSTTVRTTEINFKVGEEFEE
Slow NH exchange
dαN (i,i+1)
dNN (i,i+1)
dαN (i,i+2)
dNN (i,i+3)
Secondary structures
βB
βC
βD
βE
80
90
100
QTVDGRPCKSLVKWESENKMVCEQKLLKGEGPKT
Slow NH exchange
dαN (i,i+1)
dNN (i,i+1)
dαN (i,i+2)
dNN (i,i+2)
dNN (i,i+3)
Secondary structures
βF
βG
110
120
130
SWTMELTNDGELILTMTADDVVCTRVYVRE
Slow NH exchange
dαN (i,i+1)
dNN (i,i+1)
dαN (i,i+2)
dNN (i,i+2)
dNN (i,i+3)
Secondary structures
βH
βI
βJ

(Leu28–Ile31) of the second helix appeared to be partially unwound. The strength and number of the NOEs observed for residues Leu28-Ile31were well below what is expected for a regular α-helix. However, the chemical shifts of these residues indicated that they are predominately in a helical conformation. No slow exchanging amide protons were observed for the residues in the second helix. The loop (Gly23–Val26) between the two helices could be detected but the link (Ser37–Pro39) connecting αB and βB appeared to undergo conformational exchange in solution.

Structural restraints. A total of 2302 structurally useful distance restraints were obtained by the analysis of the homonuclear 2D NOESY and heteronuclear 3D ^{1}H-^{15}N NOESY-HMQC data. Of the 2302 distance restraints, 704 were from long-range NOEs (Table 4.2). The distributions of the distance restraints along the amino acid sequence were quite uniform except for the two N-terminal residues, the residues in turns or loops and in the second α-helix (Figure 4.5). The number of NOE restraints for each of them was well below the average. A total of 77 dihedral angle restraints were obtained from the 2D ^{1}H-^{15}N HMQC-J data. A total of 98 hydrogen bond restraints from 49 hydrogen bonds were included in the structural refinement.

Quality of the structures. Figure 4.6A shows the superposition of 28 refined structures with no NOE violation exceeding 0.25 Å and no dihedral angle violation larger than 5.0° The statistics of the NMR structures is summarized in Table 4.2. The stereochemical qualities of the structures were checked by the program PROCHECK–NMR as described in Chapter 3. The Ramachandran plots and overall G-factors showed that the stereochemical qualities of the structures were comparable to those X-ray structures with 2.5 Å resolution and R-factors less than 20 % (Table 4.2).

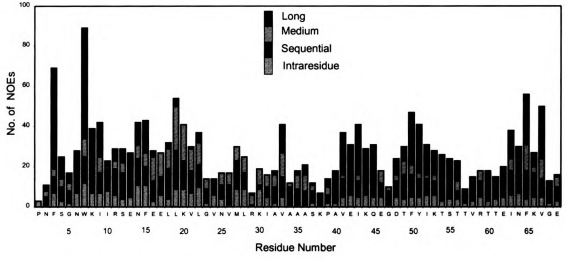
Table 4.2. Restraint and structural statistics of apo-R111M

Restraint Statistics		
Numbers of experimental NOE restraints		
Intra	711	
Sequential	635	
Medium	252	
Long	704	
Total	2302	
Number of experimental ϕ dihedral restraints	77	
Number of hydrogen bonds	98	
Structural Statistics		
	<SA>[a]	$<SA>_r$[b]
RMSD from experimental distance restraints (Å) (2302)	0.032 ± 0.001	0.024
RMSD from experimental dihedral restraints (deg) (77)	0.926 ± 0.102	0.836
Deviation from idealized covalent geometry		
Bonds (Å)	0.004 ± 0.001	0.003
Angles (deg)	0.621 ± 0.028	0.520
Impropers (deg)	0.495 ± 0.004	0.411
Measures of Structural Quality (By Procheck)		
% residues in most favorable regions of the Ramachandran plot	71.8 ± 2.4	71.0
% residues in additional allowed regions of the Ramachandran plot	25.8 ± 0.8	25.8
No. of bad contacts	6 ± 2	6
H-bond energy	0.40 ± 0.10	0.50
Overall G-factor	-0.19 ± 0.01	-0.20
Coordinate Precision		
RMSD for C_α trace (Å)	0.43 ± 0.14	
RMSD for backbone atoms (Å)	0.54 ± 0.26	
RMSD for all heavy atoms (Å)	0.98 ± 0.23	

[a] <SA> are the final 28 simulated annealing structures.

[b] $<SA>_r$ is the restrained minimized mean structure obtained by restrained regularization of the mean structure, which is obtained by averaging the coordinates of the individual SA structures best fitted to each other.

Figure 4.5. The distribution of NOEs along the amino acid sequence of CRABPII. For intra-residue NOEs only those that are structurally useful are included.


Long
Medium
Sequential
Intraresidue
No. of NOEs
Residue Number

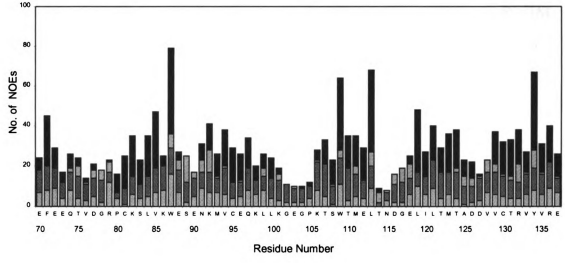
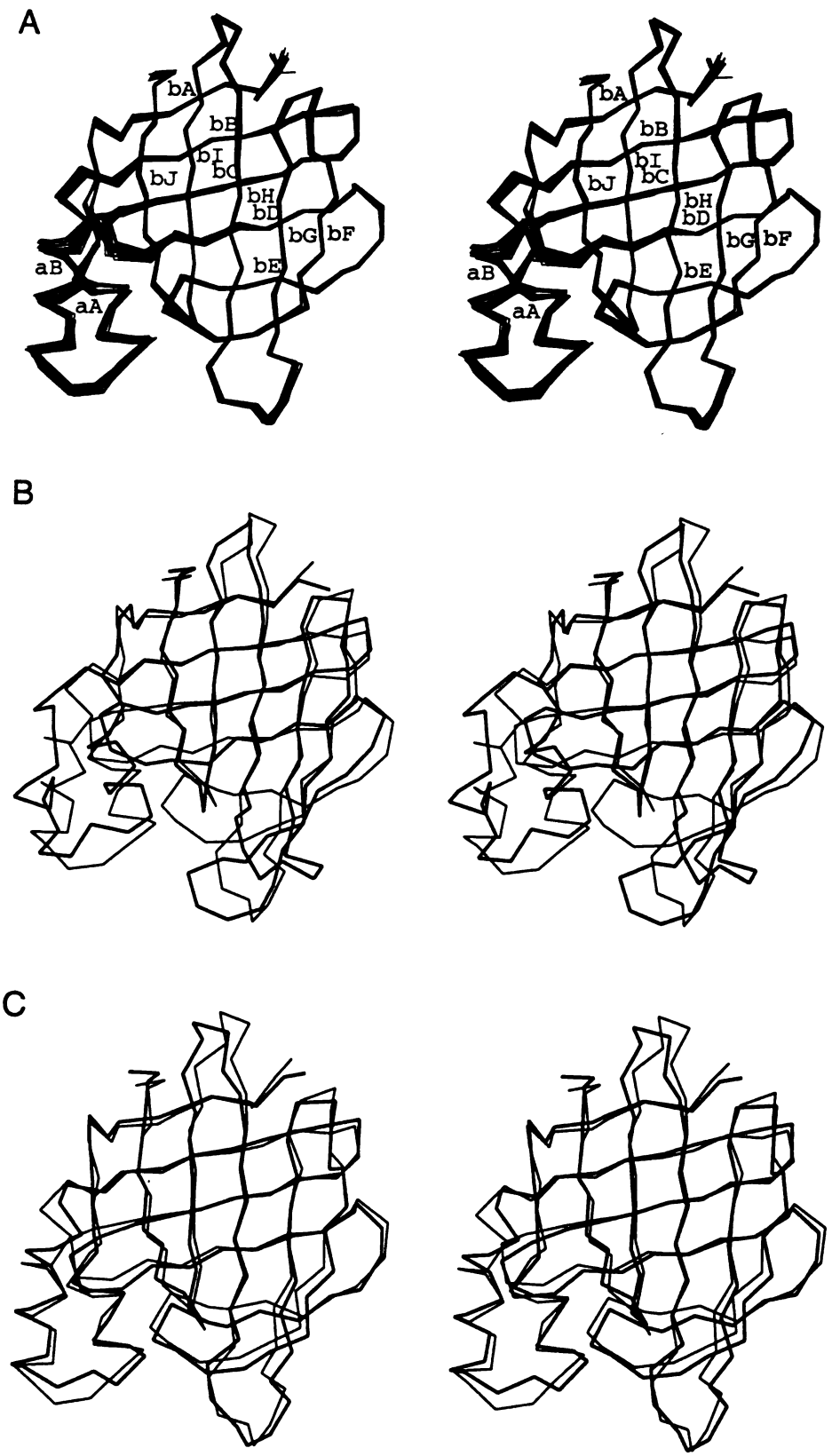

No. of NOEs
Residue Number

Figure 4.6. (A) Stereoview of the C^{α} traces of the superimposed 28 refined solution structures of apo-R111M. (B) Stereoview of the superimposed C^{α} traces of the restrained minimized mean structures of apo-R111M (thin line) and apo-CRABPII (thick line). (C) Stereoview of the C^{α} traces of the restrained minimized mean structure of apo-R111M (thin line) superimposed with the C^{α} traces of the crystal structure of holo-CRABPII (thick line).


A
bA
bB
bI
bC
bJ
bH
bD
bG
bF
aB
bE
aA
bA
bB
bI
bC
bJ
bH
bD
bG
bF
aB
bE
aA
B
C

Description of R111M solution structure. The tertiary structure of apo-R111M was very similar to the crystal structure of CRABPs (Figure 4.6C) (Kleywegt et al., 1994; Thompson et al., 1995). The protein formed a compact, single domain structure consisting of two antiparallel β-sheets and two short α-helices. Both β-sheets consisted of five β-strands and were twisted into a nearly orthogonal flattened barrel. Inside the protein a binding pocket was formed by the side-chains from eight of the ten β-strands and lined by both hydrophobic and hydrophilic residues. It was closed off from the exteriors except for a small tunnel-like opening. The entrance was formed mainly by the hydrophobic side-chains from one side of helix αB, the βC-βD loop and the βE-βF loop, as well as Leu19 and Val24. The loop connecting αA and αB extended into the solvent. In solution the ligand entrance was slightly more flexible than the other parts of the protein, but it was still relatively well defined in the solution structure.

Discussion

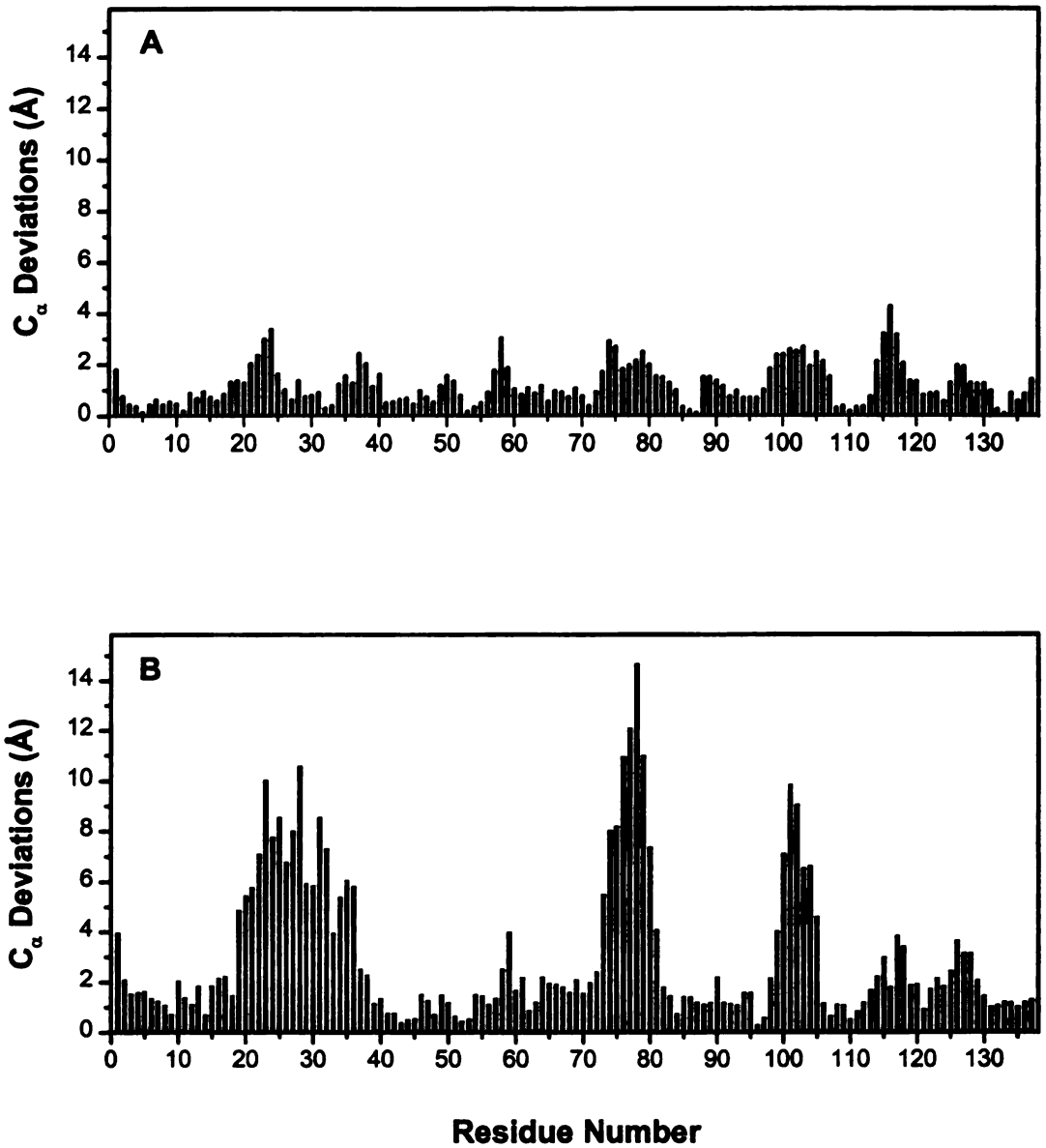
Assignment. The sequential assignment of the apo-R111M posed a challenge for amide-directed sequential assignment strategy based on a series of homonuclear 2D and heteronuclear 2D and 3D ^{15}N-edited NMR experiments. The problem was caused by the presence of three prolines in the middle of the amino acid sequence, the low percentage of the residues whose spin systems could be identified unambiguously by homonuclear 2D NMR experiments, and the relatively low intensities of a dozen cross peaks in the 2D ^{1}H-^{15}N HSQC spectrum. The intensities of these weak peaks were even lower in the 3D ^{1}H-^{15}N NOESY-HMQC spectrum and the lowest in the 3D ^{1}H-^{15}N TOCSY-HMQC spectrum. Moreover, eight residues had H^{α} protons signals and one residue had an H^{β} proton signal quite close to the solvent resonance. These protons were not observed in the two sets of 3D ^{15}N-edited data. Consequently, their sequential assignments could not be achieved by the 3D ^{1}H-^{15}N NOESY-HMQC experiment. Most of these difficulties were overcome by selective ^{15}N-labeling. Lysine, leucine, isoleucine and phenylalanine were chosen because of the abundance of these residues in the protein, and the difficulty in identifying them by the standard procedure (Wüthrich, 1986). After their identifications by selective ^{15}N-labeling, the spin systems of over sixty percent of the total residues were known, which made it possible to achieve the total sequential assignment.

Comparison of the solution structures of apo-R111M and wild-type apo-CRABPII. The backbone fold of apo-R111M is quite similar to that of apo-CRABPII.

The restrained minimized mean NMR structures of the two proteins can be superimposed with a RMSD of 3.2 Å for all backbone C^{α} atoms (Figure 4.6B and Figure 4.7C), a RMSD of 2.3 Å for the backbone C^{α} atoms of the ten β-strands and the first α-helix. However, there are significant conformational differences between the two structures, mainly localized to three segments (Leu19-Ala36, Glu73-Cys81 and Leu99-Pro105) clustered around the ligand entrance. Surprisingly, the conformational differences occur in the region far away from the site of the point mutation. If these segments are excluded, the C^{α} RMSD between the two structures is reduced to 1.6 Å. The first segment (Leu19-Ala36) encompasses the middle of the first helix to the end of the second helix, the second segment (Glu73-Cys81) the end of βE and the beginning of the βF, and the third segment (Leu99-Pro105) the long βG–βH loop. In apo-R111M, all three segments move towards the center of the ligand entrance so that the opening of the ligand binding pocket is much smaller than that in wild-type apo-CRABPII. The binding pocket in apo-R111M is much less exposed to the external solvent than in the wild-type structure. For example, the distance between the C^{α} atoms of Val58 in the βC–βD loop and Val24 in the turn connecting two helices is 7.4 Å shorter in apo-R111M than in wild-type apo-CRABPII. The distance between the C^{α} atoms of Val58 and Asp77 in the βE–βF loop is 10.8 Å shorter in apo-R111M than in wild-type apo-CRABPII.

Comparison of the solution structures of apo-R111M and the crystal structure of wild-type holo-CRABPII. The restrained minimized mean NMR structure of apo-R111M and the crystal structure of wild-type holo-CRABPII could be superimposed very well (Figure 4.6C). The RMSD between the two superimposed structures is 0.96 Å for the C_{α} atoms of all regular secondary structures and 1.38 Å for all C_{α} atoms. The

Figure 4.7. C^{α} atom deviations between apo-R111M and wild-type holo-CRABPII (A) and between apo-R111M and wild-type apo-CRABPII (B).


A
B
C_α Deviations (Å)
C_α Deviations (Å)
0
2
4
6
8
10
12
14
0
10
20
30
40
50
60
70
80
90
100
110
120
130
Residue Number

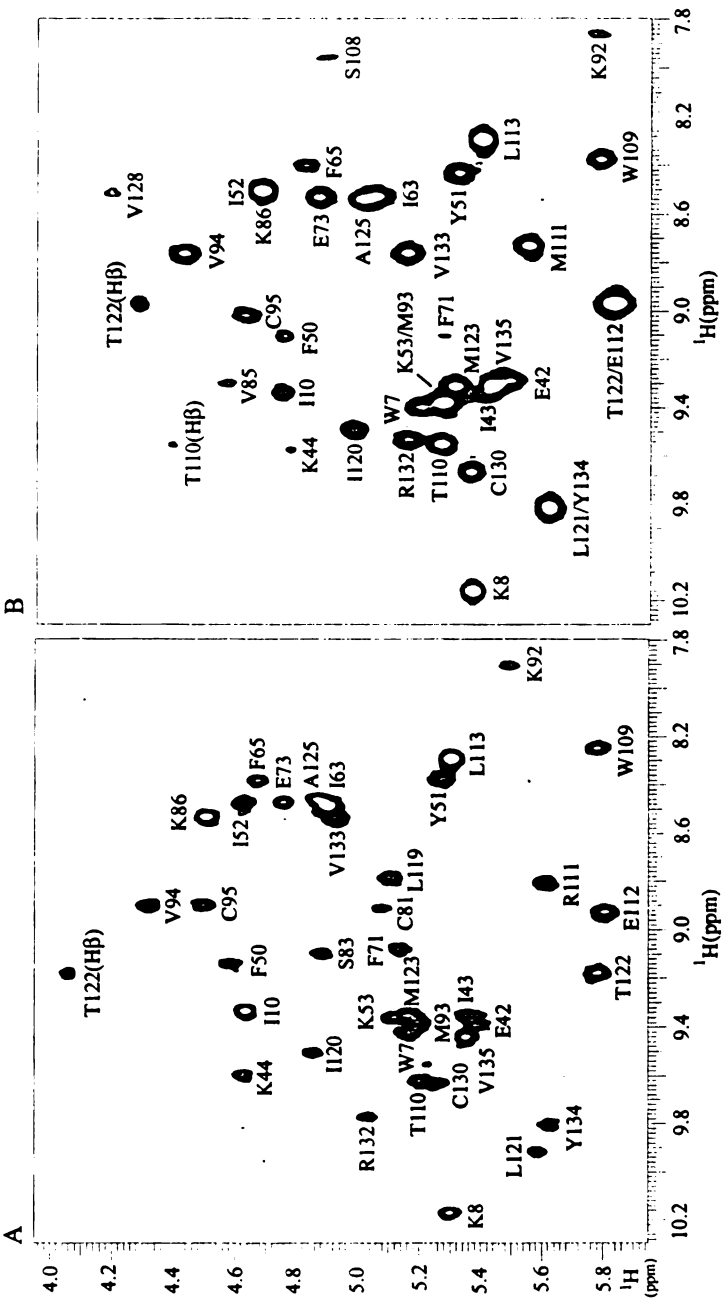
superposition also shows good alignments of the side-chains that are involved in binding of RA such as those of Arg132 and Tyr134.

It is obvious from examining Figure 4.7 that the solution structure of apo-R111M is more similar to the crystal structure of holo-CRABPII than to the solution structure of apo-CRABPII. The ligand entrance of apo-R111M is much smaller than that of the apo-CRABPII, but is still larger than that of holo-CRABPII. The distance between the C^{α} atoms of Val58 in the βC–βD loop and Val24 in the turn connecting two helices is 15.0, 21.3 and 28.8 Å in holo-CRABPII, apo-R111M and apo-CRABPII, respectively. The distance between the C^{α} atoms of Val58 and Asp77 in the βE–βF loop is 12.4, 13.4 and 24.4 Å in holo-CRABPII, apo-R111M and apo-CRABPII, respectively. It appears that the point mutation and the binding of RA cause similar conformational changes, although the magnitudes of the changes induced by the mutation are not as large as those caused by the binding of RA.

Dynamical properties of apo-R111M. Apo-R111M was quite stable in solution. The NMR sample lasted for several months with almost no changes in the spectral properties and quality. More than 50 % backbone amide protons in the first α-helix and two β-sheets remained after a day of exchange with D_2O (Figure 4.8). The hydroxyl protons of Thr107 and Ser83 were observed in the homonuclear 2D spectra recorded in H_2O, and the indole ε^1 proton of Trp7 was observed in the homonuclear 2D spectra recorded in D_2O, indicating that these protons had very low solvent accessibility. All these observation suggested that R111M is for the most part fairly stable and rigid.

The peak intensity of a 2D ^1H-^{15}N HSQC spectrum is, in general, modulated by both amide proton exchange and conformational exchange. Amide proton exchange

Figure 4.8. The fingerprint regions of the homonuclear 2D TOCSY spectra of apo-CRABPII (A) and apo-R111M (B) recorded with 40 ms mixing time. The TOCSY experiment was initiated more than 24 h after dissolving a lyophilized protein sample in PBS, pD 7.5, in D_2O.


A
B
^{1}H (ppm)
^{1}H(ppm)
T122(Hβ)
V94
C95
K86
I52
F65
E73
A125
I63
V133
L119
C81
Y51
L113
K92
S83
F71
M123
I43
K53
M93
E42
I110
F50
K44
W7
C130
V135
I120
R132
T110
K8
L121
Y134
R111
T122
E112
W109
V128
T110(Hβ)
V85
K53/M93
L121/Y134
T122/E112
M111
S108

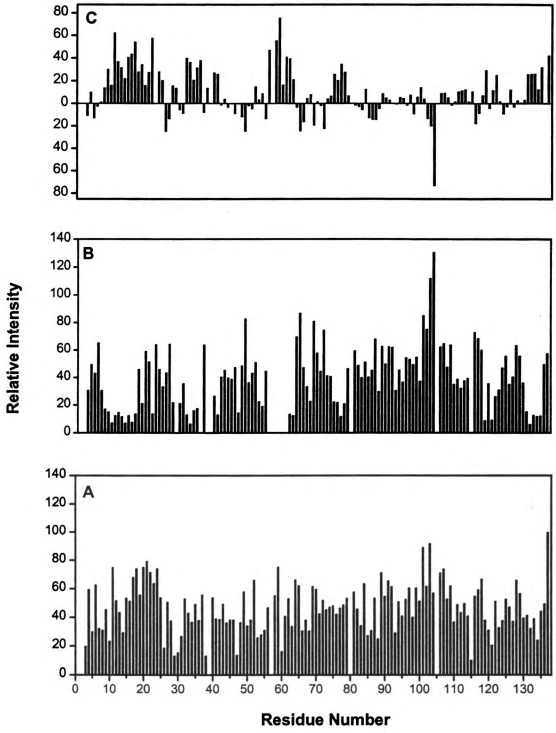
appeared to have little effects on the peak intensities of the HSQC spectra of apo-R111M and apo- CRABPII (Chapter 2), partly owing to the inclusion of a "water flip-back" pulse (Grzesiek & Bax, 1993) in the HSQC experiments that minimized water saturation and dephasing. In fact, the average peak intensity of the slow exchanging amide protons was the same as that of all amide protons. Thus, variation in peak intensity is primarily due to conformational exchange. As expected, the residues in the loops such as Gly47 and Asn115 had peak intensities lower than the average (Figure 4.1). In general, the turns or loops were more mobile than the regular secondary structures. Of the all secondary structural elements only helix αB, especially its N-terminal half Leu28–Ile31, were more flexible than the others. All the amide protons in this helix exchanged completely with the solvent in less than 12 hours. By contrast, most of the amide protons of the first helix remained after more than two days of exchange with D_2O. The peaks of Leu28, Arg29, Lys30 and Ile31 were much weaker than the average (Figure 4.1 and 4.7A). Val26 and Lys38 had very weak peaks, and Thr57 was not observed in 2D ^{1}H-^{15}N HSQC spectrum. Furthermore, the number of NOEs observed for these residues was very low. Taken together, these results suggest that conformational exchanges are occurring in the αA-αB turn, the αB-βB link and the βC-βD loop. In other words, the ligand entrance was slightly more mobile than most of the other parts of the molecule.

Did mutation change the dynamical properties of CRABPII? For the most part, apo-R111M and wild-type apo-CRABPII had similar dynamic properties. Thus, both proteins had about the same number of slow-exchanging backbone amide proteins. Furthermore, the identities of most of the amide protons were the same for the two proteins (Figure 4.7). However, apo-CRABPII appeared to be more mobile than apo-

R111M. The spectral qualities of the apo-CRABPII deteriorated in about two weeks as opposed to several months for apo-R111M. For most peaks, the wild-type protein had slightly larger line width than the mutant so that the DQF-COSY and TOCSY experiments did not work as efficiently for the wild-type as for the mutant.

To compare the dynamic properties of the two proteins in more detail, we examined the intensities of the peaks in their 2D ^{1}H-^{15}N HSQC spectra. The intensities of the 2D ^{1}H-^{15}N HSQC spectrum of the mutant were quite uniform (Figure 4.7A). As mentioned earlier, of all secondary structure elements, only three residues (Arg29-Ile31) at the flexible N-terminal half of the second helix showed weak peaks. The rest of the weak peaks invariably belonged to the residues in the loops or turns. However, the wild-type protein displayed large variations in peak intensities (Chapter 3). Most residues in the RA-binding pocket and near the ligand entrance showed weak or no peaks in the 2D ^{1}H-^{15}N HSQC spectrum, indicating that the RA-binding pocket, especially the ligand entrance, is rather dynamic. Most of these peaks became much stronger after the mutation (Figure 4.9C). In fact, most of the residues with significant intensity differences between the two spectra were clustered in these regions. Of the 21 residues that constitute the RA-binding pocket, all but three (Ile31, Thr54 and Met123) showed significant increases in peak intensities after the point mutation. In particular, the mutation stabilized the β-bulge structure composed of residues Ile9-Arg11 and Thr131-V133. This may be due to the attenuation of the flexibility of its neighboring residues Leu28-Ala36. The same phenomenon was observed for the residues Ser12-Glu17, which may be caused by the attenuation of the fluctuations of the two helices (αA and αB). The residues Lys8, Ala40,

Figure 4.9. Relative peak intensities of the 2D ^{1}H–^{15}N HSQC spectra of apo-R111M (A) and wild-type apo-CRABPII (B) as a function of the residue number. (C) was obtained by subtracting (B) from (A).


C
B
A
Relative Intensity
Residue Number
80
60
40
20
0
20
40
60
80
140
120
100
80
60
40
20
0
0
10
20
30
40
50
60
70
80
90
100
110
120
130

Lys53, Glu62, Ile63, Leu119 and Val135 also became much stronger after the point mutation. Their implications will be discussed later.

Binding of RA substantially stabilized CRABPI as measured by thermal denaturation. In particular, it stabilized the flexible parts of the ligand binding pocket in CRABPI as revealed by NMR (Rizo et al., 1994). Our NMR studies showed that the changes in dynamical properties caused by the point mutation were rather similar to those induced by RA binding (Chapter 5). Thus, the dynamic properties of apo-R111M were more similar to those of holo-CRABPI and holo-CRABPII than to apo-CRABPI and apo-CRABPII. Binding of RA and the mutation caused similar changes in the dynamical properties of CRABPII.

Roles of Arg111. Arg111 interacts with the carboxyl group of RA via a hydrogen bond mediated by a water molecule. In Chapter 2 I have shown that Arg111 contributes to the binding of RA by ~2.2 kcal/mol, indicating that Arg111 is indeed important for the binding of RA. X-ray crystallographic and molecular modeling studies suggest that Arg111 is one of the four arginine residues that generate a positive potential that may facilitate the binding of RA (Zhang et al., 1992). The NMR results presented in this chapter suggest an additional role for Arg111 related to the structure and dynamics of CRABPII.

As mentioned earlier, RA binding induces significant but localized conformational and dynamical changes in both CRABPI and CRABPII. The ligand entrance, in particular, is much more open and flexible in the apo-forms than in the holo-forms. The large opening of the ligand entrance and its high flexibility are probably essential for the unhindered entry of RA into the deep binding pocket and the subsequent

closure of the ligand entrance. Our NMR results show that the point mutation also causes significant but localized changes in the conformation and dynamics of CRABPII. Strikingly, apo-R111M is more similar to wild-type holo-CRABPII than to wild-type apo-CRABPII in both structure and dynamics, suggesting that Arg111 may play a major role in maintaining the ligand entrance in an open and dynamic state for the entry of RA.

The next question is how Arg111 affects the internal motions of CRABPII. At first thought, the interactions leading to the high flexibility of apo-CRABPII was the electrostatic repulsion between the positively charged guanidinium groups of two conserved arginines, Arg111 and Arg132 in CRABPII and Arg111 and Arg131 in CRABPI (Zhang et al., 1992). However, the two guanidinium groups in apo-CRABPII were ~7.6 Å apart. Furthermore, if the electrostatic repulsion played a critical role substitution of Arg132 with methionine would most probably change the dynamical properties to the same extent. However, this was not supported by our NMR characterization of apo-R132M. The qualities of the homonuclear 2D spectra of apo-R132M were slightly better than those of the wild-type but worse than those of the apo-R111M, suggesting that apo-R132M is less flexible than apo-CRABPII but not as rigid as apo-R111M. For most peaks apo-R111M had the smallest line widths among all mutants that we generated (V24A, R59G, L121A, M123A and R132M). Thus the electrostatic repulsion between the two arginines is, at most, only partially responsible for the conformational flexibility of apo-CRABP.

Other interactions contributing to the dynamical properties of apo-CRABPII were suggested by the solution structures of apo-R111M and apo-CRABPII (Chapter 3). In both structures, the guanidinium group of the Arg132 is partially exposed and surrounded

by the side-chains of Ser12, Pro39, Thr54 (γ-methyl group), Cys130 and Tyr134. The hydroxyl groups of Ser12 and Tyr134 and the thiol group of Cys130 provide the favorable polar interactions with the guanidinium group of the Arg132. In particular, the hydroxyl group of the phenolic ring of Tyr134 is within the hydrogen bond distance to the Nε nitrogen of the guanidinium group of Arg132. By contrast, in the solution structure of apo-CRABPII, the guanidinium group of Arg111 is almost completely buried and in an environment much more hydrophobic than that of Arg132. It is surrounded by the side-chains of Val41, Ile43, Ile52, Thr54, Ile63, Trp109, Leu119 and Leu121. The hydrophobic side-chains of Val41, Ile63, Leu119 and Leu121 had close contacts with the guanidinium group of Arg111. The only favorable interactions may be between the guanidinium group of Arg111 and the aromatic ring of Trp109 and between the guanidinium group of Arg111 and the hydroxyl group of Thr54. As mentioned earlier, the peaks of the residues Val41, Lys53, Glu62, Ile63, Leu119 and Leu121 became much stronger in the 2D ^{1}H-^{15}N HSQC spectrum and/or 2D homonuclear spectra after the point mutation. It is tempting to speculate that the charged guanidinium group of Arg111 is mobile because of its hydrophobic environment and possibly the repulsive interaction with Arg132. The mobility of the guanidinium group of Arg111 drives the conformational flexibility of apo-CRABPII and results in a more open conformation for the ligand entrance. The hydrophobic side-chain of Met111 is preferred by the hydrophobic environment. Binding of RA neutralizes, at least partially, the positive charge of Arg111 and therefore reduces the mobility of its guanidinium group. CRABPs are most stable in the holo-forms because there are additional hydrophobic interactions between the proteins and RA. It is noted that the guanidinium group of Arg111 is more

than 17 Å away from the ligand entrance, the most flexible region of the apo-CRABPII. It is not clear how the mobility of the guanidinium group propagates to the ligand entrance.

Conclusion

The solution structure of apo-R111M, a site-directed variant of human CRABPII, has been determined by multidimensional heteronuclear NMR. The solution structure of apo-R111M was more similar to the crystal structure of wild-type holo-CRABPII than to the solution structure of wild-type apo-CRABPII. The size of the ligand entrance, the mobility of the RA-binding pocket and the ligand entrance were reduced greatly after the point mutation Arg111 to Met111. The results suggest that Arg111 plays a major role in maintaining the ligand entrance in an open and dynamic state for the entry of RA.

References

Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1-6.

Basus, V. J. (1989) *Methods Emzymol. 177*, 132-149.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Magn. Reson. 55,* 301-315.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185-188.

Cavanagh, J., Palmer A. G., Wright, P. E., & Rance, M. (1991) *J. Magn. Reson. 91*, 429-436.

Chambon, P., Olson, J. A., & Ross, A. C. (coordinators) (1996) *The retinoid revolution, FASEB J. 10*, 938-1048.

Delaglio, F., Grzesiek, S., Vüister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995) *J. Biomol. NMR 6*, 277-293.

Garret, D. S., Powers, R., Gronenborn, A. M., & Clore, G. M. (1991) *J. Magn. Reson. 95*, 214-220.

Clore, G. M., & Gronenborn, A. G. (1989) *CRC Crit. Rev. Biochem. Mol. Biol. 24*, 479-564.

Clore, G. M., & Gronenborn, A. G. (1991*) Annu. Rev. Biophys. Chem. 21*, 29-63.

Griesinger, C., Sørensen, O. W., & Ernst, R. P. (1985) *J. Am. Chem. Soc. 107*, 6394-6396.

Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. P. (1988) *J. Am. Chem. Soc. 110*, 7870-7872.

Grzesiek, S., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593-12594.

Hodson, M. E., & Cistola, D. P. (1997) *Biochemistry 36*, 2278-2290.

Jeener, J., Meier, B. H., & Ernst, R. R. (1978) *J. Chem. Phys.* 71, 4546-4553.

Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 120-126.

Kay, L. E., Keiffer, P., & Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663-10665.

Kleywegt, G. J., Bergfors, T., Senn, H., le Motte, P., Gsell, B., Shuao, K., & Jones, T. A. (1994) *Structure 2*, 1241-1258.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989a) *J. Magn. Reson. 85*, 393-399.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfeld, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989b) *Biochemistry 28*, 6150-6156.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., Bax, A. (1989c) *J. Am. Chem. Soc. 111*, 1515-1517.

Müeller, L. (1979) *J. Am. Chem. Soc. 101*, 4481-4484.

Palmer, A. G., Cavanagh, J., Wright, P. E., & Rance, M. (1991) *J. Magn. Reson. 93*, 151-170.

Piantini, U., Sørensen, O. W., & Ernst, R. P. (1982) *J. Am. Chem. Soc. 104*, 6800-6801.

Rizo, J., Liu, Z. P., & Gierasch, L. M. (1994) *J. Biomol. NMR 4*, 741-760.

Shaka, A. J., Lee, C. J., & Pine, A. (1988) *J. Magn. Reson. 77*, 274-293.

Shaka, A. J., Barker, P. B., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547-552.

Sporn, M. B., Roberts, A. B. & Goodman, D. S. (1994) *The Retinoids: Biology, Chemistry, and Medicine*, 2nd ed. Raven Press, New York.

States, D. J., Haberkon, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.

Wüthrich, K. (1986) *NMR of Proteins and Nuclear Acids*, Wiley, New York.

Wüthrich, K., Billletter, M., & Braün, W. (1983) *J. Mol. Biol. 169*, 949-961.

Zhang, J., Liu, Z., Jones, T. A., Gierasch, L. M., & Sambrook, J. F. (1992) *Proteins: Structure, Function, and Genetics 13*, 87-99.

CHAPTER 5

NMR Studies of Human CRABPII in Complex with All-*trans*-Retinoic Acid

Introduction

The crystal structures of holo-CRABPs suggest that RA cannot enter or exit the deep binding pocket in the absence of major conformational changes in the protein (Kleywegt et al., 1994). Indeed, significant conformational differences between the crystal structure of holo-CRABPII and the solution structure of apo-CRABPII have been observed, especially in the ligand entrance region (Chapter 3). In comparison to the holo crystal structure, apo-CRABPII exhibits a concerted movement of the second helix, the βC-βD loop and the βE-βF loop, so that the ligand entrance of apo-CRABPII is greatly enlarged and readily accessible to RA. Furthermore, the ligand-binding pocket of apo-CRABPII is rather disordered. The results suggest that binding of RA induce significant changes in both conformation and dynamics of CRABPII.

This chapter describes the sequential resonance assignments and preliminary

results showed that RA binds to CRABPII in solution in the same manner as determined by crystallography and the ligand entrance of holo-CRABPII is much less flexible than that of apo-CRABPII. Relative motions between the bound RA and human CRABPII in solution, particularly at the ligand entrance, were detected. Such motions may provide a transient pathway for the exit of RA.

Experimental Procedures

Sample preparation. All-*trans*-RA and D_2O were purchased from Sigma (St. Louis, Missouri), dimethyl sulfoxide-d_6 (DMSO-d_6) were purchased from Aldrich.

Expression of CRABPII gene and purification of the recombinant protein have been described in Chapter 2. The NMR samples of holo-CRABPII were prepared as follows. CRABPII was first dissolved in ~50 ml PBS buffer (20 mM sodium phosphate, 150 mM sodium chloride with 99.6% D_2O, pD 7.5, uncorrected). The protein concentration was ~40 μM. All-*trans*-RA was dissolved in DMSO-d_6. The concentration of the RA solution was ~3 mM. The RA solution was then added to the protein solution by aliquots of ~5 μl. The solution was quickly mixed by stirring after each addition. The entire procedure was carried out in a dark room. The final DMSO concentration was less than 1.5%. The sample was concentrated to ~0.6 ml. H_2O sample of uniformly ^{15}N-labeled holo-CRABPII was prepared similarly except that the D_2O in PBS buffer was replaced by H_2O. The final protein concentrations of the NMR samples were ~2 mM.

NMR spectroscopy. Homonuclear 2D spectra of the unlabeled D_2O sample and heteronuclear 2D and 3D spectra of the uniformly ^{15}N-labeled H_2O sample were acquired and processed as detailed in Chapter 3 and 4. The temperature was set at 25 °C. Homonuclear 2D spectra were acquired with a spectral width of 8000 Hz in both dimensions. The carrier frequency was 4.70 ppm (solvent resonance). The relaxation delay was 1.2 s. The residual water resonance was suppressed by low-power

presaturation. One DQF-COSY spectrum, three clean-TOCSY spectra with mixing times of 20, 33 and 50 ms, and two NOESY spectra with mixing times of 100 and 150 ms were acquired on the D_2O sample. The time domain data were composed of 2048 × 320 complex points for the DQF-COSY (128 transients) and NOESY (64 transients) experiments and 2048 × 256 complex points for the TOCSY experiments (64 transients).

The carrier frequencies were 4.70 ppm for ^{1}H and 121.0 ppm for ^{15}N for the heteronuclear experiments. A 2D ^{1}H-^{15}N HSQC spectrum was acquired with the following acquisition times and complex data points: ^{1}H (F2) 121.9 ms, 1024 and ^{15}N (F1) 53.3 ms, 128 (32 transients). A 3D ^{1}H-^{15}N NOESY-HSQC spectrum (16 transients, 150 ms mixing time) and a 3D ^{1}H-^{15}N TOCSY-HSQC spectrum (8 transients, 35 ms mixing time) were recorded with the same acquisition times and complex data points: ^{1}H (F3) 121.9 ms, 1024, ^{15}N (F2) 15.2 ms, 32, ^{1}H (F1) 15.2 ms, 128.

Results and Discussion.

Protein ^{1}H and ^{15}N resonance assignments and secondary structures. The protein ^{1}H and ^{15}N resonances (Table 5.1) were assigned by the same strategy as detailed previously (Chapter 4). Briefly, spin system identification was achieved by the homonuclear 2D and 3D ^{15}N-edited TOCSY experiments. The sequential linking was accomplished by the analyses of the 2D and 3D NOESY spectra. The assignment was helped greatly by comparison with the sequential assignments of the wild type apo-CRABPII and the mutant apo-R111M (Chapters 3 and 4). Secondary structural elements were identified as described in Chapter 4. They were essentially the same as those observed in the crystal structure except that the first helix αA appeared to be mobile in solution (Figure 5.1 and 5.2).

Conformational and dynamical changes in CRABPII induced by the binding of RA. The majority of the NOEs observed between the bound RA and the protein were in agreement, at least qualitatively, with those predicated from the crystal structure of holo-CRABPII, suggesting that the tertiary structure of holo-CRABPII in solution is most likely very similar to that in the crystalline state. In particular, the contacts between the bound RA and the residues at the ligand entrance observed by crystallography were also detected by NOEs, indicating that the conformation of the ligand entrance of holo-CRABPII in solution is similar to that in the crystalline state. These NMR results further strengthened the conclusion based on the comparison of the holo crystal structure and the

Table 5.1. ^{1}H and ^{15}N chemical shifts of human CRABPII in complex with all-*trans*-RA in PBS at 25°C.

No	^{15}N	H^N	H^α	H^β	Others
P1			4.35	2.39, 1.92	H^γ 2.11, 2.11; H^δ 3.35, 3.28
N2			5.08	2.98, 2.90	γN 112.5; γNH_2 7.75, 7.02
F3	121.9	9.86	4.35	3.34, 3.28	2, 6H 7.43; 3, 5H 7.38; 4H 7.17
S4	112.4	8.48	4.70	4.26, 3.99	
G5	110.5	9.22	4.04, 3.82		
N6	118.5	8.19	5.69	2.72, 2.55	
W7	122.9	9.31	5.21	3.23, 2.99	2H 6.80; 4H 7.07; 5H 7.26; 6H 6.81;7H 7.14 N^ε 124.8; H^ε 9.17
K8	123.3	10.17	5.26	1.98, 1.88	H^γ 1.68, 1.63; H^δ 1.37, 1.33; H^ε 2.87, 2.83
I9	125.8	8.60	3.99	1.91	$H^{\gamma m}$ 0.83
I10	118.0	9.24	4.62		
R11	120.8	7.49	4.58	1.70, 1.70	
S12			5.07	3.59, 3.33	
E13			4.83	1.94, 1.94	H^γ 2.09, 2.19
N14			4.99	3.35, 3.16	
F15	120.5	8.52	4.46	2.99, 2.72	2, 6H 7.06; 3, 5H 6.64; 4H 6.59
E16			3.30	1.85, 1.85	H^γ 1.98, 2.17
E17			3.74	1.87, 2.16	H^γ 2.27, 2.27
L18	121.2	7.59	3.74	1.86, 1.69	H^δ 0.81, 0.81
L19	113.8	7.25	3.60	1.66, 1.60	H^γ 1.17; Hδ 0.32, 0.22
K20	118.9	8.24	3.83	1.93, 1.63	H^γ 1.56, 1.35; H^ε 2.82, 2.93
V21	121.3	7.84	3.77	2.15	H^γ 1.10, 0.94
L22	117.5	7.15	3.96	1.84, 1.58	H^δ 0.92, 0.70
G23	106.6	7.68	3.96, 3.61		
V24	121.4	7.60	3.74	1.47	H^γ 0.82, 0.82
N25	128.2	8.80	4.42	3.17, 2.64	γN 112.4; γNH_2 7.71, 6.92
V26	120.5	8.57	3.49	2.07	H^γ 0.97
M27	119.8	7.94	4.13	2.13, 1.98	H^γ 2.60, 2.45
L28	117.8	7.80	4.09	1.80, 1.64	H^γ 1.33; H^δ 0.88, 0.77
R29	118.8	8.81	3.58	1.95	
K30	116.2	7.63	3.90	1.91, 1.91	H^γ 1.72 1.53; H^δ 1.45, 1.45; H^ε 2.90, 2.90
I31	118.9	7.00	3.75	1.91	$H^{\gamma m}$ 0.83
A32	120.5	8.56	3.88	1.18	
V33	116.5	8.85	3.50	2.07	H^γ 1.03, 0.96
A34	120.4	7.19	4.15	1.47	
A35	121.2	8.18	4.24	1.51	
A36	116.6	8.69	3.96	0.86	
S37	113.5	7.38	4.22	3.98	
K38	119.0	7.05	4.80	1.90, 1.61	H^γ 1.42, 1.31; H^γ 2.95, 2.95
P39			5.09	1.92, 1.87	H^δ 3.65, 3.55
A40	123.0	8.77	5.03	1.37	
V41	124.5	7.95	5.22	1.88	H^γ 0.76, 0.44
E42	130.0	9.29	5.35	2.08, 2.02	H^γ 2.27, 2.27
I43	122.9	9.29	5.38	2.26	H^γ 1.67, 1.34; H^δ 0.48; $H^{\gamma m}$ 0.85
K44	128.8	9.57	4.65	1.87, 1.76	H^γ 1.49, 1.36
Q45	127.9	8.60	4.51	1.84, 1.79	H^γ 2.04, 2.04

No	^{15}N	H^N	H^α	H^β	Others
E46	128.7	8.53	4.37	1.80, 1.66	H^γ 2.01, 2.01
G47	117.5	9.30	3.96, 3.61		
D48	126.5	8.56	5.21	3.16, 2.72	
T49	116.6	8.08	4.66	4.15	H^γ 1.01
F50	124.1	9.10	4.63	2.57, 2.57	2, 6H 6.46; 3, 5H 6.81; 4H 7.09
Y51	122.5	8.35	5.33	3.03, 2.69	2, 6H 6.93; 3, 5H 6.61
I52	123.3	8.35	4.62	1.89	H^γ 1.35, 1.35; $H^{\gamma m}$ 0.28; H^δ 0.37
K53	133.4	9.31	5.14	1.71, 1.67	H^γ 1.60, 1.48; H^γ 1.17; H^ε 2.81, 2.81
T54	124.7	8.84	4.79	3.83	H^γ 0.85
S55	123.1	9.56	5.25	3.82, 3.77	
T56	113.5	7.74	4.93	4.33	H^γ 1.18
T57			3.98	4.18	H^γ 1.32
V58	109.5	6.71	4.48	2.18	H^γ 0.81, 0.81
R59	119.0	7.09	4.76	2.02, 1.91	H^γ 1.65, 1.49
T60	122.7	8.73	5.32	3.93	H^γ 1.06
T61	117.2	8.72	4.64	3.98	H^γ 0.97
E62	122.9	8.77	4.85	1.94, 1.94	H^γ 2.21, 2.10
I63	116.8	8.29	4.93	1.90	H^γ 1.33, 0.75; H^δ 0.84; $H^{\gamma m}$ 1.06
N64	120.2	8.10	5.59	2.58, 2.53	γN 113.0; γNH_2 7.62, 6.93
F65	116.8	8.36	4.70	2.13, 2.13	2, 6H 6.20; 3, 5H 6.51; 4H 6.28
K66	120.8	9.35	5.30	1.65, 1.60	H^γ 1.53, 1.47; H^ε 2.94, 2.94
V67	127.3	8.98	3.24	1.88	H^γ 0.66, 0.45
G68	114.8	8.93	4.42, 3.62		
E69	120.1	7.96	4.76	2.04, 1.92	H^γ 2.16, 2.16
E70	130.3	9.25	4.80	2.00, 1.91	H^γ 2.22, 2.17
F71	121.3	9.06	5.14	3.57, 3.18	2, 6H 7.24; 3, 5H 6.83; 4H 6.47
E72	119.4	8.72	5.05	1.96, 1.92	H^γ 2.33, 2.14
E73	120.3	8.44	4.76	2.04, 1.92	H^γ 2.16, 2.16
Q74	118.1	7.68	5.32	1.69, 1.73	H^γ 2.24, 1.99
T75	111.7	9.25	4.61	4.15	H^γ 1.07
V76	122.4	9.25	3.95	2.12	H^γ 1.18, 0.86
D77	112.3	8.75	4.46	2.82, 2.13	
G78	107.0	7.96	4.06, 3.59		
R79	120.7	7.91	4.70	1.80,1.56	
P80			4.65		
C81	118.8	8.85	5.09	2.62, 2.17	
K82	119.8	8.89	4.75	1.58, 1.50	H^γ 1.27, 1.27; H^ε 2.80, 2.80
S83	124.7	9.05	4.86	2.78, 1.54	
L84	120.5	7.34	4.25	1.60, 1.30	H^γ 1.17; H^δ 0.73, 0.69
V85	130.0	9.24	4.38	1.37	H^γ 0.24, 0.00
K86	123.6	8.47	4.51	1.64, 1.45	H^γ 1.52, 1.27; H^δ 1.15
W87	120.1	8.69	5.14	3.26, 3.00	2H 7.32; 4H 7.98; 5H 6.85; 6H 7.23; 7H 7.63 N^ε 129.5; H^ε 10.68
E88	127.7	9.25	4.37	1.98, 1.69	H^γ 2.16, 2.02
S89	112.4	8.60	4.57	4.14, 3.99	
E90	117.1	8.72	3.68	1.87, 1.80	H^γ 2.09, 2.09
N91	111.0	8.05	4.97	3.18, 2.88	γN 112.6; γNH_2 7.49, 6.84
K92	119.9	7.83	5.52	2.01, 1.95	H^γ 1.49, 1.49
M93	127.0	9.40	5.21	1.80, 1.80	H^γ 2.07, 1.99; H^ε 0.43

No	^{15}N	H^N	H^α	H^β	Others
V94	118.8	8.80	4.30	1.80	H^γ 0.86, 0.75
C95	127.1	8.87	4.51	0.41 –0.27	
E96	127.5	7.67	4.35	1.90, 1.80	H^γ 2.14, 2.14
Q97	120.7	7.41	4.93	1.77, 1.77	H^γ 2.43, 2.04
K98	122.8	8.96	4.59	1.53, 1.53	H^γ 1.26, 1.26; H^ε 2.79, 2.79
L99	125.9	8.84	4.02	1.64, 1.47	H^γ 1.38; H^δ 0.67, 0.67
L100	121.5	7.74	4.10	1.65, 1.48	H^γ 1.28; H^δ 0.78
K101	117.5	8.03	4.42	1.80, 1.80	H^γ 1.53, 1.37
G102	108.7	8.27	4.06, 3.75		
E103	116.8	8.17	4.61	1.80, 1.80	H^γ 2.13, 1.98
G104	108.8	8.16	4.14, 3.93		
P105			4.51	2.03, 2.03	H^γ 1.85, 1.73; H^δ 3.61, 3.25
K106	123.6	8.58	4.47	1.80, 1.80	H^γ 1.53, 1.42
T107	119.0	8.80	5.39	4.49	H^γ 1.15
S108	112.1	7.80	4.70	3.93, 3.77	
W109	116.2	8.26	5.79	3.53, 3.33	2H 6.92; 4H 6.90; 5H 6.35; 6H 6.44; 7H 6.67 N^ε 138.1; H^ε 11.38
T110	110.8	9.58	5.24	3.96	H^γ 1.00
R111	119.9	8.81	5.62	1.80, 1.50	H^γ 1.32
E112	123.1	8.91	5.80	1.95, 1.95	H^γ 2.04, 2.04
L113	126.4	8.23	5.31	1.28, *0.75*	H^γ 1.12, H^δ 0.46, 0.46
T114	116.9	9.38	4.60	4.761	H^γ 1.26
N115			4.46	2.82	γN 112.6; γNH_2 7.65, 6.95
D116	115.6	7.74	4.76	2.81, 2.42	
G117	107.6	7.98	4.19, 3.70		
E118	117.4	7.42	4.98	1.83, 1.72	H^γ 2.04, 1.91
L119	124.5	8.92	5.13	0.87	H^γ 0.46; H^δ 0.15, –0.25
I120	125.1	9.51	4.98	1.82	
L121	135.4	9.98	5.61	1.90, 1.34	H^δ 0.78, 0.71
T122	119.3	9.15	5.85	4.05	H^γ 1.09
M123	121.2	9.28	5.16	1.84, 1.80	H^γ 2.64, 2.33 H^ε 1.63
T124	115.6	8.27	5.48	3.96	H^γ 1.00
A125	125.4	8.40	4.88	1.04	
D126	122.2	9.26	4.16	2.75, 2.75	
D127	121.1	8.49	4.52	2.68, 2.68	
V128	123.3	8.41	3.95	2.14	H^γ 0.92, 0.69
V129	125.6	8.42	4.42	1.84	H^γ 0.78, 0.70
C130	129.7	9.59	5.20	2.73	
T131	124.5	8.45	5.08	3.76	H^γ 1.00
R132	128.6	9.85	5.09	1.75, 1.61	
V133	122.7	8.56	4.98	1.71	H^γ 0.86, 0.86
Y134	126.7	9.79	5.61	3.03, 3.03	2, 6H 6.83; 3, 5H 6.46
V135	112.0	9.52	5.35	2.39	H^γ 1.06, 0.86
R136	122.3	8.44	3.88	1.32, 1.09	H^γ 0.58, 0.04; H^δ 2.81, 2.73
E137	126.5	8.10	3.87	1.81 1.69	H^γ 2.07 2.07

Figure 5.1. Strip plots extracted from the 150 ms mixing time ^{15}N-edited NOESY-HSQC spectrum of the holo-form CRABPII showing the NOE connectivies for residues Met27-Lys38.

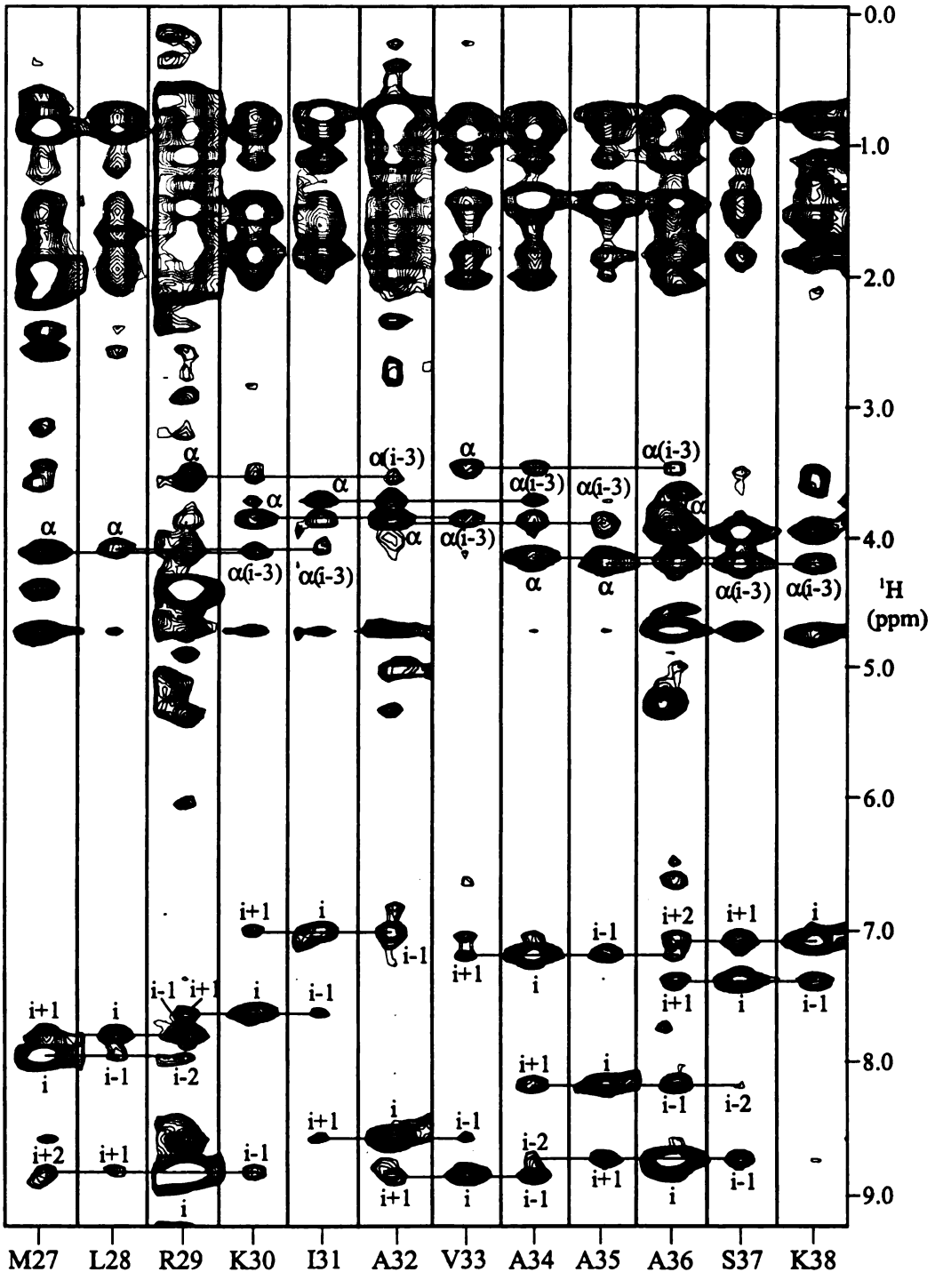
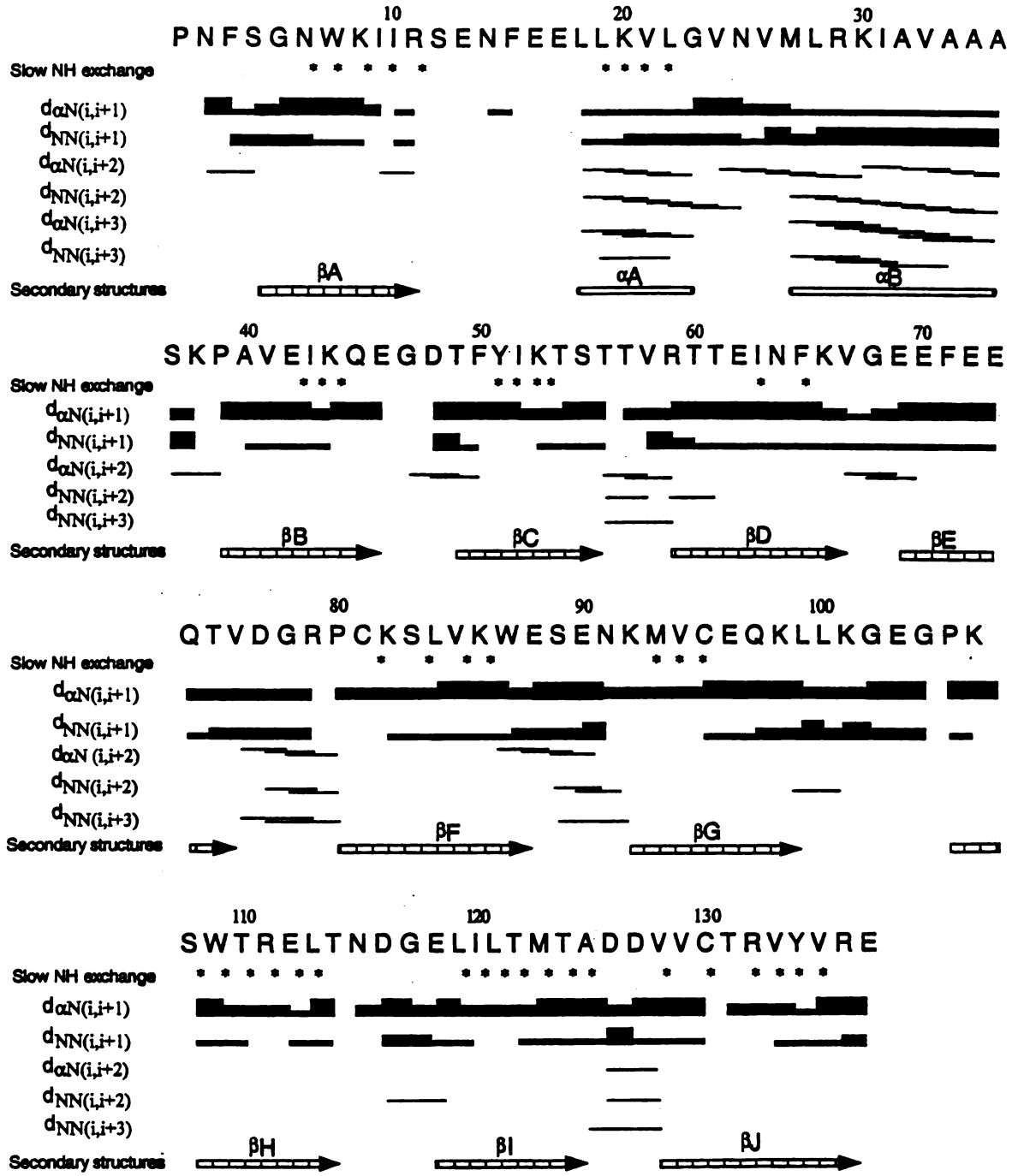

0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
¹H (ppm)
M27
L28
R29
K30
I31
A32
V33
A34
A35
A36
S37
K38

Figure 5.2. Summary of the sequential and medium-range NOEs involving backbone H^N and H^α atoms and the deduced secondary structures of holo-CRABPII. Line thickness for $d_{\alpha N}$ and d_{NN} sequential NOE distances reflects the intensities of the cross peaks.


10
20
30
PNFSGNWKIIRSENFEELLKVLGVNVMLRKIAVAAA
Slow NH exchange
dαN(i,i+1)
dNN(i,i+1)
dαN(i,i+2)
dNN(i,i+2)
dαN(i,i+3)
dNN(i,i+3)
Secondary structures
βA
αA
αB
40
50
60
70
SKPAVEIKQEGDTFYIKTSTTVRTTEINFKVGEEFEE
Slow NH exchange
dαN(i,i+1)
dNN(i,i+1)
dαN(i,i+2)
dNN(i,i+2)
dNN(i,i+3)
Secondary structures
βB
βC
βD
βE
80
90
100
QTVDGRPCKSLVKWESENKMVCEQKLLKGEGPK
Slow NH exchange
dαN(i,i+1)
dNN(i,i+1)
dαN (i,i+2)
dNN(i,i+2)
dNN(i,i+3)
Secondary structures
βF
βG
110
120
130
SWTRELTNDGELILTMTADDVVCTRVYVRE
Slow NH exchange
dαN(i,i+1)
dNN(i,i+1)
dαN(i,i+2)
dNN(i,i+2)
dNN(i,i+3)
Secondary structures
βH
βI
βJ

apo solution structure that RA binding induces significant conformational changes at the ligand entrance (Chapter 3).

Dynamical changes upon binding of RA were evidenced by the differences in peak intensity between the 2D ^{1}H-^{15}N HSQC spectra of apo- and holo-CRABPII (Figure 5.3 and 5.4) and the differences in the number and intensity of the observed NOEs. Remarkable differences were observed in the second helix, which is part of the ligand entrance. It had a typical α-helical NOE pattern in the holo-form (Figure 5.1 and 5.2) but no such pattern was observed in the apo-form. Furthermore, the residues in the second helix had much more intense cross peaks in the 2D ^{1}H-^{15}N HSQC spectrum of the holo form. The second helix is apparently disordered in the apo form (Chapter 3). Other residues at the ligand binding pocket also showed more intense cross peaks in the 2D ^{1}H-^{15}N HSQC spectrum upon binding of RA, including Thr56-Ile63, Val76-Gly78, Leu121, Arg132 and Arg134. Thr56-Ile63 and Val76-Gly78 constitute the βC-βD loop and the βE–βF loop, respectively, at the ligand entrance. Leu121, Arg132 and Arg134 are located at the bottom of the binding pocket. Their side-chains interact with the carboxyl group of RA. The results suggest that while the ligand binding pocket of apo-CRABP undergoes conformational exchange, the conformational exchange is greatly reduced upon binding of RA. Similar ligand-induced dynamical changes have been observed in the NMR studies of CRABPI (Rizo et al., 1994) and intestinal fatty acid binding protein (I-FABP) (Hodsdon et al., 1995; Hodsdon et al., 1996; Hodsdon & Cistola, 1997a, b).

The stability of the ligand entrance in holo-CRABPs may be brought about by the hydrophobic interactions between the isoprene tail of RA and the protein. However, not all HSQC cross peaks in CRABPII became stronger after the binding of RA. The cross

Figure 5.3. Gradient- and sensitivity-enhanced 2D ^{1}H–^{15}N HSQC spectrum of uniformly ^{15}N enriched of the wild type holo-CRABPII in complex with all-*trans*-retinoic acid. Sequential assignments are indicated with one-letter amino acid code and residue number.

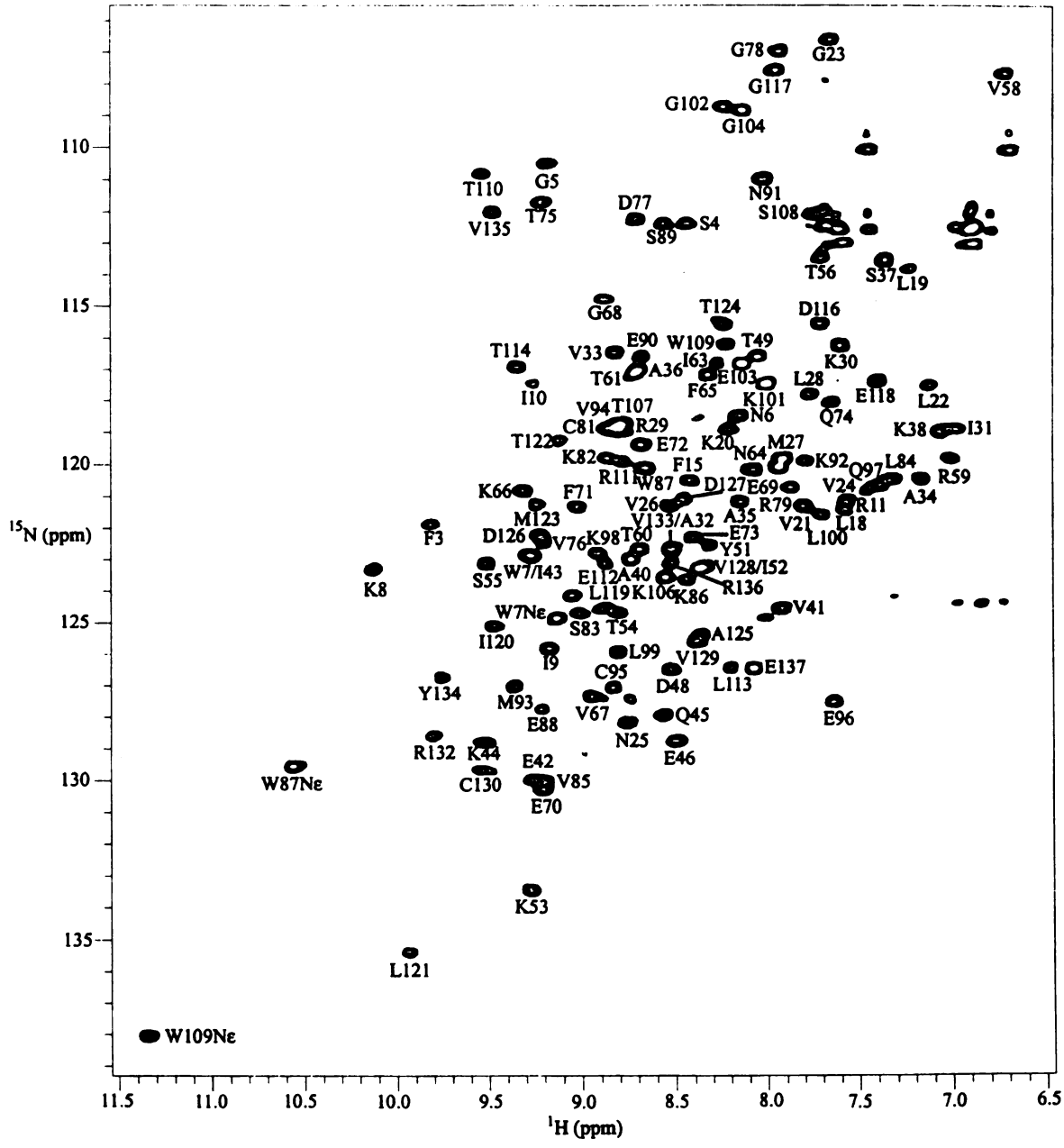
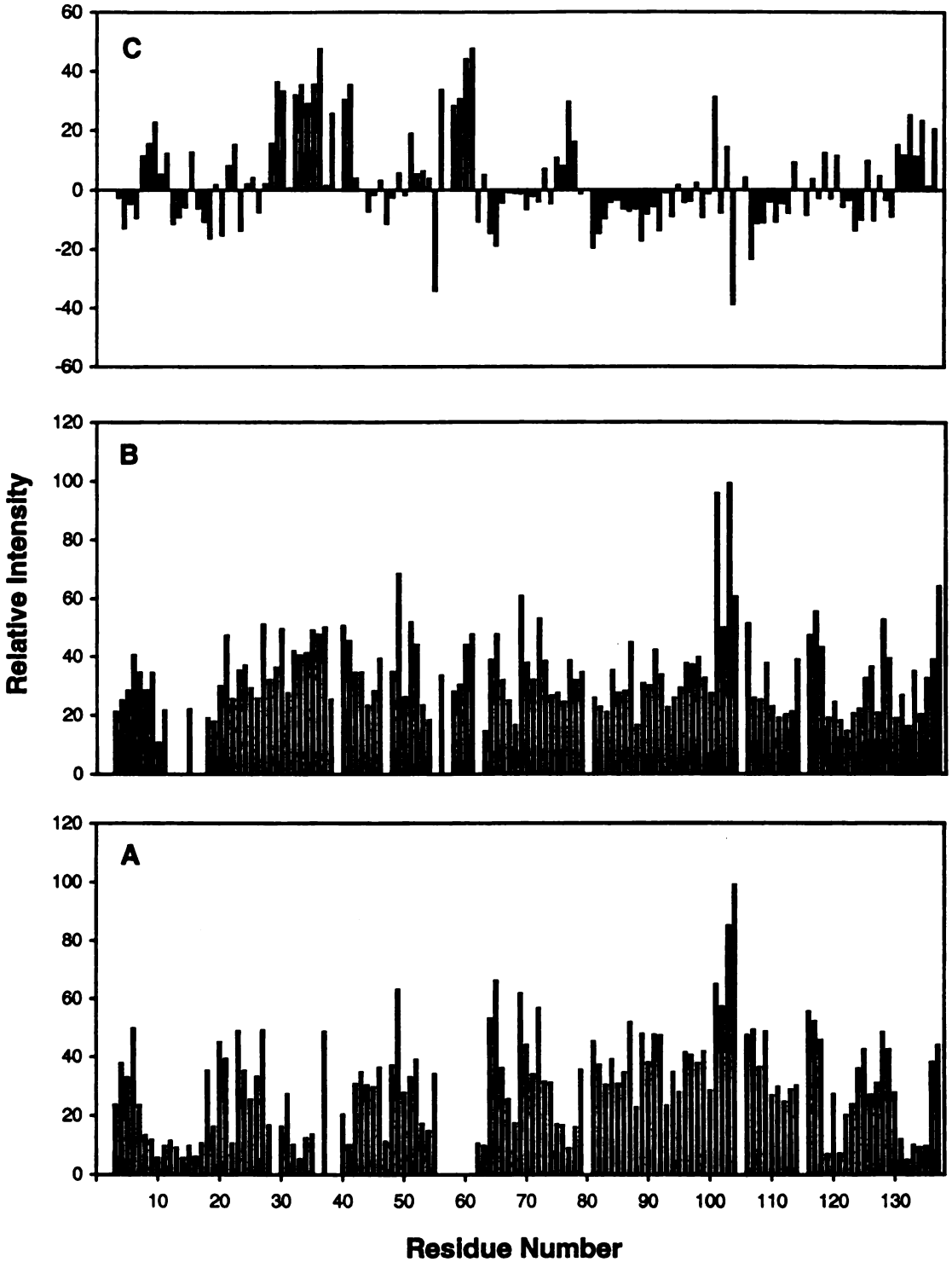

G78
G23
G117
V58
G102
G104
T110
G5
N91
D77
T75
S108
V135
S4
S89
T56
S37
L19
G68
T124
D116
E90
W109
T49
T114
V33
K30
I63
T61
A36
E103
L28
I10
F65
K101
E118
L22
V94
T107
Q74
C81
R29
N6
K38
I31
T122
E72
K20
M27
K82
N64
K92
L84
F15
Q97
R59
R11
W87
D127
E69
V24
K66
F71
V26
R79
R11
A34
M123
V133/A32
A35
V21
L18
F3
D126
V76
K98
T60
E73
L100
Y51
S55
W7/I43
E112
A40
V128/I52
K8
L119
K106
K86
R136
W7Nε
V41
S83
T54
I120
A125
I9
L99
V129
E137
C95
D48
L113
Y134
M93
V67
Q45
E96
E88
N25
R132
K44
E42
E46
W87Nε
C130
V85
E70
K53
L121
W109Nε
110
115
120
125
130
135
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
15N (ppm)
1H (ppm)

Figure 5.4. Relative peak intensities of the 2D ^{1}H–^{15}N HSQC spectra of apo-CRABPII (A) and holo-CRABPII (B) as a function of the residue number. (C) was obtained by subtracting (B) from (A).


C
B
A
Relative Intensity
Residue Number
60
40
20
0
-20
-40
-60
120
100
80
60
40
20
0
10
20
30
40
50
60
70
80
90
100
110
120
130

peaks of Ser12, Glu13, Asn14, Glu16 and Glu17 could not be observed in the 2D ^{1}H-^{15}N HSQC spectra. The cross peaks of Ile10, Arg11, Phe15, Leu19 were very weak. Leu18 probably was also weak. These residues became more flexible after the binding of RA. Their high mobility compensates, at least partially, the decrease in entropy caused by ligand binding.

^{1}H resonance assignments of CRABPII-bound RA. The ^{1}H resonances of the bound RA were assigned by analysis of the homonuclear 2D DQF-COSY, TOCSY and NOESY spectra of holo-CRABPII recorded in D_2O. The olefinic protons were well resolved from the protein resonances. Through-bond correlations could be clearly identified for CH(7), CH(8), CH(10), CH(11) and CH(12) protons (Figure 5.5). CH(11) was assigned by its distinct through-bond correlation pattern. The remaining olefinic protons were assigned based on their NOE patterns (Figure 5.6). The methyl protons of CH_3(18), CH_3(19) and CH_3(20) were assigned based on their NOEs to olefinic protons. The assignments of the methyl protons of CH_3(16) and CH_3(17) were tentative. None of the methylene protons of CH_2(2), CH_2(3) and CH_2(4) of β-ionone ring was assigned at present. The ^{1}H chemical shifts of the RA bound to human CRABPII (Table 5.2) were almost identical to those reported for the RA bound to mouse CRABPII (Norris, et al., 1994) except for CH(12) which differed by ~0.2 ppm. The chemical shifts of the CRABPII-bound RA protons were also similar to those of free RA in CCl_4 except for CH(8), CH(11) and CH_3(20) protons, which were 6.06 ppm, 6.96 ppm and 2.36 ppm in CCl_4, respectively (Waterhous & Muccio, 1990).

Conformation and dynamics of bound RA. Most of the NOEs observed for the bound RA (Table 5.3, Figure 5.6 and 5.7) were consistent with the distances measured

Table 5.2. ^{1}H resonance assignment of retinoic acid bound to human CRABPII in PBS at 25°C.

Olefinic proton (CH)	CH(7)	CH(8)	CH(10)	CH(11)	CH(12)	CH(14)
Chemical shift (ppm)	6.08	6.36	6.22	6.47	6.14	5.62
Methyl protons (CH_3)	CH_3(16)*	CH_3(17)*	CH_3(18)	CH_3(19)	CH_3(20)	
Chemical shift (ppm)	0.90	0.99	1.66	1.92	1.97	

* the individual assignments of CH_3(16) and CH_3(17) were assigned tentatively.

Table 5.3. The relative intramolecular NOE volumes [a] of RA bound to human CRABPII in PBS at 25°C (the mixing time was 100 ms).

Proton pair	7-8	7-10	7-16	7-17	7-18	7-19
Relative NOE volume	4.8	1.9	34.0*	40.1*	26.9	45.0
Proton pair	8-10	8-12	8-14	8-16	8-17	8-18
Relative NOE volume	40.0	16.1	10.7	31.8*	31.6*	45.6
Proton pair	10-11	10-12	10-14	10-18	10-19	10-20
Relative NOE volume	38.3	43.1	36.7	33.2	30.5	6.0
Proton pair	8-19	11-14	11-12	11-19	11-20	12-19
Relative NOE volume	33.6	30.0	9.1	strong**	strong**	12.3
Proton pair	12-20	12-14	14-20			
Relative NOE volume	18.5	100.0	strong**			

[a] The NOEs were picked by the program NMRCOMPASS 2.5 (Molecular Simulations) from the homonuclear 2D NOESY spectrum with a mixing time of 100 ms. The relative volumes were obtained by integration using NMRCOMPASS.

*These intramolecular NOE peaks may be overlapped with intermolecular ones.

** Strong peaks but no accurate values could be obtained because of overlapping.

Figure 5.5. Part of the 2D TOCSY spectrum with 33 ms mixing time of the RA bound to human CRABPII at 25°C. The CH(7), CH(8), CH(10), CH(11), CH(12) and CH(14) protons are labeled.

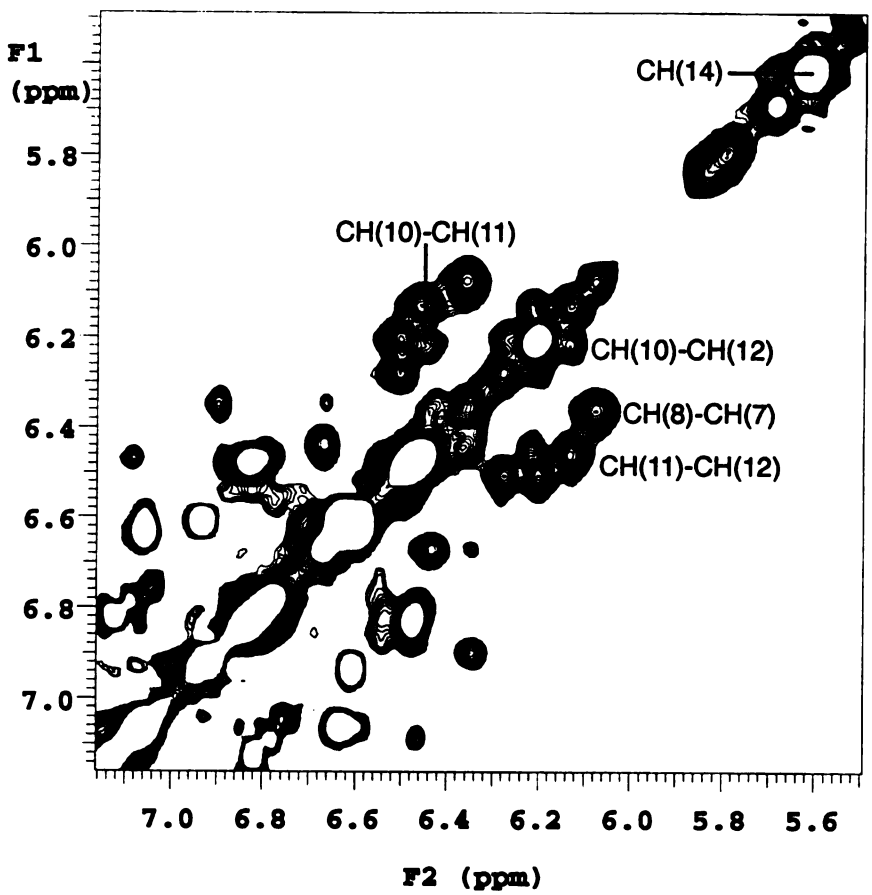
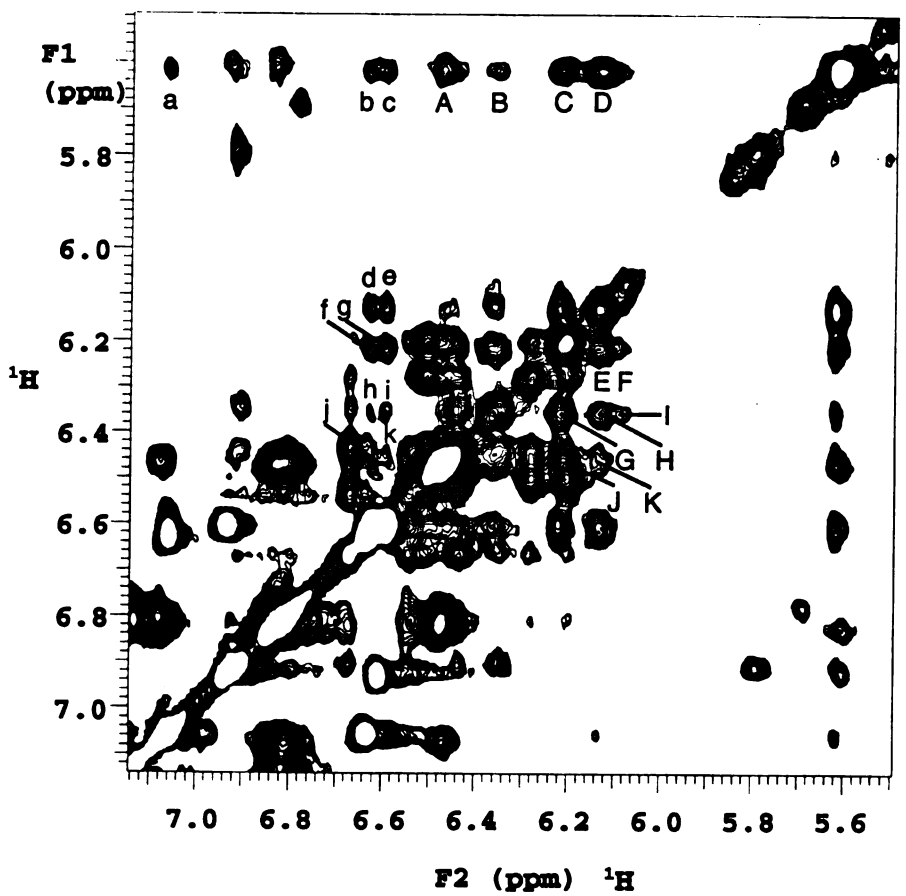

F1
(ppm)
5.8
6.0
6.2
6.4
6.6
6.8
7.0
CH(14)
CH(10)-CH(11)
CH(10)-CH(12)
CH(8)-CH(7)
CH(11)-CH(12)
7.0 6.8 6.6 6.4 6.2 6.0 5.8 5.6
F2 (ppm)

Figure 5.6. Part of the NOESY spectrum with 100 ms mixing time of the holo-CRABPII in complex all-*trans*-RA at 25°C showing its intramolecular NOEs and the intermolecular NOEs between RA and Phe15 of the protein. The identities of the NOEs are as follows. A, CH(14)-CH(11); B, CH(14)-CH(8); C, CH(14)-CH(10); D, CH(14)-CH(12); E, CH(10)-CH(12); F, CH(10)-CH(7); G, CH(8)-CH(10); H, CH(8)-CH(12); I, CH(8)-CH(7); J, CH(11)-CH(12); K, CH(11)-CH(10); a, 2, 6H(F15)-CH(14); b, 4H(F15)-CH(14); c, 3, 5H(F15)-CH(14); d, 4H(F15)-CH(12); e, 3, 5H(F15)-CH(12); f, 4H(F15)-CH(10); g, 3, 5H(F15)-CH(10); h, 4H(F15)-CH(8); i, 3, 5H(F15)-CH(8); j, 4H(F15)-CH(11); k, 3, 5H(F15)-CH(14).


F1
(ppm)
a
b c
A
B
C D
5.8
6.0
d e
f g
6.2
^{1}H
h i
j
k
E F
I
G H
J K
6.4
6.6
6.8
7.0
7.0
6.8
6.6
6.4
6.2
6.0
5.8
5.6
F2 (ppm) ^{1}H

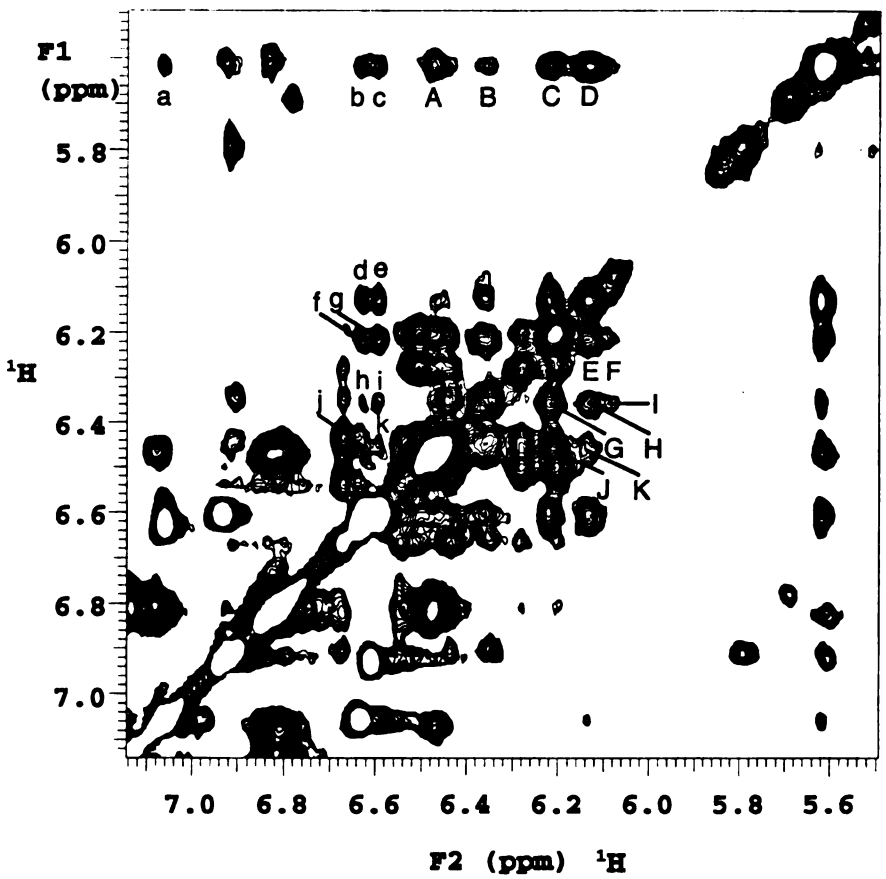
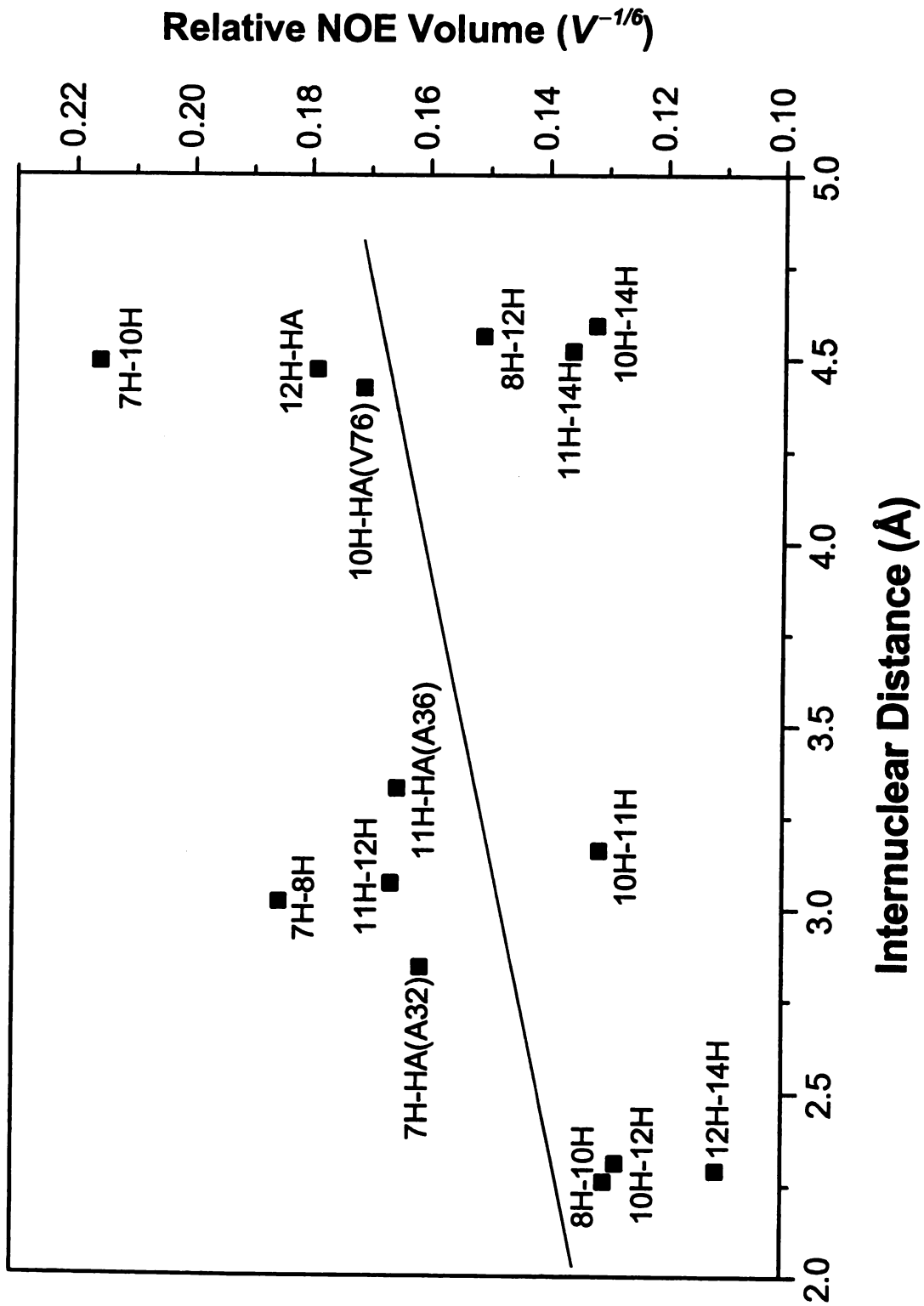

F1
(ppm)
a
bc
A
B
C
D
de
fg
hi
j
k
EF
I
G
H
J
K
5.8
6.0
6.2
6.4
6.6
6.8
7.0
^{1}H
7.0
6.8
6.6
6.4
6.2
6.0
5.8
5.6
F2 (ppm) ^{1}H

Figure 5.7. The relative volumes ($V^{-1/6}$) of some of the NOEs observed for the bound RA *vs.* the internuclear distance (r) measured from the crystal structure of the holo-CRABPII in complex with all-*trans*-RA. The straight line represents a linear relationship between the relative volume and the distance.


Relative NOE Volume ($V^{-1/6}$)
0.22
0.20
0.18
0.16
0.14
0.12
0.10
7H-10H
12H-HA
10H-HA(V76)
8H-12H
11H-14H
10H-14H
7H-8H
11H-12H
11H-HA(A36)
7H-HA(A32)
10H-11H
8H-10H
10H-12H
12H-14H
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Internuclear Distance (Å)

from the crystal structure of holo-CRABPII (Kleywegt et al., 1994), suggesting that the conformation of the bound RA in solution is similar to that in the crystalline state. The few discrepancies between the observed NOEs and the distances measured from the holo crystal structure were probably caused by the internal motions of the bound RA in solution rather than a different conformation as discussed below.

It is well known that NOE between an isolated pair of protons H_1-H_2 is proportional to cross-relaxation rate σ, which is determined by the internuclear distance r between the two protons and the spectral densities $J^{(q)}(\omega)$ at two frequencies (Chapter 1).

$$\sigma (H_1 \leftrightarrow H_2) = \gamma^4_H \hbar^2 / 4r^6 \{6J^{(2)}(2\omega_H) - J^{(0)}(0)\}$$

where γ_H is the proton gyromagnetic ratio. However, for a pair of protons belonging to a protein or a ligand bound to a protein the situation is usually much more complicated. In particular, internal motions may have significant effects on the spectral density. Furthermore, cross-correlation with additional spins may also change the apparent cross-relaxation rate (Werbelow & Grant, 1977). While these factors may complicate the accurate measurement of inter-proton distances, they are also part of the basis for studying internal motions of biomacromolecules. If two pairs of protons in the same molecule have the same inter-nuclear distance but different NOE intensities, it suggests that their motions relative to other parts of the molecule are different.

Analysis of the intramolecular NOEs of the bound RA indicated that all NOEs involving CH(7) are weaker than expected. For example, the distances between CH(7) and CH(8), between CH(10) and CH(11) and between CH(11) and CH(12) protons all are ~3.0 Å. Medium NOE cross-peaks are expected for all three pairs of protons if their

relaxation times (T_1 and T_2) are much longer than the mixing time (100 ms) and their positions in the binding pocket are fixed. However, the volume of the NOE cross-peak between CH(7) and CH(8) was significantly lower that those of the other two pairs of protons (Table 5.3 and Figure 5.6 and 5.7). By contrast, the corresponding peaks in both 2D DQF-COSY and TOCSY spectra were all relative strong (Figure 5.5), indicating that the relaxation times of these protons were similar to the others, and were longer than the 100 ms mixing time. The weakness of the NOE cross-peaks involving CH(7) was most likely due to the relative motions of the proton.

The nature of such relative motions could be either intramolecular or intermolecular. For RA, if disregarding the methyl groups, the possible intramolecular motion is about the 6-s bond (Honig et al., 1971), which affects the NOEs of the protons, among the others, CH(7) and CH(8). The intermolecular motions could be either the whole RA molecule moves as a unit relative to the protein or local motions restricted to certain parts of the RA molecule. The motions of the whole RA molecule would attenuate the intensities of all the NOEs involving the RA protons to similar extent. Thus, the attenuation of the peak intensity is most likely due to the local motions relative to a portion of RA. The weakness of the CH(7)-CH(8) cross-peak can be caused by either intramolecular motions about 6-s bond or local intermolecular motions restricted to their locations. In either case it suggested that the ionone ring of the bound RA is not fixed as snugly in the protein as suggested by the crystal structure of holo-CRABPII.

The NOEs involving CH(14), on the other hand, were all stronger than the average. In fact, the NOE peaks between CH(8) and CH(14) and between CH(10) and CH(14) protons were stronger than that between CH(7) and CH(8) protons (Table 5.3). In

the crystal structure of holo-CRABPII, the distances between the CH(10) and CH(14) protons, between the CH(8) and CH(14) protons and between the CH(7) and CH(8) protons were ~4.8 Å, ~6.8 Å and ~3.0 Å, respectively. Obviously, spin diffusion contributed to the appearance of the NOE cross-peak between CH(8) and CH(14) protons and to the strength of the NOE cross-peak between CH(10) and CH(14) protons. The spin diffusions could be mediated by protons on the protein and/or RA. Since only the methyl group of Met123 is within 3 Å distance of CH(14) and the methyl group is ~7.5 Å away from CH(8), the spin diffusions are most likely through CH(12). Since only a structure rigid on the time scale relevant to the NOESY experiment can provide efficient pathways for spin diffusion, it suggests that the carboxylate end of the isoprene tail was essentially immobile relative to the protein.

Very interestingly, the rigidity of the carboxylate end of the isoprene tail and the flexibility of the β-ionone of the bound RA molecule correlate well with the binding affinities and specificity of CRABPs for the retinoids. It has been shown that the carboxylic group is critical for RA to bind to CRABPS, while the modifications at the β-ionone ring and the C-7 and C-8 positions can be tolerated.

Our conclusion of the flexibility of the bound RA relative to the protein appears to be inconsistent with what was reported for mouse holo-CRABPII (Norris, et al., 1994). Their conclusion of a single, static conformation of the bound RA at the 6-s bond was based solely on intramolecular NOEs between the olefinic protons and the methyl protons. It was derived under the assumptions that both the protein and RA are rigid, and the NOE volume is proportional to r^{-6}. Furthermore, only three NOEs picked from an NOESY spectrum with 30 ms mixing time had values much higher than the noise. All

others have intensities similar to the noise. In fact, most NOEs involving CH(7) were also lower than expected by their model, suggesting that there are also motions in the RA bound to mouse CRABPII. Interestingly, similar intermolecular motions of the bound RA relative to the protein was observed in mouse holo-CRABPI.

A mechanism for the exit of RA. The intensities of most of the intermolecular NOEs between the bound RA and the protein agreed, at least qualitatively, with those predicted from the crystal structure (Table 5.3). However, differences between the solution and crystalline state were also observed. The predicted NOEs between the H^{δ} protons of Arg59 and CH(7) of the bound RA, between the H^{δ} protons of Arg59 and CH(8) and between the H^{δ} protons of Leu19 and CH(8) were not observed, and the intensity of the NOE between the H^{β} protons of Ala32 and CH(7) was much lower than predicted. The intensities of these NOEs must be attenuated by intermolecular motions. In the holo crystal structure, the side-chain of Arg59 extended from the βC-βD loop to the Glu74 at the C-terminus of βE. The motion of this side-chain relative the bound RA could temporarily open up the ligand entrance. Such flexibility may be necessary for the exit of the bound RA.

Conclusion.

The ^{1}H and ^{15}N resonance assignments of holo-CRABPII have been completed. Most of the observed intermolecular NOEs between the protein and bound RA and the intramolecular NOEs of the bound RA are, at least qualitatively, consistent with those predicted from the crystal structure of holo-CRABPII. The results suggested that RA binds to CRABPII in solution in the same manner as determined by crystallography. Comparison of the intensities of backbone amide resonances between apo- and holo-CRABPII indicated that the ligand entrance of holo-CRABPII is much less flexible than that of apo-CRABPII. Relative motions between the bound RA and human CRABPII in solution, particularly at the ligand entrance, were observed. Such motions may provide a transient pathway for the exit of RA.

References

Chambon, P., Olson, J. A., & Ross, A. C. (coordinators) (1996) *The retinoid revolution, FASEB J. 10*, 938-1048.

Hodson, M. E., Toner, J. J., & Cistola, D. P. (1995) *J. Biomol. NMR 6*, 198-210.

Hodson, M. E., Ponder, J. W., & Cistola, D. P. (1996) *J. Mol. Biol. 264*, 585-602.

Hodson, M. E., & Cistola, D. P. (1997a) *Biochemistry 36*, 1450-1460.

Hodson, M. E., & Cistola, D. P. (1997b) *Biochemistry 36*, 2278-2290.

Honig, B., Hudson, B., Sykes, B. D., & Karplus, M. (1971) *Proc. Natl. Acad. Sci. U. S. A.* 68, 1289-1293.

Jamison, R. S., Newcomer, M. E., & Ong, D. E. (1994) *Biochemistry 33*, 2873-2879.

Kleywegt, G. J., Bergfors, T., Senn, H., le Motte, P., Gsell, B., Shuao, K., & Jones, T. A. (1994) *Structure 2*, 1241-1258.

Norris, A. W., Rong, D., d'Avignon, D. A., Rosenberger, M., Tasaki, K., & Li. E. (1995) *Biochemistry 34*, 15564-15573.

Rizo, J., Liu, Z. P., & Gierasch, L. M. (1994) *J. Biomol. NMR 4*, 741-760.

Sporn, M. B., Roberts, A. B. & Goodman, D. S. (1994) *The Retinoids: Biology, Chemistry, and Medicine*, 2nd ed. Raven Press, New York.

Sporn, M. B., Roberts, A. B. & Goodman, D. S. (1994) *The Retinoids: Biology, Chemistry, and Medicine*, 2nd ed. Raven Press, New York.

Waterhous, D. V., & Muccio, D. D. (1990) *Magn. Reson. Chem. 28*, 223-226.

Werbelow, L. G., & Grant, D. M. (1977) *Adv. Magn. Reso. 9*, 189-299.