



3 1293 01771 8366

LIBRARY
Michigan State
University

This is to certify that the
thesis entitled
NMR Studies of Human Annexin I and
Yeast Guanylate Kinase

presented by

Jinhai Gao

has been accepted towards fulfillment
of the requirements for

M.S. degree in Biochemistry



Major professor

Date

5/13/99

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
MAY 17 2002		

**NMR STUDIES OF
HUMAN ANNEXIN I AND YEAST GUANYLATE KINASE**

By

Jinhai Gao

A THESIS

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

MASTER OF SCIENCE

Department of Biochemistry

1999

ABSTRACT

NMR STUDIES OF HUMAN ANNEXIN I AND YEAST GUANYLATE KINASE

By

Jinhai Gao

Annexins are excellent models for studying the folding mechanisms of multidomain proteins because they have 4-8 domains with high similarity in folding but low identity in sequence. The solution structure of an isolated domain 1 of human annexin I has been determined by NMR spectroscopy. The root-mean-square deviation of the ensemble of 20 refined conformers was 0.57 ± 0.14 Å for the backbone atoms. The NMR structure of domain 1 could be superimposed with an RMSD of 1.36 Å for all backbone atoms with the corresponding part of the crystal structure of a truncated human annexin I containing all four domains. The result suggests that isolated domain 1 constitutes an autonomous folding unit and interdomain interactions may play critical roles in the folding of annexin I. A sequential working model was proposed for the folding of annexin I

Guanylate kinase (GK) is a suitable model enzyme for NMR studies of structural and dynamic properties of nucleoside monophosphate kinases. A series of 2D and 3D NMR data have been collected for free and GMP-bound forms of GK. Sequential backbone resonance assignments for the GK complex with GMP have been made. The results obtained in this work provide the basis for the NMR studies of the structure-function relationships of GK. Proposals for the further efforts towards elucidating dynamic and structural changes that control kinase catalysis were also discussed.

To my parents

ACKNOWLEDGEMENTS

First of all, I would like to express my appreciation to my thesis advisor, Dr. Honggao Yan, for his guidance and support for my graduate study at MSU. Thanks are also due to my guidance committee members, Dr. Zachary Burton and Dr. Michael Garavito.

I would like to thank Dr. Yue Li for establishing the expression system for human annexin I and yeast guanylate kinase. Thanks also go to other lab members, Dr. Yanling Zhang for her helpful discussion at the beginning of this project, Dr. Lincong Wang for his critical but insightful suggestions during my research, and Dr. Genbin Shi for preparing the inhibitor GP₅A.

I Thank Mr. Kermit Johnson, Dr. George Gray, and Dr. Ouwen Zhang for their help in NMR experiments. I Thank Drs. Charles Cottrell, Clemens Anklin and George Gray for assistance in acquiring the NMR data for annexin I. I Thank Drs. Xiangwei Weng and Sung-Hou Kim for providing us the unpublished coordinate of the refined crystal structure of annexin I. I Thank Dr. Zhiheng Huang for collecting mass spectrometry data for isotope-labeled GK samples.

Finally, I am grateful to many of my friends and classmates whose friendship helped me go through my graduate life at MSU.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF ABBREVIATIONS	viii
INTRODUCTION	1
CHAPTER 1	
NMR SOLUTION STRUCTURE OF DOMAIN 1 OF HUMAN ANNEXIN I	
1.1 Introduction	4
1.2 Materials and Methods	9
1.3 Results	13
1.4 Discussion	32
References	42
CHAPTER 2	
SEQUENTIAL BACKBONE RESONANCE ASSIGNMENTS OF YEAST	
GUANYLATE KINASE IN COMPLEX WITH GMP	
2.1 Introduction	46
2.2 Materials and Methods	55
2.3 Results	57
2.4 Discussion	69
2.5 Summary and Perspective	70
References	72
APPENDICES	75

LIST OF TABLES

CHAPTER 1

Table 1.1	^{15}N , ^{13}C and ^1H resonance assignments for domain 1 of human annexin I	21
Table 1.2	Statistics of NMR solution structures of domain 1 of human annexin I	27

CHAPTER 2

Table 2.1	^{15}N , ^{13}C and ^1H backbone resonance assignments for yeast guanylate kinase in complex with GMP	65
-----------	--	----

LIST OF FIGURES

CHAPTER 1

Fig. 1.1 (a) Ribbon diagram of the X-ray structure of a truncated human annexin I that lacks the N-terminal 31 residues.(b) Superposition of the four domains of annexin I. (c) Sequence alignment of the four domains.	6
Fig. 1.2 Strip plot of 3D HNCACB and 3D CACB(CO)NH spectra of Domain 1 of annexin I.	15
Fig. 1.3 Strip plot of ^{15}N -edited NOESY-HSQC spectrum of domain 1 of annexin I.	17
Fig. 1.4 ^1H - ^{15}N HSQC spectrum of domain 1 of human annexin I.	19
Fig. 1.5 Summary of sequential and short-range NOEs and chemical shift indices for H^α and C^α observed for domain 1 of annexin I.	25
Fig. 1.6 (a) Superposition of the final 20 calculated NMR structures of domain 1 of annexin I.(b) Superposition of the minimized average NMR structure and the X-ray structure of domain 1.	29
Fig. 1.7 Distributions of the average backbone RMSDs of the ensemble of the NMR structures from its mean coordinate and from the X-ray crystal structure.	31
Fig. 1.8 The hydrophobic core structure of domain 2 and the interface between domain 2 and domain 4 of annexin I.	35
Fig. 1.9 A working model for the folding process of annexin I.	40

CHAPTER 2

Fig. 2.1 The crystal structure of yeast guanylate kinase in complex with GMP.	49
Fig. 2.2 Domain movements correlated with substrate binding to adenylate kinase.	52
Fig. 2.3 Strip plot of 3D HNCACB (A) and 3D CACB(CO)NH (B) spectra of yeast GK in complex with GMP.	59
Fig. 2.4 Strip plot of ^{15}N -edited NOESY-HSQC spectrum of yeast GK in complex with GMP.	62
Fig. 2.5 ^1H - ^{15}N HSQC spectrum of yeast GK in complex with GMP.	64

LIST OF ABBREVIATIONS

2D:	two-dimensional;
3D:	three-dimensional;
DQF-COSY:	double quantum filtered correlation spectroscopy;
HSQC:	heteronuclear single quantum coherence;
IPTG:	isopropyl-1-thio- β -D-galactopyranoside;
NMR:	nuclear magnetic resonance;
NOESY:	nuclear Overhauser effect spectroscopy;
PAGE:	polyacrylamide gel electrophoresis;
RMSD:	root-mean-square deviation;
SDS:	sodium dodecyl sulfate;
TOCSY:	total correlation spectroscopy;
HCCH-TOCSY:	proton-carbon-proton correlation using carbon total correlated spectrum;
HNCA:	amide proton to nitrogen to α -carbon correlation;
HNCO:	amide proton to nitrogen to carbonyl carbon correlation;
HNCACB:	amide proton to nitrogen to α/β -carbon correlation;
CBCA(CO)NH:	α/β proton to α/β carbon (via carbonyl carbon) to nitrogen to amide proton correlation;
(HB)CBCACO(CA)HA:	α/β carbons to α/β carbonyl carbon (via α -carbon) to a proton correlation;
GK:	guanylate kinase;

AK:	adenylate kinase;
UK:	uridylate kinase;
GP ₅ A:	(P ¹ -(5'-adenosyl) P ⁵ -(5'-guanosyl) pentaphosphate);
AP ₅ A:	(P ¹ ,P ⁵ -(5'-diadenosyl) pentaphosphate;
UP ₅ A:	(P ¹ -(5'-adenosyl) P ⁵ -(5'-uridyl) pentaphosphate.

INTRODUCTION

The developments of modern molecular biology and multidimensional nuclear magnetic resonance (NMR) spectroscopy have increased explosively the use of NMR spectroscopy for studying the structure-function relationships of biological molecules. NMR spectroscopy and X-ray crystallography are complementary methods for studying biomolecular structure and dynamics. While X-ray crystallography is more productive and can be applied to very large biomolecules, NMR data can be interpreted in terms of dynamic models in solution. The most important dynamics for biological function are those with time constants on the order of a nanosecond to second, and it is in this time range that NMR relaxation measurement is most powerful. Although it is still not easy to interpret relaxation data with full confidence, NMR relaxation experiments have been applied successfully in a number of dynamic studies of proteins (1).

In this thesis, NMR spectroscopy has been applied to study two proteins, human annexin I and yeast guanylate kinase. With the well-defined domains and the symmetric structure, annexins are excellent models for studying the folding mechanisms of multidomain proteins (2). Our approach to dissect the folding mechanism of annexin I is to compare the folding properties of the intact protein and the four isolated domains. The results showed that domain 1 of human annexin I constitutes an autonomous folding unit and interdomain interactions may play critical roles in the folding of annexins. The results allowed us to propose a possible scenario for the folding process of annexin I. Guanylate kinase (GK) belongs to a family of nucleoside monophosphate (NMP) kinases.

GK is required for the metabolic activation of the anti-herpes drugs acyclovir and gancyclovir and anti-HIV agent carbovir (3-5). Thus it is of biomedical significance to study the catalytic mechanism of this enzyme. It has been suggested that the substrate-induced domain closure and the dynamic relocation are important for the catalysis of NMP kinase (6,7). Guanylate kinase is an excellent model enzyme to study these conformational and dynamic changes by NMR spectroscopy. This thesis presents the multinuclear multidimensional NMR experiments and sequential backbone assignments of the GK complex with GMP. The results provide the basis for further NMR studies of the structure-function relationship of this enzyme.

References

1. Wagner, G. (1997) *Nature Struct. Biol. (NMR supplement)* 4, 841-844.
2. Liemann, S., and Huber, R. (1997) *Cell. Mol. Life Sci.* 53, 516-521
3. Miller, W. H. and Miller R. L. (1980) *J. Biol. Chem.* 255,7204-7207.
4. Boehme, R. E. (1984) *J. Bio. Chem.* 259, 12346-12349.
5. Miller, M. H., Daluge, S. M., Garvey, E. P., Hopkins, S., Reardon, J. E., Boyd, F. L., and Miller, R. L. (1992) *J. Biol. Chem.* 267, 21220-21224.
6. Vonrhein, C., Schlauderer, G. L., and Schulz, G. E. (1995) *Structure* 3, 483-490.
7. Müller, C., Schlauderer, G. L., Reinstein, J., and Schulz, G. E. (1996) *Structure* 4, 147-156.

CHAPTER 1

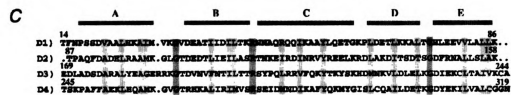
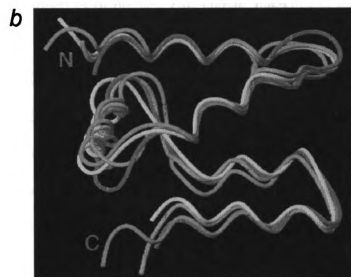
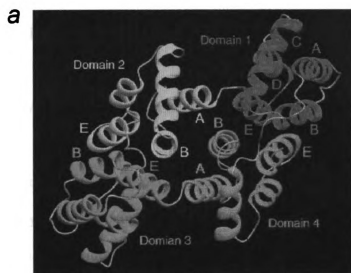
NMR SOLUTION STRUCTURE OF DOMAIN 1 OF HUMAN ANNEXIN I

1.1 Introduction

Most proteins in nature are large multidomain proteins (1). While a great deal of knowledge on the folding properties of small single-domain proteins has been acquired (2), our understanding of the folding of multidomain proteins is still poor. To date, the folding mechanisms of few multi-domain proteins has been studied. It has been suggested that the domains of large proteins fold independently and subsequently assemble to form the native structures (3-5).

Annexins are a large family of ubiquitous proteins that bind to phospholipids in the presence of calcium ions (6,7). Although their physiological functions are not clear, these proteins are implicated in many important cellular processes (8) such as exocytosis (9,10) and ion channeling (11). All annexins contain four homologous repeats of ~70 residues (Fig. 1.1a and 1.1c) and a variable N-terminus, with the exception of annexin VI which has four additional repeats. The crystal structures of annexins I, II, III, IV, V, VI, VII and XII have been determined (12). As revealed by X-ray crystallography, each repeat forms a compact domain consisting of five helix segments, named A to E, organized in a typical super-helix topology. All the domains are highly similar in structure, as illustrated in Fig. 1.1b with the four domains of annexin I. The four domains of each annexin are arranged in a planar-cyclic manner with domain 4 in contact with

Figure 1.1 (a) Ribbon diagram of the X-ray structure of a truncated human annexin I that lacks the N-terminal 31 residues (18). The four homologous domains are indicated in different colors: domain 1, green; domain 2, yellow; domain 3, cyan; and domain 4, magenta. Except domain 1, only the helices involved in the interdomain interactions are labeled. (b) Superposition of the four domains of annexin I: domain 1 (17-86), domain 2 (87-158), domain 3 (169-246) and domain 4 (247-319). Domains 1 to 4 are colored as in (a). Only the helices of each domain were used for the structural alignment. (c) Sequence alignment of the four domains. The numbering is according to the crystal structure of the truncated annexin I (18). The hydrophobic core residues are shown in yellow, and other conserved residues in blue. Fig. 1a and 1b were generated using the program MOLMOL (43).



domain 1, as depicted in Fig. 1.1a. Domains 1 and 4 as well as domains 2 and 3 have many tight hydrophobic contacts, mainly involving helices B and E, constituting two two-domain modules. The interactions between these two modules are mostly hydrophilic via helices A and B of domains 2 and 4, forming a central hydrophilic channel.

With the well-defined domains and the simple and elegant structure, annexins are excellent models for studying the folding mechanisms of multidomain proteins. They are composed of four domains with almost identical topologies but only limited sequence homology of approximately 30%. Using synthetic peptides and more recently recombinant peptides, Sanson and collaborators have been systematically studying the folding properties of domain 2 of human annexin I (13-16). They have clearly shown, with CD and NMR, that isolated domain 2 of annexin I is largely unfolded in aqueous solution (15). A preliminary study on the folding properties of domain 1 has also been reported (17).

Our approach to dissect the folding mechanism of annexin I is to compare the folding properties of the intact protein and the four isolated domains. We have expressed the entire annexin I and the four individual domains in *Escherichia coli*. Expression of separated domain 3 and 4 in *Escherichia coli* results in inclusion bodies. Domain 2 was found to be largely unfolded in solution, although it contains a significant amount of secondary structure in solution. Using multidimensional NMR techniques, we have determined the solution structure of domain 1 (residues 14-86, according to the numbering of the crystal structure of an N-terminally truncated human annexin I (18)). The NMR structure of the isolated domain 1 is highly similar to the corresponding part of

the crystal structure of a truncated human annexin I containing all four domains (18). The result shows that in contrast to isolated domain 2, isolated domain 1 constitutes an autonomous folding unit. Comparative structural analysis suggests that inter-domain interactions may play critical roles in the folding of annexin I.

1.2 Materials and Methods

Materials. The *Escherichia coli* clone containing the cDNA encoding human annexin I was purchased from ATCC (ATCC number 65114, deposited by Joel Ernst). The expression vector pET-17b was purchased from Novagen. DNA sequencing kit was obtained from United States Biochemical. Enzymes for recombinant DNA experiments were purchased from Gibco BRL or New England Biolabs. $^{15}\text{NH}_4\text{Cl}$ and $[^{13}\text{C}_6]$ D-glucose were purchased from ISOTECH. Other chemicals were analytical or reagent grade from commercial sources.

Cloning. The amino acid sequence of domain 1 of human annexin I is shown in Fig. 1c. The portion of human annexin I cDNA that encodes domain 1 was cloned into the expression vector pET-17b by PCR and other standard recombinant DNA techniques. The primers used for the PCR cloning were 5'-GGAATTCCATATGACCTTCAATCCATCCTCG-3' (forward) and 5'-CCGGATCCTTATTTTAGCAGAGCTAAAACAAC-3' (reverse). The correct amino acid sequence was verified by double stranded DNA sequencing of the DNA insert in the expression construct pET-17b-ANX1D1.

Expression and purification. Unlabeled protein was produced by growing the *Escherichia coli* strain BL21(DE3) containing the expression construct pET-17b-ANX1D1 in LB media in the presence of 100 $\mu\text{g/ml}$ ampicillin at 37 °C without IPTG induction. Uniformly ^{15}N -labeled protein was produced by growing the same expression strain in M9 media with $^{15}\text{NH}_4\text{Cl}$ as the sole nitrogen source, and uniformly $^{15}\text{N}/^{13}\text{C}$ -labeled protein in M9 media with $^{15}\text{NH}_4\text{Cl}$ and $[^{13}\text{C}_6]$ D-glucose as the sole nitrogen and carbon sources. Protein production in the M9 media was induced by addition of IPTG to

a final concentration of 0.4 mM when the cultures reached an OD₆₀₀ of ~1.0. The culture was incubated for four more hours after addition of IPTG. The bacterial cells were harvested by centrifugation and suspended in buffer A (40 mM acetate, pH 5.3). The bacterial suspension was sonicated on ice and centrifuged (27,000 g) at 4 °C for 30 min. The supernatant was applied to a CM-cellulose column equilibrated with buffer A. The column was washed with buffer A until OD₂₈₀ of the eluent was less than 0.05. Elution of the column was achieved by a linear NaCl gradient (0-500 mM in buffer A) and monitored by OD₂₈₀ and 15% SDS-PAGE. The fractions containing domain 1 of annexin I were pooled and concentrated by an Amicon ultrafiltration cell using a YM 3 membrane. The protein preparations were >95% pure as judged by SDS-PAGE. Isotopically labeled proteins were further purified by a Sephadex G-50 column. The protein solutions were dialyzed against double distilled water, lyophilized and stored at – 80 °C.

NMR spectroscopy. NMR samples were prepared by dissolving the lyophilized protein in 20 mM acetate-d₃, pH 5.2 (pH meter reading without correction for isotope effects), in H₂O/²H₂O(9/1) or ²H₂O. The protein concentrations of the NMR samples were 2-5 mM. NMR spectra were acquired at 25 °C on a Bruker DMX 600 spectrometer at The Ohio State University, a Bruker DRX 600 spectrometer at Bruker USA, or a Varian INOVA 600 spectrometer at Varian Application Laboratories. Homonuclear 2D spectra recorded were DQF-COSY (D₂O) (19,20), TOCSY (D₂O) (21-23), and NOESY (H₂O) (24,25). Heteronuclear double and triple resonance spectra acquired included 2D ¹H-¹⁵N HSQC (26,27), 3D ¹H-¹⁵N TOCSY-HSQC (28), 3D ¹H-¹⁵N NOESY-HSQC (28,29), HNCACB (30,31), CBCA(CO)NH (31,32), and HCCH-TOCSY (33,34). The

acquisition sweep widths and numbers of complex points for these experiments were as follow: 2D DQF-COSY, TOCSY and NOESY, $^1\text{H}(\text{F1})$ 7183 Hz, 512, $^1\text{H}(\text{F2})$ 7183 Hz, 512; 2D ^1H - ^{15}N HSQC, $^{15}\text{N}(\text{F2})$ 2500 Hz, 256, $^1\text{H}(\text{F2})$ 7000 Hz, 102; ^{13}C - ^1H HSQC, $^{13}\text{C}(\text{F2})$ 27163 Hz, 512, $^1\text{H}(\text{F2})$ 7000 Hz, 962; ^{15}N -edited NOESY-HSQC with a 150 ms mixing time and a ^{15}N -edited TOCSY-HSQC experiments with a 47.3 ms mixing time, $^1\text{H}(\text{F1})$ 7183Hz, 256, $^{15}\text{N}(\text{F2})$ 2074Hz, 64, $^1\text{H}(\text{F3})$ 7183 Hz, 1024; 3D HNCACB, $^{15}\text{N}(\text{F1})$ 2200 Hz, 48, $^{13}\text{C}(\text{F2})$ 9000 Hz, 256, $^1\text{H}(\text{F3})$ 8000 Hz, 1024; 3D CBCA(CO)NH, $^{15}\text{N}(\text{F1})$ 2310 Hz, 62, $^{13}\text{C}(\text{F2})$ 8000 Hz, 94, $^1\text{H}(\text{F3})$ 8000 Hz, 1024; 3D HCCH-TOCSY, $^1\text{H}(\text{F1})$ 6238 Hz, 128, $^{13}\text{C}(\text{F2})$ 10000 Hz, 128, $^1\text{H}(\text{F3})$ 8000 Hz, 1024.

The spectra were processed with the program NMRPipe (35) and analyzed with the program PIPP (36). Briefly, solvent suppression was improved by convolution of time domain data (37). The data size in each indirectly detected dimension of the 3D data was extended by backward-forward linear prediction (38). A 45°-shifted sine bell and single zero-filling were generally applied before Fourier transformation in each dimension.

Derivation of structural restraints. Approximate interproton distance restraints were derived from sequentially assigned NOEs. NOE cross peaks between aliphatic protons were picked from the homonuclear 2D NOESY spectrum, and those involving amide protons from the 3D ^1H - ^{15}N NOESY-HSQC spectrum. The NOE intensities obtained by the program PIPP were converted into approximate interproton distances by normalizing them against the calibrated intensities of NOE peaks between backbone amide protons (d_{NN}) within the identified α -helices. The upper limits of the interproton distances were calibrated according to the equation $V_a = V_b (r_b/r_a)^6$, where V_a , V_b were the NOE intensities and r_a , r_b the distances. The distance bounds were then 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 (39). An additional 0.5 Å was added to the upper bounds for methyl protons.

Structure calculation. NMR structures were calculated with a hybrid distance geometry-simulated annealing protocol (40) using the program X-PLOR (version 3.1) (41) on an SGI Indigo II 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 interactions, with atomic radii set to 0.80 times their CHARMM values (42) and a force constant of 4.0 kcal mol⁻¹ Å⁻⁴. A total of fifty structures were generated using this protocol. The structures were inspected by the programs MOMOL (43), QUANTA96 (Molecular Simulations) and analyzed by PROCHECK-NMR (version 3.4.4) (44,45). An iterative strategy was used for the structure refinement. In each round of structure refinement, newly computed NMR structures were employed to assign more NOE restraints, to correct wrong assignments, and to loosen the NOE distance bounds if spectral overlapping was deduced. Then another round of structure refinement was carried out with the modified NMR restraints. All structures were converged after several rounds of such refinement. An ensemble of 20 structures was selected according to their best fit to the experimental NMR restraints and the low values of their total energies.

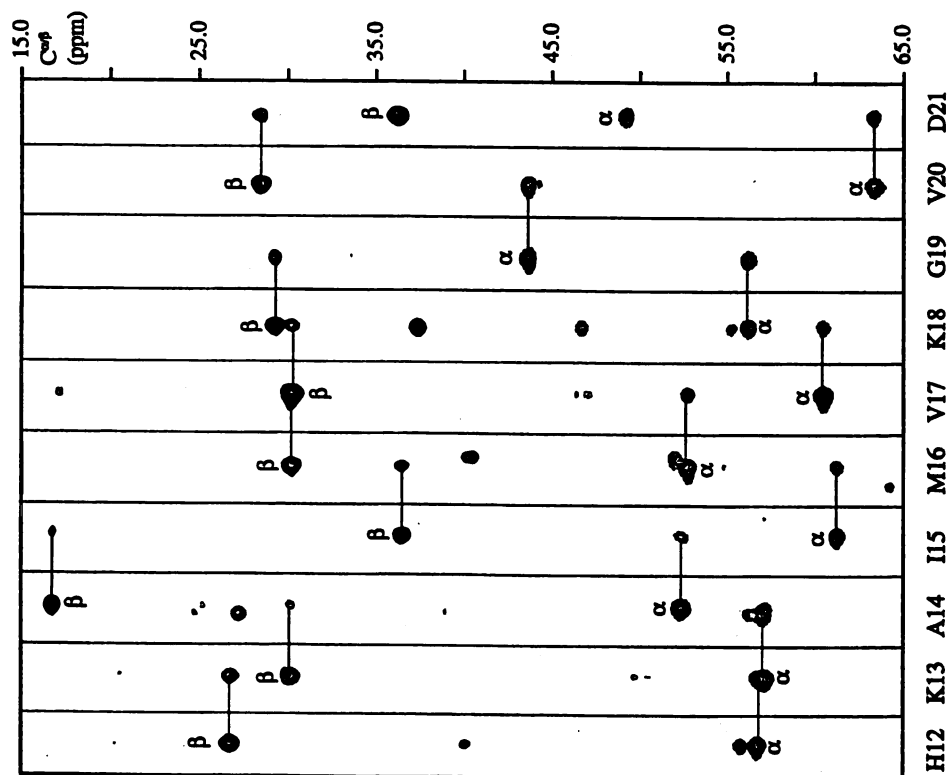
1.3 Results

Sequential backbone resonance assignments. Total sequential resonance assignments of the isolated domain 1 were achieved by the combined analysis of 2D and 3D NMR data, including 3D HNCACB, CBCA(CO)NH and HCCH-TOCSY. The combination of HNCACB and CBCA(CO)NH provided most of the sequential linkage of domain 1. Figure 1.2 shows the sequential connectivities from His12 to Asp21. In some cases, the triple-resonance spectra were incomplete because of a lack of C^α and C^β chemical shifts due to the low sensitivity of the HNCACB experiment, such as Thr30, Val67, Val68, and Leu71. The sequential connectivities for these residues could be made through sequential NOE analysis from ^{15}N -edited NOESY-HSQC spectrum. Almost all the $\text{H}^{\text{N}}\text{-H}^\alpha$ correlations could be obtained from the 3D ^{15}N -edited TOCSY-HSQC experiment. Then the sequential assignments were made from the $\text{H}^\alpha\text{-H}^{\text{N}}$ and $\text{H}^{\text{N}}\text{-H}^{\text{N}}$ NOE connectivities in the 3D ^{15}N -edited NOESY-HSQC experiment. Examples of $\text{H}^\alpha_i\text{-H}^{\text{N}}_{i+1}$ and $\text{H}^{\text{N}}_i\text{-H}^{\text{N}}_{i+1}$ connectivities from Arg37 to Thr48 of domain 1 are shown in Figure 1.3. The assignments obtained from triple-resonance experiments are in good agreement with the sequential NOE analysis. The sequential assignments of the backbone and side-chain amide resonances are shown in a $^{15}\text{N}\text{-}^1\text{H}$ HSQC spectrum in Fig. 1. 4

Side-chain resonance assignments. Most H^α , as well as some H^β and H^γ resonances were assigned in ^{15}N -edited TOCSY-HSQC spectrum. Extensions of assignments further along the side chain were made by the use of a 3D HCCH-TOCSY experiment. A 3D HCCH-TOCSY experiment yielded sequence-specific assignments of side chain proton resonances and their attached ^{13}C resonances for nearly all the aliphatic

Fig. 1.2 Strip plot of 3D HNCACB (A) and 3D CACB(CO)NH (B) spectra of domain 1 of annexin I, showing the sequential J connectivities of C^α and C^β for the residues His-12 to Asp-21. Distinction between C^α and C^β resonances is aided by their opposite phases in HNCACB strips.

a. HNCACB



b. CBCA(CO)NH

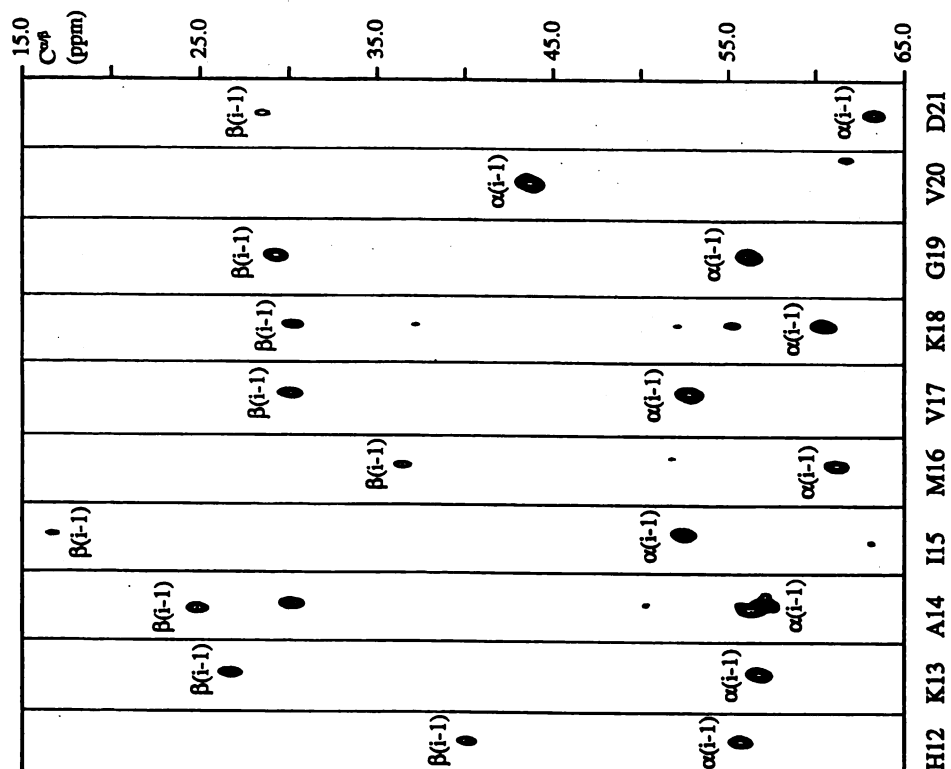


Fig. 1.3 Strip plot of ^{15}N -edited NOESY-HSQC spectrum of domain 1 of human annexin I, showing characteristic sequential d_{NN} and $d_{\alpha\text{N}}$ connectivities of the residues Arg-37 to Thr-48.

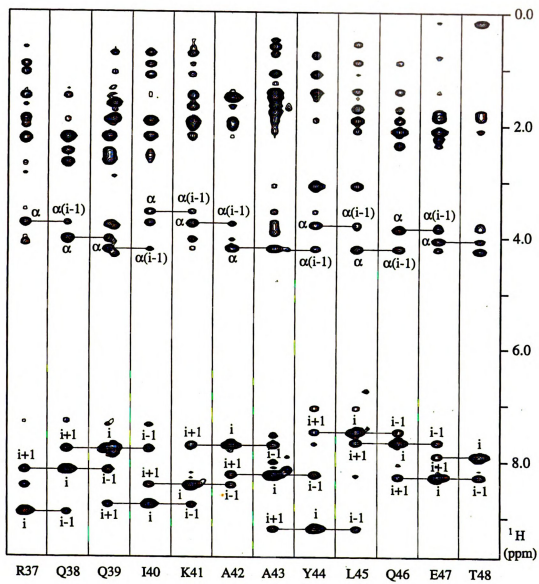


Fig. 1.4 ^{15}N - ^1H HSQC spectrum of domain 1 of human annexin I. Sequential assignments are indicated with one-letter amino acid codes and residue numbers. Pairs of cross peaks resulting from the side-chain NH_2 groups of asparagine and glutamine residues are connected by horizontal lines. The amino acid numbering is according to the isolated domain 1 with residue 1 corresponding to residue 14 in the crystal structure numbering.

residues except residue Leu71, which was assigned by 3D ^{15}N -edited TOCSY-HSQC and 2D TOCSY experiments.

Phenylalanine and tyrosine spin systems were assigned using 2D TOCSY and 2D DQF-COSY in D_2O . Aromatic side chain protons were then matched with the sequential assigned Phe3 and Tyr42 residues by the observation of NOEs between the ring protons and H^β protons. The two His residues, His12 and His63, were assigned using a combination of 2D TOCSY and NOESY spectra. Assignment of side chain amide resonances from three asparagine and four glutamine residues was made from the ^{15}N -edited NOESY-HSQC experiment, where NOEs from the amide to side chain protons were found.

Stereospecific assignments were made for about 70% of β -methylene protons and the methyl groups of valine and leucine residues based on qualitative estimations of $^3J_{\alpha\beta}$ constants from the DQF-COSY spectrum in conjunction with the NOE data (46). The complete ^1H , ^{15}N and ^{13}C assignments for domain 1 are listed in Table 1.1.

Secondary structure determination. The secondary structures were deduced from the characteristic NOE patterns and chemical shift indices. Figure 1.5 summarizes the sequential and medium-range NOEs and H^α and C^α secondary shifts for domain 1. As expected, many residues in domain 1 are found to possess features that are characteristic of an α -helix: positive C^α secondary shifts, negative H^α secondary shifts and strong $d_{\text{NN}}(i, i+1)$, $d_{\alpha\text{N}}(i, i+3)$ and $d_{\beta\text{N}}(i, i+3)$ NOE connectivities. Five helices were identified in domain 1: helix A (residues 5-15), helix B (22-30), helix C (34-47), helix D (52-58) and helix E (63-70).

Table 1.1 ^{15}N , ^{13}C and ^1H resonance assignments for domain 1 of human annexin I.

Residue	$^{15}\text{N}(\text{H}^{\text{N}})$	$^{13}\text{C}^{\alpha}(\text{H}^{\alpha})$	$^{13}\text{C}^{\beta}(\text{H}^{\beta})$	Others
T1	119.4(7.11)	59.9(4.33)	67.6(4.07)	C^{γ} 19.5(1.19)
F2	127.5(8.66)	55.3(4.63)	37.7(3.22,2.88)	C^{δ} 130.9(7.25); C^{ϵ} 130.4(7.35); C^{ξ} 127.8(7.36)
N3	128.5(8.50)	46.9(4.79)	37.6(2.85,2.52)	NH_2 7.53,6.96
P4		61.2(3.95)	30.1(1.88)	C^{γ} 24.2(2.91,1.92); C^{δ} 48.3(3.74)
S5	115.8(7.91)	60.0(4.00)	60.8(3.80)	
S6	120.1(7.96)	59.4(4.21)	60.9(3.88)	
D7	125.1(7.75)	55.7(4.47)	38.6(2.76,2.37)	
V8	121.6(8.44)	65.9(3.55)	29.4(2.18)	C^{γ} 23.0(1.10); 19.3(0.79)
A9	122.5(7.75)	52.9(4.13)	15.8(1.47)	
A10	122.3(7.87)	52.9(4.12)	16.8(1.45)	
L11	121.1(8.81)	55.9(3.83)	40.3(2.21,1.16)	C^{γ} 24.6(1.95); C^{δ} 22.8(0.75); 24.4(0.70)
H12	118.3(9.02)	57.1(4.50)	26.9(3.29,3.19)	C^{δ} 135.1(8.29); C^{ϵ} 117.5(7.17)
K13	118.7(7.84)	57.4(3.83)	30.4(1.90)	C^{γ} 22.4(1.49,1.34); C^{δ} 27.0(1.66); C^{ϵ} 39.8(2.91)
A14	121.4(7.76)	52.7(4.33)	16.9(1.62)	
I15	117.5(7.89)	61.6(3.74)	36.8(1.87)	C^{γ} 27.6(1.78); $\text{C}^{\gamma\text{Me}}$ 15.9(0.74); C^{δ} 12.1(0.61)
M16	117.8(7.30)	52.9(4.40)	30.5(2.04,1.91)	C^{γ} 29.5(2.35)
V17	122.9(7.28)	60.8(3.86)	30.4(2.02)	C^{γ} 19.9(1.06); 18.8(0.89)
K18	129.3(8.47)	50.5(4.00)	29.6(1.76,1.65)	C^{γ} 22.6(1.43); C^{δ} 26.8(1.64); C^{ϵ} 39.8(3.08,2.95)
G19	117.9(8.67)	43.9(4.13,3.65)		
V20	116.5(8.07)	63.6(3.56)	28.7(2.75)	C^{γ} 20.0(1.01); 19.7(0.70)
D21	127.7(8.37)	49.5(4.87)	36.6(2.93,2.60)	
E22	125.4(7.92)	57.4(3.57)	30.0(1.92,1.52)	C^{γ} 33.5(2.20,2.00)
A23	121.4(8.31)	53.4(3.57)	16.3(1.47)	
T24	117.8(7.30)	64.5(3.80)	65.4(3.68)	C^{γ} 19.9(1.10)
I25	120.1(6.81)	63.5(3.25)	36.3(1.73)	C^{γ} 26.5(1.71); $\text{C}^{\gamma\text{Me}}$ 14.9(0.64); C^{δ} 12.3(0.54)
I26	117.7(7.95)	61.4(3.38)	36.5(1.63)	C^{γ} 26.4(1.16); $\text{C}^{\gamma\text{Me}}$ 15.6(0.75); C^{δ} 10.5(0.54)
D27	124.9(8.20)	56.1(4.11)	39.8(2.75,2.62)	
I28	116.2(7.66)	62.9(3.53)	36.6(1.64)	C^{γ} 27.0(1.83,1.01); $\text{C}^{\gamma\text{Me}}$ 13.9(0.66); C^{δ} 12.0(0.75)
L29	113.2(7.77)	55.9(3.74)	40.0(1.72,1.12)	C^{γ} 24.3(1.94); C^{δ} 19.5(0.62); 24.6(0.62)
T30	105.3(7.95)	60.2(4.04)	66.6(4.22)	C^{γ} 20.1(1.15)
K31	122.9(7.33)	53.4(4.36)	30.5(1.91,1.81)	C^{γ} 23.4(1.46,1.34); C^{δ} 27.1(1.57); C^{ϵ} 39.7(2.85)
R32	120.8(6.96)	50.5(4.60)	28.9(1.66,1.48)	C^{γ} 22.7(1.32); C^{δ} 41.3(2.95)
N33	119.8(8.09)	49.1(4.62)	36.3(3.23,2.75)	NH_2 7.30,6.73
N34	117.5(8.52)	56.1(4.13)	36.8(2.87,2.73)	NH_2 7.66,7.15
A35	123.8(8.18)	53.4(4.01)	15.7(1.37)	

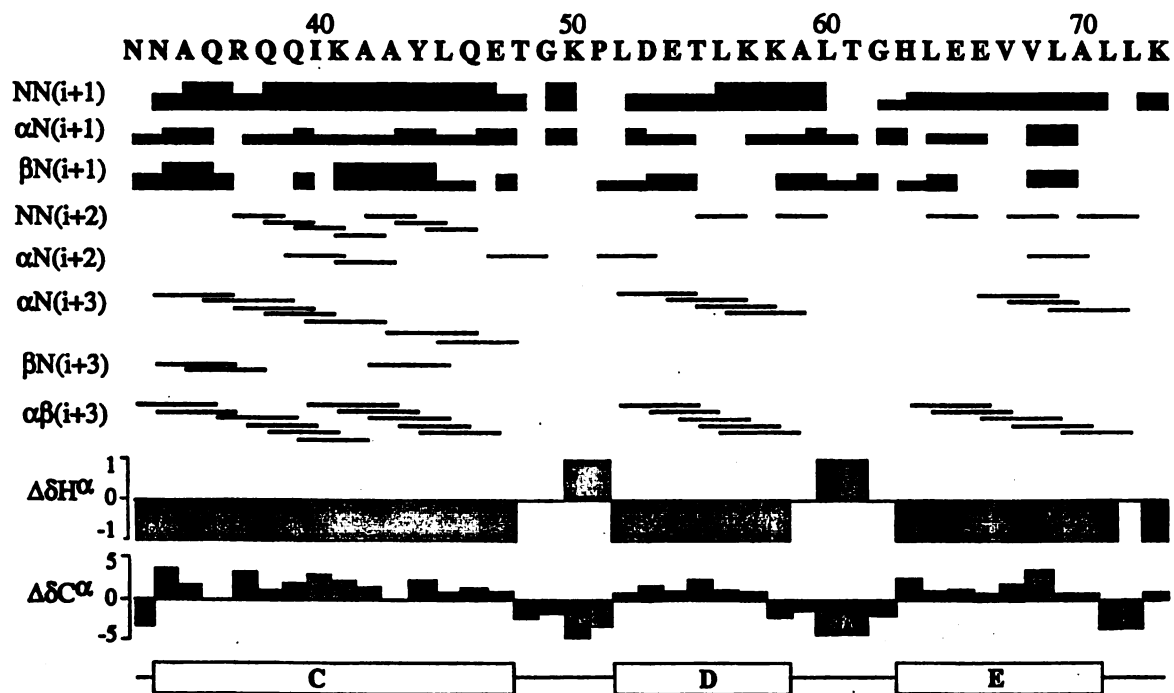
Table 1.1 (Continued)

Residue	$^{15}\text{N}(\text{H}^{\text{N}})$	$^{13}\text{C}^{\alpha}(\text{H}^{\alpha})$	$^{13}\text{C}^{\beta}(\text{H}^{\beta})$	Others
Q36	119.6(8.44)	56.2(3.80)	26.9(2.03)	C^{γ} 30.6(2.29,0.95); NH_2 6.57,6.56
R37	120.4(8.91)	59.1(3.76)	28.5(2.27)	C^{γ} 22.8(1.47); C^{δ} 41.6(3.52,2.91)
Q38	118.8(8.16)	56.5(4.06)	25.0(2.22)	C^{γ} 30.5(2.67,2.47); NH_2 7.29,6.83
Q39	121.5(7.79)	57.3(4.24)	27.3(2.57,2.20)	C^{γ} 32.8(2.65,2.48); NH_2 7.39,6.83
I40	123.4(8.78)	64.1(3.56)	36.0(1.94)	C^{γ} 28.1(2.21,1.13); $\text{C}^{\gamma\text{Me}}$ 14.9(0.73); C^{δ} 12.3(0.92)
K41	120.9(8.42)	58.9(3.77)	30.5(2.02,1.97)	C^{γ} 23.6(1.65); C^{δ} 28.1(1.55); C^{ϵ} 38.2(2.98)
A42	123.8(7.71)	53.1(4.20)	16.0(1.52)	
A43	123.0(8.24)	52.4(4.22)	16.5(1.43)	
Y44	123.0(9.20)	60.5(3.80)	37.6(3.08)	C^{δ} 117.4(6.53); C^{ϵ} 131.9(7.04)
L45	123.3(7.48)	55.3(4.24)	39.8(1.92,1.69)	C^{γ} 244.7(1.39); C^{δ} 23.4(0.89); 20.1(0.56)
Q46	120.1(7.66)	56.6(3.87)	26.5(2.11,1.97)	C^{γ} 31.2(2.36); NH_2 7.42,6.75
E47	116.8(8.28)	56.5(4.06)	27.9(1.87,1.78)	C^{γ} 33.1(2.22,2.08)
T48	107.4(7.90)	60.0(4.26)	68.8(3.86)	C^{γ} 15.3(0.19)
G49	112.6(7.91)	43.3(4.20,3.75)		
K50	122.8(8.15)	49.9(4.87)	32.6(1.70,1.60)	C^{γ} 22.3(1.31); C^{δ} 24.6(1.39); C^{ϵ} 39.9(3.09,2.94)
P51		59.4(4.67)	30.5(2.50)	C^{γ} 26.3(2.11,2.02); C^{δ} 48.4(3.82,3.53)
L52	129.1(9.93)	55.2(3.48)	39.2(1.58,1.04)	C^{γ} 24.0(1.14); C^{δ} 24.1(0.51); 18.2(-0.11)
D53	118.6(8.88)	54.9(3.73)	35.9(2.57,2.51)	
E54	121.2(7.96)	56.8(3.83)	28.1(2.00,1.85)	C^{γ} 33.4(2.30,2.18)
T55	116.0(7.55)	64.7(3.85)	66.4(4.21)	C^{γ} 19.4(1.25)
L56	123.0(8.24)	55.8(3.93)	38.9(1.75,1.09)	C^{γ} 24.3(1.64); C^{δ} 20.9(0.61); 24.3(0.48)
K57	120.4(8.02)	57.0(4.01)	30.6(1.81)	C^{γ} 22.4(1.33); C^{δ} 27.4(1.54); C^{ϵ} 39.7(2.74)
K58	115.0(6.92)	54.5(4.28)	30.8(1.91,1.81)	C^{γ} 22.8(1.52,1.42); C^{δ} 26.8(1.65); C^{ϵ} 39.8(2.95)
A59	121.6(7.52)	51.2(4.33)	18.8(1.38)	
L60	116.9(7.68)	51.0(4.68)	41.6(1.56,1.23)	C^{γ} 24.5(1.70); C^{δ} 20.2(0.79); 23.4(0.70)
T61	110.9(8.28)	57.9(4.64)	69.9(4.12)	C^{γ} 18.4(1.08)
G62	110.0(8.49)	43.6(4.00,3.81)		
H63	122.1(8.86)	57.5(4.54)	26.6(3.24,3.06)	C^{δ} 134.8(8.52); C^{ϵ} 118.4(7.35)
L64	120.7(8.62)	55.6(3.93)	39.3(1.71)	C^{γ} 24.8(1.53); C^{δ} 22.3(0.85); 22.2(0.80)
E65	119.4(6.99)	57.8(3.43)	24.5(2.33)	C^{γ} 30.3(2.41,2.01)
E66	118.0(7.39)	57.3(3.74)	27.4(2.16,1.97)	C^{γ} 33.3(2.37,2.22)
V67	117.3(7.91)	63.6(3.63)	29.3(1.94)	C^{γ} 22.8(0.76); 19.9(0.87)
V68	121.8(8.14)	64.7(3.36)	28.7(1.87)	C^{γ} 22.2(0.79); 20.1(0.60)
L69	119.8(8.09)	55.3(3.83)	37.2(1.78,1.39)	C^{γ} 25.7(1.78); C^{δ} 22.7(0.80); 20.1(0.73)

Table 1.1 (Continued)

Residue	$^{15}\text{N}(\text{H}^{\text{N}})$	$^{13}\text{C}^{\alpha}(\text{H}^{\alpha})$	$^{13}\text{C}^{\beta}(\text{H}^{\beta})$	Others
A70	122.3(7.36)	52.7(4.11)	15.9(1.46)	
L71	123.0(9.66)	51.9(3.07)	39.5(2.50)	C^{γ} 24.9(1.19)
L72	117.0(7.24)	52.2(4.19)	40.6(1.74,1.56)	C^{γ} 24.0(1.72); C^{δ} 20.5(0.62); 24.0(0.59)
K73	127.8(7.00)	57.0(3.90)	31.0(1.75)	C^{γ} 22.6(1.43); C^{δ} 26.9(1.64,1.50); C^{ϵ} 39.9(3.08,2.91)

Fig. 1.5 Summary of sequential and short-range NOEs and chemical shift index for H^α and C^α observed for domain 1 of annexin I. The derived helices are shown at the bottom.



Solution structure calculation. A total of 1099 structurally useful distance restraints were obtained from the analyses of the homonuclear 2D NOESY (D₂O) and 3D ¹H-¹⁵N NOESY-HSQC spectra (Table 1.2), 707 of which were medium- and long-range NOEs. In average, each residue had ~15 NOE restraints. A superposition of 20 calculated structures with no NOE restraint violations above 0.5 Å is shown in Fig. 1.6a. The structural analysis are summarized in Table 1.2. The precision of the structures (RMSD of the ensemble of the 20 NMR structures from its mean coordinate) was 0.57 Å for the backbone (N, C^α, C', O) and 1.11 Å for all heavy atoms. The distribution of the average backbone RMSDs is shown in Fig. 1.7a. The structure of domain 1 consists of five helices: helix A, residues 5-15; helix B, residues 22-30; helix C, residues 34-47; helix D, residues 52-58; and helix E, residues 63-70 (numbering according to the isolated domain 1). Helices A, B, D and E are assembled in a bundle with two nearly parallel helix-loop-helix motifs. Helix C lies approximately perpendicular to the helical bundle with one end close to the N-terminus and the other to the C-terminus of domain 1. The ensemble of the NMR structures and constraints have been deposited at the Protein Data Bank (<http://www.pdb.bnl.gov>) under PDB code 1bo9.

Table 1.2 Statistics of NMR solution structures of domain 1 of human annexin I.

Restraints for structure calculations			
Total NOE restraints		1099	
Intraresidue		392	
Medium range($1 \leq i-j \leq 4$)		549	
Long range($ i-j > 4$)		158	
Statistics for structure calculations		$\{SA\}^1$	$\langle SA \rangle_r$
NOE violations ($>0.5 \text{ \AA}$)		0	0
R.m.s.d. from distance restraints ($\text{\AA}$)		0.026 ± 0.001	0.027
R.m.s.d. from idealized geometry			
Bonds ($\text{\AA}$)		0.0034 ± 0.0001	0.0033
Angle ($^\circ$)		0.56 ± 0.01	0.54
Impropers ($^\circ$)		0.39 ± 0.02	0.36
X-PLOR potential energies (kcal/mol) ²			
E _{total}		220.3 ± 11.9	213.5
E _{noe}		38.8 ± 3.4	41.3
E _{repel}		53.7 ± 4.7	52.1
E _{impr}		12.8 ± 1.6	11.3
Ramachandran plot statistics ³			
Residues in most favored regions		78%	81.8%
Residues in additionally allowed regions		17.4%	15.2%
Residues in generously allowed regions		3.9%	3%
Residues in disallowed regions		0.7%	0%
R.m.s.d. of atomic coordinates ($\text{\AA}$)		backbone	heavy atoms
{SA} vs. $\langle SA \rangle$	All residues	0.57 ± 0.14	1.11 ± 0.19
	Helices only	0.47 ± 0.18	1.02 ± 0.23
{SA} vs. X-ray	All residues	1.36 ± 0.11	2.12 ± 0.13
	Helices only	1.01 ± 0.13	1.82 ± 0.16

¹ {SA} is the ensemble of 20 NMR solution structures of domain 1. $\langle SA \rangle$ is the mean atomic structure obtained by averaging the individual structures following a superimposition of the backbone heavy atoms. $\langle SA \rangle_r$ is the energy-minimized average structure.

² The distance constraints were used with a square-well potential ($F_{\text{noe}} = 50 \text{ kcal mol}^{-1} \text{\AA}^{-2}$). The F_{repel} function was used to simulate van der Waals interactions with a force constant of $4.0 \text{ kcal mol}^{-1} \text{\AA}^{-4}$ and atomic radii set to 0.8 times their CHARMM values.

³ The Ramachandran plot statistics were obtained from the PROCHECK-NMR analysis.

Fig. 1.6 (a) Superposition of the final 20 calculated NMR structures of domain 1 of annexin I. Only the backbone atoms (N, C^α and C') are superimposed and colored according to the secondary structure: helices A (5-15) in red, B (22-30) in green, C (34-47) in cyan, D (52-58) in magenta and E (63-70) in yellow and the loops in gray. The amino acid numbering is according to the isolated domain 1 with residue 1 corresponding to residue 14 in the crystal structure numbering. (b) Superposition of the minimized average NMR structure (red) and the X-ray structure (cyan) of domain 1.

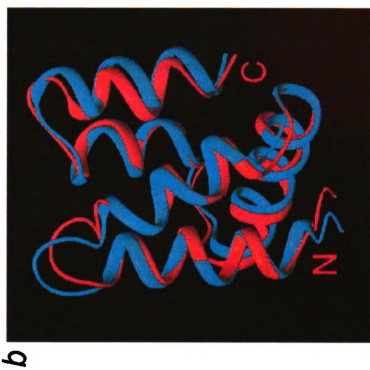
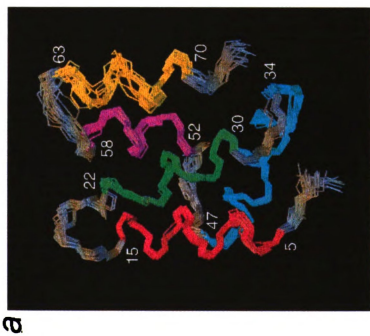
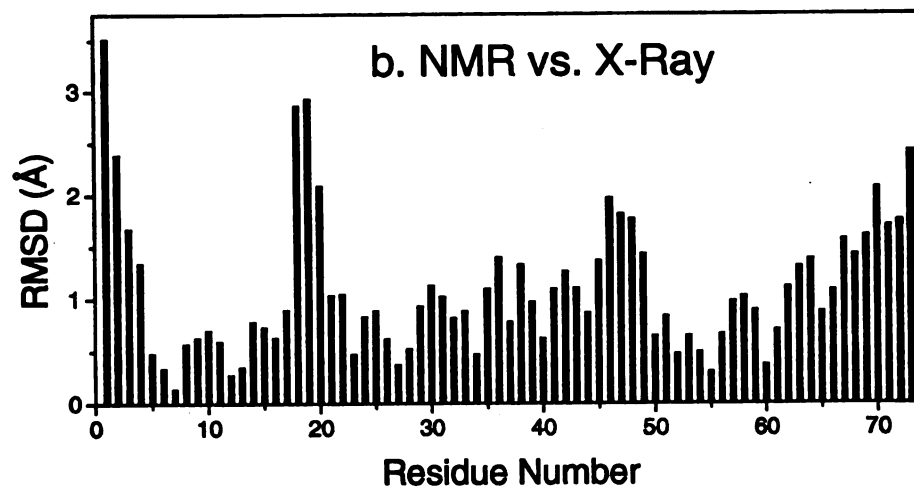
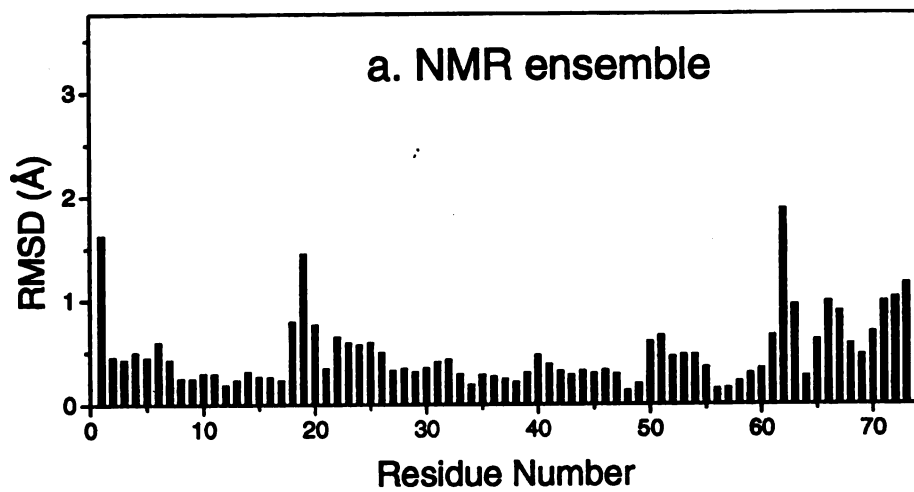


Fig. 1.7 Distributions of the average backbone RMSDs of the ensemble of the NMR structures from its mean coordinate (a, top) and from the X-ray crystal structure (b, bottom). The amino acid numbering is according to the isolated domain 1 with residue 1 corresponding to residue 14 in the crystal structure numbering.



1.4 Discussion

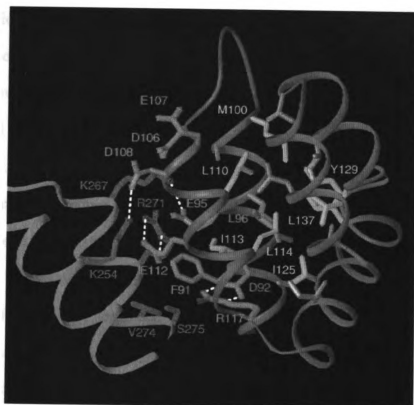
Comparison with the crystal structure of human annexin I. The structure of a truncated human annexin I has been determined by X-ray crystallography in the presence of 10 mM CaCl_2 (18). The truncated annexin I lacks the N-terminal 32 residues but has four domains all intact (Fig. 1.1a). Six calcium ions are found to bind to the truncated annexin I, two each in domains 1 and 4 and one each in domains 2 and 3. The solution structure of the isolated domain 1 is highly similar to the corresponding part of the crystal structure of the truncated annexin I containing all four domains. Thus, the minimized average NMR structure of the isolated domain 1 can be superimposed very well with the corresponding X-ray structure as shown in Fig. 1.6b. There are 1-2 residues differences in the lengths of some helices but the five helices are assembled in the same way. The distribution of the average backbone RMSDs of the ensemble of the 20 NMR structures from the corresponding X-ray structure is shown in Fig. 1.7b. The largest differences are found at the N-terminus and in the AB loop. It should be noted that the NMR structure of the isolated domain 1 was determined in the absence of Ca^{2+} . The difference in the conformations of the AB loop could be due to binding of Ca^{2+} because the carbonyls of Gly-32 and Val-33 in the AB loop along with the carboxylate of Glu-35 at the N-terminus of helix B form a calcium-binding site. However, binding of Ca^{2+} to the second calcium-binding site apparently does not cause any significant conformational change because the conformation of the DE loop that constitutes the second site is essentially the same as that found in the crystal structure, probably because the second site has lower affinity for Ca^{2+} than the first site.

Implications for protein folding. As described earlier, the four domains of annexin I are highly homologous in structure when folded together (Fig. 1.1a and b). The hydrophobic cores are highly conserved among all annexin domains. Surprisingly, isolated domain 2 is largely unfolded in aqueous solution and thus is not an independent folding unit (15). Its helical content is less than 25% compared to ~80% when the domain is folded together with the rest of the protein. In contrast to domain 2, our work presented here clearly demonstrates that the isolated domain 1 is fully folded in solution with little change in structure from that in the native state, and thus constitutes an autonomous folding unit. The results present the interesting question of why the domains with high sequential and structural homologies exhibit totally different folding behaviors.

The failure of the isolated domain 2 to form its native structure is likely due to the removal of the interdomain interactions that exist in the whole protein. As mentioned earlier, according to the crystal structure of annexin I (18), domains 2 and 3 form a modular structure with many hydrophobic interactions, and so do domains 1 and 4. Thus, it is unlikely that the removal of the hydrophobic contacts with domain 3 is the cause for the folding failure of the isolated domain 2. By default, then, the removal of the interactions with domain 4 may be the cause for the failure of the isolated domain 2 to fold to its native structure. Indeed, there are many interactions between domain 2 and domain 4 as shown in Fig. 1.8. This explanation is supported by the NMR studies of the isolated domain 2 and its components helices A and B (14,15).

It has been shown by NMR that a stable nonnative N-terminal cap, with the sequence F₉₁D₉₂A₉₃D₉₄E₉₅L₉₆ (numbering according to the crystal structure of the truncated annexin I), is formed in helix A in a peptide fragment containing helices A and

Fig. 1.8 The hydrophobic core structure of domain 2 and the interface between domain 2 and domain 4. The drawing is based on the X-ray structure of the truncated human annexin I containing four domains (18). The main-chains of domain 2 and domain 4 (partial) are represented by blue and cyan ribbons, respectively. The residues involved in the nonnative cap and the cluster of acidic residues as well as Arg-117 in domain 2 are shown in magenta. The residues within 5 Å distance of Leu-96 are shown in yellow, and other core residues in gray. The residues of domain 4 are in green. Hydrogen bonds are indicated by dotted lines. The amino acid numbering is according to the crystal structure of the truncated annexin I.



B of domain 2 (14). With the carboxyl groups of Asp-92 and Glu-95 hydrogen-bonded to their reciprocal backbone amides and many hydrophobic contacts between Phe-91 and Leu-96, it is a canonical N-terminal cap (47,48). Furthermore, the nonnative cap persists in isolated domain 2 (15,16). It has been suggested that the nonnative N-terminal cap serve as a very potent initiation site for folding (14). However, it may be more likely that the formation of the nonnative N-terminal cap prevents the isolated domain 2 from reaching the native state for two reasons, although its role in the folding of entire annexin I is not known. First, it disrupts a pair of hydrogen bonds between the carboxyl group of Asp-92 and the guanidinium group of Arg117 that helps to lock helices A and B in place (18) (Fig. 1.8). The breakage of the hydrogen bond also makes it possible for Arg117 to form nonnative salt bridges as found in the isolated domain 2 (16). Second, as shown in Fig. 8, in the native structure, Leu96 is roughly at the center of the hydrophobic core. It is surrounded by as many as seven core residues: Met-100 from helix A, Leu110, Ile113 and Ile114 from helix B, Ile125 and Tyr129 from helix C, and Leu137 from helix D. On the other hand, the side-chains of Phe91 and Leu96 are >10 Å apart. Thus, the nonnative hydrophobic interactions between Phe91 and Leu96 in the isolated domain may not only take the side-chain of Leu96 out of the hydrophobic core structure but also disrupt the packing of the other hydrophobic core residues. The nonnative conformation of the isolated domain 2, however, may not necessarily have a lower energy than the native conformation. The nonnative N-terminal cap may act as a kinetic trap that keeps the isolated domain 2 from reaching the native structure.

Why does the nonnative N-terminal cap form in the isolated domain 2? The separation of domain 2 from the rest of the protein has two structural consequences that

may bear on the formation of the nonnative N-terminal cap as shown in Fig. 1.8. First, it breaks four hydrogen bonds between domains 2 and 4, namely Glu95/Lys267, Asp108/Lys254 and Glu112/Arg271 (two hydrogen bonds). The salt bridge between Glu107 of domain 2 and Lys235 of domain 3 is also broken. This leaves a cluster of negatively charged residues without positively charged partners, including Glu95, Asp106, Glu107, Asp108, and Glu112. The carboxyl group of Glu95 is ~ 6.7 Å away from that of Asp106 and ~ 7.1 Å away from that of Glu-112. It is likely that the negative charge potential generated by the cluster of acidic residues may push away the carboxyl group of Glu95 so that it forms a hydrogen bond to the backbone amide of Asp92. Second, Phe91 is almost completely buried in the whole protein but its side-chain becomes mostly exposed to solvent in the isolated domain 2. Thus, Phe91 in the isolated domain 2 seeks hydrophobic partners and it finds Leu96. It is noted that Phe91 and Glu95 are replaced by a serine and an alanine, respectively, in domain 1 (Fig. 1.1c). Therefore, the nonnative N-terminal cap is unlikely to form in the folding process of the isolated domain 1. The hypothesis may be tested by replacing Phe91 and Glu95 of domain 2 with the corresponding amino acids of domain 1 by site-directed mutagenesis. Refolding at a higher salt concentration may also help the isolated domain 2 to reach the native conformation by reducing the effects of the negative charges of the cluster of acidic residues and strengthening the hydrophobic interactions to drive formation of the hydrophobic core.

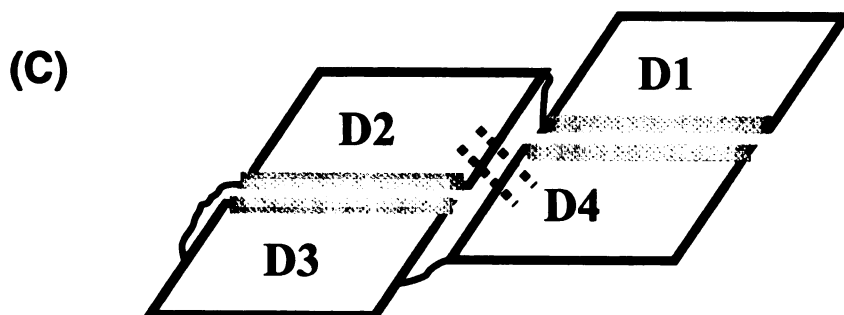
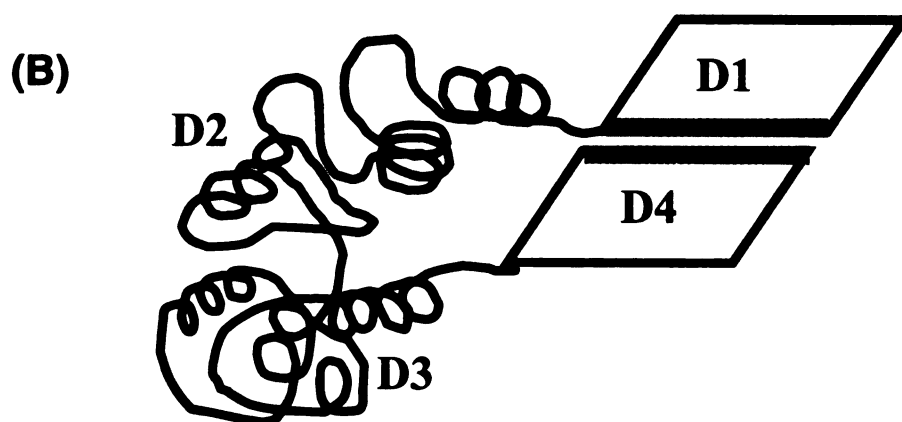
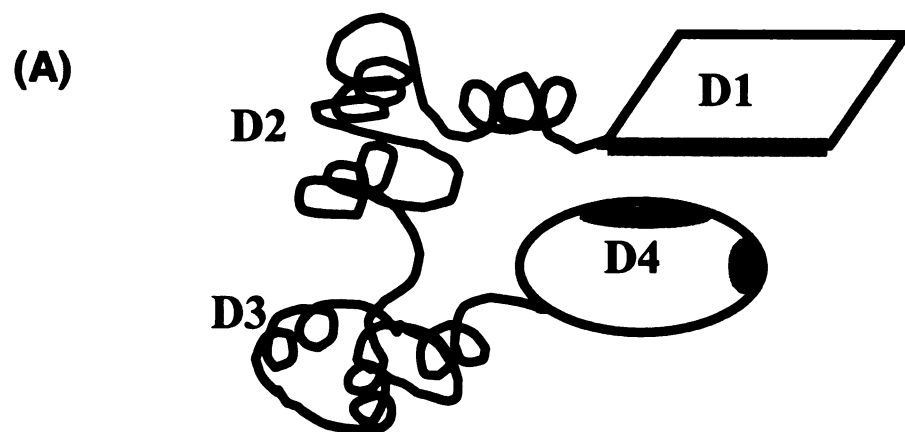
A sequential working model for annexin folding. For multidomain proteins, the formation of a native structure requires not only the correct folding of each domain but also the appropriate assembly of the domains via interdomain interactions. However,

little is known about the roles of interdomain interactions during the folding process. As discussed above, interdomain interactions may play a critical role in the folding of domain 2 of annexin I. It is interesting to note that among the four domains of annexin I, only domain 1 is folded and soluble when expressed in *Escherichia coli*. Domain 2 is soluble but largely unfolded. Expression of separated domain 3 and 4 in *Escherichia coli* results in inclusion bodies (data not shown). It has been reported that domain 3 is easily degraded but domain 4 forms inclusion bodies when expressed as fusion proteins of glutathione transferase (17). It appears that only domain 1 is an autonomous folding unit, although it is not known at present whether domains 3 and 4 can be solubilized and refolded to their native structures.

As described earlier, annexin I is composed of two modules. One module consists of domains 1 and 4, and the other domains 2 and 3. Each module has a hydrophobic interface between its constituents. The two modules are assembled with mostly hydrophilic interactions between domains 2 and 4. Several possible scenarios can be proposed for the folding process of this multi-domain protein such as a general model proposed by Fink (49), in which the D2-D3 module constitutes an autonomous folding unit that brings domain 1 and 4 together. Apparently, Our experimental data did not agree this model. We therefore propose another model in which the folding of annexin I follows a sequential process with domain 1 as an autonomous initial folding unit.

The sequence of the events in our proposed working model is depicted in Fig. 1.9. (1) First, domain 1 folds independently, domains 2 and 3 are maintained partly unfolded by local non-native interactions. As discussed above, the inter-domain interactions between domains 2 and 4 are critical for the complete folding of domain 2. We may

Fig. 1.9 A working model for the folding process of annexin I. In this model the protein folds sequentially by three principle steps. (A) Domain 1 folds first as an autonomous unit. Domains 2 and 3 are remained partly unfolded by local nonnative interactions to facilitate the docking of domain 4 to domain 1. (B) In a second step, Domain 4 is docked to domain 1 by the hydrophobic interactions (gray bars) between these two domains, which will also facilitate the complete folding of domain 4. (C) Finally, the hydrogen bonds and hydrophobic interactions (dash lines) between domains 4 and 2 help domain 2 to get rid of the nonnative cap and reach the native structure. Domain 2, in turn, assists the folding of domain 3 through many hydrophobic inter-domain interactions (gray bars).



reasonably assume that domain 4 must dock to domain 1 in order to establish the hydrophilic interface between domains 2 and 4. The flexibility of unfolded domain 2 and 3 allows domain 4 to search for domain 1. (2) In a second step, domains 1 and 4 are docked together and the folded structure of domain 1 facilitates the folding of domain 4 through the hydrophobic interface. (3) Finally, the hydrogen bonds and hydrophobic interactions between domains 4 and 2 help domain 2 to get rid of the nonnative cap and reach the native structure. The hydrophilic core is formed between domains 2 and 4. Domain 2, in turn, assists the folding of domain 3 through many hydrophobic inter-domain interactions. Our model emphasizes the inter-domain interactions in the folding of annexins. This proposal can be tested by systematic studies of the folding properties of the entire protein and separated domains of annexin I.

References

1. Srere, P. A. (1984) *Trends Biochem. Sci.* 9, 387-390
2. Creighton, T. E. (ed) (1992) *Protein Folding*, W. H. Freeman and Co., New York
3. Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol.* 49, 117-237
4. Jaenicke, R. (1991) *Biochemistry* 30, 3147-3161
5. Jaenicke, R. (1996) *Curr. Top. Cell. Reg.* 34, 209-314
6. Barton, G. J., Newman, R. H., Freemont, P. S., and Crumpton, M. J. (1991) *Eur. J. Biochem.* 198, 749-760
7. Morgan, R. R., and Fernández, M.-P. (1995) *Mol. Biol. Evol.* 12, 967-979
8. Raynal, P., and Pollard, H. B. (1994) *Biochem. Biophys. Acta* 1197, 63-93
9. Creutz, C. E. (1992) *Science* 258, 924-931
10. Donnelly, S. R., and Moss, S. E. (1997) *Cell. Mol. Life Sci.* 53, 533-538
11. Voges, D., Berendes, R., Demange, P., Benz, J., Gottig, P., Liemann, S., Huber, R., and Burger, A. (1995) *Adv. Enzymol. Relat. Areas Mol. Biol.* 71, 209-39
12. Liemann, S., and Huber, R. (1997) *Cell. Mol. Life Sci.* 53, 516-521
13. Macquaire, F., Baleux, F., Huynh Dinh, T., Rouge, D., Neumann, J. M., and Sanson, A. (1993) *Biochemistry* 32, 7244-54
14. Odaert, B., Baleux, F., Huynh-Dinh, T., Neumann, J. M., and Sanson, A. (1995) *Biochemistry* 34, 12820-9
15. Cordier-Ochsenbein, F., Guerois, R., Baleux, F., Huynh Dinh, T., Chaffotte, A., Neumann, J. M., and Sanson, A. (1996) *Biochemistry* 35, 10347-57

16. Cordier-Ochsenbein, F., Guerois, R., Baleux, F., Huynh-Dinh, T., Lirsac P.-N, Russo-Marie, F., J.-M., N., and Sanson, A. (1998) *J. Mol. Biol.* 279, 1163-1175
17. Cordier-Ochsenbein, F., Guerois, R., Russo-Marie, F., Neumann, J.-M., and Sanson, A. (1998) *J. Mol. Biol.* 279, 1177-1185
18. Weng, X., Luecke, H., Song, I. S., Kang, D. S., Kim, S. H., and Huber, R. (1993) *Protein Sci.* 2, 448-58
19. Piantini, U., Sørensen, O. W., and Ernst, R. R. (1982) *J. Am. Chem. Soc.* 104, 6800-6801
20. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 117, 479-85
21. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson.* 53, 521-528
22. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* 65, 355-360
23. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. P. (1988) *J. Am. Chem. Soc.* 110, 7870-7872
24. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys.* 71, 4546-4553
25. Macura, S., and Ernst, R. R. (1980) *Mol. Phys.* 41, 95-117
26. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett.* 69, 185-188
27. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc.* 114, 10663-10665
28. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry* 28, 6150-6
29. Fesik, S. W., and Zuiderweg, E. R. P. (1988) *J. Magn. Reson.* 78, 588-593
30. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson. Ser. B* 101, 201-205

31. Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson. Ser. B* 103, 203-216
32. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* 114, 6291-6293
33. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) *J. Magn. Reson.* 88, 425-431
34. Kay, L. E., Xu, G.-Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993) *J. Magn. Reson. Ser. B* 101, 333-337
35. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol NMR* 6, 277-93
36. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson.* 95, 214-220
37. Marion, D., Ikura, M., and Bax, A. (1989) *J. Magn. Reson.* 84, 425-430
38. Zhu, G., and Bax, A. (1992) *J. Magn. Reson.* 100, 202-207
39. Wüthrich, K., Billeter, M., and Braun, W. (1983) *J. Mol. Biol.* 169, 949-61
40. Nilges, M., Gronenborn, A. M., and Clore, G. M. (1988) *FEBS Lett.* 229, 317-324
41. Brünger, A. T. (1992) *X-PLOR Version 3.1: A System for Crystallography and NMR*, Yale Univeristy Press, New Haven, CT
42. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., and Karplus, M. (1983) *J. Comput. Chem.* 4, 187-217
43. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graph.* 14, 51-5, 29-32
44. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283-291
45. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1997) *J. Biomol. NMR* 8, 477-86
46. Basus, V. J. (1989) *Methods Enzymol.* 177, 132-149

47. Harper, E. T., and Rose, G. D. (1993) *Biochemistry* 32, 7605-9
48. Seale, J. W., Srinivasan, R., and Rose, G. D. (1994) *Protein Sci.* 3, 1741-5
49. Fink, A. L., (1995) *Annu. Rev. Biophys. Biomol. Struct.* 24, 495-522

CHAPTER 2

SEQUENTIAL BACKBONE RESONANCE ASSIGNMENTS OF YEAST GUANYLATE KINASE

2.1 Introduction

Guanylate kinase (GK) belongs to a family of nucleoside monophosphate (NMP) kinases, including adenylylate kinase (AK), uridylylate kinase (UK), cytidylylate kinase (CK), and thymidylylate kinase (TK). All NMP kinases catalyze the phosphoryl transfer from ATP to NMP to form nucleoside diphosphates, which are then activated by nucleoside diphosphate kinase to nucleoside triphosphates as building blocks for DNA or RNA synthesis. Guanylate kinase catalyzes the following reversible reaction: $\text{MgATP} + \text{GMP} \leftrightarrow \text{MgADP} + \text{GDP}$. It plays an essential role in the cGMP cycle and may be involved in guanine nucleotide-mediated signal transduction pathways by regulating the ratio of GTP to GDP (1,2). It is also required for the metabolic activation of the anti-herpes drugs acyclovir and gancyclovir and the anti-HIV agent carbovir (3-5). Thus it is of biomedical significance to study the catalytic mechanism and the nucleotide specificity of this enzyme. The functional significance of GK is also highlighted by the discovery of membrane-associated GK homologues (MAGUK), including the *Drosophila* discs-large tumor suppressor protein (*dlg-A*), the protein encoded by *C. elegans* vulvaless gene *lin-2*, the mammalian zonula occludens or tight junction proteins ZO-1 and ZO-2, the erythrocyte membrane protein p55 and several synapse-associated proteins (PSD-95/SAP90, SAP97

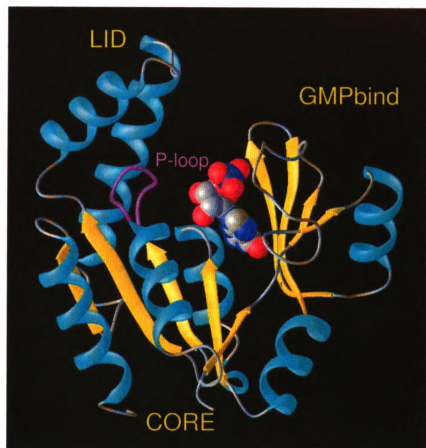
and SAP102). However, these GK homologues are unlikely to be enzymatically active. It has been suggested that the GK domains in these proteins may be involved in protein-protein interactions (6).

The GK activity was first reported by Klenow & Lichtler in 1957 (7). GK has been purified from several sources, but detailed characterization has been hampered by its low abundance. It was not until 1989 that yeast GK was purified to homogeneity and its amino acid sequence was determined. The yeast GK gene was cloned by Konradi in 1992 (9), followed by cloning of the *E. coli* GK gene and bovine GK gene (10). Yeast GK shares 45% identity with *E. coli* GK and 55% with bovine GK. The crystal structure of a yeast GK complex with GMP was first reported in 1990 and refined at 2Å resolution in 1992 (11). Human GK has recently been cloned and shares ~50% amino acid identity with yeast GK. The human enzyme is inactive when it is expressed in *E. coli* or produced by cell-free translation (12). We have been interested in the catalytic mechanism of yeast GK and have completed extensive kinetic and mutagenesis studies in our lab (13-15).

To date, about 26 crystal structures have been determined for NMP kinases (16). All structures are highly similar; containing three domains termed CORE, LID and NMPbind (Fig. 2.1). A typical five-stranded parallel β -sheet with helices on both sides constitutes the rigid CORE domain. The CORE domain contains a “glycine-rich loop” (P-loop) which forms a giant anion hole and binds ATP. The NMPbind domain forms the NMP binding site. The LID domain covering the phosphates at the active site carries many of the catalytically important residues.

Among these NMP kinases, AK has been extensively studied by X-ray, NMR and site-directed mutagenesis (17). By comparison of different forms of homologous AKs,

Fig. 2.1 The crystal structure of yeast guanylate kinase in complex with GMP
(11).



apo-form AK1, AMP complex with AK3 and AP₅A (P¹,P⁵-(5'-diadenosyl)-pentaphosphate) complex with AKe, the substrate-induced conformational changes have been established in a gradual manner (Fig. 2.2) (18,19): Binding of AMP induces movement of the AMP binding domain, while binding of ATP causes closure of the LID domain. Binding of the second substrate causes further closure of both domains. The formation of a “ternary” complex with AP₅A results in the closure of both LID and NMPbind domains. These domain movements have been summarized in a movie that represents an interpolation of different structures of NMP kinases (20). It has been suggested that these substrate-induced domain movements are important for preventing the enzyme from hydrolytic activity and stabilizing the transition state. However, no detailed descriptions have been possible because no NMP kinase structures have been determined in all forms.

It has been widely accepted that dynamics of enzymes play an important role in catalysis. However, a direct correlation has not been clearly demonstrated between dynamics and catalysis. From studies of several different enzymes, it has been found that the dynamic properties of binding sites are important for substrate binding. In the case of AK, it has been suggested that the flexibility of P-loop, which mainly binds the phosphate chain of ATP, is required for efficient substrate binding by allowing different isomers of ATP to convert to a productive isomer (21,22). Assuming that B-factors reflect the relative mobility, it has been found that the LID and NMPbind domains in AK are mobile in free form and the rest of the enzyme is relatively well fixed. Upon binding of AP₅A, these two domains become immobilized and the two loops between $\alpha 4$ - $\beta 3$ and $\alpha 5$ - $\beta 4$ in the CORE domain become mobilized. The mobility of these two loops is proposed to be

Fig. 2.2 Domain movements correlated with substrate binding to adenylate kinase (18,19). (a) Model of AK1 without bound substrates. (b) Model of AK3 with bound AMP. (c) Model of AKy mutant (D89V, R165I) with an ATP analogue (AMPPCF₂P). (d) Model of AKe with bound AP₅A. In all depicted models, the CORE, LID and NMPbind domains are shown in cyan, green and yellow, respectively. All the substrates (AMP, AMPPCF₂P and AP₅A) are shown in red.



an “energetic counterweight” that keeps the ternary complex from dropping into an energy well (23). It is an interesting idea and also an important model because in this model the dynamics is directly correlated to the catalytic mechanism. However, the observed B-factor distributions of other NMP kinases show significant differences (24,25).

The catalytic mechanism and the structural basis of nucleotide specificity of GK are still largely unknown. Work in our lab has shown that GK catalyzes the phosphoryl transfer via a sequential mechanism and the chemical step is the major rate-limiting step (13). GK has the highest specificity at the NMP binding site among the NMP kinases. Compared with AK, the CORE domain and the putative ATP binding domain of GK are similar to those of AK (Fig. 2.1). However, the GMP binding domain of GK and the AMP binding domain of AK are quite different. While the GMP binding domain consists of a mixed β -sheet and a short helix, the AMP binding domain is completely α -helical. GK has not been cocrystallized with ATP, because the ATP binding site is partly covered by the crystal contact. The ATP binding site was tentatively assigned on the basis of the structural homology to AK and GTP-binding proteins (EF-Tu and H-*ras*-p21). However, there are two problems with the proposed ATP binding model: (1) The distance (6Å) between the γ -phosphate of ATP and the nearest oxygen of GMP is too far for a nucleophilic attack. (2) The γ -phosphate of ATP is so much exposed to the solvent that hydrolysis cannot be avoided.

Because of its high specificity at both ATP and GMP binding sites and the favorable properties for NMR study (soluble, stable and ~ 20 kDa monomeric), yeast GK

is an excellent model enzyme for studying the substrate specificity and catalytic mechanism. Besides the development of multidimensional multinuclear NMR spectroscopy which has greatly facilitated the structural studies of proteins, NMR relaxation experiments have been applied successfully to elucidate the dynamic properties of proteins in solution (26). We aim to study the structures and dynamics of GK in solution by NMR spectroscopy and to evaluate how the conformational and dynamic changes are correlated to the catalytic mechanism. This thesis presents the multinuclear multidimensional NMR experiments and sequential backbone assignments of the GK complex with GMP. The results provide the basis for further NMR studies of the structure-function relationship of this enzyme.

2.2 Materials and Methods

Protein expression and purification. The GK gene from a yeast genomic library has been amplified by PCR and then cloned into the expression vector pET17b designated pET-YGK (13). The unlabeled protein samples were expressed in the BL21 (DE3) *E. coli* strain containing pET-YGK in a LB medium without IPTG induction. For samples labeled uniformly with ^{15}N or $^{15}\text{N}/^{13}\text{C}$, $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -glucose were substituted for their unlabeled counterparts in a variation of M9 minimal medium with IPTG induction. Proteins selectively labeled with specific ^{15}N amino acid were expressed from the *E. coli* strain DL49PS in a medium supplemented with appropriate unlabeled amino acids (27). ^{15}N -labeled amino acids were substituted for their unlabeled counterparts and the expression was induced with IPTG. Proteins were expressed and purified according to the protocols in the Appendices.

NMR sample preparation and experiments. Approximate 0.6 ml protein samples for NMR experiments were prepared in 20 mM predeuterated Tris-HCl buffer and 100 mM KCl in 90% H_2O / 10% D_2O solution at pH 7.5. The final protein concentration was ~2.0 mM for all samples, except for the selectively labeled samples, which were ~1.0 mM. The GMP complex samples contained 5-fold excess of GMP. GMP titration by 2D HSQC experiments demonstrated that 5-fold excess of GMP is sufficient to saturate the enzyme. NMR experiments were conducted at 22 °C on a Varian Inova 600 MHz spectrometer. All the NMR data were acquired in the States-TPPI mode.

Heteronuclear double and triple resonance spectra acquired for both free and GMP-bound forms of GK included 2D ^1H - ^{15}N HSQC (28,29), 3D ^1H - ^{15}N TOCSY-HSQC (30), 3D ^1H - ^{15}N NOESY-HSQC (30,31), HNCB (32), CBCA(CO)NH (33,34), and

HCCH-TOCSY (35,36). The 2D HSQC spectra for specific labeled proteins (including ^{15}N -Gly, ^{15}N -Leu, ^{15}N -Ile, ^{15}N -Lys, ^{15}N -Phe and ^{15}N -Val) have been collected for both free and GMP-bound forms of GK. 3D HNCA (37) spectrum was collected for free GK. 3D HNCACB (33,38) and (HB)CBCACO(CA)HA (39) spectra were collected for GK complex with GMP. The acquisition sweep widths and numbers of complex points for these experiments were as follow: 2D ^1H - ^{15}N HSQC, ^{15}N (F2) 3600 Hz, 256, ^1H (F2) 8000 Hz, 102; ^{15}N -edited NOESY-HSQC (150 ms mixing time) and ^{15}N -edited TOCSY-HSQC (48.6 ms mixing time), ^1H (F1) 7200Hz, 128, ^{15}N (F2) 2200Hz, 32, ^1H (F3) 8000 Hz, 1024; 3D HNCACB, ^{15}N (F1) 2200 Hz, 32, ^{13}C (F2) 11000 Hz, 96, ^1H (F3) 8000 Hz, 1024; 3D CBCA(CO)NH, ^{15}N (F1) 2200 Hz, 32, ^{13}C (F2) 11000 Hz, 64, ^1H (F3) 8000 Hz, 1024; 3D HCCH-TOCSY (23.4 ms mixing time), ^1H (F1) 7200 Hz, 128, ^{13}C (F2) 12070 Hz, 64, ^1H (F3) 8000 Hz, 1024; 3D HNCA, ^{15}N (F1) 2200 Hz, 32, ^{13}C (F2) 4900 Hz, 64, ^1H (F3) 8000 Hz, 1024; 3D HNCO, ^{15}N (F1) 2200 Hz, 32, ^{13}C (F2) 2200 Hz, 32, ^1H (F3) 8000 Hz, 1024; 3D (HB)CBCACO(CA)HA, ^{13}C (F1) 12001 Hz, 86, ^{13}CO (F2) 3000 Hz, 45, ^1H (F3) 8000 Hz, 1024.

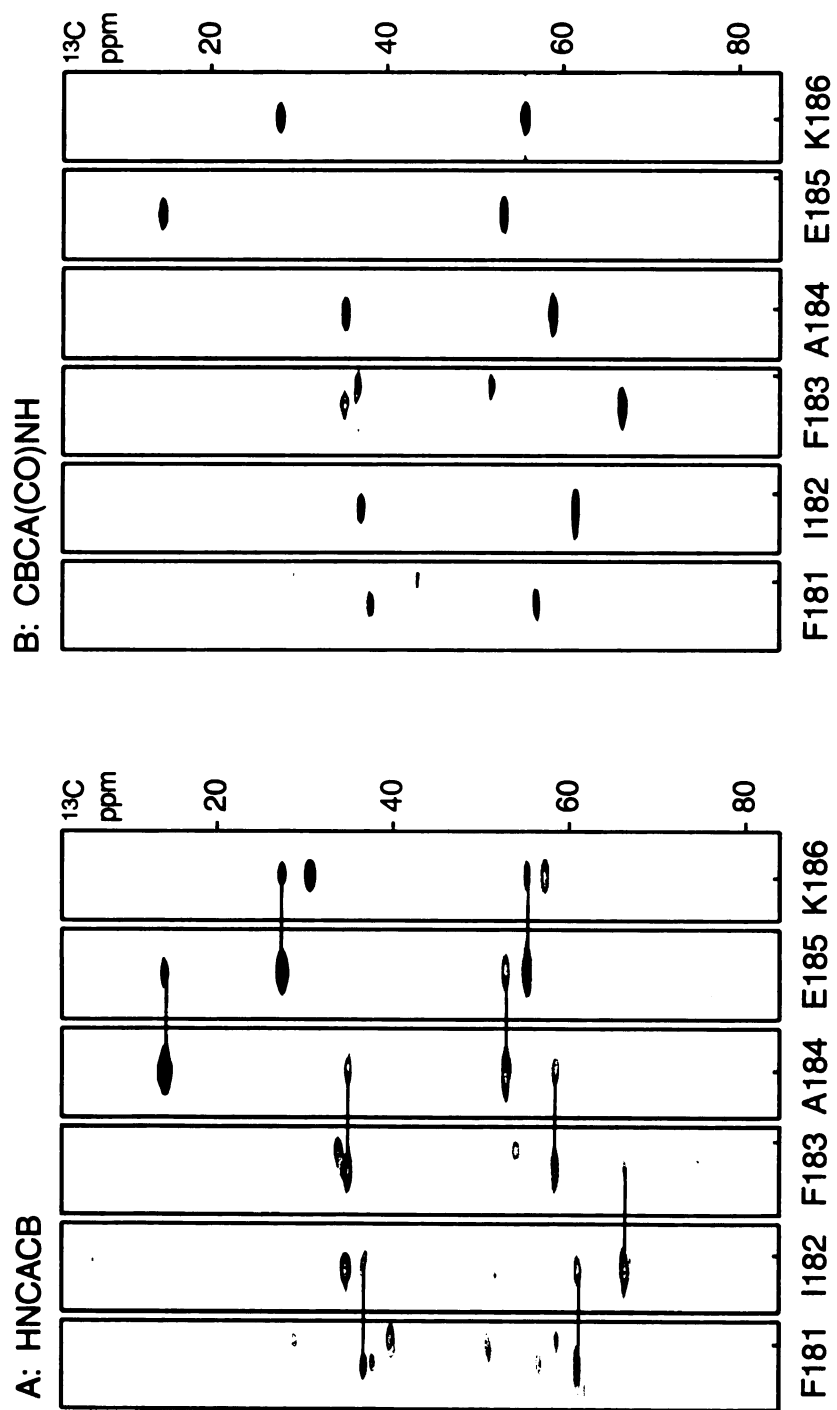
The spectra were processed with the program NMRPipe (40) and analyzed with the program NMRView (41). Briefly, solvent suppression was improved by convolution of time domain data (42). The data size in each indirectly detected dimension of the 3D data was extended by backward-forward linear prediction (43). A 45°-shifted sine bell and single zero-filling were generally applied before Fourier transformation in each dimension.

2.3 Results

The general strategy for the sequential assignments of the GK•GMP complex was to link all the spin systems sequentially by HNCACB/CBCA(CO)NH experiments via C^α and C^β chemical shifts, HNCO/(HB)CBCACO(CA)HA experiments via CO chemical shifts, and 3D ^{15}N -edited NOESY-HSQC via sequential NOEs. The assignments procedure was carried out using the program NMRView (41). According to the ^1H - ^{15}N cross peaks in the HSQC spectrum, strips along the carbon dimension were extracted from all ^{15}N -edited 3D NMR spectra. Spin systems for some residues with typical chemical shifts, such as Ala, Ser, Thr and Gly, could be identified. Six GK samples labeled with one type of ^{15}N -amino acid were used to aid spin system identifications (44), including Gly, Ile, Leu, Lys, Phe and Val. These spin systems were used as the starting points in the subsequent sequential assignments. The HNCACB and CBCA(CO)NH spectra were first analyzed to obtain sequential connectivities. Most of the assignments were made from these two spectra. Figure 2.3 shows the sequential connectivities from Phe181 to Lys186. The HNCO and (HB)CBCACO(CA)HA spectra provided additional independent links through C' resonances, which could be assigned directly from 3D HNCO spectrum. The relative higher sensitivity for these two experiments not only confirmed all linkages established from HNCACB/CBCA(CO)NH analysis, but also provided additional linkages.

When the triple-resonance spectra were incomplete because of a lack of C^α and C^β chemical shifts due to the low sensitivity of the HNCACB experiment, the sequential

Fig 2.3 Strip plot of 3D HNCACB (A) and 3D CACB(CO)NH (B) spectra of yeast GK in complex with GMP, showing the sequential J connectivities of ^{13}C nuclei for the residues F181 to K186.



connectivities could be made through sequential NOE analysis from the ^{15}N -edited NOESY-HSQC spectrum. Examples of $\text{H}^{\alpha}_i\text{-H}^{\text{N}}_{i+1}$ and $\text{H}^{\text{N}}_i\text{-H}^{\text{N}}_{i+1}$ connectivities from Leu21 to Tyr25 of domain 1 are shown in Figure 2.4.

Sequence-specific assignments of ^1H , ^{15}N and ^{13}C backbone resonances have been obtained for 70% of 179 non-proline residues in GMP-bound GK. A ^1H - ^{15}N HSQC spectrum a uniformly ^{15}N -labeled GK complex with GMP is shown in Fig. 2.5. The backbone ^1H , ^{15}N and ^{13}C chemical shifts assigned for the GK complex with GMP are listed in Table 2.1.

Fig. 2.4 Strip plot of the ^{15}N -edited NOESY-HSQC spectrum of yeast GK in complex with GMP, showing characteristic sequential d_{NN} and $d_{\alpha N}$ connectivities of the residues L21 to Y25.

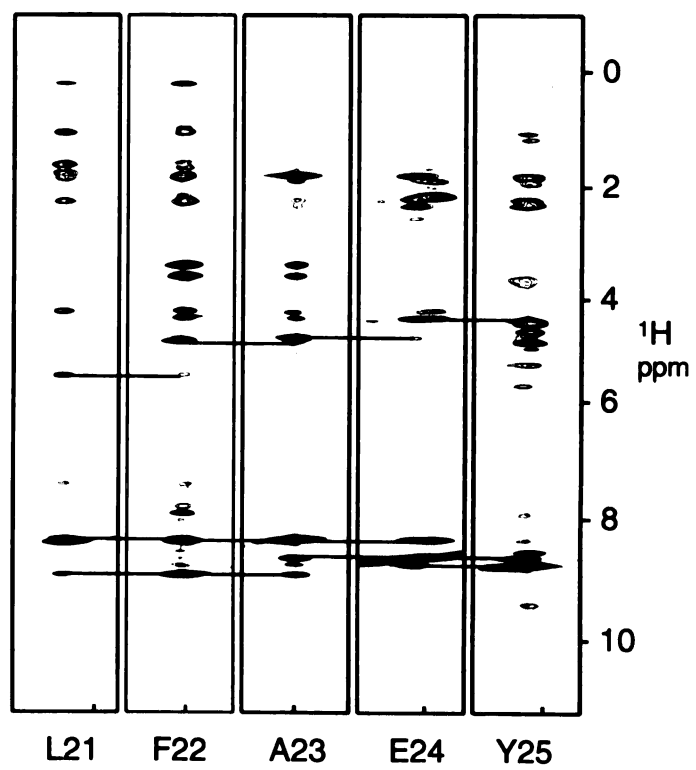


Fig. 2.5 ^{15}N - ^1H HSQC spectrum of yeast guanylate kinase in complex with GMP. Sequential assignments are indicated with one-letter amino acid codes and residue numbers.



Table 2.1 Sequential backbone ^1H , ^{15}N and ^{13}C resonance assignments for yeast guanylate kinase in complex with GMP

Residues	^{15}N	H^{N}	^{13}CO	$^{13}\text{C}^{\alpha}$	$^{13}\text{C}^{\beta}$	H^{α}
S1						
R2						
P3			176.6	61.7	34.9	
I4	129.9	9.09		60.8	42.0	4.31
V5						
I6						
S7						
G8						
P9						
S10						
G11						
T12						
G13						
K14						
S15						
T16			177.0	67.1	69.1	
L17	122.9	7.83	178.7	59.1	42.9	
L18	118.3	7.80	178.8	58.3	41.7	3.88
K19	120.5	8.09	181.0	60.8	33.1	4.02
K20	121.9	7.79	179.1	60.8	32.9	3.90
L21	123.4	7.86	179.8	58.8	43.4	5.01
F22	117.9	8.41	178.5	60.8	39.1	4.41
A23	121.5	7.81	180.3	54.8	19.3	3.90
E24	120.6	8.13	177.2	59.5	30.9	4.01
Y25	118.1	8.24		56.5	39.7	5.16
P26			178.9	65.9	32.9	
D27	117.7	8.63	177.6	55.4	42.6	4.92
S28	114.9	7.40	173.8	61.5	65.7	4.48
F29	120.4	7.99	174.8	57.5	44.1	5.51
G30	108.4	8.45	172.1	45.4		
F31	123.9	8.72	177.0	56.9	41.2	5.45
S32	120.3	8.57	172.6	58.8	64.4	4.18
V33			178.1	63.4	31.7	
S34	131.4	9.62	175.6	61.9	64.2	
S35	121.0	9.80	174.3	59.3	65.6	6.18
T36	117.7	8.75	172.6	59.1	70.5	5.58
T37	115.8	7.79	176.8	61.7	70.7	5.48
R38	125.9	8.02	175.4	57.2	30.4	4.34
T39	118.5	8.11		61.7	70.2	4.41
P40			177.8	64.0	32.7	
R41	125.9	9.54	177.8	55.3	33.2	4.53
A42	126.2	8.59	181.9	55.0	18.7	4.79
G43	112.9	8.80	175.8	46.0		
E44	121.0	7.94	178.1	57.5	33.1	4.63
V45	125.3	10.05	178.0	62.7	35.8	4.51
N46	131.9	9.53	177.2	55.8	39.7	4.80

Table 2.1 (Continued)

G47	118.9	9.41	173.4	45.9		
K48	121.7	8.44	175.6	57.9	35.2	4.63
D49	119.2	8.47	175.6	58.3	44.0	4.51
Y50	113.5	7.31		58.4	44.6	4.67
N51			174.7	54.0	40.0	
F52	127.4	8.68	177.1	60.0	38.9	4.94
V53	120.8	8.94	175.9	59.4	36.7	5.03
S54	118.0	8.83	176.7	57.7	66.6	4.83
V55	123.5	9.04	178.3	68.2	32.2	
D56	120.0	7.96	177.1	58.6	41.2	4.33
E57	117.9	8.22	179.4	60.0	30.6	4.41
F58	124.5	8.77	179.4	62.4	40.3	
K59	118.0	8.71	180.7	61.2	32.7	
S60	119.2	7.89	177.0	62.5	63.4	4.17/3.99
M61	124.6	8.05	179.5	60.2	35.4	3.80
I62	121.6	8.12	181.7	66.8	39.3	
K63	122.9	7.64	177.7	59.7	32.9	4.09
N64	116.9	7.82	174.9	53.3	39.7	4.86
N65	116.6	8.13	176.5	55.4	37.5	4.86
E66	114.7	7.92	178.9	58.9	31.4	4.05
F67	118.9	8.43	178.0	59.9	39.9	5.01
I68	124.8	9.51	176.1	64.8	40.5	3.97
E69	114.0	7.76	173.3	54.6	31.5	4.85
W70	119.2	8.06	175.1	57.1	32.9	5.54
A71	125.6	10.13	176.0	52.3	24.6	4.40
Q72	118.9	8.37	176.2	55.1	33.1	5.30
F73	127.8	9.05		58.5	42.0	4.74
S74			175.1	59.0	63.2	
G75	105.9	8.57	174.3	45.8		4.14/3.49
N76	119.2	7.79	172.9	52.7	42.9	4.20/4.84
Y77	121.1	8.05	175.9	58.6	37.5	4.82
Y78	122.0	9.02	176.3	57.8	43.2	5.54
G79	109.2	9.14	172.8	49.2		
S80	118.8	8.53		57.7	65.2	5.54
T81			175.4	61.0	71.0	
V82	123.9	8.04	179.3	67.3	33.0	4.24
A83	120.5	8.85	180.9	56.0	19.2	
S84	115.2	8.13	177.9	62.4	64.2	
V85	122.3	7.47	180.1	67.1	32.9	
K86	121.7	8.44	179.6	60.5	33.0	4.11
Q87	121.0	8.20	179.9	59.9	29.1	4.05
V88	120.6	7.93	180.3	67.1	32.9	3.76
S89	117.4	8.09	178.8	62.1	63.7	4.61
K90	124.3	8.46	178.4	59.3	32.8	4.23
S91	115.5	7.77	175.4	60.2	65.1	4.53
G92	111.2	7.99	174.8	46.2		4.34/3.82
K93	120.4	7.11	176.1	56.2	35.8	4.21

Table 2.1 (Continued)

T94	120.0	7.96	173.9	63.9	70.2	4.17
C95	131.8	9.19	173.4	59.7	29.2	4.87
I96				59.6	40.0	
L97	130.1	9.16	174.5	55.9	44.8	
D98	127.4	8.81	176.1	53.2	42.0	5.42
I99	115.8	7.38	180.7	59.4	43.4	4.13
D100						
M101						
Q102						
G103						
V104						
K105						
S106						
V107						
K108			175.58	59.8	33.7	
A109	120.0	7.20	177.0	52.6	19.5	4.53
I110	123.0	7.74		59.7	38.7	
P111			180.0	65.7	32.4	
E112	118.9	9.28	178.4	59.6	29.5	4.17
L113	120.9	7.65	177.9	57.1	43.3	4.24
N114	117.7	7.99	175.5	53.7	38.2	4.58
A115	120.3	7.85		58.8	20.6	4.35
R116						
F117						
L118						
F119						
I120						
A121						
P122						
P123						
S124						
V125						
E126						
D127						
L128			178.7	58.5	42.4	
K129	121.1	8.43	178.4	61.3	33.4	3.78
K130	119.2	7.79		59.9	33.1	4.84/4.20
R131						
L132						
E133						
G134						
R135			177.9	57.8	30.9	
G136	111.1	8.35	175.3	46.3		4.04
T137	114.9	8.14	175.8	62.5	70.6	4.42
E138	123.4	8.44	177.4	56.0	31.9	4.69
T139	114.6	8.61	175.7	61.2	72.3	4.57
E140	123.6	8.90		61.0	30.2	4.79
E141	119.6	8.51	179.4	60.0	30.4	

Table 2.1 (Continued)

S142	118.9	8.18		62.2	63.5	4.28
I143			178.3	66.5	38.3	
N144	120.8	8.36	179.4	57.1	38.7	4.53
K145	122.6	8.10		60.3	33.2	4.23
R146						
L147						
S148			178.3	62.5	63.7	
A149	126.5	8.02	180.6	55.7	19.2	4.43
A150	122.3	8.40	180.2	55.8	19.2	4.25
Q151	118.9	8.37	179.4			
A152	124.8	7.82		55.7	18.9	4.32
E153						
L154						
A155						
Y156						
A157			177.8	53.0	20.4	
E158	122.5	8.32	177.6	57.4	30.9	
T159	116.3	8.26	176.8	63.4	71.4	
G160	112.7	8.72		46.5		
A161						
H162						
D163						
K164						
V165			175.6	63.4	33.2	
I166	120.9	8.68	175.2	60.6	41.1	4.44
V167	130.4	9.14	177.1	62.2	32.6	
N168	127.5	8.70	174.7	52.5	37.6	4.95
D169	123.7	8.50	176.3	54.4	41.5	4.56
D170	122.3	7.72	176.4	55.0	43.7	4.65
L171	130.7	8.81		59.3	43.3	4.79
D172			180.2	58.5	41.4	
K173	125.3	8.22	179.6	60.2	33.2	4.11
A174	123.5	8.84	180.1	55.9	19.7	4.13
Y175	119.8	8.63		62.0	38.6	4.27
K176	121.1	7.67	179.7	60.9	32.8	
E177	119.7	8.12	180.2	60.6	30.8	4.09
L178	125.1	8.66	178.7	59.5	42.8	
K179	121.0	8.20	178.6	62.2	32.6	3.55
D180	120.2	8.40	179.5	57.9	40.9	4.39
F181	119.5	7.84	179.0	61.8	40.0	4.39
I182	122.6	8.73	178.7	66.5	38.3	
F183	116.9	8.16	176.2	59.4	38.4	4.67
A184	124.0	7.30	178.4	54.5	19.8	4.30
E185	118.2	7.57	175.5	56.6	31.8	4.37
K186	128.5	7.89		58.5	34.6	4.20

2.4 Discussion

Most of the backbone resonances of the GK•GMP complex have been assigned, mainly from the analysis of HNCACB and CBCA(CO)NH spectra. As shown in Fig. 2.5, almost all cross peaks except seven to eight backbone and some side chain amide peaks in the ^1H - ^{15}N HSQC spectrum have been assigned. Although the complete assignments have not been achieved and we have not exhausted all possibilities at this stage of the analysis, it has been found that some residues might show very weak or no cross peaks in the ^1H - ^{15}N HSQC spectrum. The residues that remained unassigned mainly locate at the ATP binding site including the LID domain and the P-loop. In general, the peak intensities of the HSQC spectra are affected by amide proton exchange and conformational exchange. Because the incorporation of “water flip-back” pulse were used in all NMR experiments to avoid attenuation of the amide resonances due to exchange with water, we can reasonably assume that the weak or missing HSQC peaks are mainly due to the variation in the line width of amide resonances. In the absence of any ligand, the ATP binding fragments may be involved in intermediate conformational exchange on the NMR time scale, resulting in broader lines and low-intensity peaks. NMR relaxation measurements will provide more direct information on the dynamic properties of these fragments in the enzyme.

2.5 Summary and Perspective

All the necessary NMR data for the assignments of free and GMP-bound forms of GK have been collected. About 70% of the ^1H , ^{15}N and ^{13}C backbone resonances of the GK complex with GMP have been assigned. Further analysis of 3D ^{15}N -edited NOESY-HSQC and 3D ^{15}N -edited TOCSY-HSQC will provide more linkages. The side chain assignments will be made by the analysis of 3D ^{15}N -edited TOCSY-HSQC and 3D HCCH-TOCSY experiments. The identification of complete side chain spin systems will provide additional information for the sequence-specific assignments.

The structure-function studies of GK will provide a deeper insight into the functional roles of conformational and dynamic changes in enzyme catalysis. This analysis is also of practical interest because it will provide another aspect for drug design. The results presented in this thesis provide the basis for the further structure and dynamic studies of GK by NMR spectroscopy:

Structure and dynamics of apo-GK. So far no crystal structure has been available for free GK, presumably because the crystallization has not yet been successful. It is noteworthy that there are two crystal structures reported for apo-AK from pig muscle. Which conformation relates to the structure of apo-AK in solution is not known (45). The determination of structure and dynamics of apo-GK in solution will lay a reliable basis for the further comparison studies.

Structures and dynamics of the binary and ternary complexes. The crystal structure of GK complexed with GMP has been available in good quality. This provides us an opportunity to compare the structural and dynamic properties of the enzyme in crystal and solution. The determination of solution structure of GK•ATP complex will

provide information on the ATP binding site. Like in AK and UK, the $\text{MgGP}_5\text{A} \text{ (P}^1\text{-(5'-adenosyl) P}^5\text{-(5'-guanosyl) pentaphosphate)}$ complex can be used to mimic the ternary complex of GK.

Structure and dynamics of the transition state. The transition state is an important complex for studying the mechanism of an enzyme. It is also a critical experiment to test our proposed model in this project. Aluminum fluoride has been successfully used to mimic the phosphate in a transition state related structure (25). As demonstrated in UK, AlF_3 complex would be expected to mimic the geometry of the transition state in GK.

Comparative studies of different forms of GK. The substrate-induced domain movements have been identified in AK. However, the structures for all free and complexed forms of GK determined in solution will make it possible to compare the gradual conformational change in much more detail. The comparison dynamic studies of different forms of GK will help to elucidate how dynamics correlate with catalysis. It has been suggested that the dynamics of the P-loop as well as the two “energetic counterweight” loops in NMP kinases are important for the efficiency of enzyme catalysis (23). One interesting example is from an AK isoenzyme in mitochondria, in which the post-translational modification ($\text{Asn3} \rightarrow \text{Asp}$) causes a twofold catalytic rate acceleration (23). This modification is close to the “energetic counterweight” and possibly affects the dynamic relocation during catalysis.

References

1. Hall, S. W., and Kuhn, H. (1996) *Eur. J. Biochem.* 161, 551-556.
2. Woods, D. F., and Bryant P. J. (1991) *Cell* 66, 451-464.
3. Miller, W. H. and Miller R. L. (1980) *J. Biol. Chem.* 255,7204-7207.
4. Boehme, R. E. (1984) *J. Bio. Chem.* 259, 12346-12349.
5. Miller, M. H., Daluge, S. M., Garvey, E. P., Hopkins, S., Reardon, J. E., Boyd, F. L., and Miller, R. L. (1992) *J. Biol. Chem.* 267, 21220-21224.
6. Kim, E., Naisbitt, S., Hsueh, Y. P., Rao, A., Rothschild, A., Craig, A. M., and Sheng, M. (1997) *J. Cell Biol.* 136, 669-678.
7. Klenow, H., and Lichtler, E. (1957) *Biochem. Biophys. Acta* 23, 6-12.
8. Berger, A., Schiltz, E., and Schulz, G. E. (1989) *Eur. J. Biochem.* 184, 433-443.
9. Konradi, M. (1992) *J. Biol. Chem.* 267, 25652-25655.
10. Geentry, D., Bengra, C., Ikehara, K., and Cashel, M. (1993) *J. Biol. Chem.* 268, 14316-14323.
11. (a) Stehle, T., and Schulz, G. E. (1990) *J. Mol. Biol.* 211, 249-254; (b) Stehle, T., and Schulz, G. E. (1992) *J. Mol. Biol.* 224, 1127-1141.
12. Brady, W. A., Kokoris, M. S., Fitzgibbon, J., and Black, M. E. (1996) *J. Biol. Chem.* 271, 16734-16740.
13. Li, Y., Zhang, Y. And Yan. H. (1996) *J. Biol. Chem.* 271, 28038-28044.
14. Zhang, Y., Li, Y., and Yan, H. (1997) *J. Biol. Chem.* 272, 19343-19350.
15. Zhang, Y., Li, Y., and Yan, H. (1996) in *Techniques in Protein Chemistry VIII* (Marshak, D. R. Ed.) pp 679-689, Academic Press, San Diego.

16. Yan, H., and Tsai, M.-D. (1997) *Adv. Enzymol. Relat. Areas Mol. Biol.* (invited review).
17. Tsai, M.-D., and Yan, H. (1991) *Biochemistry* 30, 6808-6818.
18. Schulz, G. E., Müller, C. W., and Diederichs, K. (1990) *J. Mol. Biol.* 213, 627-630.
19. Gerstein, M., Schulz, G. E., and Chothia, C. (1993) *J. Mol. Biol.* 229, 494-501.
20. Vonnrhein, C., Schlauderer, G. L., and Schulz, G. E. (1995) *Structure* 3, 483-490.
21. Snaders, C. R., Tian, G., and Tsai, M.-D. (1989) *Biochemistry* 28, 9028-9043.
22. Zhao, Z., Liu, X., Shi Z., Danley, L. Jinag, R., Huang, B., and Tsai, M. -D. (1996) *J. Am. Chem. Sci.* 118,3535-3536.
23. Müller, C., Schlauderer, G. L., Reinstein, J., and Schulz, G. E. (1996) *Structure* 4, 147-156.
24. Scheffzek, K., Kliche, W., Wiesmuller, L., and Reinstein, J. (1996) *Biochemistry* 35, 9716-9727.
25. Schlichting, I., and Reinstein, J. (1997) *Biochemistry* 36, 9290-9296.
26. Wagner, G. (1997) *Nature Struct. Biol. (NMR supplenment)* 4, 841-844.
27. Muchmore, D. C., McIntosh, L. P., Russell, C. B., and Dahlquist, F. W. (1989) *Methods Enzymol.* 177, 44-73.
28. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett.* 69, 185-188
29. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc.* 114, 10663-10665
30. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry* 28, 6150-6
31. Fesik, S. W., and Zuiderweg, E. R. P. (1988) *J. Magn. Reson.* 78, 588-593
32. Kay, L. E., and Bax, A. (1990) *J. Magn. Reson. Ser.* 86, 110-126

33. Muhandiram, D. R., and Kay, L. E. (1994) *J. Magn. Reson. Ser. B* 103, 203-216
34. Grzesiek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* 114, 6291-6293
35. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) *J. Magn. Reson.* 88, 425-431
36. Kay, L. E., Xu, G.-Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D.
(1993) *J. Magn. Reson. Ser. B* 101, 333-337
37. Ikura, M., Kay, L. E., and Bax, A. (1990) *Biochemistry* 29, 4659-4667.
38. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson. Ser. B* 101, 201-205.
39. Kay, L. E. (1993) *J. Am. Chem. Sci.* 115, 2055-2057.
40. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J Biomol. NMR* 6, 277-93
41. Johnson, B. A., and Blevins, R.A. (1994) *J Biomol. NMR* 4, 603-614.
42. Marion, D., Ikura, M., and Bax, A. (1989) *J. Magn. Reson.* 84, 425-430
43. Zhu, G., and Bax, A. (1992) *J. Magn. Reson.* 100, 202-207
44. Byeon, I. J., Yan, H., Edison, A. S., Mooberry, E. S., Abildgaard F., Markley, J. L.,
and Tsai, M.-D. (1993) *Biochemistry* 32, 12508-12521.
45. Sinev, M. A., Sineva, E. V., Ittah, V., and Hass, E. (1996) *Biochemistry* 35, 6425-
6437.

APPENDICES

APPENDIX A

Protocol for Expression and Purification of Non-labeled GK

1. Inoculate 1 loop of BL21(DE3) *E. coli* cells containing pET17b-YGK in 1 liter LB medium containing 100 µg/ml ampicillin. Incubate at 37 °C with shaking at 200 rpm until the OD₆₀₀ of the culture reaches ~2.0.
2. Harvest the cells by centrifugation at 7,000 rpm for 10 min. Resuspend the cell pellet in precooled buffer A (30 mM Tris-HCl, 1mM EDTA, pH7.5) and stir at 4 °C until it becomes a homogeneous suspension.
3. Sonicate the cell suspension for 3 min in a pulse mode at 4 °C. The resulting lysate was centrifuged for 20 min at 15,000 rpm. Resuspend the pellet in the same buffer and repeat the sonication. The supernatant was loaded onto an Affi-Gel Blue column equilibrated with buffer A.
4. The column was washed with buffer A until A₂₈₀ of the effluent was <0.05. It was eluted with 5 mM GMP in buffer A. GK fractions were identified by SDS-PAGE and concentrated to ~ 15 ml by an Amicon concentrator with a YM10 membrane.
5. The protein solution was then applied to a Sephadex G-75 column equilibrated with buffer A. The column was developed with the same buffer and the fractions were monitored by A₂₈₀ and SDS-PAGE. Pure GK fractions was concentrated, dialyzed in 300 mM KCl solution and lyophilized.

APPENDIX B

Protocol for Expression and Purification of Uniformly ^{15}N - or $^{15}\text{N}/^{13}\text{C}$ -labeled GK

1. Inoculate a single colony of BL21 strain containing pET17b-YGK in 10 ml LB medium containing 100 $\mu\text{g}/\text{ml}$ ampicillin overnight at 37 °C and 200 rpm. The culture was used to incubate 2 liter of ^{15}N - or $^{15}\text{N}/^{13}\text{C}$ -labeled minimal medium at the same condition. Induce protein expression using 0.4 mM IPTG when the OD_{600} of the culture reaches ~ 0.8 . The culture was continuously grown until the OD_{600} reaches 2.0. The cells were harvested and the remainder of the preparation is as described in Appendix A.
2. ^{15}N - or $^{15}\text{N}/^{13}\text{C}$ -labeled minimal medium (1 liter):

Dissolve Na_2HPO_4 (6 g), KH_2PO_4 (3 g), NaCl (0.5 g) and $^{15}\text{NH}_4\text{Cl}$ (1 g) in 1 liter water. Dissolve 5 g glucose (2 g ^{13}C -glucose for $^{15}\text{N}/^{13}\text{C}$ -labeled protein) in 20 ml water. The above solutions were autoclaved and mixed before use with 4 ml heavy metal stock solution, 0.1 ml 1% thiamin (B_1) and 2 ml of 50mg/ml ampicillin solution.
3. Heavy metal stock solution (500 ml):

250 mg $\text{MoNa}_2\text{O}_4\cdot\text{H}_2\text{O}$; 125 mg CoCl_2 ; 88 mg $\text{CuSO}_4\cdot 5\text{H}_2\text{O}$; 0.5 g $\text{MnSO}_4\cdot\text{H}_2\text{O}$; 4.38 g $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$; 0.63 g $\text{ZnSO}_4\cdot 7\text{H}_2\text{O}$; 0.63 g $\text{FeSO}_4\cdot 7\text{H}_2\text{O}$; 1.25 g $\text{CaCl}_2\cdot 2\text{H}_2\text{O}$; 0.5 g H_3BO_3 . Dissolve all the components together in 500 ml 1N HCl, stir the mixture at room temperature and filtrate it. Store at room temperature.

APPENDIX C

Protocol for expression and purification of specific ^{15}N -labeled GK.

1. Inoculate a single colony of DL49PS strain containing pET17b-YGK in 10 ml LB medium containing 100 $\mu\text{g/ml}$ ampicillin and 20 $\mu\text{g/ml}$ chloramphenicol overnight at 37 °C and 200 rpm. The culture was used to incubate 2 liter of Selectively ^{15}N -labeled medium at the same condition. Induce protein expression using 0.4 mM IPTG when the OD_{600} of the culture reaches ~ 0.8 . The culture was continuously grown until the OD_{600} reaches 2.0. The cells were harvested and the remainder of the preparation is as described in Appendix A.

2. Selectively ^{15}N -labeled medium:

Dissolve unlabeled L-amino acids (0.5 g alanine, 0.4 g arginine, 0.4 g aspartic acid, 0.05 g cystine, 0.4 g glutamine, 0.65 g glutamic acid, 0.55 g glycine, 0.1 g histidine, 0.23 g isoleucine, 0.23 g leucine, 0.42 g lysine, 0.25 g methionine, 0.13 g phenylalanine, 0.1 g proline, 2.1 g serine, 0.23 g threonine, 0.17 g tyrosine and 0.23 g valine), as well as 0.5 g adenine, 0.65 g guanosine, 0.2 g thiamin, 0.5 g uracil, 0.2 g cytosine, 1.5 g sodium acetate, 1.5 g succinic acid, 0.5 g NH_4Cl , 0.85 g NaOH and 10.5 g K_2HPO_4 in 950 ml water. After autoclaving, 50 ml of 40% glucose, 4 ml of heavy metal solution (same as in Appendix B), 10 ml of a filter-sterilized solution containing 50 mg L-tryptophan, 50 mg thiamin (B_1) and 50 mg niacin, 2 ml of 50 mg/ml ampicillin and 0.5 ml of 40 mg/ml chloramphenical are added. The ^{15}N -labeled amino acids are substituted for their unlabeled counterparts, except the ^{15}N -labeled lysine which is added just before the IPTG induction.

MICHIGAN STATE UNIV. LIBRARIES



31293017718366