



HOKKAIDO UNIVERSITY

Title	New techniques for protein NMR spectroscopy : development and application
Author(s)	Ogura, Kenji
Degree Grantor	北海道大学
Degree Name	博士(理学)
Dissertation Number	乙第5433号
Issue Date	1999-03-25
DOI	https://doi.org/10.11501/3151594
Doc URL	https://hdl.handle.net/2115/51608
Type	doctoral thesis
File Information	000000336784.pdf



NEW TECHNIQUES FOR PROTEIN
NMR SPECTROSCOPY:
DEVELOPMENT AND APPLICATION

KENJI OGURA

**New techniques for
protein NMR
spectroscopy:
development and
application**

Dissertation for the degree of
Doctor of Science

Kenji Ogura

Contents

Abbreviations	1
Introduction	3
Chapter 1 <i>Fully ^{13}C-refocused multidimensional ^{13}C-edited pulse schemes using broadband shaped inversion and refocusing pulses</i>	
Introduction	6
Theory	8
Results and discussion	9
Conclusion	11
References	12
Chapter 2 <i>An improved double-tuned and isotope-filtered pulse scheme based on a pulsed field gradient and a wide-band inversion shaped pulse</i>	
Abstract	22
Introduction	23
Theory, experiments and results	25
Conclusion	31
References	32
Chapter 3 <i>Solution structure of human acidic fibroblast growth factor and interaction with heparin-derived hexasaccharide</i>	
Abstract	49
Introduction	50

Materials and methods	52
Results and discussion	57
References	65

Chapter 4 *Solution structure of the SH2 domain of Grb2 complexed with the Shc derived phosphotyrosine-containing peptide*

Abstract	86
Introduction	87
Materials and methods	90
Results and discussion	98
References	109
Acknowledgment	131

Abbreviations

1D, 2D, 3D	one-dimensional, two-dimensional, three-dimensional
Abl, etc.	protein coded by <i>abl</i> gene, etc.
aFGF	acidic fibroblast growth factor
bFGF	basic fibroblast growth factor
COSY	correlation spectroscopy
CSI	chemical shift index
CT	constant time
DQF	double quantum filtered
DSS	2,2-dimethyl-2-silapentane-5-sulfonic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
FAK	focal adhesion kinase
FGF	fibroblast growth factor
FID	free induction decay
GARP	globally optimized alternating phase rectangular pulse
Grb2	growth factor receptor bound protein 2
Hep6	heparin-derived hexasaccharide
HEPES	<i>N</i> -2-hydroxyethylpiperazine- <i>N'</i> -2-ethanesulfonic acid
HMQC	heteronuclear multiple-quantum coherence spectroscopy
HNCO, etc.	triple resonance experiments involving ^{15}N , ^{13}C and ^1H
HPLC	high performance-liquid chromatography
HSQC	heteronuclear single-quantum coherence spectroscopy

IRS-1	insulin receptor substrate 1
MLEV	Malcom Levitt
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
PDGFR	platelet-derived growth factor receptor
PFG	pulsed field gradient
PLC- γ 1	phospholipase C- γ 1
PMSF	phenylmethanesulfonyl fluoride
pTyr	phosphotyrosine
RMSD	root mean square difference
SH2	src homology 2
SH3	src homology 3
Shc	Src homology and collagen
Sos	Son of sevenless
TOCSY	total correlation spectroscopy
TPPI	time propotional phase incrementation
VPP-peptide	Sos-derived peptide with a sequence of VPPPVPPRRR
WURST	wideband, uniform rate, and smooth truncation
YASAP	yet another simulated annealing protocol

Introduction

Recently, solution structure and dynamics of proteins have been widely studied using high resolution nuclear magnetic resonance (NMR) spectroscopy. Progress of the high resolution NMR methodology of proteins in solution has been supported by the following three techniques: (i) overexpression of protein samples and isotope-labeling based on the genetic engineering, (ii) development of the high performance NMR spectrometers, (iii) development of the new pulse schemes.

Many pulse schemes for investigation of proteins have been developed in the past decades. With the development of new pulse schemes, we can observe the interactions that have not been detected. Hence, we will be able to get new information about structural and dynamic properties of proteins.

The author developed two NMR pulse schemes to investigate the solution structures of proteins (Chapter 1 and 2). Furthermore, as the applications of the pulse schemes, the author determined the solution structures of two kinds of proteins (Chapter 3 and 4).

This dissertation consists of four chapters. In Chapter 1, the author developed the fully ^{13}C -refocused multidimensional ^{13}C -edited pulse schemes using broadband shaped inversion and refocusing pulses. The pulse schemes can avoid the sensitivity loss due to 180° pulse imperfection. Moreover, a large unwanted signal derived from H_2O solvent can be effectively eliminated by PFGs. The pulse schemes were demonstrated with a sample of uniformly ^{13}C -labeled protein.

In Chapter 2, the author developed the double-tuned and isotope-filtered pulse schemes based on PFGs and broadband shaped inversion pulses. Isotope-filter pulse schemes can separate the NMR signals derived from unlabeled molecules and the

signals derived from isotope-labeled (^{13}C and/or ^{15}N) molecules. Hence, we can selectively observe interaction between protein and small ligand molecule. The author demonstrated the pulse schemes with the isotope-labeled protein complexed with the small unlabeled peptide.

In Chapter 3, the author determined the solution structure of aFGF with a series of multidimensional NMR experiments and simulated annealing calculation. The solution structure and the binding sites of heparin-derived oligosaccharide of aFGF were very similar to those of bFGF. The evidence shows that aFGF and bFGF recognize heparin in a similar manner.

In Chapter 4, the author determined the solution structure of Grb2 SH2 domain complexed with Shc-derived phosphotyrosine (pTyr)-containing peptide. SH2 domains are protein modules found within a wide variety of cytoplasmic signaling molecules. The pTyr binding site of Grb2 SH2 domain was similar to those of other SH3 domain. In contrast to other SH2 - peptide complexes, the residues C-terminal to pTyr did not form an extended structure but took a turn conformation because of steric hindrance by a bulky side-chain of Trp121 of Grb2 SH2 domain.

Chapter 1

Fully ^{13}C -refocused multidimensional ^{13}C -edited pulse schemes
using broadband shaped inversion and refocusing pulses

Introduction

^{13}C -edited multidimensional NOESY pulse schemes are essentially important for spectral assignment and structural determination of ^{13}C -enriched proteins. Two types of pulse schemes have been proposed, either based on the HMQC (Müller, 1979) or HSQC (Bodenhausen and Ruben, 1980) type magnetization transfer. Pulsed field gradients (PFGs) have been also incorporated into these pulse schemes to reduce experimental phase cycling, to eliminate undesirable artifacts, and to suppress the H_2O resonances (Bax and Pochapsky, 1992). The HSQC type is more preferable to accommodate PFGs into the pulse schemes than the HMQC type, so that the HSQC-type experiments have been widely used for protein samples in H_2O solution (Majumdar and Zuiderweg, 1993; Muhandiram et al., 1993a; Pascal et al., 1994). However, mainly due to additional 180° pulses in the HSQC-type experiments, which are not ideal near the edge of the ^{13}C spectrum, the sensitivity of the HSQC-type experiments is much reduced, especially for aromatic and methyl signals (Vuister et al., 1993). Although the HMQC-type scheme covers a wider spectral range, it has been used for the ^{13}C -edited NOESY experiments in $^2\text{H}_2\text{O}$ solution, owing to less efficient suppression of water resonances (Ikura et al., 1990; Clore et al., 1991; Vuister et al., 1993).

In both ^{13}C HMQC- and HSQC-type experiments, the spectra in the ^{13}C dimension are usually folded to obtain higher spectral resolution (Bax et al., 1991). In the actual experiments, for minimizing the baseline distortion, delayed acquisition is made in the ^{13}C dimension so that the resulting ^{13}C -edited NOESY spectrum shows positive and negative peaks according to the order of the folding of the ^{13}C signals

(Kay et al., 1990a; Muhandiram et al., 1993b). This complicates the spectral view and moreover, due to the presence of the side lobe signals, it is difficult to pick the negative and positive peaks in the spectra. When ^{13}C magnetization can be refocused, both the folded and the unfolded peaks appear in the same phase, as was shown in the case of HCCH-COSY (Kay et al., 1990b; Ikura et al., 1991; Bax et al., 1992) and HCCH-TOCSY (Bax et al., 1990; Kay et al., 1993) experiments. Because of the lack of a refocusing pulse which can cover the whole the chemical shift range of ^{13}C signals, this scheme has not been applied.

The author propose here ^{13}C -edited two-, three-, and four-dimensional pulse schemes which make full inversion and full refocusing of all the ^{13}C signals possible.

Theory

Figure 1 shows the fully ^{13}C -refocused two-dimensional ^1H - ^{13}C HSQC pulse scheme. In this pulse scheme, two kind of shaped pulses are used; one is a hyperbolic-secant inversion pulse (Silver et al., 1984) and another is a broadband refocusing pulse described in Table 2.C by Abramovich and Vega (1993), denoted by "V1" in this chapter. The hyperbolic-secant pulse inverts the longitudinal ^{13}C magnetization and V1 pulse refocuses the ^{13}C transverse magnetization for the wide chemical shift range. Hence, these pulse schemes are effective for the measurement of the ^{13}C HSQC-type experiments without sensitivity loss. The excitation profiles of the hyperbolic-secant and the V1 pulses are shown in Figure 2. The ^{13}C longitudinal magnetization is efficiently inverted by the hyperbolic-secant pulse, and the ^{13}C transverse magnetization is efficiently refocused by the V1 pulse. Figure 3 shows the fully ^{13}C -refocused, ^{13}C -edited three-dimensional NOESY and fully ^{13}C -refocused $^{13}\text{C}/^{13}\text{C}$ -edited four-dimensional NOESY pulse schemes. These schemes are similar to the recently developed schemes (Muhandiram et al., 1993a; Vuister et al., 1993; Pascal et al., 1994) except for incorporating the fully ^{13}C -refocused ^1H - ^{13}C HSQC pulse scheme (Figure 1) instead of the conventional HMQC or HSQC scheme. It should be noted here the hyperbolic-secant inversion pulse was applied to the ^1H - ^{13}C CT-HSQC (Hallenga and Lippens, 1995) and ^{13}C half-filtered NOESY experiment (Folmer and Hallenga, 1995). However, it required higher-order phase correction in the ^{13}C dimension.

Results and discussion

These pulse schemes were applied to the N-terminal SH3 domain of Grb2 complexed with a Sos-derived peptide (Terasawa et al., 1994). The $^{13}\text{C}/^{15}\text{N}$ uniformly labeled N-terminal SH3 domain (NSH3) was overexpressed by *E.coli*. The unlabeled peptide with a sequence of VPPPVPPIRRR (VPP peptide) was synthesized by a solid-phase method. NSH3 complexed with the VPP peptide were dissolved at 2 mM concentration in 90 % H_2O /10 % $^2\text{H}_2\text{O}$ or $^2\text{H}_2\text{O}$ in 20 mM potassium phosphate buffer containing 150 mM NaCl, 10 mM DTT-*d*₁₀, and 8 mM NaN_3 (pH 6.7, direct meter reading). All NMR experiments were carried out on a Varian UNITY500 or a Varian UNITY*plus*600 spectrometer (^1H resonance frequencies were 500 and 600 MHz, respectively) equipped with a 5 mm $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple-resonance probe with a shielded z-axis gradient coil. The WURST (Kupce and Freeman, 1995a, b) decoupling sequence was used instead of GARP-1 (Shaka et al., 1985) to decouple all the ^{13}C resonances efficiently without sample heating. PFGs were used for the purge of undesirable transverse magnetization and the cancellation of the 180° pulse imperfection.

In Figure 4, the spectra of conventional ^1H - ^{13}C HMQC (Müller, 1979), fully ^{13}C -refocused ^1H - ^{13}C HSQC (the pulse scheme in Figure 1), and ^{13}C -refocused ^1H - ^{13}C HSQC-type experiments where all the pulses are replaced by rectangular shapes are compared. Use of the hyperbolic-secant and V1 pulses significantly improves the intensities of the aromatic and the methyl signals in the HSQC-type experiment, and the sensitivity of the fully ^{13}C -refocused HSQC experiment is now

similar to that of the HMQC experiment. Figure 5 shows the first transient of the two-dimensional spectrum obtained with the pulse scheme in Figure 1. In this experiment, a presaturation pulse was not applied since the HSQC-type experiment has an advantage in solvent suppression over the HMQC-type experiment. Due to application of the V1 pulse, all the signals are refocused and have the same phases. The two skyline projected spectra of a fully ^{13}C -refocused three-dimensional ^{13}C -edited NOESY experiment are shown in Figure 6, obtained with the pulse scheme of Figure 3a. The spectra are similar to the normal two-dimensional NOESY and HSQC experiments because all the signal of three-dimensional spectra have the same phase irrespective of the folded and unfolded peaks. Finally, the author shows two sliced subspectra from a fully ^{13}C -refocused four-dimensional $^{13}\text{C}/^{13}\text{C}$ -edited NOESY spectrum in Figure 7 recorded with the pulse scheme of Figure 3b. Though the ^{13}C signals are split due to ^{13}C - ^{13}C homonuclear couplings, the spectra obtained by the present three- and four-dimensional pulse schemes offer no obstacles for the analysis, considering the digital resolution of the spectra in the ^{13}C dimension.

Conclusion

In the ^{13}C -edited three- and four-dimensional NOESY experiments, used of the hyperbolic-secant inversion pulse and the V1 refocusing pulse in the HSQC-type experiments has the following advantages. One is uniform excitation over the wide ^{13}C chemical shift range, which gives all NOE correlations with a single experiment. Another advantage is that the folded and the unfolded signals appear in the same phase which makes both the data processing and the automatic peak-picking simple and convenient. Furthermore, such spectral features also make it possible to apply the maximum entropy method in the data processing (Laue et al., 1986).

References

- Abramovich,D., and Vega,S. (1993) *J. Magn. Reson. A* 105, 30-48.
- Bax,A., Clore,M., and Gronenborn,A.M. (1990) *J. Magn. Reson.* 88, 425-431.
- Bax,A., Ikura,M., Kay,L.E., and Zhu,G. (1991) *J. Magn. Reson.* 91, 174-178.
- Bax,A., Clore,G.M., Driscoll,P.C., Gronenborn,A.M., Ikura,M., and Kay,L.E. (1992) *J. Magn. Reson.* 87, 620-627.
- Bax,A. and Pochapsky,S.S. (1992) *J. Magn. Reson.* 99, 638-643.
- Bodenhausen,G. and Ruben,D.J. (1980) *Chem. Phys. Lett.* 69, 185-189.
- Clore,G.M., Kay,L.E., Bax,A., and Gronenborn,A.M. (1991) *Biochemistry* 30, 12-18.
- Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J., and Bax,A. (1995) *J. Biomol. NMR* 6, 277-293.
- Folmer,R.H.A., and hallenga,K. (1995) *J. Biomol. NMR* 5, 427-432.
- Hallenga,K., and Lippens,G.M. (1995) *J. Biomol. NMR* 5, 59-66.
- Ikura,M., Kay,L.E., and Bax,A. (1990) *J. Magn. Reson.* 86, 204-209.
- Ikura,M., Kay,L.E., and Bax,A., (1991) *J. Biomol. NMR* 1, 299-304.
- Kay,L.E., Clore,G.M., Bax,A., and Gronenborn,A.M. (1990a) *Science* 249, 411-414.
- Kay,L.E., Ikura,M., and Bax,A. (1990b) *J. Am. Chem. Soc.* 112, 888-889.
- Kay,L.E., Xu,G.Y., Singer,A.U., Muhandiram,D.R., and Forman-Kay,J.D. (1993) *J. Magn. Reson. B* 101, 333-337.
- Kupce,E., and Freeman,R. (1995a) *J. Magn. Reson. A* 115, 273-276.
- Kupce,E., and Freeman,R. (1995b) *J. Magn. Reson. A* 117, 246-256.
- Laue,E.D., Mayger,M.R., Skilling,J., and Staunton,J. (1986) *J. Magn. Reson.* 68, 14-29.
- Majumdar,A and Zuiderweg,E.R.P. (1993) *J. Magn. Reson. B* 102, 242-244.
- Marion,D., Ikura,M., Tschudin,R., and Bax,A. (1989) *J. Magn. Reson.* 85, 393-399.

- Muhandiram,D.R., Farrow,N., Xu,G.Y., Smallcombe,S.H., and Kay,L.E. (1993a)
J. Magn. Reson. B 102, 317-321
- Muhandiram,D.R., Xu,G.Y., and Kay,L.E. (1993b) *J. Biomol. NMR* 3, 463-470.
- Müller,L. (1979) *J.Am.Chem.Soc.* 101, 4481-4484.
- Pascal,S.M., Muhandiram,D.R., Yamazaki,T., Forman-Kay,J.D., and Kay,L.E.
(1994) *J. Magn. Reson. B* 103, 197-201
- Shaka,A.J., Barker,P.B., and Freeman,R. (1985) *J. Magn. Reson.* 64, 547-552.
- Silver,M.S., Joseph,R.I., and Hoult,D.I. (1984) *J. Magn. Reson.* 59, 347-353.
- Terasawa,H., Kohda,D., Hatanaka,H., Tsuchiya,S., Ogura,K., Nagata,K., Ishii,S.,
Mandiyan,V., Ullrich,A., Schlessinger,J., and Inagaki,F. (1994)
Nature Struct. Biol. 1 891-897.
- Vuister,G.W., Clore,G.M., Gronenborn,A.M., Powers,R., Garrett,D.S., Tschudin,R.,
and Bax,A. (1993) *J. Magn. Reson. B* 101, 210-213.

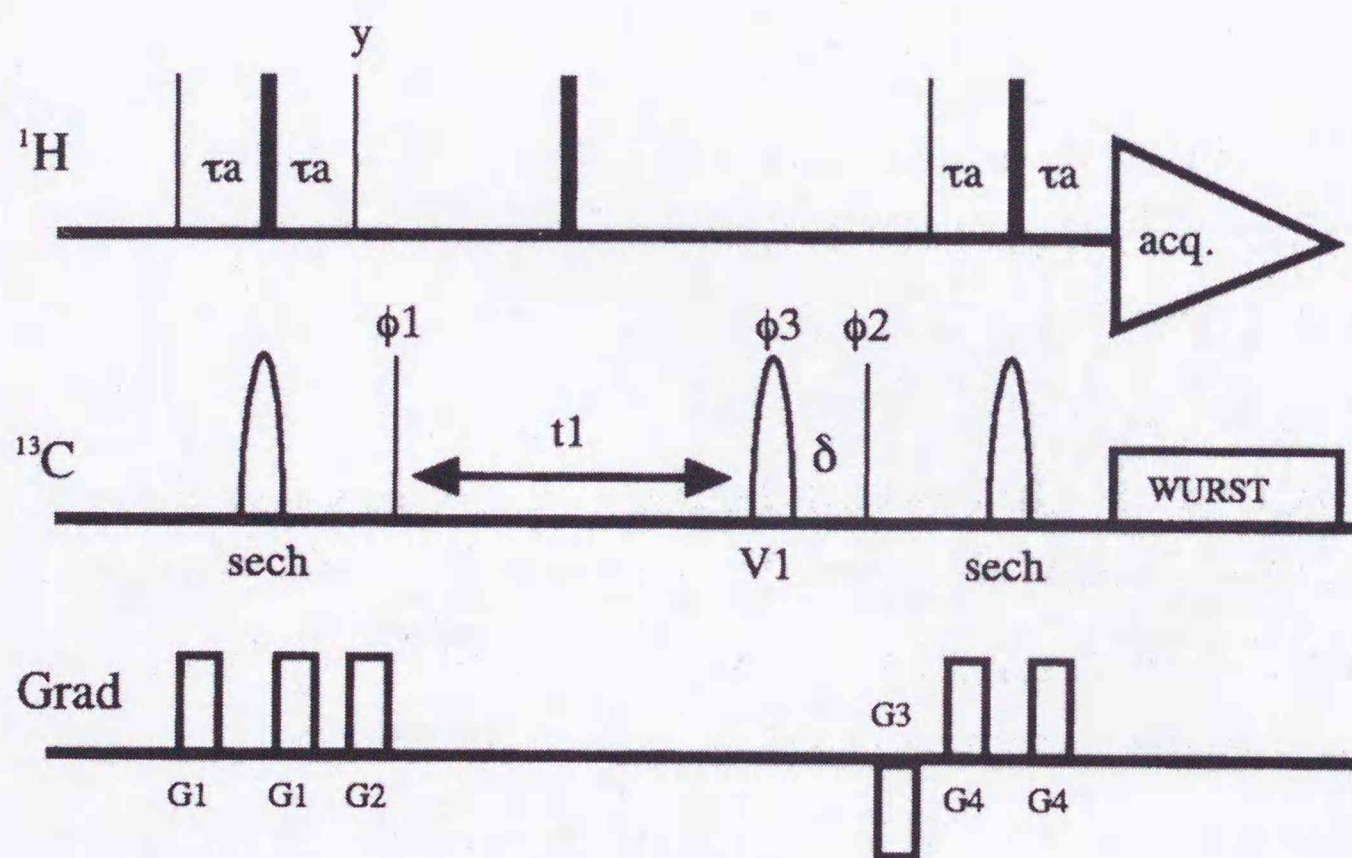


Figure 1. Pulse scheme of the fully ^{13}C -refocused 2D ^1H - ^{13}C HSQC experiment. Wide lines and narrow lines are rectangular 90° pulses and 180° pulses, respectively. The shaped pulses labeled by sech and V1 denote the hyperbolic-secant inversion pulse and the V1 refocusing pulse, respectively. All pulses are applied with 20 kHz RF strength. τ_a is set to 1.6 ms. δ is a refocusing delay, set to be equal to $\tau_{\text{H}180^\circ}$. Phase cyclings are $\phi_1 = x, -x$, $\phi_2 = 2(x), 2(-x)$, $\phi_3 = 4(x), 4(-x)$, receiver = $x, 2(-x), x$. Unlabeled pulses are applied from the x axis. Hypercomplex data acquisition is carried out in the States-TPPI manner (Marion et al., 1989). ^{13}C broadband decoupling at the detection period is achieved with a 3 kHz WURST sequence (Kupce and Freeman, 1995a, b). All PFGs have rectangular shapes. PFG delays and strength are $G1 = (0.5 \text{ ms}, 8 \text{ G/cm})$, $G2 = (1.0 \text{ ms}, 25 \text{ G/cm})$, $G3 = (1.0 \text{ ms}, -30 \text{ G/cm})$, $G4 = (0.5 \text{ ms}, 10 \text{ G/cm})$.

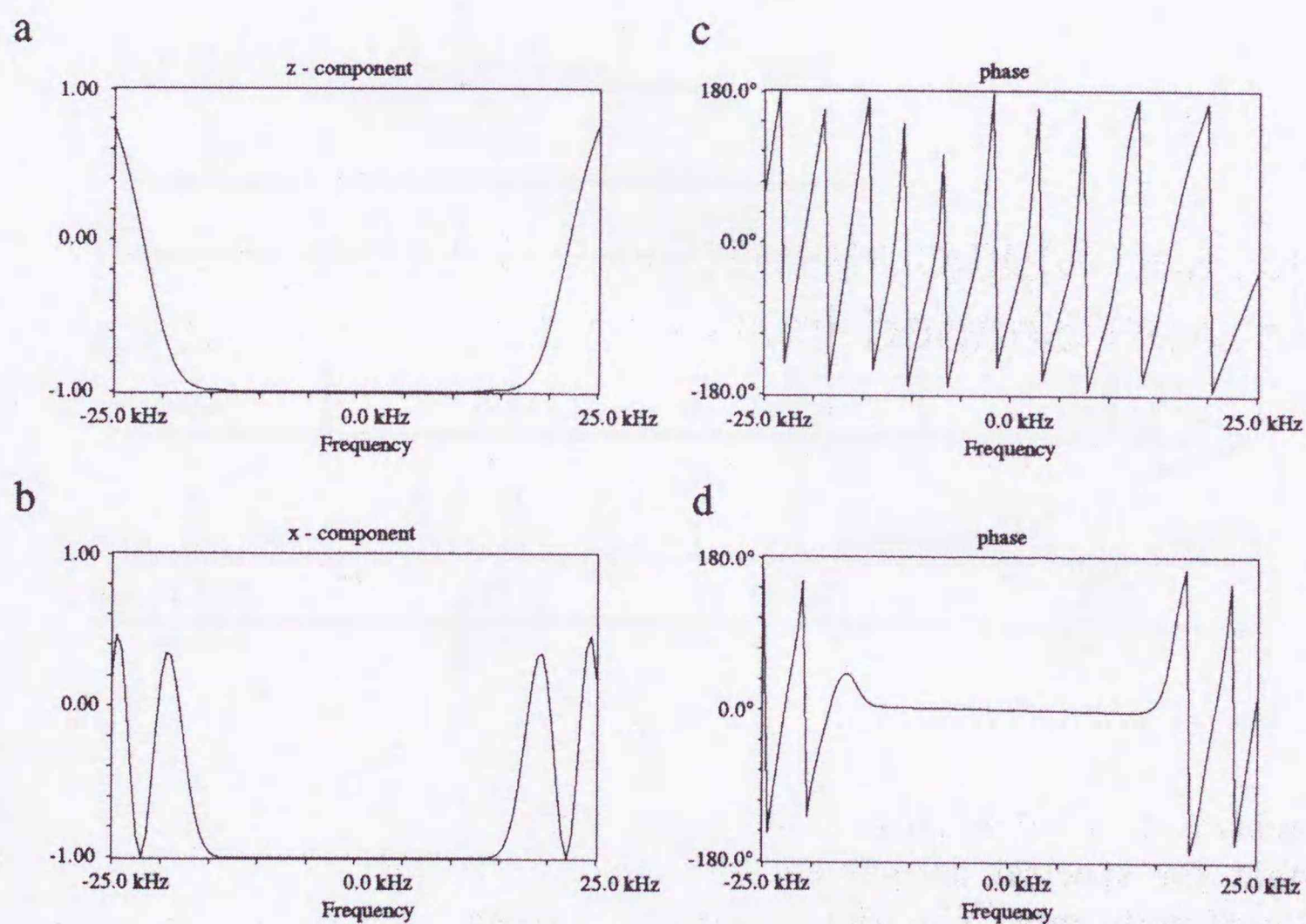


Figure 2. Simulated excitation profiles of the hyperbolic-secant pulse (max. $B_1 = 20$ kHz, pulse width = $360 \mu\text{s}$) (a, c) and the V1 refocusing pulse (max. $B_1 = 20$ kHz, pulse width = $313 \mu\text{s}$) (b, d) applied from the y axis. The initial magnetization directions are +z and +x for the hyperbolic-secant pulse and the V1 refocusing pulse, respectively. (a) and (d) show the intensities of the magnetization. (c) and (d) show the phase of the magnetization. The excitation profiles were calculated with the Pulsetool program provided by the VNMR software package from Varian Instruments (USA).

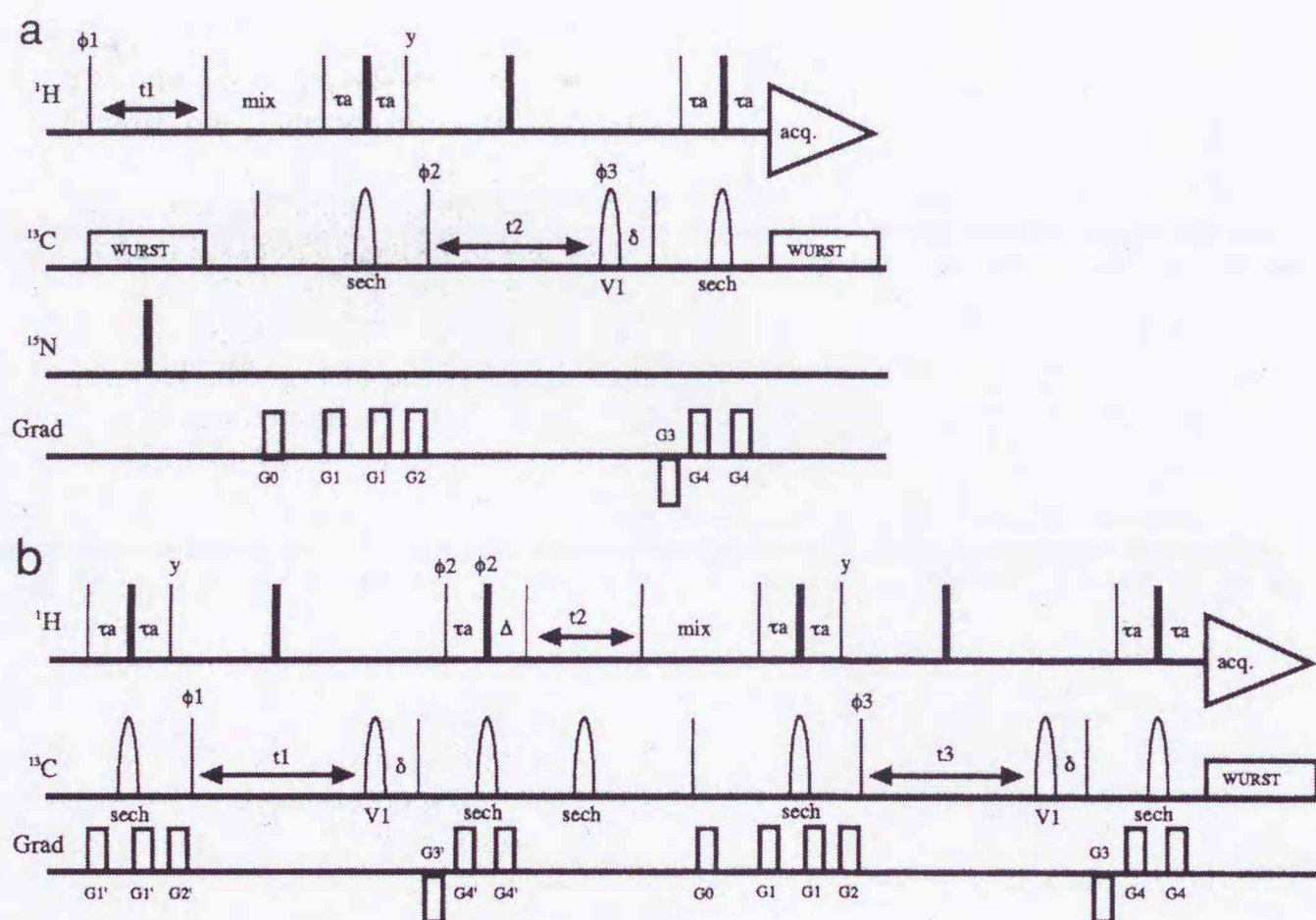


Figure 3. Pulse schemes of the fully ^{13}C -refocused 3D ^{13}C -edited NOESY (a) and 4D $^{13}\text{C}/^{13}\text{C}$ -edited NOESY (b) experiments. Phase cyclings are $\phi_1 = x, -x$, $\phi_2 = 2(x), 2(-x), \phi_3 = 4(x), 4(-x)$, receiver = $x, 2(-x), x$ for pulse sequence (a), and $\phi_1 = x, -x$, $\phi_2 = 2(y), 2(-y), \phi_3 = x$, receiver = $x, 2(-x), x$ for pulse sequence (b). Δ is set to $\tau_a - \tau_s$ (τ_s is the width of thesech pulse) for refocusing ^1H magnetization in the t_2 period. PFG delays and strength are $G_1' = (0.5 \text{ ms}, 8 \text{ G/cm})$, $G_2' = (1.0 \text{ ms}, -20 \text{ G/cm})$, $G_3' = (1.0 \text{ ms}, 15 \text{ G/cm})$, $G_4' = (0.5 \text{ ms}, 10 \text{ G/cm})$, and $G_0 = (2.0 \text{ ms}, 10 \text{ G/cm})$. Other PFGs and experimental parameters are set to be same as defined in the legend to Figure 1.

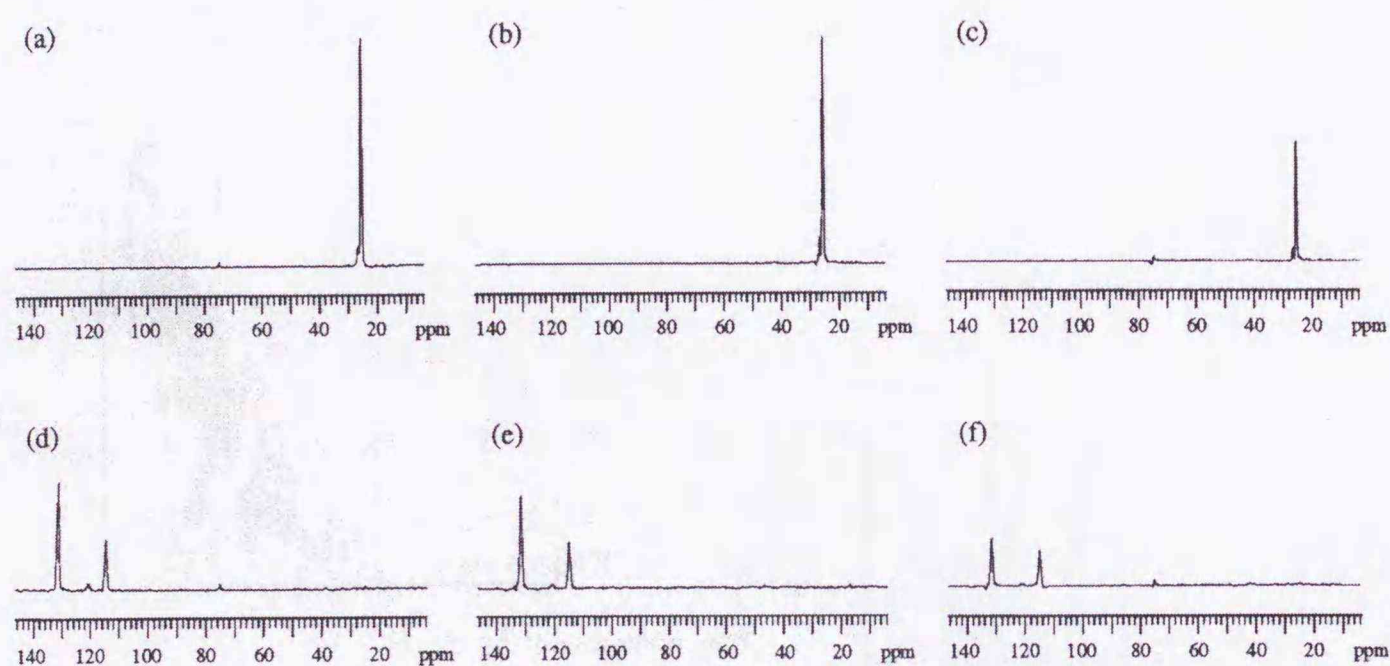


Figure 4. A comparison of the sliced subspectra on the F_1 axis from the three kinds of 2D ^1H - ^{13}C correlation experiments at 500 MHz. All spectra were measured for the sample of NSH3 complexed with VPP peptide in 90 % H_2O solution. In all the experiments, 2D spectra were taken with 256 increments on the F_1 axis and eight scans for each FID. The ^{13}C carrier frequency was set at 75.2 ppm. The HMQC spectra (a, d) were obtained with the pulse scheme described by Müller (1979). The fully ^{13}C -refocused HSQC spectra (b, e) were obtained with the pulse scheme described in the legend to Figure 1. The ^{13}C -refocused HSQC spectra (c, f) were obtained with the pulse scheme in which all 180° pulses were replaced by rectangular shaped pulse. The upper and lower columns show the methyl and the aromatic carbon signals, respectively. In the HMQC experiment, the initial delay for the t_1 evolution period was adjusted to be zero-order phase = 90° and first-order phase = -180° for F_1 axis. In the ^{13}C -refocused HSQC experiments, the phase correction along the F_1 axis was 0° for both zero- and first-order.

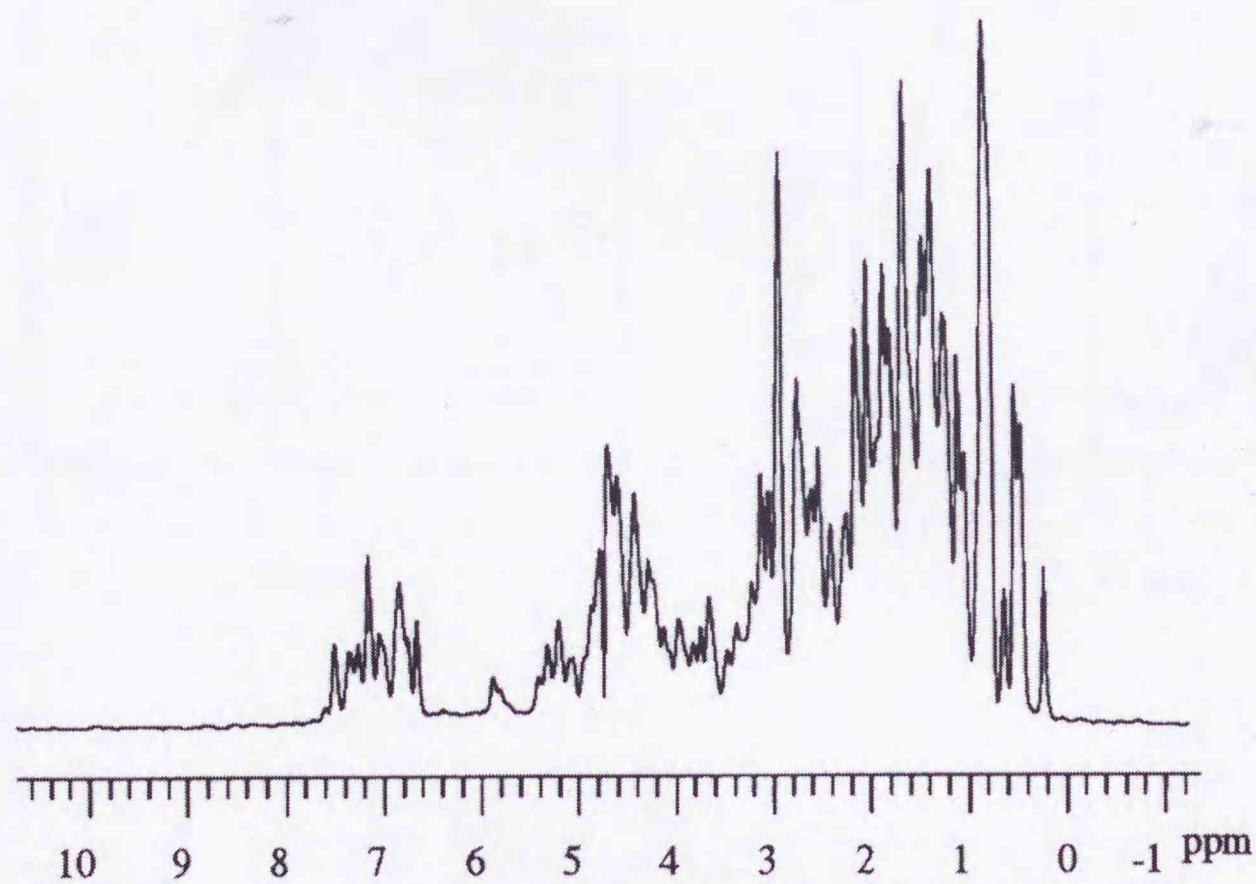


Figure 5. First transient of the 500 MHz fully ^{13}C -refocused HSQC spectrum obtained by the pulse scheme of Figure 1 for the sample of NSH3 complexed with the VPP peptide in 90 % H_2O solution. The number of scans was eight, and ^{13}C was decoupled with the WURST sequence during the detection period. Presaturation pulse and post-acquisition data manipulation were not applied because the H_2O signal was efficiently eliminated by PFG.

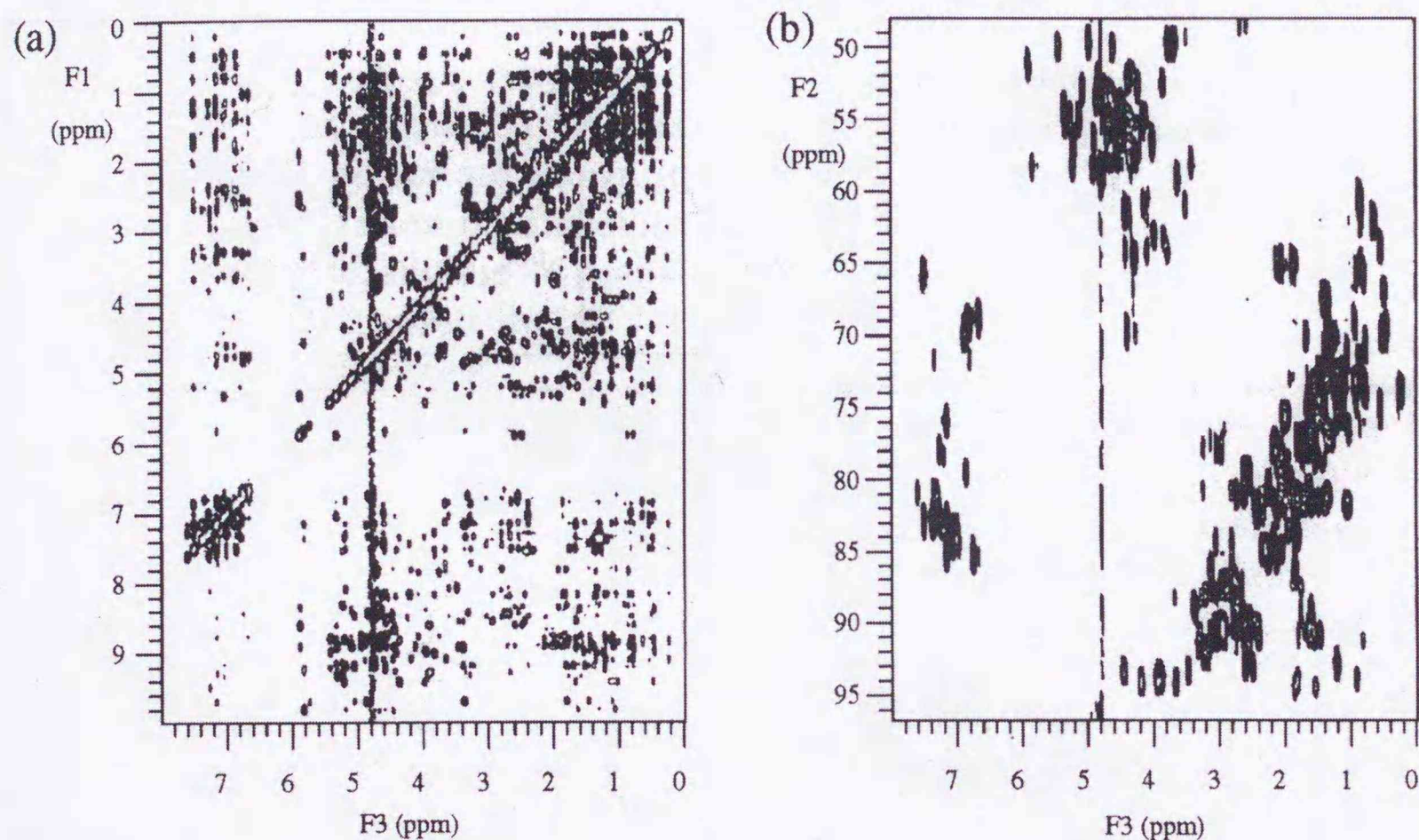


Figure 6. Skyline-projected spectra for F_1 - F_3 planes (a) and F_2 - F_3 planes (b) taken from the 600 MHz 3D fully ^{13}C -refocused ^{13}C -edited NOESY experiment obtained using the pulse scheme of Figure 3a with the sample of NSH3 complexed with VPP in 90 % H_2O /10 % $^2\text{H}_2\text{O}$. The weak (~ 5 Hz) presaturation pulse was applied at the repetition delay (1.0 s). The ^{13}C carrier frequency was set at 72.5 ppm. The NOE mixing time was set to 150 ms. The 3D spectrum was taken with 128 increments in F_1 , 32 increments in F_2 , and eight scans for each FID. Spectral width for each axis was 7000 Hz for F_1 , 7334 Hz for F_2 , and 7000 Hz for F_3 . The total measuring time was 42 h. Data processing was carried out with the VNMR software package. A cosine-shaped apodization function was used for the FID on the F_1 and F_2 axes. The final matrix size was $(F_1, F_2, F_3) = (512 \times 64 \times 1024)$.

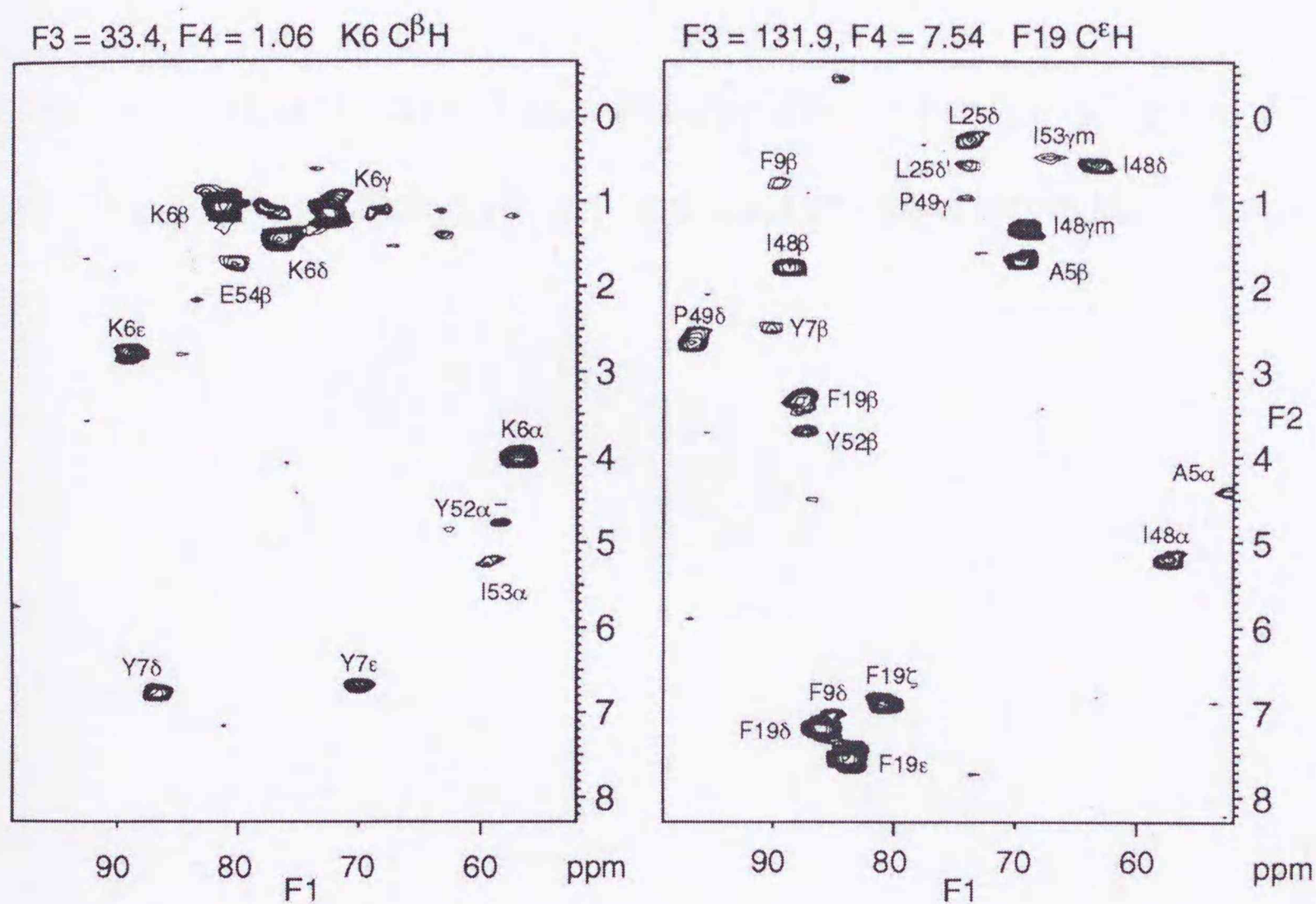


Figure 7. (F₃, F₄) cross sections through the 500 MHz 4D fully ¹³C-refocused ¹³C/¹³C-edited NOESY spectrum obtained using the pulse scheme of Figure 3b with the sample of NSH3 complexed with VPP peptide in ²H₂O. Cross sections are taken at the F₃/F₄ resonance frequencies of (a) K6C β H and (b) F19C ϵ H. NOE mixing time was set to 150 ms. The 4D spectrum was taken with 12 increments in F₁ and F₃, 64 increments in F₂, and four scans for each FID. The spectral widths were 4500 Hz for F₁ and F₃, and 6000 Hz for F₄, respectively. Total measuring time was 103 h. The ¹³C carrier position was set at 75.2 ppm. Data matrix size was (F₁, F₂, F₃, F₄) = (64 x 128 x 64 x 512). All signals appear in positive.

Chapter 2

An improved double-tuned and isotope-filtered pulse scheme based on a pulsed field gradient and a wide-band inversion shaped pulse

Abstract

The author has developed an improved isotope-filtered pulse scheme in combination with a double-tuned filter, a hyperbolic secant inversion pulse, and a z-filter with a pulsed field gradient. These filtering pulse schemes have been incorporated into several one-, two-, and three-dimensional experiments, which were applied to the $^{13}\text{C}/^{15}\text{N}$ uniformly labeled N-terminal SH3 domain of Grb2 complexed with the unlabeled Sos-derived peptide. The proton resonances of the Sos-derived peptide were unambiguously assigned using isotope-filtered DQF-COSY, TOCSY and NOESY spectra. Furthermore, in the isotope-filtered, isotope-edited 3D-NOESY spectrum, intermolecular NOEs between the labeled protein and the unlabeled peptide could be identified. Through these applications, the author demonstrated the high filtering efficiency of the present pulse scheme.

Introduction

Isotope-filtering experiments detect proton signals attached to $^{12}\text{C}/^{14}\text{N}$ nuclei and remove $^{13}\text{C}/^{15}\text{N}$ attached proton signals for selective observation of interactions between $^{13}\text{C}/^{15}\text{N}$ isotope-labeled and unlabeled molecules (Fesik et al., 1991; Weber et al., 1991; Burgering et al., 1993a,b; Lee et al., 1994a; Zhang et al., 1994). These experiments are classified into half-filter (Otting et al., 1986) and purged-filter experiments (Gemmecker et al., 1992; Ikura et al., 1992). Wüthrich and co-workers initially proposed the X-half-filter scheme (Otting et al., 1986), based on the addition or subtraction of FIDs with or without a 180° pulse on X nuclei. During the last few years, improved and modified versions of the X-half filter schemes have been proposed (Otting and Wüthrich, 1989; Burgering et al., 1993a,b; Folkers et al., 1993). Ikura and Bax (1992) developed the purged-filter schemes based on the magnetization transfer of the $^{13}\text{C}/^{15}\text{N}$ attached protons to heteronuclear multiple-quantum coherence, while Fesik and co-workers (Gemmecker et al., 1992; Petros et al., 1992) developed the double-tuned filter to attenuate leakage from aliphatic protons attached to the ^{13}C nuclei.

In these pulse schemes, both aromatic and aliphatic protons are not generally detected, partly because tuning to both coupling constants for the aliphatic (125~140Hz) and aromatic groups (~170Hz) is difficult and partly because there was no suitable pulse that simultaneously inverted the magnetization of aliphatic and aromatic carbons. Hallenga and co-workers (Folmer et al., 1995) have introduced wide-band inversion (Rosen and Zener, 1932; Silver et al., 1984a,b) and band-selective inversion pulses (Geen and Freeman, 1994) to the conventional X-half-filter. In this way, the aromatic and aliphatic protons of $^{13}\text{C}/^{15}\text{N}$ -labeled molecules

can be filtered out simultaneously.

In the present study, the author proposes a novel purged-filter pulse scheme based on a PFG z-filter with a wide-band hyperbolic secant inversion pulse.

Theory, experiments and results

Figure 1(a) shows the basic pulse scheme of the PFG z-filter, where the transverse magnetization of protons (I_y) attached to X-nuclei (X denotes ^{13}C or ^{15}N) is converted to antiphase magnetization, $2I_X S_Z$ after $2\tau = 1/2J_{\text{HX}}$ (^1H and X-nuclei are denoted by I and S, respectively). The transverse magnetization of protons attached to non-X-nuclei is refocused to in-phase magnetization, I_y . The second ^1H 90°_x pulse returns the magnetization of the non-X-attached protons to the z-axis, but leaves the magnetization of the X-attached protons in the x-y plane. The subsequent PFG pulse dephases the transverse magnetization but does not affect the longitudinal magnetization (Bax and Pochapsky, 1992). The final ^1H 90°_x pulse only detects the ^1H magnetization not attached to X-nuclei. Hence, this pulse scheme works as an isotope-purged filter.

The basic PFG z-filter pulse scheme is tandemly linked with different delays (τ_a and τ_b) tuned to the aliphatic and aromatic ^1H - ^{13}C coupling constants, respectively.

The hyperbolic secant pulse is incorporated for wide-band inversion of the ^{13}C magnetization. Thus, this ^{13}C -filtered pulse scheme can be applied to selectively observe aromatic and aliphatic protons bound to ^{12}C nuclei and to eliminate those bound to ^{13}C nuclei (Figure 1(b)). A ^{15}N filter can also be incorporated into this pulse scheme (Figure 1(b)) to eliminate the proton resonances attached to ^{13}C or ^{15}N nuclei (Figure 1(c)). In this case, τ_c and τ_d are set to $1/4J_{\text{NH}}$. After $2\tau_c$ ($2\tau_d$),

the magnetization of protons attached to ^{13}C or ^{15}N becomes antiphase and the subsequent ^1H 90°_x pulse leaves this magnetization in the x-y plane. Thus, these protons are efficiently eliminated with the PFG z-filter, while the protons attached to either ^{12}C or ^{14}N nuclei remain unaffected.

A high suppression efficiency of the proton signals attached to ^{13}C or ^{15}N nuclei by the PFG z-filtered pulse scheme was demonstrated with the $^{13}\text{C}/^{15}\text{N}$ doubly labeled N-terminal SH3 domain of Grb2 complexed with the unlabeled Sos-derived peptide (VPPPVPPIRRR, abbreviated as VPP peptide) (Terasawa et al., 1994). Figure 2(a) shows the conventional 1D ^1H NMR spectrum of the complex of N-SH3 and the VPP peptide in $^2\text{H}_2\text{O}$ solution, where proton resonances of both labeled protein and unlabeled peptide were simultaneously observed. In contrast, Figure 2(b) shows the spectrum measured by the PFG z-filtered pulse scheme (Figure 1(b)), which extracted only the unlabeled VPP peptide signals. As the VPP peptide has no aromatic amino acid residues, the author could observe negligibly small proton signals in the aromatic region. Thus, the ^{13}C -bound proton resonances in both aromatic and aliphatic regions were efficiently suppressed.

The PFG z-filtered pulse scheme was incorporated into 2D NMR pulse sequences to selectively observe the proton resonances of the unlabeled ligand complexed with the $^{13}\text{C}/^{15}\text{N}$ labeled protein. The proton resonances of the unlabeled ligand can be unambiguously assigned through the use of PFG z-filtered 2D NMR. Figure 3 shows ^{13}C -filtered DQF-COSY, ^{13}C -filtered TOCSY and ^{13}C [F1, F2]-filtered NOESY pulse sequences using the PFG z-filter. The ^{13}C purged-filter 2D pulse sequences were initially proposed by Ikura and Bax (1992) and Fesik

and co-workers (Gemmecker et al., 1992). However, in the original schemes the filtering efficiency was limited to either aliphatic or aromatic regions. Therefore, it is an advantage of the present pulse schemes that the ^{13}C -attached proton resonances in the aromatic and the aliphatic regions are simultaneously filtered. The pulse schemes were applied to the complex of $^{13}\text{C}/^{15}\text{N}$ -labeled Grb2 N-SH3 and unlabeled VPP peptide in $^2\text{H}_2\text{O}$ solution. The 2D ^{13}C -filtered DQF-COSY, ^{13}C -filtered TOCSY and ^{13}C [F1, F2]-filtered NOESY spectra are shown in Figure 4. The author only observed the peaks due to proton resonances of the unlabeled peptide in these spectra, demonstrating again the high filtering efficiency of the PFG z-filtered pulse schemes.

For samples dissolved in $^1\text{H}_2\text{O}$, suppression of the solvent resonance is essential. The author applied band-selective shaped 'S' and 'SS' pulses (Smallcombe, 1993) as the observing pulses, which do not excite the $^1\text{H}_2\text{O}$ resonance and simultaneously avoid saturation of the amide proton resonances. Figure 5(a) shows the basic pulse scheme of the $^{13}\text{C}/^{15}\text{N}$ -filter for a sample containing 90% $^1\text{H}_2\text{O}$. Since the amide proton resonances overlap those of the aromatic protons, a ^{13}C -filter for the aromatic carbons, but not for the aliphatic carbons, is sufficient to selectively observe the amide protons of unlabeled molecules. Figure 5(b and c) shows the $^{13}\text{C}/^{15}\text{N}$ -filtered 2D-TOCSY and 2D-NOESY pulse schemes, respectively, incorporating the SS and S pulses as observing pulses. These pulse schemes were applied to the complex of Grb2 N-SH3 and the VPP peptide in $^1\text{H}_2\text{O}$ solution, as shown in Figure 6. In these experiments, solvent presaturation pulses

were not applied. Figure 6(a) shows that the present filter scheme efficiently eliminates the $^{13}\text{C}/^{15}\text{N}$ -attached amide and aromatic proton resonances, leaving only those from ^{14}N attached protons from the VPP peptide. The 2D $^{13}\text{C}/^{15}\text{N}$ -filtered TOCSY and NOESY spectra are shown in Figure 6(b and c) respectively, using the pulse schemes of Figure 5. In both spectra only peaks derived from the VPP peptide are clearly observed. Using the PFG z-filtered pulse schemes for the 2D measurements in $^1\text{H}_2\text{O}$ and $^2\text{H}_2\text{O}$ solution, the proton resonances of the unlabeled peptide complexed with the labeled protein were unambiguously assigned using conventional sequential assignment methods, as shown in Figure 4 and Figure 6.

The filtering experiments in combination with editing experiments such as HMQC or HSQC, are especially powerful in detecting intermolecular NOEs between $^{13}\text{C}/^{15}\text{N}$ -labeled proteins and unlabeled ligands, which enables the unambiguous assignment of the protons involved in the molecular interface. The ^{13}C -edited and ^{13}C -filtered 3D NOESY pulse scheme was initially reported by Ikura and Bax (Ikura et al., 1992). They used a simple HMQC-type method (Müller, 1979). Kay and co-workers (Lee et al., 1994b) modified the HMQC scheme, incorporating a semi-constant time evolution (Grzesiek and Bax, 1993). In the present work, the author used the HSQC-type scheme (Bodenhausen and Ruben, 1980) as the ^{13}C or ^{15}N editor instead of HMQC. Although the HSQC-type scheme has a narrower band range compared with that of HMQC, the incorporation of the ^{13}C hyperbolic secant inversion pulse into the HSQC-type experiment significantly improved the sensitivity of the conventional HSQC-type experiments around the edges of the ^{13}C aromatic and methyl spectral regions (Hallenga et al., 1995; Ogura et al., 1996). The author

modified the ^{13}C -edited, ^{13}C -filtered 3D NOESY pulse scheme, incorporating the hyperbolic secant pulses and the PFG z-filter (Figure 7(a)). The magnetization of the ^{13}C -attached protons is transferred to the ^{12}C -attached protons during the mixing time and is detected using the ^{13}C -filtered pulse scheme.

The author can arbitrarily detect a particular path of the NOESY magnetization transfer using a combination of the isotope-filter and the isotope-editor. In Figure 7(b), the $^{13}\text{C}/^{15}\text{N}$ -filter selects the proton magnetization derived from the unlabeled ligand. During the mixing time, the proton magnetization from the unlabeled ligand transfers to that of the labeled protein. Finally the ^1H - ^{15}N HSQC scheme detects the ^{15}N -labeled amide protons. Thus, all the cross-peaks in the 3D spectra obtained by this pulse scheme are intermolecular NOEs from protons attached to ^{12}C and ^{14}N nuclei to those attached to ^{15}N nuclei.

Figure 8 shows sliced strips of the 3D intermolecular NOESY spectrum obtained with the pulse scheme of Figure 7(a). The measurements were performed using a sample of $^{13}\text{C}/^{15}\text{N}$ -labeled Grb2 N-SH3 and the unlabeled VPP peptide dissolved in $^2\text{H}_2\text{O}$. Many NOE correlations were observed from both the aromatic and aliphatic protons of $^{13}\text{C}/^{15}\text{N}$ -labeled N-SH3 to the aliphatic protons of the unlabeled VPP peptide. Obviously, the hyperbolic secant pulse efficiently and simultaneously inverted not only the aliphatic carbons, but also the aromatic carbons.

Figure 9 shows the sliced spectrum obtained with the pulse scheme of Figure 7(b). The spectrum was recorded with a sample of $^{13}\text{C}/^{15}\text{N}$ labeled Grb2 N-SH3 and the unlabeled VPP peptide dissolved in 90% $^1\text{H}_2\text{O}$ without solvent

presaturation. The intermolecular NOE cross peaks between the side-chain amide protons of Asn-51 and the Val-5, Pro-3 and Pro-6 proton resonances of the VPP peptide were found to be presented. Because no presaturation pulses were used in this experiment, the amide protons were observed with high sensitivity, even the fast exchangeable amide protons in the side chains.

Conclusion

The author has shown that a combination of PFG and the hyperbolic secant pulse is useful for isotope-filtering experiments. By incorporating the basic filtering pulse scheme into multidimensional experiments, the author has presented a suite of filtering pulse experiments that unambiguously assign the proton resonances of unlabeled molecules and elucidate intermolecular NOEs between the labeled and unlabeled molecules.

References

- Bax, A. and Pochapsky, S.S. (1992) *J. Magn. Reson.* 99, 638-643.
- Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.* 69, 185-189.
- Burgering, M.J.M., Boelens, R., Caffrey, M., Breg, J.N. and Kaptein, R. (1993a) *FEBS Lett.* 330, 105-109.
- Burgering, M.J.M., Boelens, R. and Kaptein, R. (1993b) *J. Biomol. NMR* 3, 709-714.
- Fesik, S.W., Gampe Jr., R.T., Eaton, H.L., Gemmecker, G., Olejniczak, E.T., Neri, P., Holtzman, T.F., Egan, D.A., Edalji, R., Simmer, R., Helfrich, R., Hochlowski, J. and Jackson, M. (1991) *Biochemistry* 30, 6574-6583.
- Folkers, P.J.M., Folmer, R.H.A., Konings, R.N.H and Hilbers, C.W. (1993) *J. Am. Chem. Soc.* 115, 3798-3799.
- Folmer, R.H.A., Hilbers, C.W., Konings, R.N.H. and Hallenga, K. (1995) *J. Biomol. NMR* 5, 427-432.
- Geen, H. and Freeman, R. (1991) *J. Magn. Reson.* 93, 93-141.
- Gemmecker, G., Olejniczak, E.T. and Fesik, S.W. (1992) *J. Magn. Reson.* 96, 199-204.
- Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. (1988) *J. Am. Chem. Soc.* 110, 7870-7872.
- Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR* 3, 185-204.
- Hallenga, K and Lippens, G.M. (1995) *J. Biomol. NMR* 5, 59-66.
- Ikura, M. and Bax, A. (1992a) *J. Am. Chem. Soc.* 114, 2433-2440.
- Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. and Bax, A. (1992b) *Science* 256, 632-638.
- Lee, W., Harvey, T.S., Yin, Y., Yau, P., Litchfield, D. and Arrowsmith, C.H. (1994a) *Nature Struct. Biol.* 1, 877-890.
- Lee, W., Revington, M.J., Arrowsmith, C. and Kay, L.E. (1994b) *FEBS Lett.* 350,

87-90.

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

Ogura, K., Terasawa, H. and Inagaki, F. (1996) *J. Magn. Reson. B* 112, 63-68.

Otting, G., Senn, H., Wagner, G. and Wüthrich, K. (1986) *J. Magn. Reson.* 70, 500-505.

Otting, G. and Wüthrich, K. (1989) *J. Magn. Reson.* 85, 586-594.

Petros, A.M., Kawai, M., Luly, J.R. and Fesik, S.W. (1992) *FEBS Lett.* 308, 309-314.

Rosen, N. and Zener, C. (1932) *Phys. Rev.* 40, 502-507.

Silver, M.S., Joseph, R.I. and Hoult, D.I. (1984a) *J. Magn. Reson.* 59, 347-353.

Silver, M.S., Joseph, R.I. and Hoult, D.I. (1984b) *Phys. Rev. A* 31, 2753-2755.

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

Terasawa, H., Kohda, D., Hatanaka, H., Tsuchiya, S., Ogura, K., Nagata, K., Ishii, S.,

Mandiyan, V., Ullrich, A., Schlessinger, J. and Inagaki, F. (1994)

Nature Struct. Biol. 1, 891-897.

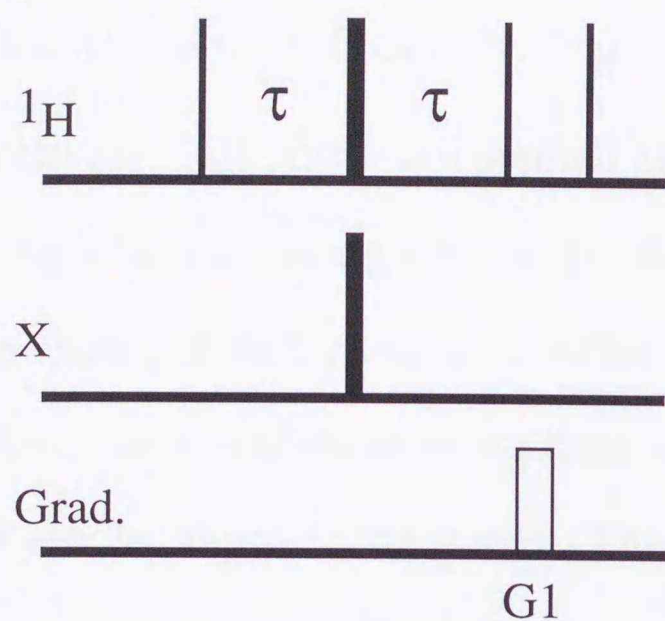
Weber, C., Wider, G., Von Freyberg, B., Traber, R., Braun, W., Widmer, H. and

Wüthrich, K. (1991) *Biochemistry* 30, 6563-6574.

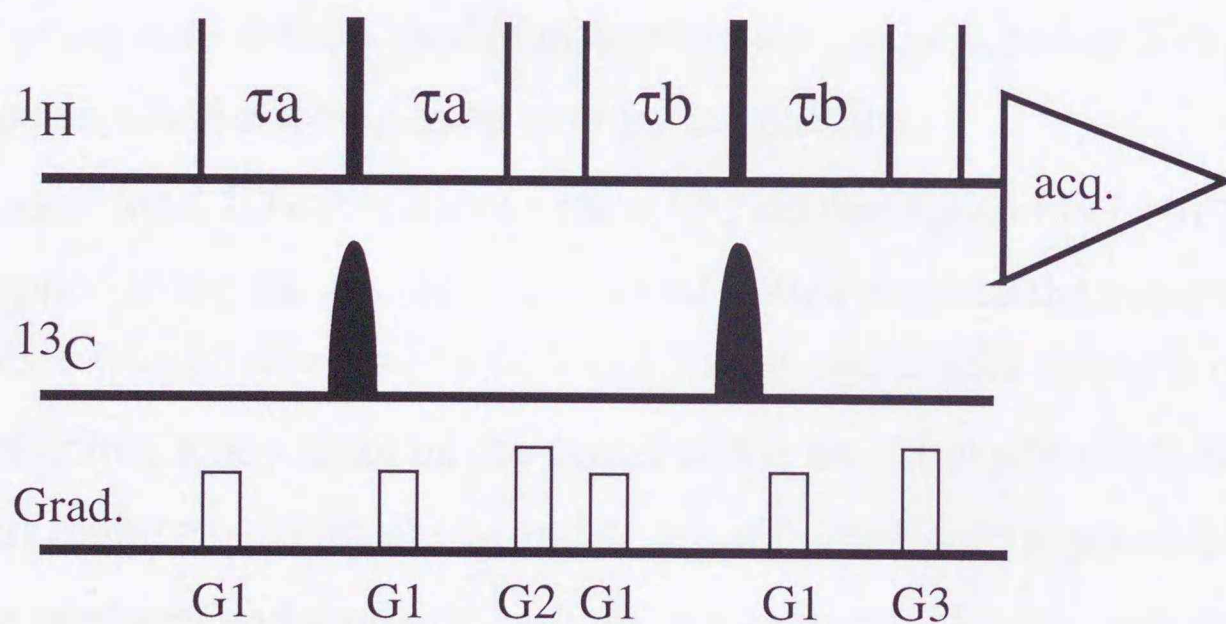
Zhang, H., Zhao, D., Revington, M., Lee, W., Jia, X., Arrowsmith, C. and Jardetsky, O.

(1994) *J. Mol. Biol.* 238, 592-614.

(a)



(b)



(c)

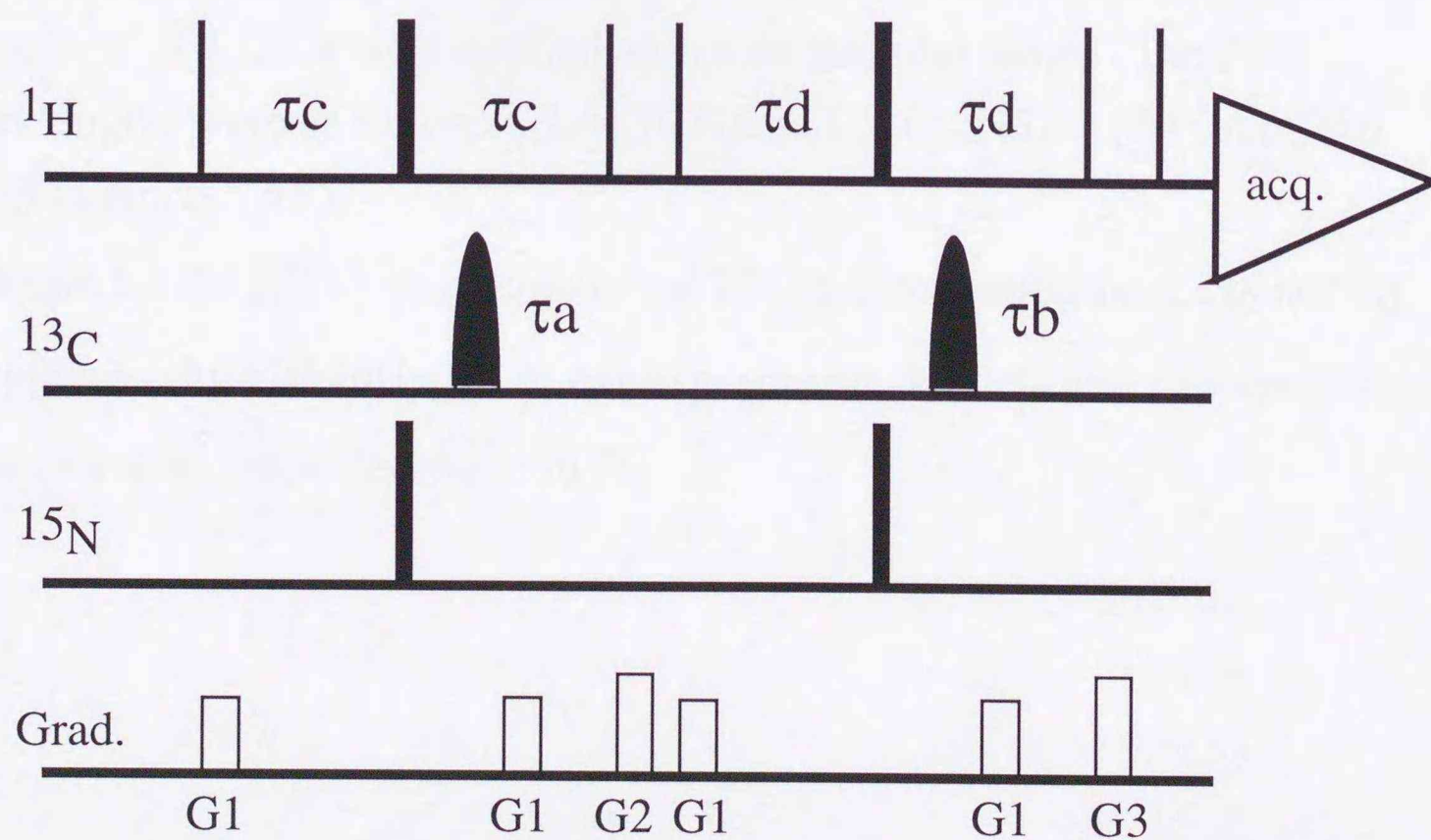


Figure 1. (a) The basic part of the PFG z-filter pulse scheme, in which τ is set to $1/4J_{HX}$. X denotes either ^{13}C or ^{15}N . Narrow and wide lines indicate 90° and 180° pulses, respectively. All pulses are applied along the x-axis. First a ^1H 90° pulse excites ^1H magnetization along the y-axis. A pair of ^1H and X-nuclei 180° pulses refocuses the chemical shift evolution. After the delay 2τ , the ^1H magnetization attached to X-nuclei is antiphase along the x-axis, and the ^1H magnetization not attached to X-nuclei aligns to the y-axis. The second ^1H 90° pulse returns the ^1H magnetization not attached to X-nuclei to the z-axis, and leaves the ^1H magnetization attached to X-nuclei in the x-y plane. A pulsed field gradient G1 purges all the transverse magnetization and does not affect the longitudinal magnetization. The final ^1H 90° pulse only detects the ^1H magnetization not attached to X-nuclei. Hence, this pulse scheme works as an isotope-purged filter.

(b) Pulse scheme for a 1D experiment with a ^{13}C double-tuned PFG z-filter. All pulses are applied along the x-axis. The shaped pulses indicate the hyperbolic secant inversion pulses which have a pulse width of $360\ \mu\text{s}$ and a field strength of $20\ \text{kHz}$. The ^{13}C carrier frequency is set on the center of the aromatic and aliphatic ^{13}C chemical shift regions ($\sim 75\ \text{ppm}$). τ_a and τ_b are set to typical values of $1/4J_{\text{CH}}$, tuned for the aliphatic and aromatic carbons, respectively. Under our experimental conditions, τ_a and τ_b were set to 1.85 and $1.45\ \text{ms}$, respectively. PFGs of G2 and G3 eliminate the unwanted transverse magnetization. PFG G1 cancels the imperfection of the 180° pulses. All PFGs were applied with a rectangular shape. The PFG strengths and lengths were as follow: $G1 = (10\ \text{G/cm}, 0.5\ \text{ms})$, $G2 = (20\ \text{G/cm}, 0.5\ \text{ms})$, $G3 = (25\ \text{G/cm}, 0.5\ \text{ms})$.

(c) Pulse scheme for the $^{13}\text{C}/^{15}\text{N}$ double-tuned PFG z-filter experiment. τ_c and τ_d were set to typical values of $1/4J_{\text{NH}}$. In our experiments, both τ_c and τ_d were $2.7\ \text{ms}$. Other parameters were set as described in (b).

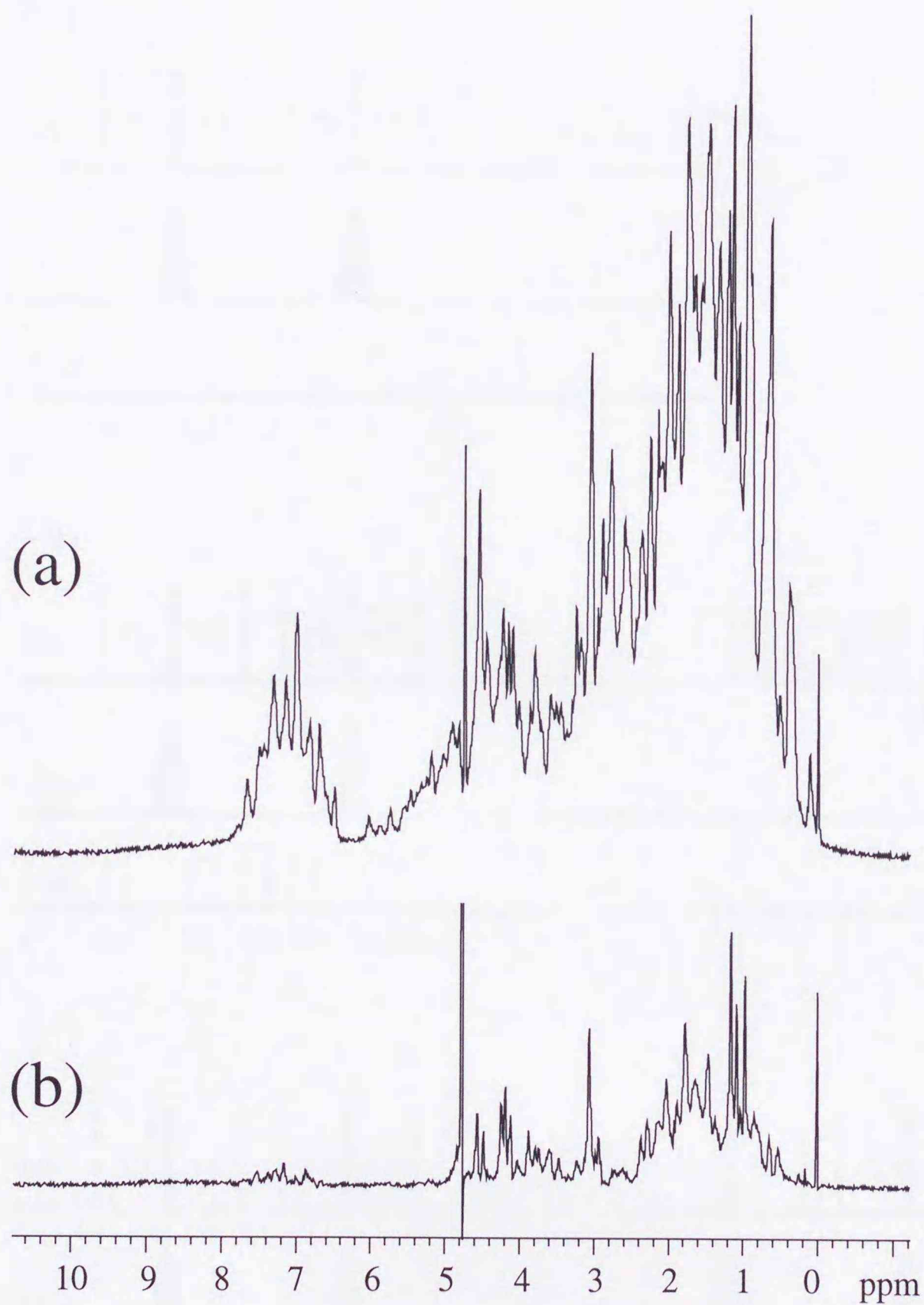


Figure 2. Spectral comparison of the 2 mM unlabeled VPP peptide complexed with the $^{13}\text{C}/^{15}\text{N}$ uniformly enriched N-SH3 protein dissolved in $^2\text{H}_2\text{O}$, recorded with a Varian UNITY 500, ^1H 500MHz NMR spectrometer. The vertical scaling and the number of scans (64, with 5984 complex points) were identical for each spectrum. The ^{13}C carrier frequency was set at 75.2ppm. (a) normal unfiltered spectrum; (b) ^{13}C double-tuned PFG z-filtered spectrum.

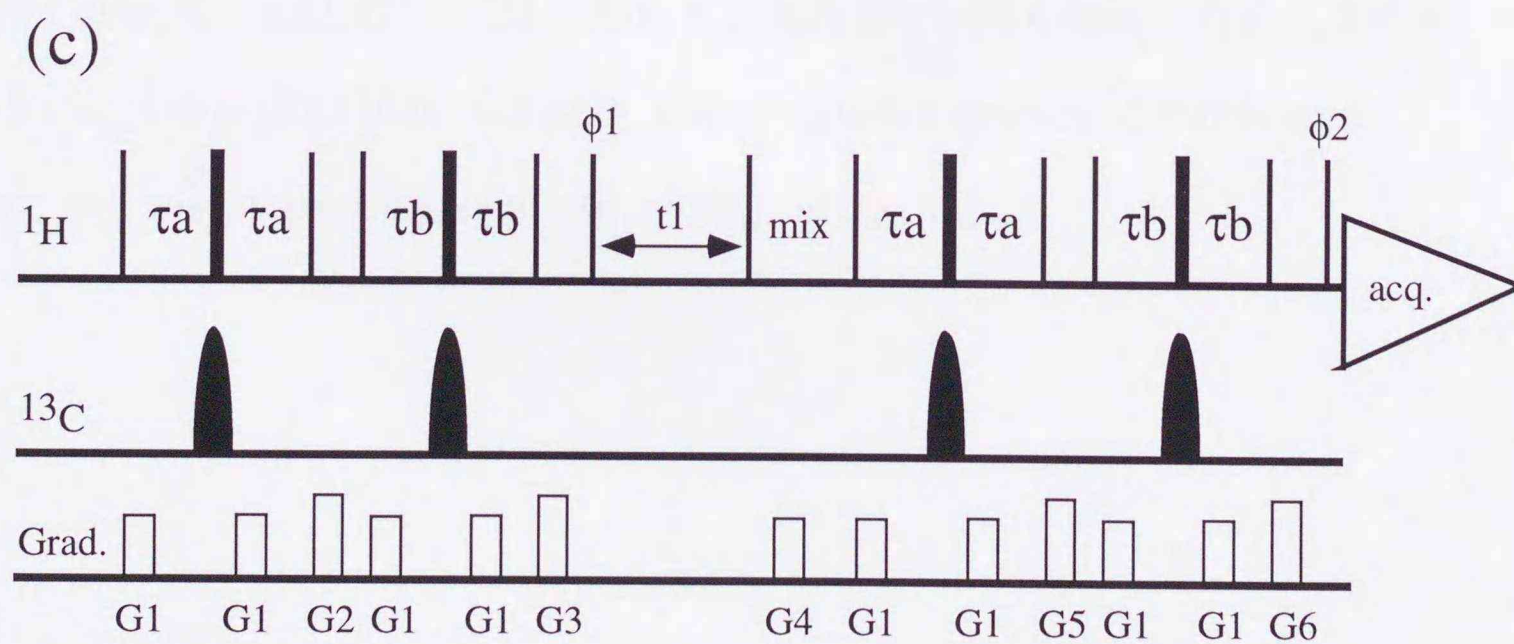
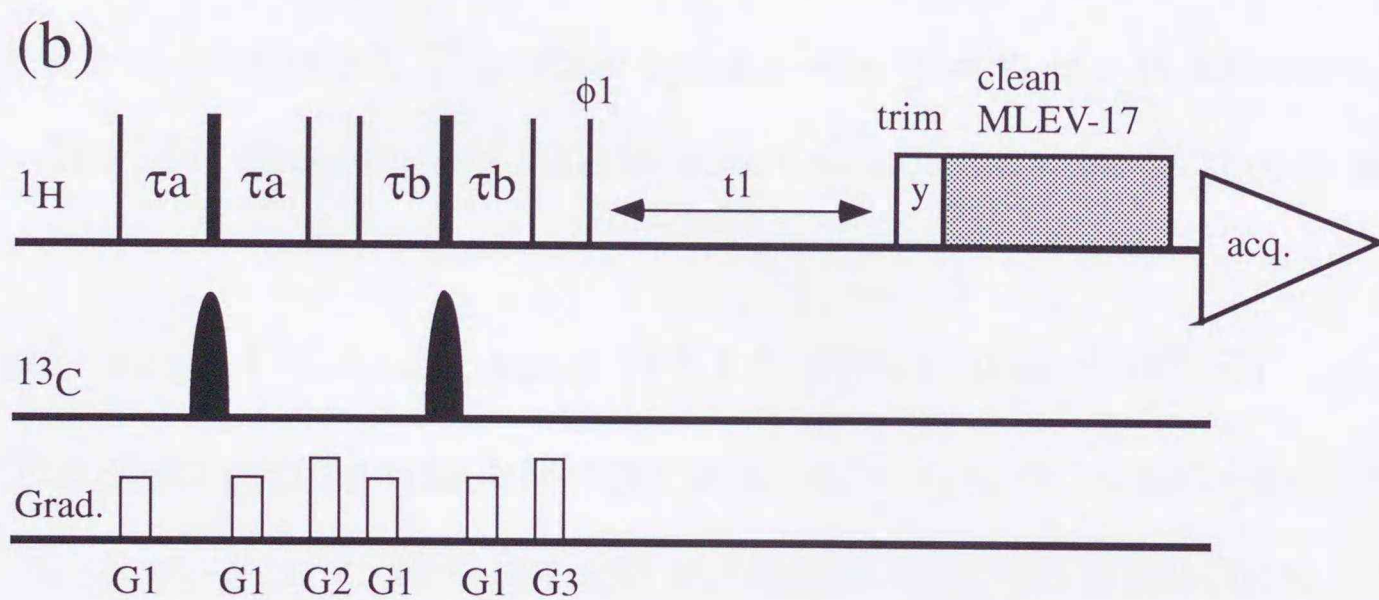
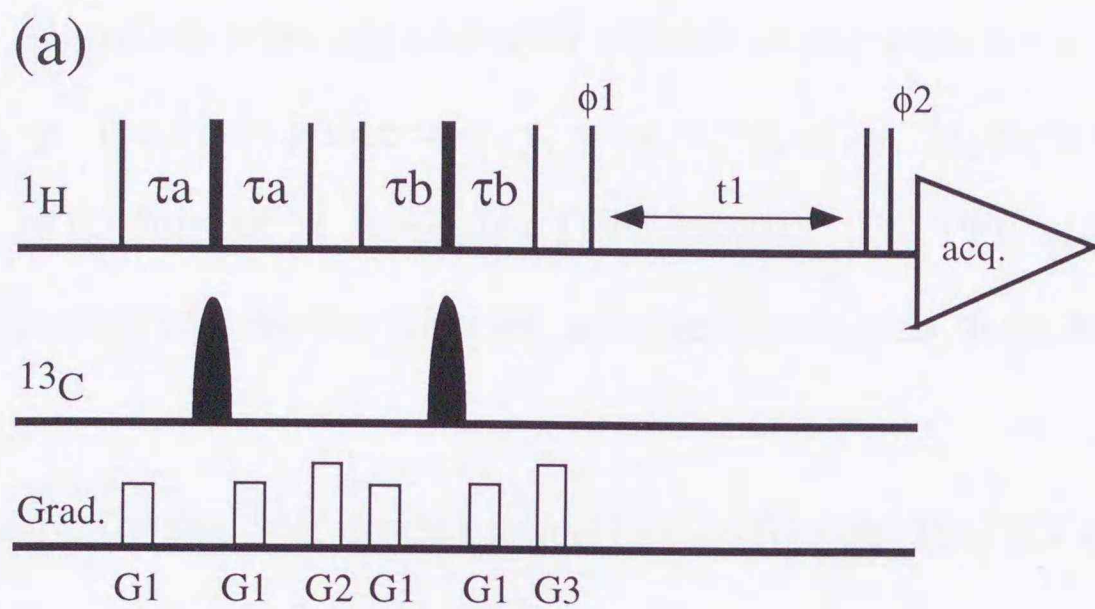
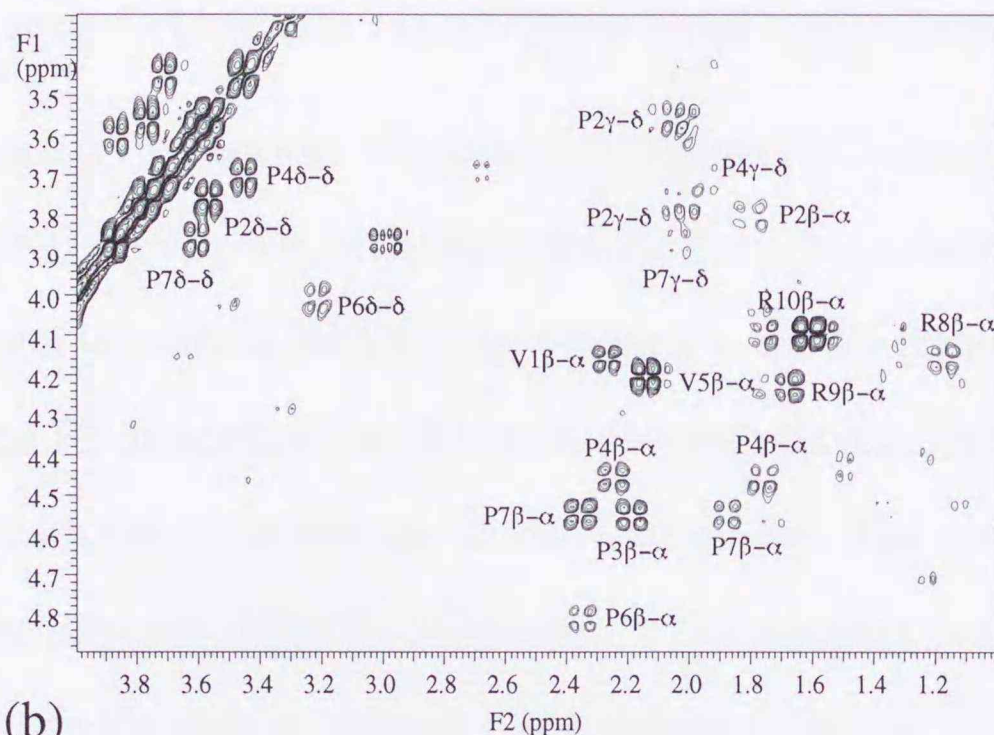


Figure 3. (a) Pulse scheme for the ^{13}C double-tuned PFG z-filtered DQF-COSY experiment. All pulses were applied with a phase of x except for $\phi_1 = 4(x), 4(-x)$ and $\phi_2 = x, y, -x, -y$. Receiver phase = $-x, y, x, -y, x, -y, -x, y$. Hypercomplex data acquisition was performed in the States-TPPI manner. The PFG strengths and lengths and the intervals for the filtering scheme were set as described in the legend to Fig.1(b).

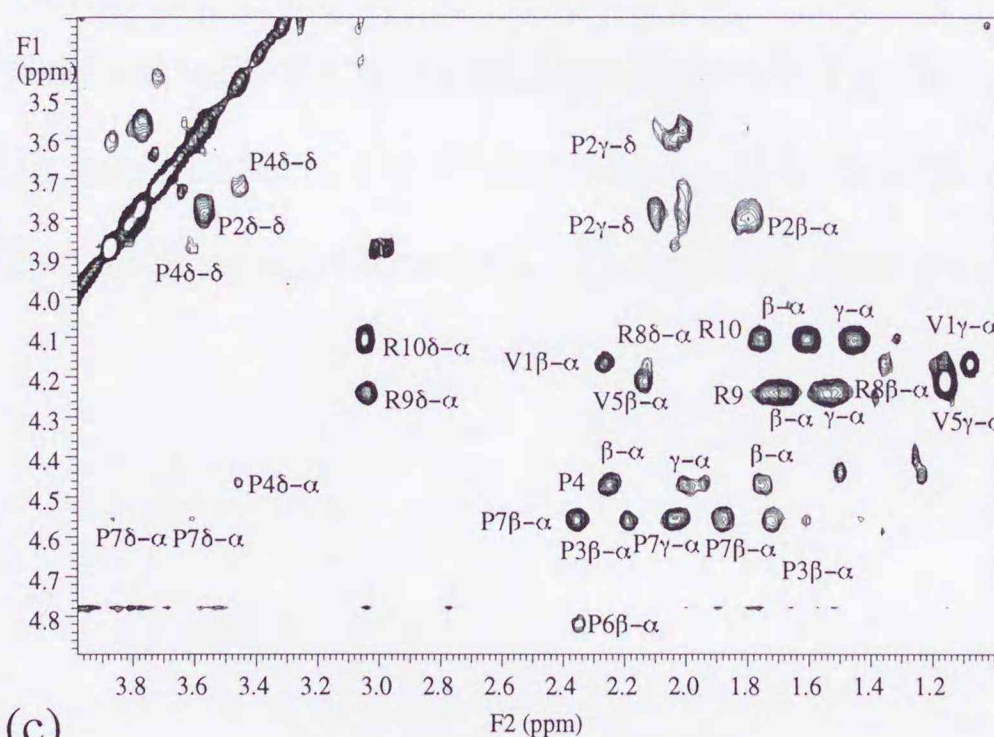
(b) Pulse scheme for the ^{13}C double tuned PFG z-filtered TOCSY experiment. TOCSY mixing is achieved by the clean MLEV-17 (Griesinger et al., 1988) pulse sequence with a 2 ms trim pulse. The phase cycling was: $\phi_1 = x, -x$ and Receiver phase = $x, -x$. The PFG strengths and lengths were set to be the same as shown in Fig.1(b).

(c) Pulse scheme for the ^{13}C double-tuned, [F1, F2] PFG z-filtered NOESY experiment. The phase cycling was: $\phi_1 = 4(x), 4(-x), \phi_2 = x, y, -x, -y$ and receiver phase = $x, y, -x, -y, -x, -y, x, y$. PFG strength and length were: $G_1 = (10 \text{ G/cm}, 0.5 \text{ ms}), G_2 = (20 \text{ G/cm}, 0.5 \text{ ms}), G_3 = (25 \text{ G/cm}, 0.5 \text{ ms}), G_4 = (8 \text{ G/cm}, 3.0 \text{ ms}), G_5 = (18 \text{ G/cm}, 0.5 \text{ ms}), G_6 = (22 \text{ G/cm}, 0.5 \text{ ms})$. G_4 eliminates unwanted transverse magnetization, and therefore reduces phase cycling of the pulses.

(a)



(b)



(c)

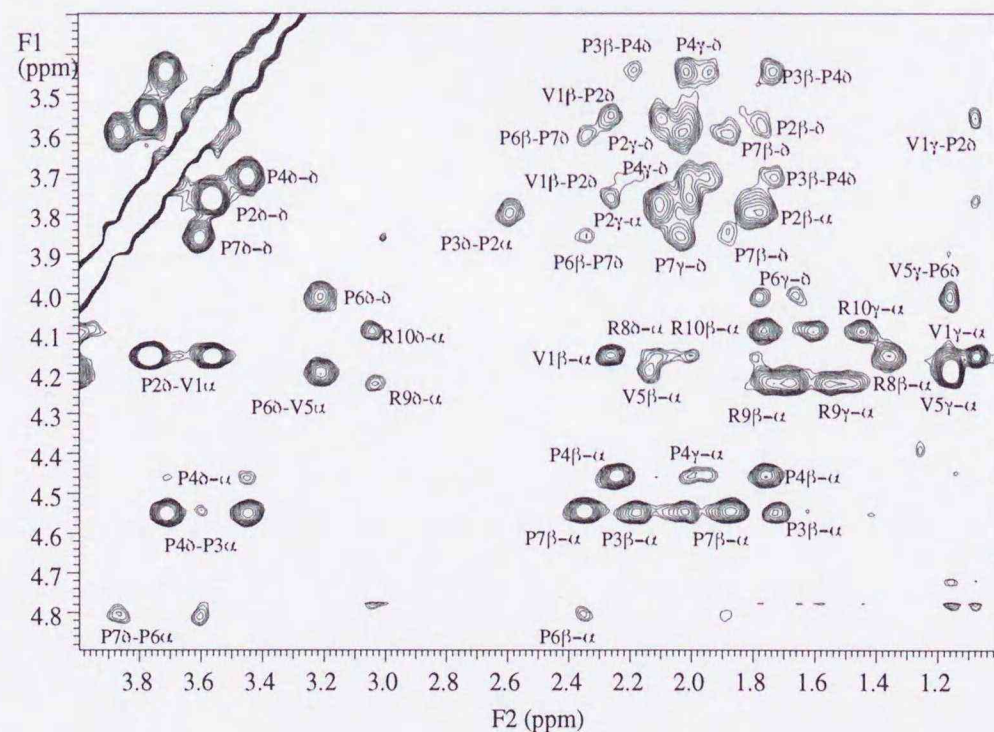


Figure 4. ^{13}C double-tuned PFG z-filtered 2D spectra of a 2mM sample of unlabeled VPP peptide complexed with the $^{13}\text{C}/^{15}\text{N}$ uniformly enriched N-SH3 protein dissolved in $^2\text{H}_2\text{O}$, recorded with the ^1H 500 MHz NMR spectrometer. The cross peaks derived from the VPP peptide, which were assigned by Terasawa et al. (1994) are labeled. (a) DQF-COSY spectrum, obtained with the pulse scheme shown in Fig.3(a). The spectrum was taken with 512 increments in the F1 dimension, 1024 complex points in the F2 dimension and 32 scans for each increment. (b) TOCSY spectrum, obtained with the pulse scheme shown in Fig.3(b). The spectrum was recorded with 512 increments in the F1 dimension, 1024 complex points in the F2 dimension and 16 scans for each increment. The mixing time was 50 ms. (c) NOESY spectrum, obtained with the pulse scheme shown in Fig.3(c). The spectrum was recorded with 512 increments in the F1 dimension, 1024 complex points in the F2 dimension and 32 scans for each increment. The mixing time was 150 ms.

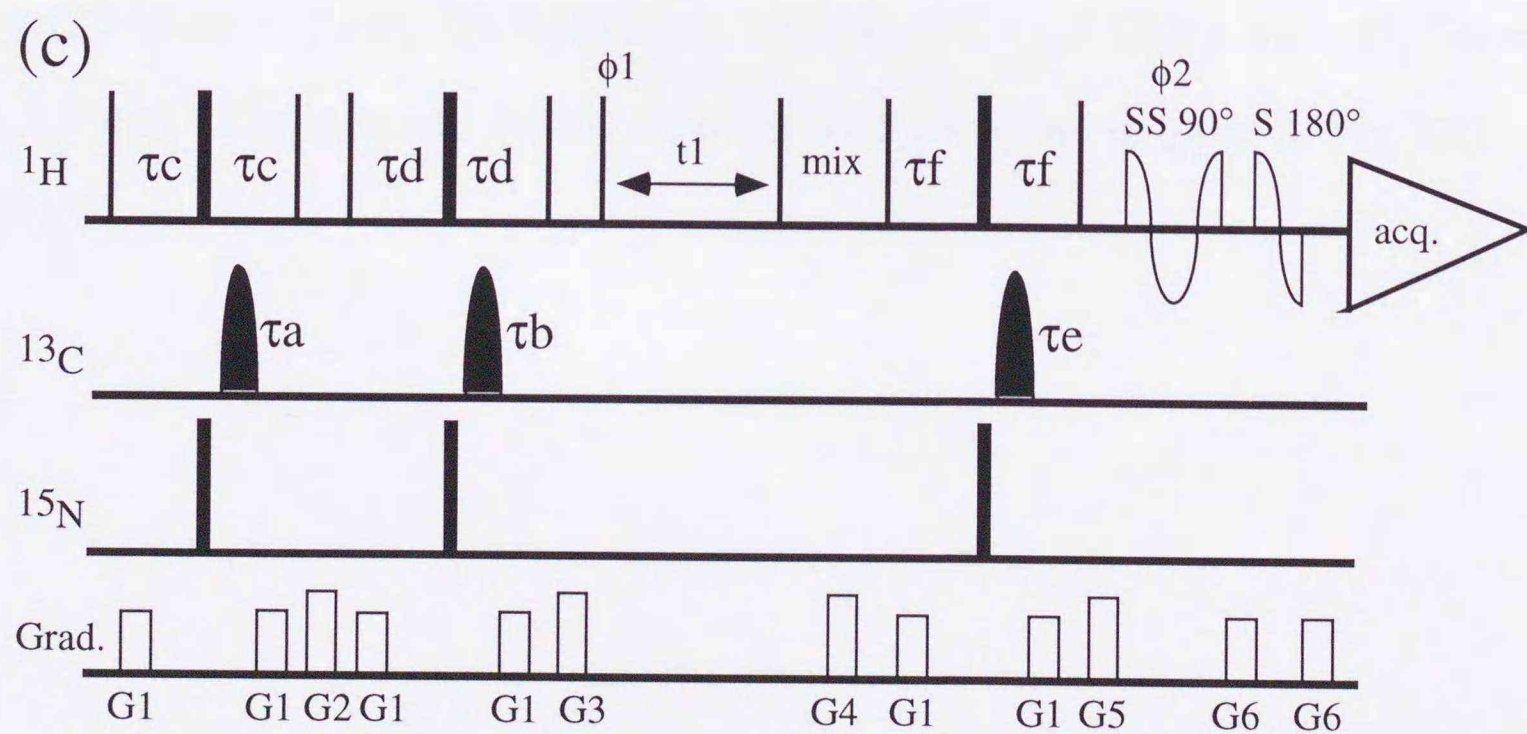
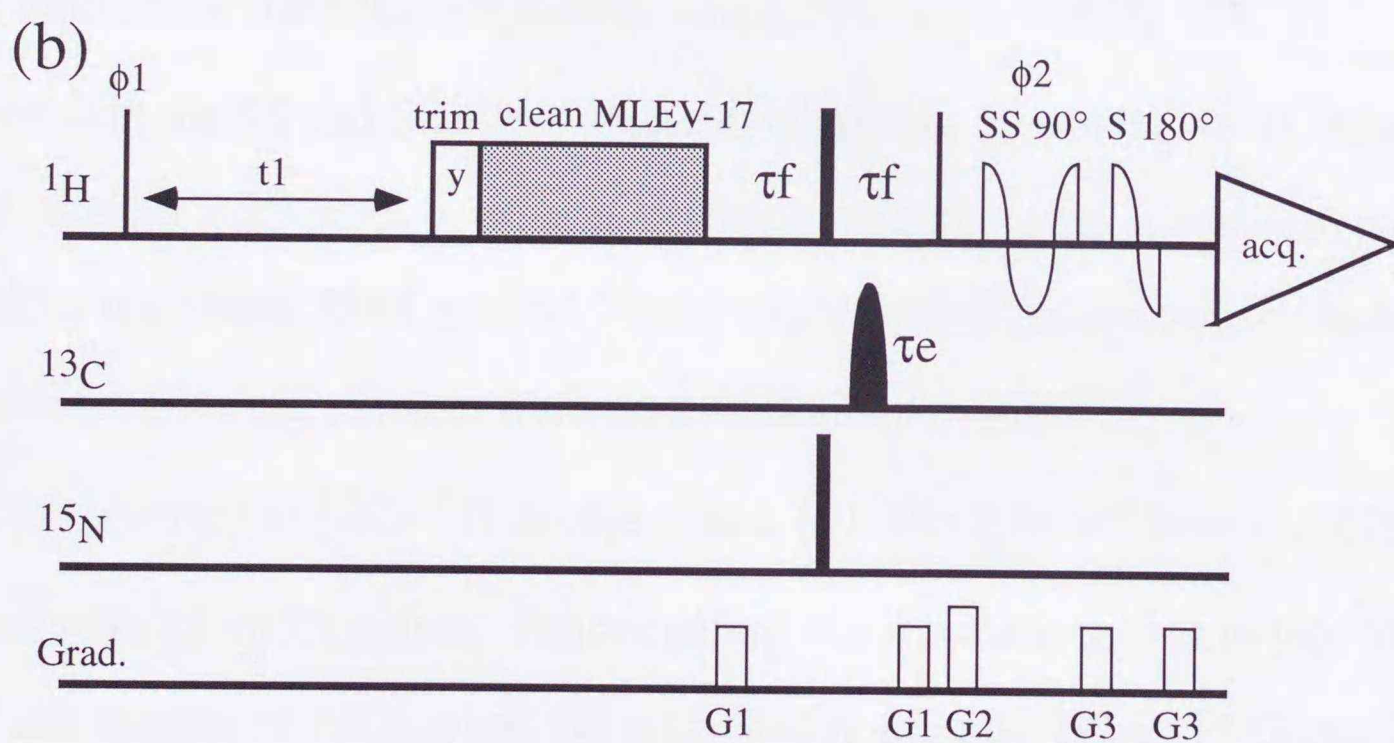
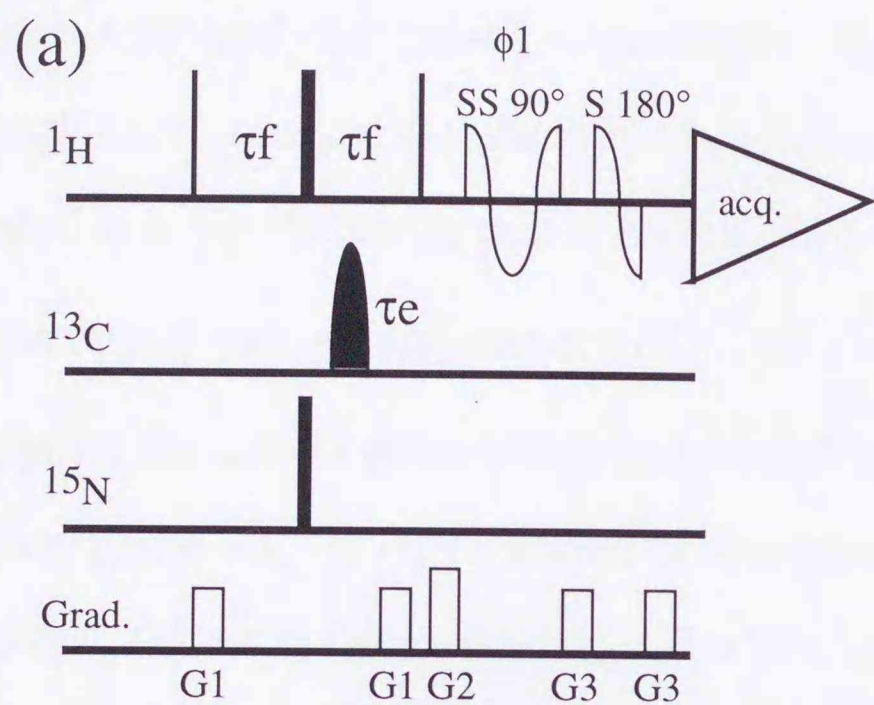
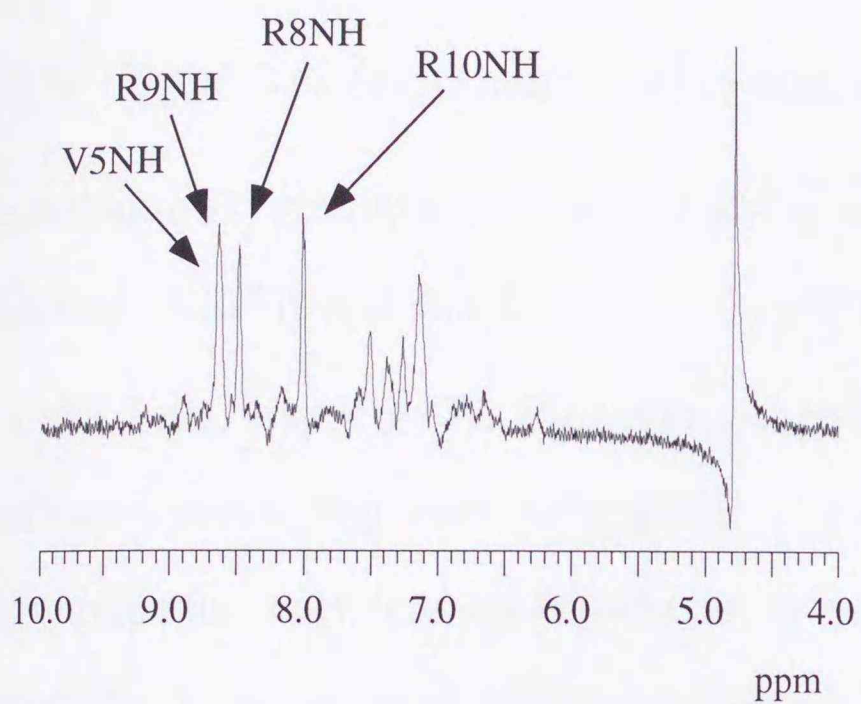


Figure 5. (a) Pulse scheme for the $^{13}\text{C}/^{15}\text{N}$ double-tuned, PFG z-filtered 1D experiment with SS and S shaped pulses. The SS and S shaped pulses function as band-selective 90° and 180° pulses, respectively. The carrier of the transmitter was set at the solvent resonance, and the effective excitation range of the SS and S pulses was adjusted to cover the amide proton region. τ_e and τ_f were set to typical values of $1/4J_{\text{CH}}$ (aromatic) and $1/4J_{\text{NH}}$, respectively. All pulses were applied along the x-axis, except for the last SS pulse which had the following phase cycling: $\phi_1 = x, y, -x, -y$. Receiver phase = $x, -y, -x, y$. The PFG strengths and lengths were: $G_1 = (10 \text{ G/cm}, 0.5 \text{ ms})$, $G_2 = (25 \text{ G/cm}, 0.5 \text{ ms})$, $G_3 = (8 \text{ G/cm}, 1 \text{ ms})$.

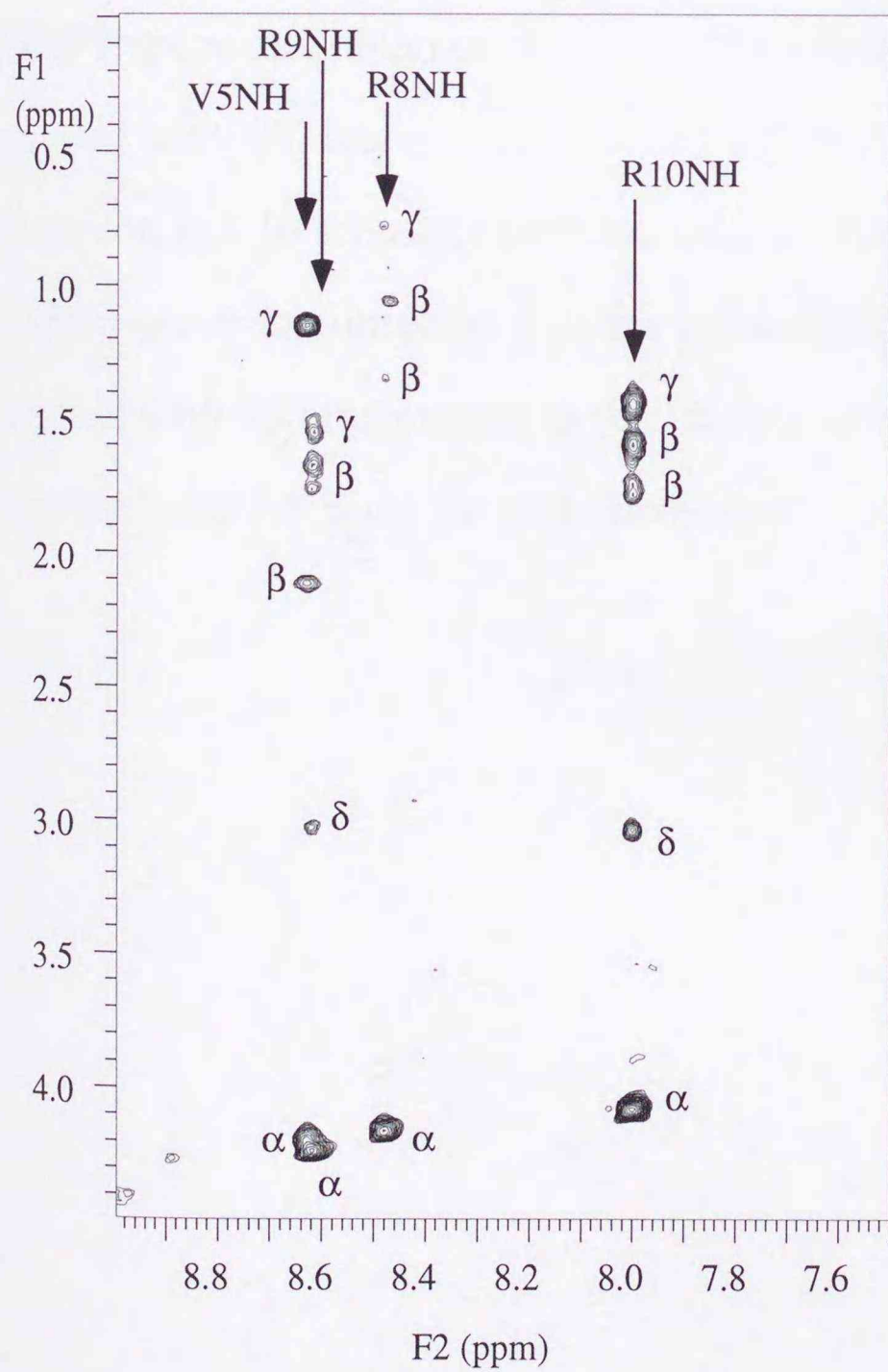
(b) Pulse scheme for the $^{13}\text{C}/^{15}\text{N}$ double-tuned, PFG z-filtered 2D TOCSY experiment with the SS and S pulses. Phase cycling was: $\phi_1 = 4(x), 4(-x)$, $\phi_2 = x, y, -x, -y$ and receiver = $x, -y, -x, y, -x, y, x, -y$. Hypercomplex data acquisition was performed by the States-TPPI method. Other experimental parameters, including the intervals for the filtering scheme, were set as described in (a).

(c) Pulse scheme for the $^{13}\text{C}/^{15}\text{N}$ double-tuned, $[\text{F1}, \text{F2}]$ PFG z-filtered NOESY experiment with SS and S pulses. Phase cycling was the same as that in (b). The strengths and lengths of PFGs were: $G_1 = (10 \text{ G/cm}, 0.5 \text{ ms})$, $G_2 = (15 \text{ G/cm}, 0.5 \text{ ms})$, $G_3 = (22 \text{ G/cm}, 0.5 \text{ ms})$, $G_4 = (8 \text{ G/cm}, 3.0 \text{ ms})$, $G_5 = (25 \text{ G/cm}, 0.5 \text{ ms})$, $G_6 = (8 \text{ G/cm}, 1.0 \text{ ms})$. The intervals for the filtering were the same as those in Fig.1(c) and Fig.5(a).

(a)



(b)



(c)

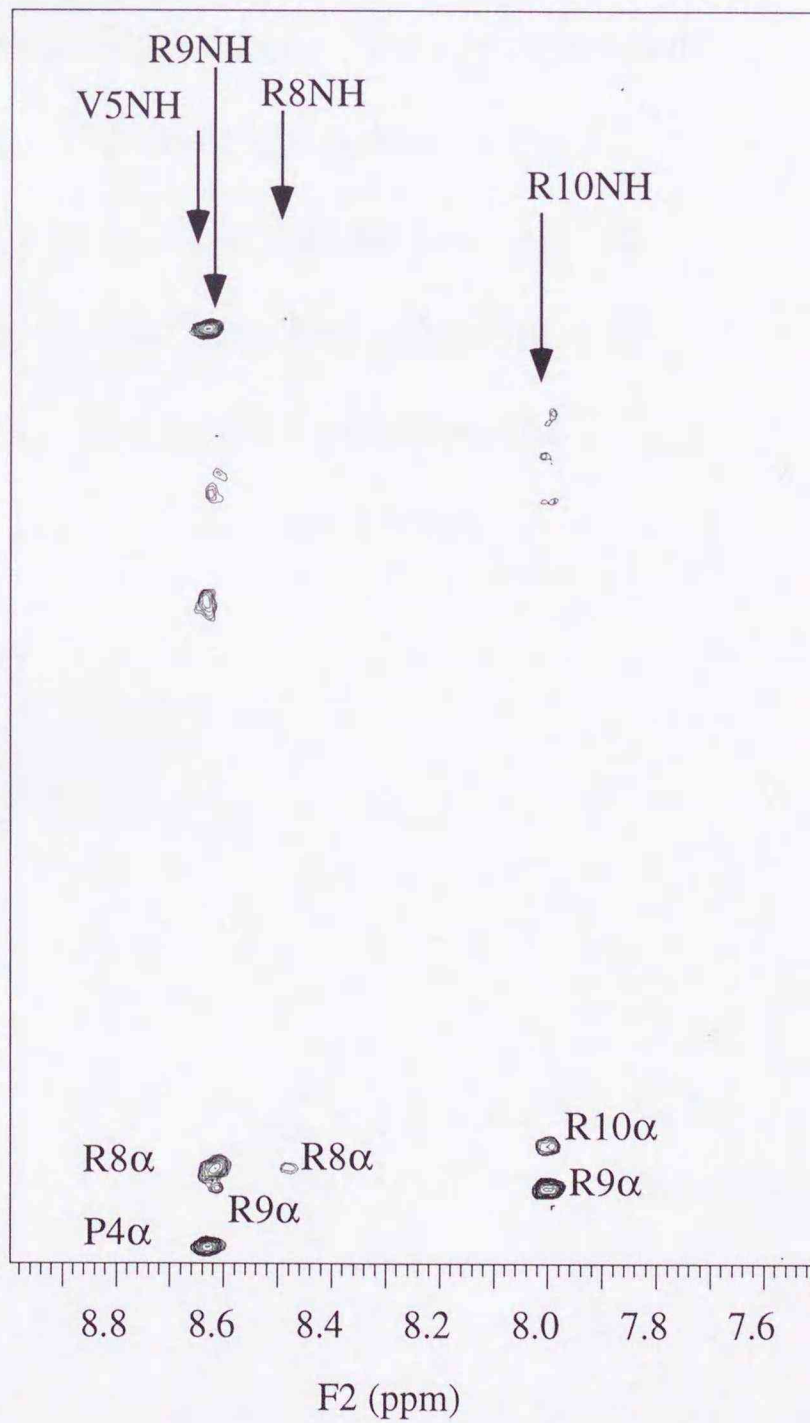


Figure 6. $^{13}\text{C}/^{15}\text{N}$ PFG z-filtered 1D and 2D spectra of the 2 mM unlabeled VPP peptide complexed with the $^{13}\text{C}/^{15}\text{N}$ uniformly enriched N-SH3 protein, dissolved in 90 % $^1\text{H}_2\text{O}$ /10 % $^2\text{H}_2\text{O}$ and recorded with the ^1H 500 MHz NMR spectrometer. The maximum rf strengths and pulse widths were 813 Hz and 601.2 μs for the SS pulse, and 2886 Hz and 361.2 μs for the S pulse, respectively. τ_e and τ_f were set to 1.4 and 2.7 ms, respectively. Presaturation pulses and post-acquisition solvent suppression processing were not applied. The amide proton regions are shown. The VPP peptide has only four observable amide protons; Those peaks are labeled as assigned by Terasawa et al. (1994). (a) 1D spectrum obtained with the pulse scheme of Figure 5(a). The number of scans was 64 with 5984 complex points. (b) 2D TOCSY spectrum, obtained with the pulse scheme in Figure 5(b). The spectrum was recorded with 300 increments in the F1 dimension, 512 complex points in the F2 dimension and 16 scans for each increment. The mixing time was 50 ms. (c) 2D NOESY spectrum, obtained with the pulse scheme in Fig.5(c). The spectrum was recorded with 500 increments in the F1 dimension, 512 complex points in the F2 dimension and 64 scans for each increment. The mixing time was 150ms.

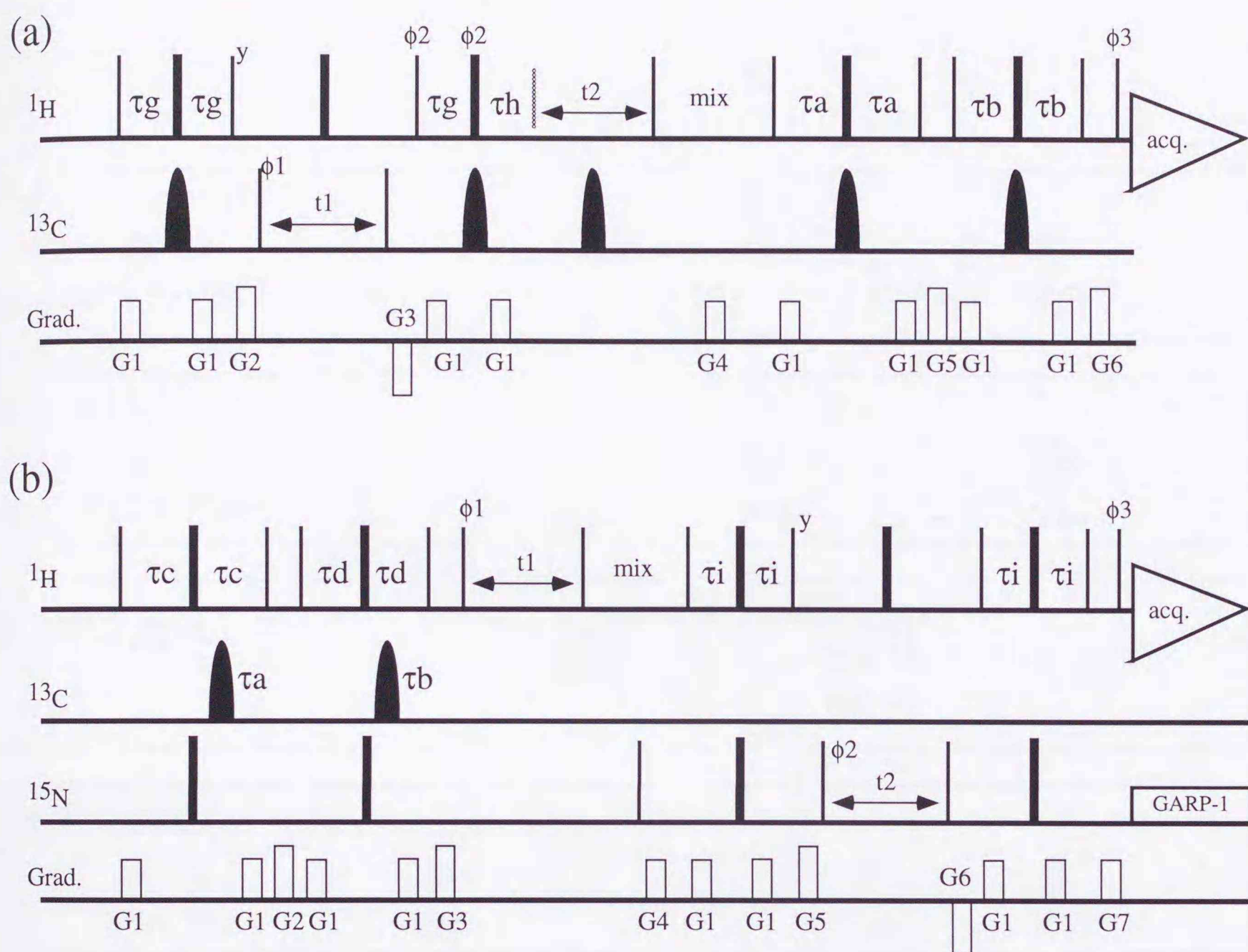


Figure 7(a). Pulse scheme for the improved 3D F1 ^{13}C edited- F3 ^{13}C filtered HSQC-NOESY experiment. Phase cycling was: $\phi_1 = x, -x$, $\phi_2 = 2(y), 2(-y)$, $\phi_3 = 4(x), 4(y), 4(-x), 4(-y)$, and receiver = $x, 2(-x), x, y, 2(-y), y, -x, 2(x), -x, -y, 2(y), -y$. $\tau_g = 1.6$ ms ($1/4J_{\text{CH}}$), $\tau_h = \tau_g - \delta$ (δ is pulse width of the hyperbolic secant inversion pulse). τ_a and τ_b were set as described in the legend to Figure 1(b). The strengths and lengths of PFGs were: $G_1 = (10 \text{ G/cm}, 0.5 \text{ ms})$, $G_2 = (30 \text{ G/cm}, 1.0 \text{ ms})$, $G_3 = (-25 \text{ G/cm}, 1.0 \text{ ms})$, $G_4 = (8.0 \text{ G/cm}, 3.0 \text{ ms})$, $G_5 = (18 \text{ G/cm}, 0.5 \text{ ms})$, $G_6 = (22 \text{ G/cm}, 0.5 \text{ ms})$.

(b). Pulse scheme for the improved 3D F1 $^{13}\text{C}/^{15}\text{N}$ filtered- F3 ^{15}N edited NOESY-HSQC experiment. Phase cyclings were: $\phi_1 = x, -x$, $\phi_2 = 2(x), 2(-x)$, $\phi_3 = 4(x), 4(-x)$, and receiver = $x, 2(-x), x, -x, 2(x), -x$. τ_a , τ_b , τ_c , and τ_d were the same as those in Fig.1(c). $\tau_i = 2.25 \text{ ms}$ (slightly smaller than the actual $1/4J_{\text{NH}}$). The strengths and lengths of PFGs were: $G_1 = (10 \text{ G/cm}, 0.5 \text{ ms})$, $G_2 = (18 \text{ G/cm}, 0.5 \text{ ms})$, $G_3 = (22 \text{ G/cm}, 0.5 \text{ ms})$, $G_4 = (8.0 \text{ G/cm}, 3.0 \text{ ms})$, $G_5 = (30 \text{ G/cm}, 1.0 \text{ ms})$, $G_6 = (-25 \text{ G/cm}, 1.0 \text{ ms})$.

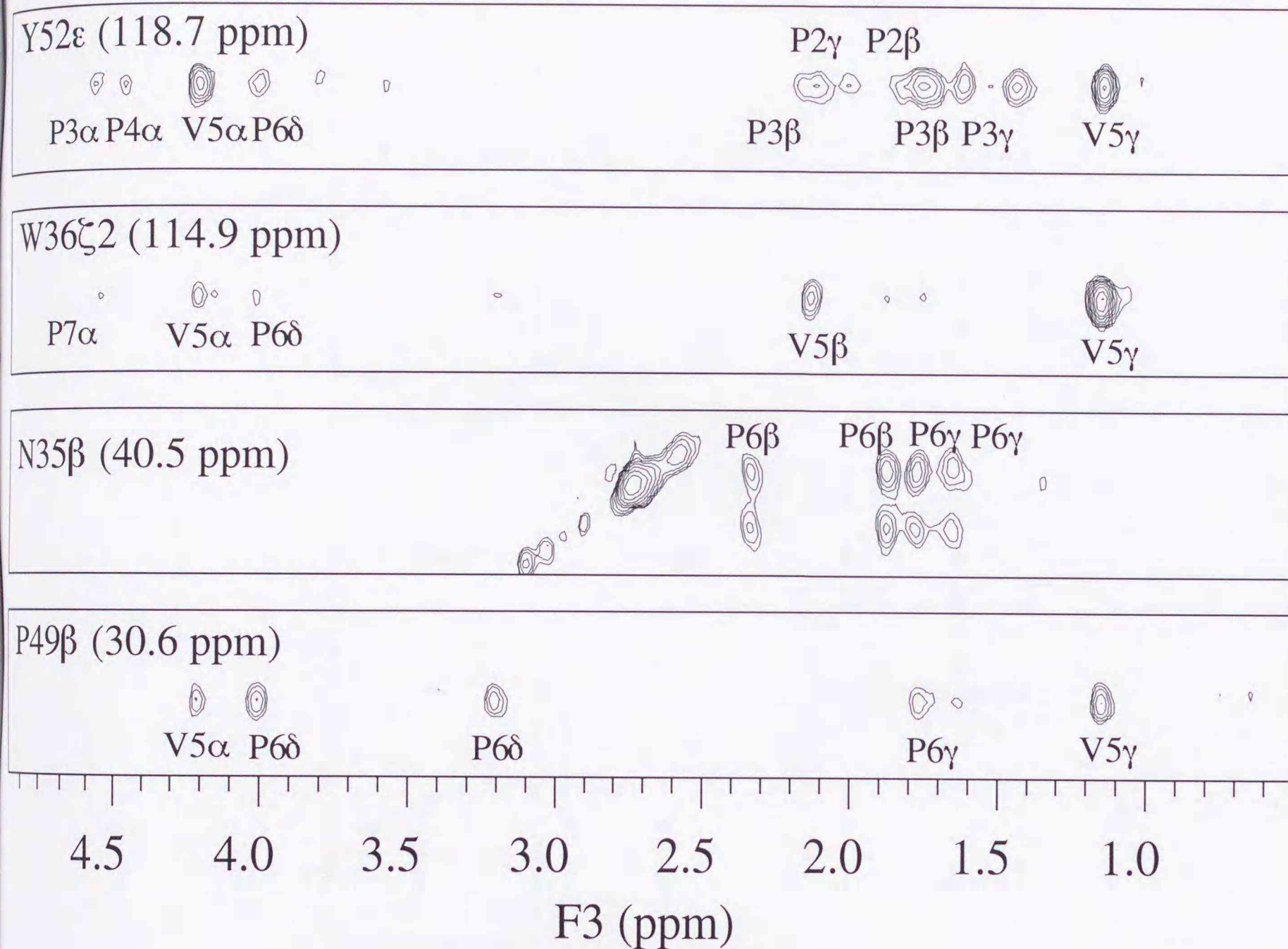


Figure 8 Strip plots of the 3D F1 ^{13}C edited-F3 ^{13}C filtered HSQC-NOESY spectrum of the 2 mM complex of the $^{13}\text{C}/^{15}\text{N}$ uniformly enriched N-SH3 protein and the unlabeled VPP peptide dissolved in $^2\text{H}_2\text{O}$. The spectrum was recorded on a UNITYplus 600 ^1H 600 MHz NMR spectrometer. The labeled peaks are assigned intermolecular NOE signals. The 3D spectrum was obtained with the pulse scheme in Fig.7(a). The NOE mixing delay was set to 150 ms. The spectrum was recorded with 32 increments in the F1 dimension, 84 increments in the F2 dimension, 416 complex points in the F3 dimension and 16 scans for each increment. Spectral width for each axis was 7334 Hz for F1 dimension, 7000 Hz for F2 dimension, and 7000 Hz for F3 dimension. The ^{13}C carrier frequency was set at 75.2 ppm. Total measuring time was 61 h.

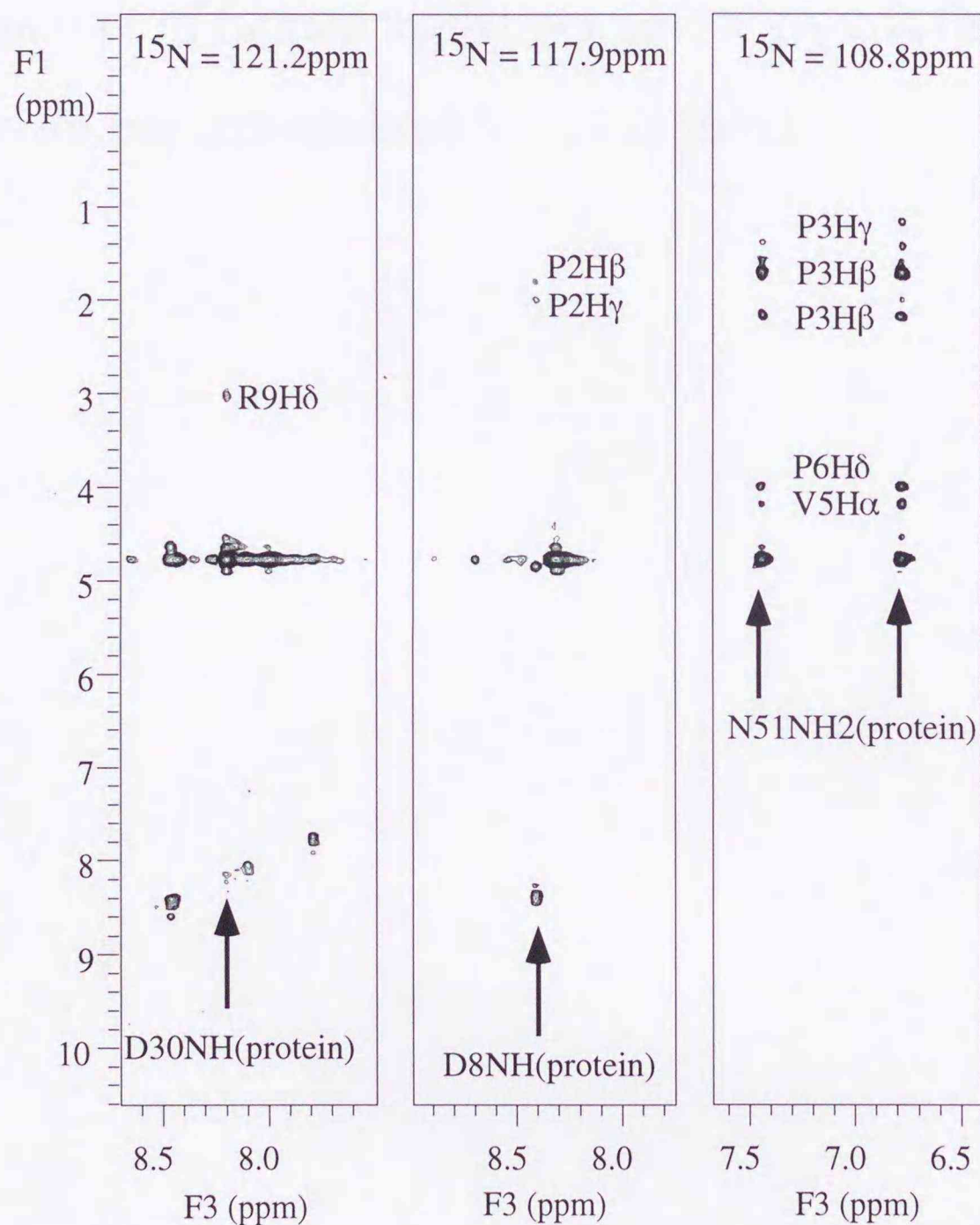


Figure 9 Strip plots of the 3D ^1H 600 MHz F1 $^{13}\text{C}/^{15}\text{N}$ -filtered F3 ^{15}N -edited NOESY-HSQC spectrum, in which the peaks were assigned to intermolecular NOE signals. The spectrum was obtained with the pulse scheme of Figure 7(b). The NOE mixing time was set to 150 ms. The spectrum was recorded with 128 increments in the F1 dimension, 32 increments in the F2 dimension, 416 complex points in the F3 dimension and 8 scans for each increment. The spectral widths for each axis were 7000 Hz for F1, 2100 Hz for F2, and 7000 Hz for F3. The ^{15}N carrier frequency was set at 118.0 ppm. Total measuring time was 57.5 hours.

Chapter 3

Solution structure of human acidic fibroblast growth factor and interaction with heparin-derived hexasaccharide

Abstract

Fibroblast growth factors (FGFs) bind to extracellular matrices, especially heparin-like carbohydrates of heparansulfate proteoglycans which stabilizes FGFs to protect against inactivation by heat, acid, proteolysis and oxidation. Moreover, binding of FGFs to cell surface proteoglycans promotes to form oligomers, which is essential for receptor oligomerization and activation.

In the present study, the author determined the solution structure of acidic FGF using a series of the triple resonance multi-dimensional NMR experiments and simulated annealing calculations. Furthermore, the author prepared the sample complexed with a heparin-derived hexasaccharide which is a minimum unit for aFGF binding. From the chemical shift differences between free aFGF and aFGF-heparin complex, the author concluded that the major heparin binding site was located on the region of 110-131 and 17-21. The binding sites are quite similar to those observed for bFGF-heparin hexasaccharide complex, showing that both FGFs recognize heparin-oligosaccharides in a similar manner.

Introduction

Acidic fibroblast growth factor (aFGF) and basic fibroblast growth factor (bFGF) are potent mitogenic polypeptides which play fundamental roles in regulation of cell proliferation, differentiation and promote wound repair and angiogenesis (Burges and Maciag, 1989; Davidson et al., 1985; Gospodarowicz, 1987; McGee et al., 1988; Nabel et al., 1993; O'Keefe et al., 1988; Shipley et al., 1989). FGFs functionally bind to extracellular matrices, especially heparin-like carbohydrates of heparan sulfate proteoglycans (Ruoslahti and Yamaguchi, 1991, Lindhal et al., 1994). Binding to heparin or heparin-like carbohydrates stabilizes FGFs and diminishes inactivation by heat, acid, proteolysis and mild oxidation *in vitro* (Gospodarowicz & Cheng, 1986; Saksela et al., 1988; Sommer and Rifkin, 1989). Therefore, heparin binding may provide storage and regulated release mechanism for FGFs from cell surfaces (Flaumenhaft et al., 1990). Furthermore, binding of FGFs to cell surface heparan sulfate proteoglycans promotes to form oligomers (Spivak-Kroizman et al., 1994), which is essential for receptor binding and activation (Yayon et al., 1991, Spivak-Kroizman et al., 1994).

Structural studies on free FGFs and the complexes with anti-ulcer agents such as inositol hexasulfate and sucrose octasulfate have been reported (Zhu et al., 1991; Zhu et al., 1993; Pineda-Lucena et al., 1994; Moy et al., 1995; Moy et al., 1996; Blaber et al., 1996). Zhu et al. (1991) initially reported the three-dimensional structures of free aFGF and bFGF using X-ray crystallography. Both molecules consist of 12 β -strands, and overall structures form typical β -trefoil structure like IL-1 β . Subsequently, Zhu et al. (1993) determined the structure of aFGF complexed with sucrose-octasulfate. While, Pineda-Lucena et al. (1994) reported the low resolution solution structure of aFGF complexed with inositol-hexasulfate using ^1H -

NMR spectroscopy. These studies concluded that drugs bound to a positively charged region of aFGF, largely composed of residues Lys112-Lys128. High-resolution solution structure of bFGF was determined by multidimensional heteronuclear magnetic resonance spectroscopy (Moy et al., 1995, 1996), which showed that the solution structure of free bFGF was quite similar to that of the crystalline state. However, there was no apparent β XI strand which was detected in the crystal structure. Most recently, the crystal structures of bFGF complexed with heparin-derived tetra- and hexa-oligosaccharides have been reported (Faham et al., 1996). Heparin formed a helical structure and bound to the residues Asn27, Arg120, Lys125, and Gln134 of bFGF. The author aligned the sequence of both FGFs based on the structural basis (Figure 1). The author also show the residues responsible for heparin binding, receptor binding and central cavity formation in magenta, cyan and green, respectively. Here, we took considerations on the structural studies together with mutagenesis and alanine scanning experiments (Faham et al., 1996; Blaber et al., 1996; Thompson et al., 1994). Most of these residues are well conserved among FGF family proteins, suggesting that FGFs bind to heparin-like oligosaccharides and receptors in a similar manner.

In the present study, the author determined the solution structure of free aFGF using multi-dimensional triple resonance NMR spectroscopy. From the analysis of the chemical shift perturbation upon binding of heparin-derived hexasaccharide, the author mapped the heparin-binding site on aFGF and discussed the recognition mechanism of heparin by aFGF.

Materials and methods

aFGF Preparation

A polypeptide comprising amino acids 7 - 140 of human aFGF was expressed in *Escherichia coli* JM103 bearing the recombinant plasmid pMJ26. Uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled aFGF was expressed in M9 medium containing 1g/L $^{15}\text{NH}_4\text{Cl}$ and 2g/L ^{13}C -glucose as nitrogen and carbon sources supplemented with 1g/L $^{13}\text{C}/^{15}\text{N}$ -labeled amino acid mixture from alga (Celltone-CNTM; Martek, Columbia, MD). Uniformly ^{15}N -labeled aFGF was also prepared in M9 medium containing 1g/L $^{15}\text{NH}_4\text{Cl}$ and 1g/L ^{15}N -labeled amino acid mixture (Celltone-NTM), while unlabeled aFGF was expressed in LB medium. The lysate was applied to heparin-Sepharose (Pharmacia, Upssala) column. aFGF was eluted at 1.5M NaCl concentration and was further purified as described previously (Jaye et al., 1987; Tate et al., 1992). To the sample solution of aFGF, $(\text{NH}_4)_2\text{SO}_4$ was added to a concentration of 100 mM, which prevented the protein from precipitation.

NMR Sample Preparation

The NMR sample solution was concentrated and solvent-exchanged by a Centriprep 10 (Amicon, Beverly, MA) concentrator and was finally prepared in 100 mM $(\text{NH}_4)_2\text{SO}_4$ and 25 mM sodium phosphate buffer in 90% $\text{H}_2\text{O}/10\%$ $^2\text{H}_2\text{O}$ containing 1.5 mM aFGF. Sample volume was 230 μL in a restricted-volume Shigemi NMR tube (Shigemi, Tokyo). pH of the protein solutions was adjusted to 6.6 by adding 0.1N HCl or NaOH solution. pH meter readings were not corrected for isotope effects.

Heparin-Derived Hexasaccharide Preparation

Heparin oligosaccharides were prepared by partial deaminative cleavage of heparin as described previously (Thunberg et al., 1982). These oligosaccharides were subjected to a column of Sephadex G-50 (1.2 x 127 cm) equilibrated with 0.2 M ammonium acetate. The hexasaccharide fraction was pooled and purified by chromatography on the same column. The fraction was applied to a column of aFGF-conjugated Sepharose (16 x 35 mm, 7 mL), which had been prepared as described previously (Habuchi et al., 1992) and equilibrated with 10 mM Tris HCl, pH 7.2 and 5 mM dithiothreitol (Solution A) containing 0.15 M NaCl at 4 °C. The column was washed with 35 mL of solution A containing 0.2 M NaCl and then eluted with 28 mL of Solution A containing 1.5 M NaCl. The eluate was desalted using PD 10 column (Pharmacia, Upssala,). The compositional analysis of the bound hexasaccharide was performed as described previously (Habuchi et al., 1992 ; Bienkowski and Conrad, 1985).

NMR Spectroscopy

All NMR experiments were performed at 30 °C on a Varian UNITY 500 NMR spectrometer equipped with a triple resonance 5 mm probe with a self-shielded z-axis gradient coil. Since the author used a single sample dissolved in 90% H₂O/10% ²H₂O for all NMR experiments, efficient water suppression was essential and the author incorporated several water suppression schemes in the pulse sequences as described in the followings.

For homonuclear two-dimensional experiments, water suppression was achieved using the solvent presaturation pulses during the relaxation delay, while water signal in heteronuclear experiments was suppressed either with a combination

of presaturation pulses during the relaxation delay and pulsed field gradients as a zz-filter (Bax and Pochapsky, 1992) or with the gradient enhanced HSQC pulse scheme (Kay et al, 1992).

Homonuclear two-dimensional DQF-COSY and TOCSY (Cavanagh et al., 1996) experiments were acquired on unlabeled aFGF dissolved in H₂O solution. In order to observe the aromatic proton signals with suppression of the overlapping amide proton signals in ¹⁵N-labeled protein, the author measured the ¹⁵N-filtered NOESY spectrum (Ogura et al., 1996a).

Two-dimensional ¹H-¹⁵N HSQC spectrum was acquired on the uniformly ¹⁵N-labeled protein. Two-dimensional ¹H-¹³C CT-HSQC (Cavanagh et al., 1996) spectra of the uniformly ¹³C/¹⁵N-labeled protein were acquired.

The following three-dimensional spectra were acquired; HNCO, HNCA, HN(CO)CA, CBCANH, CBCA(CO)NH, HN(CA)HA, HBHA(CO)NH, C(CO)NH, H(CCO)NH and HCCH-TOCSY (Cavanagh et al., 1996). For ¹³C isotropic mixing in the HCCH-TOCSY, C(CO)NH and H(CCO)NH experiments, 18 ms DIPSI-3 pulse sequences (Shaka et al., 1988) were employed.

Three-dimensional ¹³C-edited NOESY (Cavanagh et al., 1996) spectrum of uniformly ¹³C/¹⁵N-labeled protein was acquired as described previously by Ogura et al. (1996b) with a mixing time of 80 ms. Three-dimensional ¹⁵N-edited NOESY (Cavanagh et al., 1996) spectrum was acquired on uniformly ¹³C/¹⁵N labeled protein with a mixing time of 80 ms. Three-dimensional HNHA (Vuister and Bax, 1993) spectrum was acquired on uniformly ¹⁵N labeled protein. In Table 1, the author summarized acquisition parameters for the present NMR experiments.

Acquired homonuclear two-dimensional FID data were processed with a standard VNMR (Varian Instruments, Palo Alto, CA) software package. Other heteronuclear and three-dimensional FID data were Fourier-transformed with the NMRPipe (Delaglio et al., 1995) software package on a SUN SPARC Station 20 computer. The transformed spectra were converted to FELIX 95.0 (Molecular Simulations, San Diego, CA) format with an in-house format-converting software. The resonances were picked using an in-house shoulder peak picking program on the FELIX 95.0 format. Spectral browsing were carried out with the FELIX 95.0 software on the SUN SPARC Station 20 computer.

Structure Calculations

Structure calculations for free aFGF were made after the method of Hatanaka et al. (1994) on a Silicon Graphics Indigo² Extreme computer. Interproton distance constraints were derived from NOE cross peak intensities (peak height) in the NOESY spectra. The peak intensities were translated into distance constraints based on the relation of NOE intensity $\propto$ (distance)⁻⁶ and a standard distance of sequential $d_{\alpha N}$ in β -sheet = 2.2 Å (Wüthrich, 1986). The three-dimensional structures were calculated with XPLOR ver. 3.1 (Brünger, 1993) using YASAP. The final 20 converged structures were selected from 200 runs of calculations on the basis of agreement with the experimental data. A mean structure was obtained by averaging the coordinates of the structures that were superimposed in advance on the best converged structure and then minimized under the constraints.

Analysis of the Calculated Structures

For quantitative assessment of the convergence of the calculations, root-mean-square (rms), number of NOE violations, and the values of F_{repel} , F_{NOE} , and F_{cdih}

were examined. An NOE violation is defined as a deviation of the calculated interproton distance from the experimental distance constraint. For quantitative comparisons of the different structures, minimum rms distances (rmsd) were calculated for the specified atoms for all the residues. A mean structure was obtained by averaging the coordinates of the structures that were superimposed to the best converged structure with a minimum pairwise rmsd of the backbone atoms (N, C α , C').

Results and discussion

aFGF is thermolabile and is precipitated even at room temperature. But in the presence of 100 mM $(\text{NH}_4)_2\text{SO}_4$, aFGF was found to be stable and provided a high quality of NMR spectrum as revealed in the ^1H - ^{15}N HSQC spectrum (Figure 2). Thus, 1.5 mM aFGF sample dissolved in 100 mM $(\text{NH}_4)_2\text{SO}_4$ and 25 mM sodium phosphate buffer (90% H_2O /10% $^2\text{H}_2\text{O}$) was used for all NMR measurements. The author carefully designed the pulse sequences to optimize the efficiency of water suppression so that we could measure all the spectra in 90% H_2O /10% $^2\text{H}_2\text{O}$ solution.

Assignments of resonances

The sequential resonance assignments of free aFGF were achieved based on the analysis of a suite of the HNCA, HN(CO)CA, CBCANH and CBCA(CO)NH spectra. All the main-chain resonances were assigned except for N-terminal Asn7 and Glu81 (Figure 2). Other aliphatic ^1H and ^{13}C resonances were assigned using the HBHA(CO)NH, HN(CA)HA, C(CO)NH, H(CCO)NH, ^1H - ^{13}C CT-HSQC and HCCH-TOCSY spectra.

The spin systems of the aromatic side-chain proton resonances were initially identified with the conventional two-dimensional DQF-COSY and TOCSY spectra. The ^{15}N -filtered NOESY spectrum on ^{15}N -labeled protein in 90 % H_2O /10 % $^2\text{H}_2\text{O}$ provided the correlation between aromatic protons and aliphatic protons without overlap of the amide proton resonances, thus correlating the spin systems of

aromatic protons with H^β resonances. These aromatic resonance assignments were confirmed by the three-dimensional ^{13}C -edited NOESY and the two-dimensional ^1H - ^{13}C CT-HSQC spectra. The NH_2 groups of Gln and Asn residues were connected to their H^β and H^α protons with three-dimensional ^{15}N -edited NOESY spectrum.

The main-chain and side-chain ^1H , ^{13}C and ^{15}N resonance assignments are summarized in Table 2. The backbone resonances were assigned for all residues and side-chain ^1H , ^{13}C , and ^{15}N resonances were assigned for 99 % of all side-chain atoms. It should be emphasized that all the spectra were acquired in 90% H_2O /10% $^2\text{H}_2\text{O}$ solution, thus avoiding ambiguities in chemical shift values due to different solvent conditions.

Secondary structure of free aFGF

Once the sequential assignments were completed, NOE correlations from three-dimensional ^{15}N -edited NOESY and ^{13}C -edited NOESY spectra were analyzed to obtain interproton distance constraints. The analysis of sequential and medium-range NOEs are summarized in Figure 3. Qualitative amide-proton exchange rates were also measured using cross sections of amide proton-water exchange peaks taken from ^{15}N -edited NOESY at the F_1 chemical shift of the water (Pfuhl et al., 1995). Since this experiment was performed without solvent presaturation, the intensities of the amide proton-water exchange peaks can be qualitatively correlated with the exchange rates. Residues whose amide protons exchanged rapidly with solvent were identified by the presence of intense amide

proton-water cross peaks. Since these exchange peaks developed within the 80 ms mixing time, the exchange rates of these amide protons were estimated to be on the order of 10 s^{-1} . Therefore, these residues are exposed to the surface and do not participate in stable hydrogen bonds. Dihedral angle constraints for backbone ϕ angles were determined from $^3J_{\text{HNH}\alpha}$ coupling constants calculated from $\text{H}\alpha/\text{HN}$ intensity ratios in the 3D HNHA spectrum. Coupling constants greater than 8 Hz were used for dihedral angle constraints. Because the overlapped signals and the residues which have intermediate coupling constants were excluded, the total number of the dihedral angle constraints were 51 on the final stage of the structure calculation. From sequential NOE connectivities together with coupling constants and amide proton exchange information, the secondary structures of aFGF was identified (Figure 3). aFGF is a β -sheet rich protein containing eleven β -strands without α -helices. The Chemical shift index (CSI) of free aFGF was calculated for $\text{H}\alpha$, $\text{C}\alpha$, $\text{C}\beta$ and C' resonances (Wishart and Sykes, 1994) and its consensus was obtained. The consensus CSI (Figure 3) also indicated that aFGF consists of eleven β -strands connected by turns and had no apparent α -helices. Although the positions of the β -strands were shifted, truncated or elongated by one or two residues, the predicted secondary structure of aFGF from CSI was excellent agreement with the secondary structure obtained by present NMR experiments. The secondary structure and β -sheet topology of free aFGF were subsequently derived from the information summarized in Figure 4 together with inter-strand NOEs. The triple-repeat of β -sheets is remarkable. According to the crystal structure (Zhu et al., 1991), aFGF consists of the following twelve β -strands; I (11-17), II (21-26), III (30-34), IV (42-48), V (53-59), VI (63-68), VII (72-76), VIII (83-89), IX (94-100), X (107-112), XI (116-120), and XII (130-136). The ranges of β -strands in the NMR structure are

almost identical with the X-ray structure. However the residues 116 - 120 corresponding to β XI do not have apparent evidence for β -strand formation in NMR structure. Thus, the author concluded that aFGF consisted of eleven β -strands in solution; I (11-17), II (21-26), III (30-34), IV (43-48), V (53-58), VI (63- 68), VII (72-76), VIII (84-89), IX (94-99), X (107-112) and XII (131-135) joined by turns (Figure 4). Moy et al. (1995) reported that bFGF also had no apparent β XI from the analysis of the NOESY spectra. Thus, in solution, secondary structures of both aFGF and bFGF are quite similar with each other.

Tertiary Structure of free aFGF

The remaining assignments of the three-dimensional ^{15}N -edited and ^{13}C -edited NOESY spectra were obtained using an in-house programs for iterative cycles of NOESY assignments combined with structural calculations (Hatanaka et al., 1994). The structural calculations were based on 51 dihedral constraints and 1368 distance constraints which consist of 594 intraresidue and 774 interresidue containing 284 sequential ($|i - j| = 1$), 122 medium range ($2 \leq |i - j| \leq 5$), and 368 long range ($|i - j| > 5$) NOEs. A summary of the structural statistics for a set of the final structures and for the mean structure is shown in Table 3.

Figure 5 shows the overlay of the final 20 structures of aFGF obtained out of 200 runs of calculations (Figure 5a) and their mean structure (Figure 5b). In the solution structure, although β -strand regions were well converged, some loop regions were highly flexible. Figure 6 shows the number of distance constraints and the rmsd for the backbone atoms (N, C^α , C') for all residues. The residues in β -strand region have many long range constraints, and the rmsd values were 0.4 - 0.6 Å for the backbone atoms. On the other hand, the residues in the loop region have less long

range constraints, and the rmsd values were more than 1.0 Å for the backbone atoms. Especially, two loop regions around Gly20 and Glu60 were extremely large rmsd values more than 1.5 Å. Overall estimated rmsd values of the 20 final structures from the minimized averaged structure were 0.58 ± 0.08 Å for the backbone atoms (N, C α , C') and 1.09 ± 0.09 Å for the nonhydrogen atoms of residues 11 - 136 except for highly flexible loop regions of residues 17 - 20, 50 - 52, 59 - 62, 91 - 93, 100 - 106 and 113 - 130.

Figure 7 shows the Ramachandran plot for all residues in the minimized averaged solution structure of free aFGF. aFGF comprises 114 non-glycine and non-proline residues. The $\phi\psi$ -plot indicates 60 residues (52.6 %) in most favored regions, 44 residues (38.6 %) in additional allowed regions, and 9 residues (7.9 %) in generously allowed regions. Although Ser139 was in disallowed region, this residue is located on the C-terminal flexible tail. Figure 7 also shows the Ramachandran plots for four residues corresponding to β XI identified in the X-ray crystal structure of aFGF (Zhu et al., 1991). Although the NOE-based analysis (Figure 3) showed that β XI does not have apparent evidence for β -strand-formation (i.e. large H^N - H^α coupling constants, intense $H^\alpha(i) - H^N(i+1)$ NOE signals, and CSI), the ϕ and ψ maps for the solution and the crystalline structures of aFGF and the solution structure of bFGF supported that this region took β -strand-like structure.

Figure 8 shows an overlay plot of the crystal and solution structures of aFGF. The rmsd difference between the solution structure and the X-ray structure (Zhu et al., 1991) was 1.61 Å for the overall backbone atoms excluded highly flexible loop regions. aFGF takes a β -trefoil structure similar to the solution structure of bFGF and the crystal structures of aFGF and bFGF. The cavity in the center of the β -trefoil structure was surrounded by hydrophobic residues composed of Leu 14, Leu 23, Leu

44, Ile 56, Leu 65, Phe85, Tyr 97, Val 109 and Phe132 (Figure 5c), where the overlay of 20 structures of the side chains was shown on the main chain of aFGF. These residues were located with apparent three-fold axis symmetry. The central hydrophobic cavity is characteristic to the β -trefoil proteins and is found in the crystal structure of aFGF, bFGF and IL-1 β (Zhu et al., 1991; Blaber et al., 1996). The corresponding hydrophobic residues are conserved or type-conserved in the FGF family proteins (Figure 1), suggesting that the hydrophobic interaction in the cavity plays an important role for maintaining the overall structure.

Binding Site of Heparin-Derived Hexasaccharide

To identify the heparin-binding site on aFGF, the author studied the chemical shift perturbation of NMR resonances upon binding of heparin-derived hexasaccharide (Hep6). Hep6 is known to be a minimum aFGF binding unit without receptor activation. Upon addition of Hep6 to aFGF solution, the signal intensities of aFGF in ^1H - ^{15}N HSQC spectrum were significantly reduced, suggesting that aFGF was oligomerized upon complex formation with Hep6. Oligomerization of heparin-binding proteins on the heparin chain is general characteristics (Spivak-Kroizman et al., 1994; Iwasaki et al., 1997). In the case of bFGF, Moy et al. (1997) reported that bFGF bound to heparin-tetrasaccharide in dimer with trans configuration and bound to decasaccharide in tetramer with trans and cis configurations. The sequential assignments of the back-bone resonances of the Hep6-aFGF complex were completed using the series of the triple resonance experiments (data not shown). Figure 9 shows the chemical shift perturbation of H^{N} , N, C' and C^{α} resonances for each residue between free aFGF and the Hep6-aFGF complex. Because overall signals were extremely broadened in the Hep6-aFGF complex, the author could not connect the main chain assignments to those of side chains. It should be noted that

the residues with appreciable chemical shift perturbations in slow exchange condition were localized in some specific regions, supporting that aFGF did not show any significant overall conformation change upon complex formation. The main-chain amide proton resonance of Lys113 and the main-chain amide ^{15}N resonance of Lys118 were significantly perturbed by 0.72 ppm and 1.5 ppm, respectively. The following residues shifted more than 50 Hz in the sum of four (H^{N} , N, C' and $\text{C}\alpha$) resonances upon Hep6-binding; Ser17, Gly20, His21, Leu73, Glu91, Asn92, Gly110 - Gly120, Arg122 - Ala129 and Leu131. Amide proton and nitrogen resonances of Lys128 were not identified in Hep6-aFGF sample possibly due to line broadening. These residues were plotted in cyan on the main chain of aFGF which were localized on the specific surface (Figure 10). Although the magnitude of the perturbation is not directly corresponding to the measure of interaction, the author can expect that these residues are located in or in the vicinity of the contact region with Hep6. Extensive mutagenic studies for bFGF demonstrated that Asn27, Arg120, Lys125 and Gln134 are major residues responsible for heparin binding. Moreover, there is additional residues responsible for weak binding, comprising Lys26, Arg81, Lys119, Thr121, Gln123, Lys129 and Lys135. It should be noted that amino acid residues from Lys119 through Lys135 located in the C-terminal domain of bFGF collectively contribute nearly 80% of the binding free energy (Thompson et al., 1994). Recent X-ray crystal structure of bFGF complexed with heparin oligosaccharide (Faham et al., 1996) demonstrated that bFGF contained high affinity heparin binding site; Asn27, Arg120, Lys125 and Gln134 and low affinity heparin binding site; Lys26, Asn101 and Lys135 (Table 1). The amino acid residues with appreciable chemical shift perturbations plotted on the main chain of aFGF structure (Figure 10) are in good agreement with those suggested by the bFGF mutagenic studies (Thompson et al., 1994) and revealed by the structure of bFGF-heparin oligosaccharide complex. On

the main chain, the author plotted the C^α location of Asn18, Asn92, Lys113, Lys118, Gln127 and Lys128 which are counterpart residues responsible for heparin binding in bFGF (Figure 10). The locations of these residues are quite similar to those in bFGF, suggesting that aFGF and bFGF recognize heparin hexasaccharide in a similar manner. It is to be noted that all these residues are located on the highly flexible regions (Figure 6) so that the side chains adapt their locations for proper bindings to heparin oligosaccharides. Although the β_{XI} strand (116 - 120) is absent in solution structure of free aFGF, the C^α and C' chemical shift values for 116-120 in the complex have a preference for β -strand formation. In the crystalline structure of the bFGF-heparin oligosaccharide complex, β_{XI} is formed and Lys125 located on β_{XI} can interact with a heparin sulfate group (Faham et al., 1996). Since Lys 125 (Lys 118 in aFGF) is responsible for high affinity binding to heparin, the β_{XI} -strand formation upon heparin binding may be crucial for exertion of the biological function of FGF family proteins.

References

- Bax, A. and Pochapsky, S.S. (1992) *J. Magn. Reson.* 99, 638-643.
- Bienkowski, M.J. and Conrad, H.E. (1985) *J. Biol. Chem.* 260, 356-365.
- Blaber, M., DiSalvo, J. and Thomas, K.A. (1996) *Biochemistry* 35, 2086-2094.
- Brünger, A.T. (1993) *X-PLOR Version 3.1: A System for X-Ray Crystallography and NMR*, Yale University Press, New Haven
- Burgess, W.H., and Maciag, T. (1989) *Annu. Rev. Biochem.* 58, 575-606.
- Cavanagh, J., Fairbrother, W.J., Palmer III, A.G., and Skelton, N.J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego
- Davidson, J.M., Klagsburn, M., Hill, K.E., Buckley, A., Sullivan, R., Brewer, P.S., and Woodward, S.C. (1985) *J. Cell Biol.* 100, 1219-1227.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR* 6, 277-293.
- Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J. and Rees, D.C. (1996) *Science* 271, 1116-1120.
- Flaumenhaft, R., Moscatelli, D. and Rifkin, D.B. (1990) *J. Cell. Biol.* 111, 1651-1659.
- Gospodarowicz, D. and Cheng, J. (1986) *J. Cell. Physiol.* 128, 475-484.
- Gospodarowicz, D. (1987) *Methods Enzymol.* 147, 106-119.
- Habuchi, H., Suzuki, S., Saito, T., Tamura, T., Harada, T., Yoshida, K. and Kimata, K. (1992) *Biochem. J.* 285, 805-813.
- Hatanaka, H., Oka, M., Kohda, D., Tate, S., Suda, A., Tamiya, N. and Inagaki, F. (1994) *J. Mol. Biol.* 240, 155-166.
- Iwasaki, W., Nagata, K., Hatanaka, H., Inui, T., Kimura, T., Muramatsu, T., Yoshida, K., Tasumi, M., and Inagaki, F. (1997) *EMBO J.* 16, 6936-6946.
- Jaye, M., Burgess, W.H., Shaw, A.B. and Drohan, W.N. (1987) *J. Biol. Chem.* 262,

16612-16617.

- Kay, L.E., Keifer, P. and Saarinen, T. (1992) *J. Am. Chem. Soc.* 114, 10663-10665.
- Kraulis, P.J. (1991) *J. Appl. Cryst.* 24, 946-950.
- Laskowski, R.A., Rullmann, J.A., MacArther, M.W., Kaptein, R. and Thornton, J.M. (1996) *J. Biomol. NMR* 8, 477-486.
- Lindhal, U., Lindholt, K., Spillmann, D. and Kjellen, L. (1994) *Thromb. Res.* 75, 1-32.
- McGee, G.S., Davidson, J.M., Buckley, A., Sommer, A., Woodward, S.C., Aquino, A.M., Barbour, R., and Demetriou, A.A. (1988) *J. Surg. Res.* 45, 145-153.
- Moy, F.J., Seddon, A.P., Campbell, E.B., Böhlen, P. and Powers, R. (1995) *J. Biomol. NMR* 6, 245-254.
- Moy, F.J., Seddon, A.P., Bohlen, P. and Powers, R. (1996) *Biochemistry* 35, 13552-13561.
- Moy, F.J., Safran, M., Seddon, A.P., Kitchen, D., Böhlen, P., Aviezer, D., Yayon, A. and Powers, R. (1997) *Biochemistry* 36, 4782-4791.
- Nabel, E.G., Yang, Z-y., Plautz, G., Forough, R., Zhan, X., Haudenschild, C.C., Maciag, T., Nabel, G.J. (1993) *Nature* 362, 844-846.
- Ogura, K., Terasawa, H. and Inagaki, F (1996a) *J. Biomol. NMR* 8, 492-498.
- Ogura, K., Terasawa, H. and Inagaki, F (1996b) *J. Magn. Reson. B* 112, 63-68.
- O'Keefe, E.J., Chiu, M.L., and Payner, R.E.Jr. (1988) *J. Invest. Dermatol.* 90, 767-769.
- Pfuhl, M., Gautel, M., Politou, A.S., Joseph, C. and Pastore, A. (1995) *J. Biomol. NMR* 5, 48-58.
- Pineda-Lucena, A., Jiménez, M.A., Nieto, J.L, Santoro, J., Rico, M. and Giménez-Gallego, G. (1994) *J. Mol. Biol.* 242, 81-98.
- Ruoslahti, E. and Yamaguchi, Y. (1991) *Cell* 64, 867-869.
- Saksela, O., Moscatelli, D., Sommer, A. and Rifkin, D.B. (1988) *J. Cell. Biol.* 107, 743-751.

- Shaka,A.J., Lee,C.J., and Pines,A. (1988) *J. Magn. Reson.* 77, 274-293.
- Shipley,G.D., Keeble,W.W., Hendrickwson,J.E., Coffey,R.J.Jr., and Pittelkow,M.R.
(1989) *J. Cell. Physiol.* 138, 511-518.
- Sommer,A., and Rifkin,D.B. (1989) *J. Cell Physiol.* 138, 215-220.
- Spivak-Kroizman,T., Lemmon,M.A., Dikic,I., Ladbury,J.E., Pinchasi,D., Huang,J.,
Jaye,M., Crumley,G., Schlessinger,J. and Lax,I. (1994) *Cell* 79, 1015-1024.
- Tate,S., Tate,N.U., Ravera,M.W., Jaye,M. and Inagaki,F. (1992) *FEBS. Lett.* 297,
39-42.
- Thompson,L.D., Pantoliano,M.W. and Springer,B.A. (1994) *Biochemistry* 33,
3831-3840.
- Thunberg,L., Backstrom,G., and Lindhal,U. (1982) *Carbohydr. Res.* 100, 393-410.
- Vuister,G.W. and Bax,A. (1993) *J. Am. Chem. Soc.* 115, 7772-7777.
- Wishart,D.S. and Sykes,B.D. (1994) *J. Biomol. NMR* 4, 171-180.
- Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York
- Yayon,A., Klagsbrun,M., Esko,J.D., Leder,P. and Ornitz,D.M. (1991) *Cell* 64,
841-848.
- Zhu,X., Komiya,H., Chirino,A., Faham,S., Fox,G.M., Arakawa,T., Hsu,B.T. and
Rees,D.C. (1991) *Science* 251, 90-93.
- Zhu,X., Hsu,B.T. and Rees,D.C. (1993) *Structure* 1, 27-34

Table 1. Experiments for free aFGF resonance assignments and structure calculation

experiment	carrier (¹³ C, ppm)	carrier (¹⁵ N, ppm)	sw (t ₁ , ppm)	sw (t ₂ , ppm)	np (t ₁)	np (t ₂)	scans	time (h)
2D DQF-COSY			14.0		512		32	20
2D TOCSY			14.0		400		16	8
2D ¹⁵ N filtered NOESY		117.7	14.0		512		32	20
2D ¹ H- ¹⁵ N HSQC		117.7	36		256		4	0.6
3D HNCO	177.1	117.7	13.5	36	50	32	8	16
3D HN(CO)CA	55.3	117.7	28.7	36	64	32	8	20
3D HNCA	55.3	117.7	28.7	36	64	32	8	20
3D CBCA(CO)NH	41.0	117.7	70.0	36	58	32	16	36
3D CBCANH	41.0	117.7	70.0	36	58	32	32	71
3D HN(CA)HA	55.3	117.7	4.0	36	36	32	16	22
3D HBHA(CO)NH	41.0	117.7	7.0	36	64	32	16	38
3D C(CO)NH	41.0	117.7	70.0	36	64	32	32	72
3D H(CCO)NH	41.0	117.7	7.0	36	64	32	32	72
3D HCCH-TOCSY	41.0		7.0	23.3	150	32	8	47
2D ¹ H- ¹³ C CT-HSQC _{aliphatic}	41.0		70.0		200		32	4
2D ¹ H- ¹³ C CT-HSQC _{aromatic}	124.0		31.8		50		64	2
3D ¹⁵ N-edited NOESY		117.7	14.0	36	150	32	8	51
3D ¹³ C-edited NOESY	75.2		14.0	48	128	32	8	43
3D HNHA		117.7	9.0	36	64	32	32	79

sw: spectral width; np: number of point (complex); time: total measuring time;

Table 2 : ^{15}N , ^{13}C , and ^1H resonance assignments for free aFGF (7 - 140) at pH 6.6 and 30 °C

residue	N	C α	C β	C'	Others
N7		52.7 (4.80)	39.2 (2.88, 2.79)	175.0	N δ 114.3 (8.27, 7.88)
Y8	121.9 (8.58)	58.1 (4.67)	38.8 (3.40, 2.87)	176.2	C δ 132.8 (7.08), C ϵ 118.2 (6.87)
K9	120.5 (8.37)	57.5 (4.29)	33.0 (1.94, 1.94)	176.8	C γ 25.1 (1.49), C δ 29.0 (1.74), C ϵ 42.1 (3.05)
K10	119.4 (8.03)	54.2 (4.72)	32.9 (1.79, 1.79)		C γ (1.55), C δ 28.0 (2.02), C ϵ 42.1 (3.03)
P11		63.0 (4.71)	32.9 (2.15, 1.77)	174.7	C γ 27.4 (2.03), C δ 50.7 (3.90, 3.66)
K12	119.1 (9.43)	53.5 (5.06)	37.0 (1.84, 1.48)	174.2	C γ 24.6 (1.46), C δ 29.4 (1.75), C ϵ 42.7 (3.02, 2.97)
L13	117.3 (9.03)	52.8 (4.91)	43.7 (1.80, 1.61)	177.0	C γ 27.1 (1.54), C δ 25.7 (0.50), 24.0 (0.60)
L14	125.2 (10.43)	54.2 (4.99)	42.7 (1.66, 1.37)	174.7	C γ 27.9 (1.45), C δ 24.4 (0.40), 23.2 (0.64)
Y15	125.7 (8.60)	57.1 (4.58)	40.2 (3.42, 2.92)	173.6	C δ 133.8 (6.40), C ϵ 117.5 (6.54)
C16	129.9 (9.34)	57.8 (4.17)	29.2 (2.57, 2.57)	175.0	
S17	122.4 (8.26)	60.1 (3.89)	64.2 (3.89, 3.89)	174.5	
N18	119.1 (7.46)	54.7 (4.38)	37.9 (3.09, 2.98)		N δ 112.9 (7.74, 6.99)
G19	104.9 (8.48)	44.2 (4.28, 3.91)		174.2	
G20	109.3 (7.23)	46.5 (3.60, 3.19)		172.7	
H21	113.7 (6.41)	56.7 (4.08)	34.2 (2.62, 2.48)	173.6	C δ 2 117.5 (6.53), C ϵ 1 139.0 (7.65)
F22	120.4 (9.37)	56.6 (5.39)	40.0 (3.54, 2.95)	176.3	C δ 132.5 (7.25), C ϵ 131.1 (7.27), C ζ 128.9 (7.17)
L23	124.8 (8.54)	57.0 (4.34)	43.6 (1.88, 1.35)	174.1	C γ 27.1 (1.48), C δ 25.7 (0.43), 24.2 (0.66)
R24	126.8 (9.43)	55.4 (4.71)	35.0 (2.53, 1.28)	173.6	C γ 26.2 (1.22), C δ 45.6 (3.35)
I25	120.8 (7.05)	60.1 (4.56)	39.0 (1.68)	176.1	C γ 26.4 (1.51, 0.76), C δ 13.9 (0.56), C γ^m 17.1 (0.60)
L26	127.7 (9.00)	54.0 (4.14)	40.9 (2.01, 1.43)		C γ 28.2 (1.42), C δ 25.0 (0.79), 23.4 (0.67)
P27		65.1 (4.26)	31.8 (2.32, 1.96)	176.3	C γ 27.8 (1.78), C δ 49.4 (3.43)
D28	113.9 (7.49)	53.5 (4.52)	39.7 (3.09, 2.58)	176.9	
G29	109.2 (8.51)	45.5 (4.60, 3.89)		175.0	
T30	116.8 (7.96)	63.6 (4.47)	69.7 (4.19)	172.5	C γ 21.4 (1.15)
V31	123.0 (7.90)	60.5 (5.08)	34.1 (1.61)	174.0	C γ 21.4 (0.85), 21.4 (0.75)
D32	131.1 (9.02)	53.1 (4.30)	41.0 (3.01, 2.73)	172.6	
G33	102.3 (8.00)	44.3 (5.52, 3.01)		172.3	
T34	112.6 (8.63)	59.0 (5.19)	69.6 (3.99)	172.9	C γ 21.5 (1.34)
R35	127.2 (9.05)	57.2 (4.66)	29.9 (2.06, 1.73)	175.8	C γ 28.4 (1.20), C δ 43.3 (3.03)
D36	120.9 (8.43)	53.6 (4.65)	40.9 (3.22, 2.65)	175.4	
R37	126.0 (8.38)	57.1 (3.60)	30.7 (1.75, 1.75)	175.8	C γ 28.0 (1.34), C δ 43.4 (3.10)
S38	113.8 (8.50)	58.2 (4.39)	64.1 (4.01, 3.85)	174.0	
D39	124.8 (7.18)	55.3 (4.26)	43.8 (2.78, 2.70)	177.0	
Q40	125.2 (8.93)	57.6 (4.01)	28.9 (1.74, 1.52)	177.5	C γ 33.7 (2.02), N ϵ 112.1 (7.39, 6.75)
H41	117.6 (9.73)	58.9 (4.77)	28.7 (3.82, 2.94)	175.0	C δ 2 121.0 (7.03), C ϵ 1 137.4 (7.83)
I42	108.5 (6.79)	60.5 (5.12)	38.3 (2.41)	174.7	C γ 23.4 (1.56, 0.94), C δ 15.5 (0.54), C γ^m 17.4 (0.81)
Q43	119.6 (7.09)	57.2 (4.26)	27.6 (2.10, 1.92)	175.0	C γ 34.5 (2.30), N ϵ 110.2 (7.52, 6.72)
L44	127.6 (9.16)	53.6 (5.31)	44.3 (1.61, 1.01)	175.3	C γ 26.2 (1.94), C δ 26.6 (0.73), 21.8 (0.59)
Q45	120.7 (9.68)	54.1 (4.63)	31.6 (1.88, 1.78)	175.0	C γ 34.0 (2.15, 2.09), N ϵ 111.9 (7.42, 6.58)
L46	131.6 (9.07)	55.3 (5.52)	43.0 (1.65, 1.57)	176.9	C γ 30.6 (1.70), C δ 26.8 (0.87), 24.6 (0.77)
S47	115.5 (8.70)	57.2 (4.65)	65.2 (3.90, 3.83)	172.5	
A48	125.5 (8.44)	51.5 (5.03)	19.7 (1.29)	177.7	

E49	125.4 (8.55)	57.0 (4.42)	30.6 (2.02, 1.83)	176.5	CY 36.3 (2.22, 2.11)
S50	113.8 (8.13)	57.2 (4.54)	64.0 (3.81, 3.71)	173.4	
V51	121.2 (8.21)	65.1 (3.86)	31.3 (2.05)	177.9	CY 20.7 (1.00), 21.3 (1.09)
G52	115.5 (8.90)	45.9 (4.31, 3.98)		173.3	
E53	121.2 (8.04)	55.0 (5.43)	31.7 (1.83, 1.83)	175.8	CY 35.9 (2.08, 2.00)
V54	116.2 (9.64)	59.1 (5.49)	36.0 (2.30)	175.3	CY 21.8 (0.89), 19.7 (0.98)
Y55	117.6 (8.53)	56.7 (5.29)	41.5 (3.28, 2.99)	175.8	C δ 133.3 (7.15), C ϵ 117.5 (6.54)
I56	123.0 (10.64)	61.9 (4.33)	40.0 (1.62)	173.2	CY 27.3 (1.40, 0.53), C δ 12.0 (-0.20), CY ^m 18.7 (0.48)
K57	125.8 (8.65)	53.5 (4.97)	37.0 (1.30, 1.10)	176.1	CY 23.7 (0.66, 0.28), C δ 29.4 (1.13), C ϵ 41.8 (2.51)
S58	117.6 (9.01)	57.3 (4.66)	63.4 (3.88, 3.88)	177.4	
T59	123.2 (8.45)	65.2 (3.99)	67.9 (4.17)	176.1	CY 22.5 (1.15)
E60	121.9 (8.52)	58.2 (4.30)	30.7 (2.16, 1.87)	177.7	CY 35.0 (2.10)
T61	102.7 (7.80)	60.9 (4.69)	71.4 (4.54)	176.2	CY 21.9 (1.28)
G62	110.8 (7.77)	45.9 (4.14, 3.58)		173.0	
Q63	115.5 (7.24)	56.2 (4.30)	32.6 (1.67, 1.56)	174.4	CY 35.8 (2.27), N ϵ 106.6 (7.08, 6.47)
Y64	119.3 (9.37)	57.0 (4.83)	41.0 (3.16, 2.82)	176.1	C δ 132.8 (7.10), C ϵ 117.5 (6.54)
L65	123.5 (9.22)	55.6 (4.13)	42.7 (1.96, 1.03)	174.7	CY 26.4 (1.64), C δ 25.7 (0.89), 26.2 (0.19)
A66	126.9 (8.75)	50.7 (5.12)	23.5 (0.60)	175.0	
M67	116.8 (7.63)	54.6 (5.50)	37.3 (2.41, 2.20)	176.0	CY 31.2 (3.08), C ϵ 16.8 (2.04)
D68	128.0 (9.24)	53.7 (5.06)	41.6 (3.57, 2.86)	178.1	
T69	107.1 (8.04)	64.8 (4.17)	69.1 (4.54)	174.9	CY 22.3 (1.46)
D70	119.8 (8.48)	54.2 (4.96)	42.2 (2.75, 2.75)	176.1	
G71	109.2 (8.38)	45.2 (3.63, 3.34)		173.6	
L72	123.8 (8.73)	54.5 (4.53)	42.8 (2.00, 1.61)	176.8	CY 26.8 (1.74), C δ 25.0 (0.92), 24.3 (1.00)
L73	124.1 (7.53)	53.9 (5.71)	42.8 (1.96, 1.43)	176.5	CY 26.9 (2.23), C δ 27.1 (1.09), 23.8 (1.02)
Y74	120.4 (9.35)	56.4 (5.00)	41.1 (3.09, 3.09)	171.9	C δ 134.1 (7.13), C ϵ 117.5 (6.54)
G75	105.5 (8.98)	43.8 (4.67, 3.13)		172.8	
S76	119.6 (9.71)	55.8 (4.98)	65.3 (3.90, 3.53)	176.6	
Q77	127.1 (9.38)	58.4 (4.29)	29.4 (2.34, 2.20)	175.6	CY 34.3 (2.62, 2.53), N ϵ 112.2 (7.58, 6.95)
T78	108.2 (7.64)	57.5 (4.85)	70.3 (4.10)		CY 21.7 (1.21)
P79		62.3 (4.18)	29.0 (1.31, 0.31)	175.2	CY 26.5 (1.82, 1.43), C δ 49.7 (3.45)
N80	119.3 (7.45)	52.3 (4.76)	39.3 (3.21, 2.98)		N δ 113.0 (7.50, 6.59)
E81		59.4 (4.13)	29.2 (2.11, 2.11)	178.6	CY 36.0 (2.30)
E82	119.0 (8.66)	58.5 (4.20)	28.3 (2.61, 2.47)	176.2	CY 38.3 (2.52, 2.28)
C83	113.5 (7.83)	57.9 (5.42)	28.3 (3.72, 3.29)	173.1	
L84	118.0 (6.35)	54.0 (4.52)	44.5 (1.38, 1.16)	176.2	CY 27.5 (1.23), C δ 24.6 (0.61), 24.3 (0.40)
F85	122.3 (9.17)	57.2 (5.13)	42.8 (2.93, 2.79)	174.2	C δ 131.9 (7.03), C ϵ 130.8 (7.14), C ζ 128.2 (6.78)
L86	121.9 (10.05)	53.4 (5.19)	41.1 (1.89, 1.48)	176.0	CY 27.3 (1.53), C δ 25.4 (0.85), 23.1 (0.86)
E87	125.8 (8.31)	54.8 (4.80)	33.4 (2.03, 1.92)	175.7	CY 37.4 (1.76)
R88	129.6 (8.95)	54.7 (4.71)	32.7 (1.82, 1.82)	174.6	CY 26.9 (1.49), C δ 43.4 (3.20)
L89	126.0 (8.42)	55.0 (4.86)	42.5 (1.73, 1.53)	176.9	CY 27.2 (1.65), C δ 25.2 (0.86), 23.9 (0.86)
E90	125.6 (9.16)	54.5 (4.75)	31.1 (2.27, 2.03)	177.2	CY 35.7 (2.34)
E91	120.1 (8.86)	58.0 (4.11)	29.6 (2.11, 2.11)	176.6	CY 36.4 (2.35)
N92	114.6 (8.48)	54.0 (4.35)	37.5 (3.09, 2.98)	174.5	N δ 111.6 (7.67, 6.47)
H93	108.8 (8.16)	57.8 (4.24)	27.4 (3.41, 3.27)	173.6	C δ 2 119.6 (6.99), C ϵ 1 136.8 (8.39)

Y94	117.6 (7.66)	58.4 (4.83)	40.9 (2.96, 2.96)	174.3	C δ 133.8 (7.06), C ϵ 118.0 (6.84)
N95	118.7 (9.93)	52.3 (5.84)	42.5 (2.66, 2.47)	174.8	N δ 107.0 (6.84, 6.54)
T96	108.5 (8.50)	59.7 (5.13)	73.2 (4.57)	172.9	CY 22.8 (1.31)
Y97	119.1 (10.37)	57.4 (5.47)	41.3 (2.47, 2.47)	174.7	C δ 131.9 (6.64), C ϵ 117.5 (6.53)
I98	122.4 (8.86)	59.4 (4.63)	41.8 (1.51)	177.3	CY 29.0 (1.31, 0.06), C δ 13.9 (0.56), CY ^m 17.3 (0.64)
S99	120.2 (8.29)	58.5 (4.02)	64.6 (4.51, 3.83)	175.1	
K100	130.0 (8.10)	60.2 (3.77)	33.2 (1.74, 1.19)	179.3	CY 26.0 (1.21), C δ 30.1 (1.58), C ϵ 42.2 (2.80)
K101	120.5 (8.70)	58.4 (3.81)	32.7 (1.45, 1.29)	177.1	CY 24.4(0.93,0.51),C δ 29.7 (1.60),C ϵ 42.1 (2.90, 2.81)
H102	115.2 (7.38)	54.0 (4.56)	30.3 (2.48, 2.48)	176.7	C δ 2 119.6 (6.99), C ϵ 1 136.8 (7.77)
A103	125.2 (7.22)	56.2 (3.80)	18.9 (1.45)	181.0	
E104	117.6 (8.90)	58.4 (4.06)	28.6 (1.96, 1.96)	177.1	CY 36.4 (2.26, 2.18)
K105	116.3 (7.33)	55.3 (4.21)	31.9 (1.56, 0.71)	175.8	CY 24.8(0.83), C δ 29.0(1.29,1.13), C ϵ 41.6(2.47,2.27)
N106	113.3 (7.95)	54.4 (3.76)	37.5 (2.99, 2.88)	172.9	N δ 112.7 (7.42, 6.69)
W107	116.0 (6.42)	55.4 (4.84)	30.5 (3.15, 2.92)	173.8	C δ 2 124.9 (6.82), C ϵ 3 121.5 (7.34), C ζ 2 113.4 (7.39), C ζ 3 121.2 (6.19), C η 124.3 (7.05), N ϵ 1 128.0 (10.36)
F108	123.8 (8.09)	57.6 (5.35)	41.5 (3.42, 2.69)	177.3	C δ 133.5 (7.46), C ϵ 131.0 (7.20), C ζ 128.8 (6.93)
V109	120.1 (8.59)	63.3 (3.94)	32.8 (2.03)	176.6	CY 24.2 (0.49), 20.5 (0.51)
G110	114.4 (9.07)	46.2 (5.26, 3.17)		171.7	
L111	119.0 (8.23)	53.5 (4.96)	45.8 (2.23, 1.08)	176.6	CY 26.8 (1.61), C δ 26.6 (0.76), 22.1 (0.61)
K112	118.8 (8.93)	56.3 (4.21)	33.6 (2.08, 1.77)	178.7	CY 25.1 (1.37), C δ 28.4 (1.64), C ϵ (3.05)
K113	121.9 (9.09)	59.9 (3.63)	32.1 (1.77, 1.72)	176.8	CY 25.9 (1.53, 1.21), C δ 29.3 (1.69), C ϵ 42.1 (3.05)
N114	112.7 (7.44)	51.8 (4.71)	37.4 (3.32, 2.77)	176.8	N δ 109.3 (7.51, 6.81)
G115	109.8 (7.75)	44.5 (3.34, 2.09)		171.7	
S116	112.2 (7.38)	57.4 (4.74)	65.0 (3.91, 3.91)	173.6	
C117	119.8 (9.15)	60.1 (4.40)	28.2 (3.11, 2.89)	174.5	
K118	128.3 (8.20)	54.4 (4.37)	33.3 (1.45, 1.15)	174.6	CY 25.3 (1.09), C δ 30.1 (1.35), C ϵ 42.0 (2.76)
R119	124.4 (8.65)	56.5 (4.26)	31.3 (1.55, 1.55)	177.5	CY 28.2 (1.53), C δ 43.3 (3.16)
G120	106.8 (9.52)	48.0 (3.87, 3.21)			
P121		64.8 (4.36)	31.4 (2.38, 1.91)	176.9	CY 26.7 (1.67), C δ 49.4 (3.15)
R122	116.0 (8.18)	54.2 (4.51)	29.3 (2.13, 1.92)	177.0	CY 26.4 (1.40), C δ 41.8 (3.20)
T123	111.8 (7.77)	61.0 (5.18)	72.4 (4.08)	174.8	CY 21.5 (1.07)
H124	112.9 (7.16)	55.7 (4.87)	31.6 (3.40, 3.20)	172.9	C δ 2 121.9 (7.21), C ϵ 1 137.0 (8.43)
Y125	122.4 (8.99)	60.9 (3.91)	38.1 (3.09, 2.95)	176.8	C δ 132.8 (7.08), C ϵ 118.2 (6.87)
G126	116.4 (8.11)	44.8 (4.35, 3.32)		174.8	
Q127	118.2 (7.10)	54.9 (4.25)	29.6 (2.05, 2.05)	177.3	CY 33.4 (2.33, 2.11), N ϵ 107.6 (7.11, 6.80)
K128	121.1 (8.55)	58.5 (3.80)	31.9 (1.53, 1.43)	178.5	CY 25.5 (1.27), C δ 29.0 (1.55), C ϵ 42.1 (2.92)
A129	116.2 (7.61)	54.6 (3.94)	20.9 (1.14)	175.1	
I130	101.1 (5.90)	59.6 (5.17)	37.8 (2.22)	175.1	CY 24.6 (1.07,0.83), C δ 13.4 (-0.12), CY ^m 17.6 (0.59)
L131	121.4 (6.83)	54.0 (4.60)	41.8 (1.57, 0.68)	175.7	CY 25.4 (1.60), C δ 25.7 (0.42), 22.0 (0.57)
F132	121.8 (9.32)	56.9 (5.63)	43.9 (3.16, 2.80)	175.1	C δ 131.8 (7.20), C ϵ 129.7 (7.30), C ζ 128.8 (6.89)
L133	125.2 (10.04)	51.2 (5.28)	45.7 (1.95, 1.46)		CY 26.8 (1.53), C δ 25.6 (0.74), 24.2 (0.74)
P134		61.2 (5.57)	31.0 (2.03, 2.03)	176.8	CY 28.1 (2.36), C δ 51.3 (4.22, 3.94)
L135	129.2 (9.80)	51.9 (5.04)	44.4 (1.74, 1.62)		CY 26.4 (1.89), C δ 26.4 (0.90), 23.1 (0.90)
P136		62.3 (4.70)	32.3 (2.37, 1.93)	176.7	CY 27.3 (2.04), C δ 50.2 (3.85, 3.65)

V137	118.0 (8.07)	63.6 (3.80)	32.0 (1.98)	176.4	CY 21.0 (0.82), 21.0 (0.82)
S138	116.0 (7.93)	57.4 (4.56)	64.3 (3.85, 3.85)	174.0	
S139	118.0 (8.46)	58.3 (4.48)	64.0 (3.89, 3.89)	173.3	
D140	127.3 (8.01)	56.0 (4.41)	42.3 (2.69, 2.55)		

In each column, ^{15}N and ^{13}C chemical shifts are listed first and the corresponding ^1H shifts follow in parentheses. All spectra were referenced to TSP at 30 °C and pH 6.7. ^{13}C and ^{15}N dimensions were referenced indirectly to TSP using the ratio of gammas ($\text{H}_\gamma/\text{C}_\gamma$ or $\text{H}_\gamma/\text{N}_\gamma$) (Bax and Subramanian, 1986).

Table 3. Structural Statistics

Parameter	<SA>	<SA> _r
rmsd from experimental distance constraints (Å) (1368)	0.025 ± 0.001	0.023
Number of distance constraints violations greater than 0.5 Å	0 (max. 0.42)	0 (max. 0.27)
rmsd from experimental dihedral constraints (deg) (51)	0.13 ± 0.24	0
Number of dihedral constraints violations greater than 5.0 deg	0 (max. 4.6)	0 (max. 0.5)
F _{NOE} (kcal/mol)	42.1 ± 2.9	35.4
F _{cdih} (kcal/mol)	0.03 ± 0.08	0
F _{repel} (kcal/mol)	25.0 ± 2.7	19.9
E _{L-J} (kcal/mol)	-350.8 ± 49.7	-223.1
rmsd from idealized geometry		
Bonds (Å) (2148)	0.003 ± 0.000	0.002
Angle (deg) (3875)	0.613 ± 0.009	0.567
Improp(deg) (1128)	0.453 ± 0.023	0.37

<SA> refers to the final set of simulated annealing structures, and <SA>_r is the mean structure. The number of terms is given in parenthesis.

The value of the square-well NOE potential, F_{NOE}, is calculated with a force constant of 50 kcal/mol per Å². The value of F_{cdih} is calculated with a force constant of 50 kcal/mol per rad². The value of F_{repel} is calculated with a force constant of 4 kcal/mol per Å² with the van der Waals radii scaled by a factor of 0.8 of the standard values used in the CHARMM empirical function.

	-9	1	10	20	30	40
bFGF	MAEGEITT	LPALPEDGGS	GAFFPGHFKD	PKRL LYCK NGG	FFLRIHPDGR	VDGV REK SDP
			7	11	21	31
aFGF			NYKK	PKL LYC SN GG	HFLRILPDGT	VDGT RDR SDQ
				$\beta 1$	$\beta 2$	$\beta 3$
	50	60	70	80	90	100
	HI KL QLQAE E	RGVVS IK GVS	ANRY LAM KED	G RL LASKSVT	DEC FF ERLE	SNN YNTYRSR
41		51	61	71	81	91
	HI QL QLSAES	VGEVY IK STE	TGQY LAM DTD	G LL YGSQTPN	EECL F LERLE	ENH YNTYISK
	$\beta 4$	$\beta 5$	$\beta 6$	$\beta 7$	$\beta 8$	$\beta 9$
	110	118	128	138		
	KY TSWY VA	L KRT GQY KL G	S KT GPGQ KAI	L FL PMSAKS		
101		111	121	131		
	KHA EKN WF VG	L KN G SCK RG	P R THYGQ KAI	L FL PLPVSSD		
	$\beta 10$	($\beta 11$)		$\beta 12$		

Figure 1. Sequence alignment of basic FGF and acidic FGF on the structural basis. Residues responsible for heparin binding, receptor binding and cavity formation are shown in magenta, cyan and green, respectively.

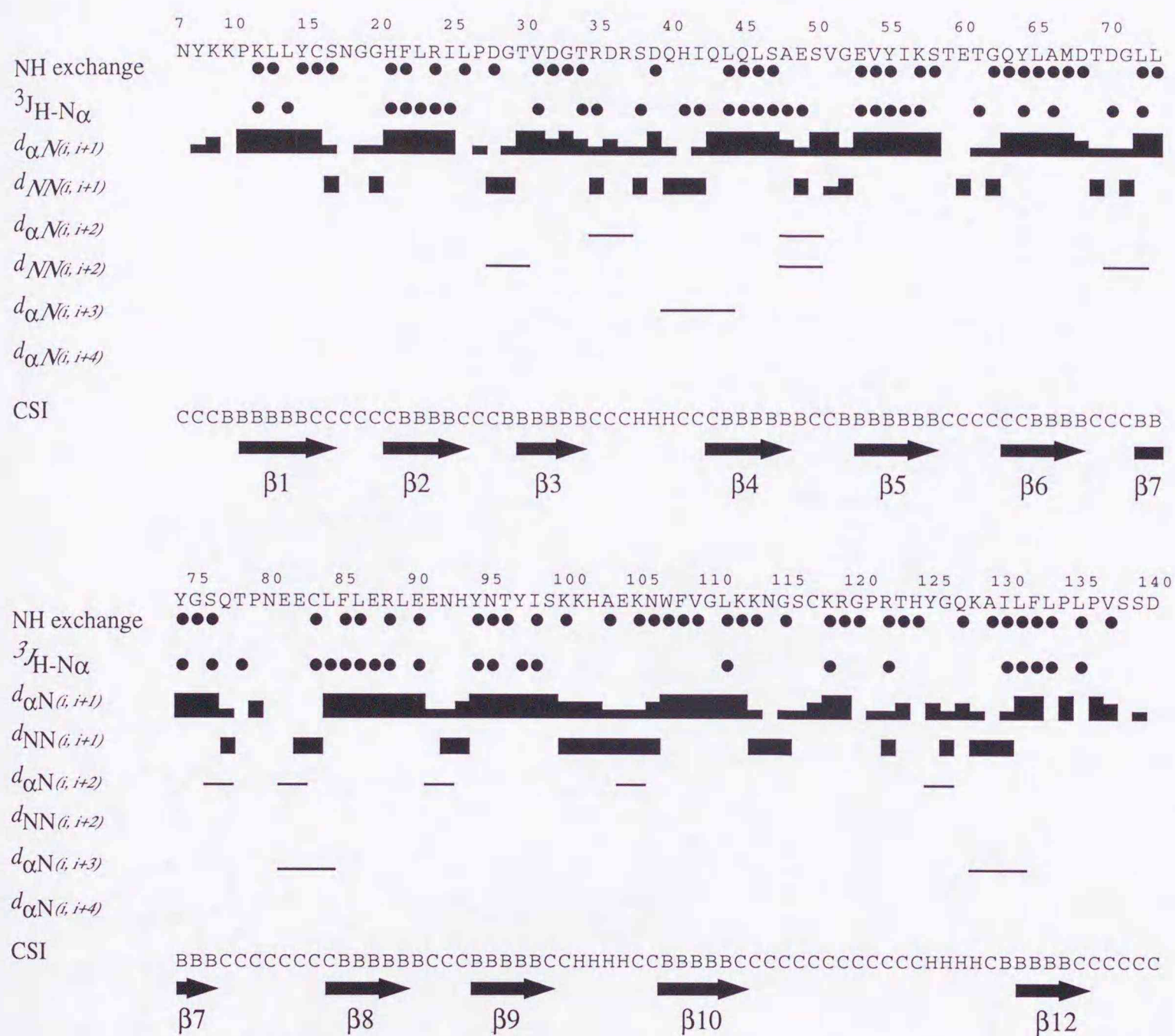


Figure 3. Diagram of sequential and short-range NOE connectivities, hydrogen exchange rates, coupling constants data, and consensus CSI for free aFGF. The NOE correlation were determined from 3D ^{15}N -edited NOESY spectrum at 30 °C. The height of the bar indicates the strength of the NOE correlation. Filled circles at NH exchange indicate slowly exchangeable amide protons estimated from the water-exchange peak intensity of 3D ^{15}N -edited NOESY spectrum. Filled circles at the $^3J_{\text{HNH}\alpha}$ indicate large (>8 Hz) coupling constants calculated from the ratio of peak intensities of 3D HNHA spectrum, respectively. Characters B, C and H at CSI indicate the stranded, coiled and helical structures estimated from the consensus CSI of C^α , C^β , C' and H^α chemical shift values, respectively.

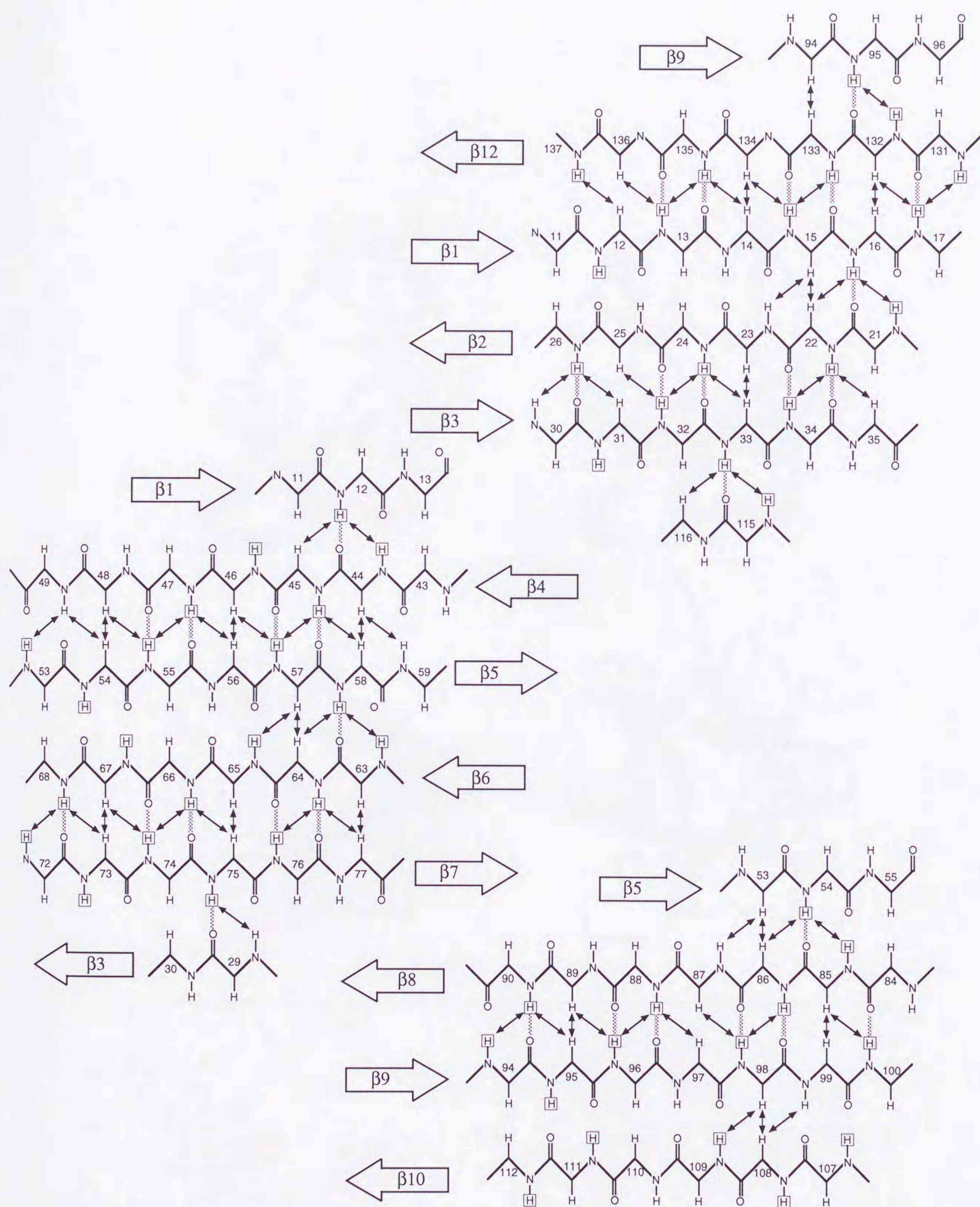


Figure 4. Schematic drawing of the secondary structure and b-sheet topology of free aFGF obtained from the inter-strands H^N-H^N , H^N-H^α and $H^\alpha-H^\alpha$ correlations observed in 3D ^{15}N -edited and ^{13}C -edited NOESY spectra. Open square indicates slowly exchangeable amide proton. Gray bars represent hydrogen bonds derived from hydrogen exchange rates and NOE patterns.

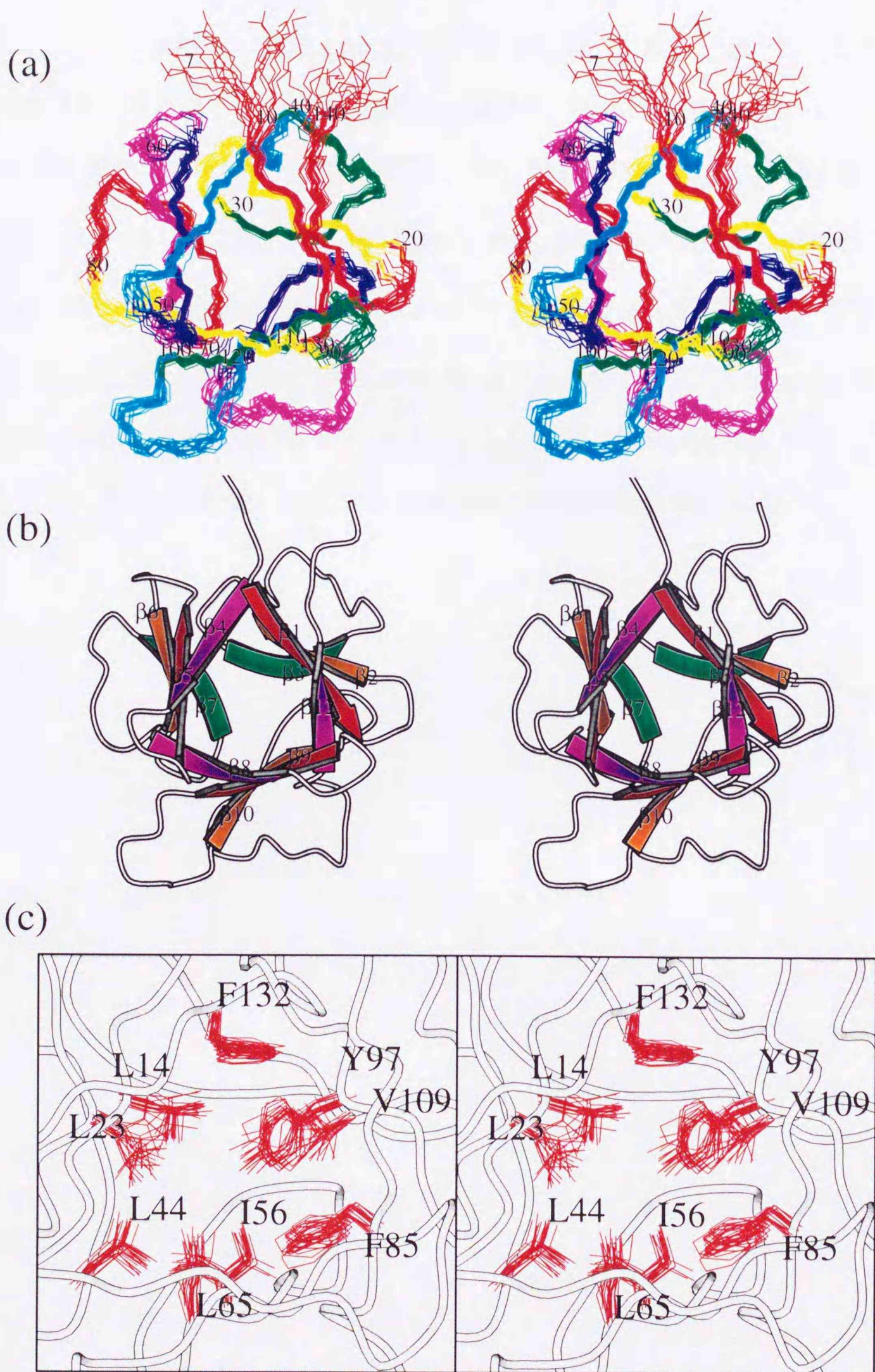


Figure 5. Stereo drawings of tertiary structures of free aFGF. (a) Best-fit superposition of the backbone (N, C α and C') atoms of the 20 final solution structures of free aFGF. Color changes in every 10 residues; Red (7 - 20 and 70 - 80), yellow (20 - 30 and 80 - 90), green (30 - 40 and 90 - 100), cyan (40 - 50 and 110 - 120), blue (50 - 60 and 120 - 130), and purple (60 - 70 and 130 - 140). (b) Ribbon diagram of the mean structure of free aFGF. The 11 β -strands are colored in red (β_I , β_V and β_{IX}), brown (β_{II} , β_{VI} and β_X), green (β_{III} , β_{VII} and β_{XII}) and violet (β_{IV} and β_{VIII}). The both figures were produced by the program MOLSCRIPT (Klauris, 1991). (c) The central cavity and its surrounding residues of aFGF. Overlay of 20 side-chain structures for Phe132, Leu14, Leu23, Leu44, Ile56, Leu65, Phe85, Tyr97 and Val109 are plotted on the main chain of the mean structure of aFGF.

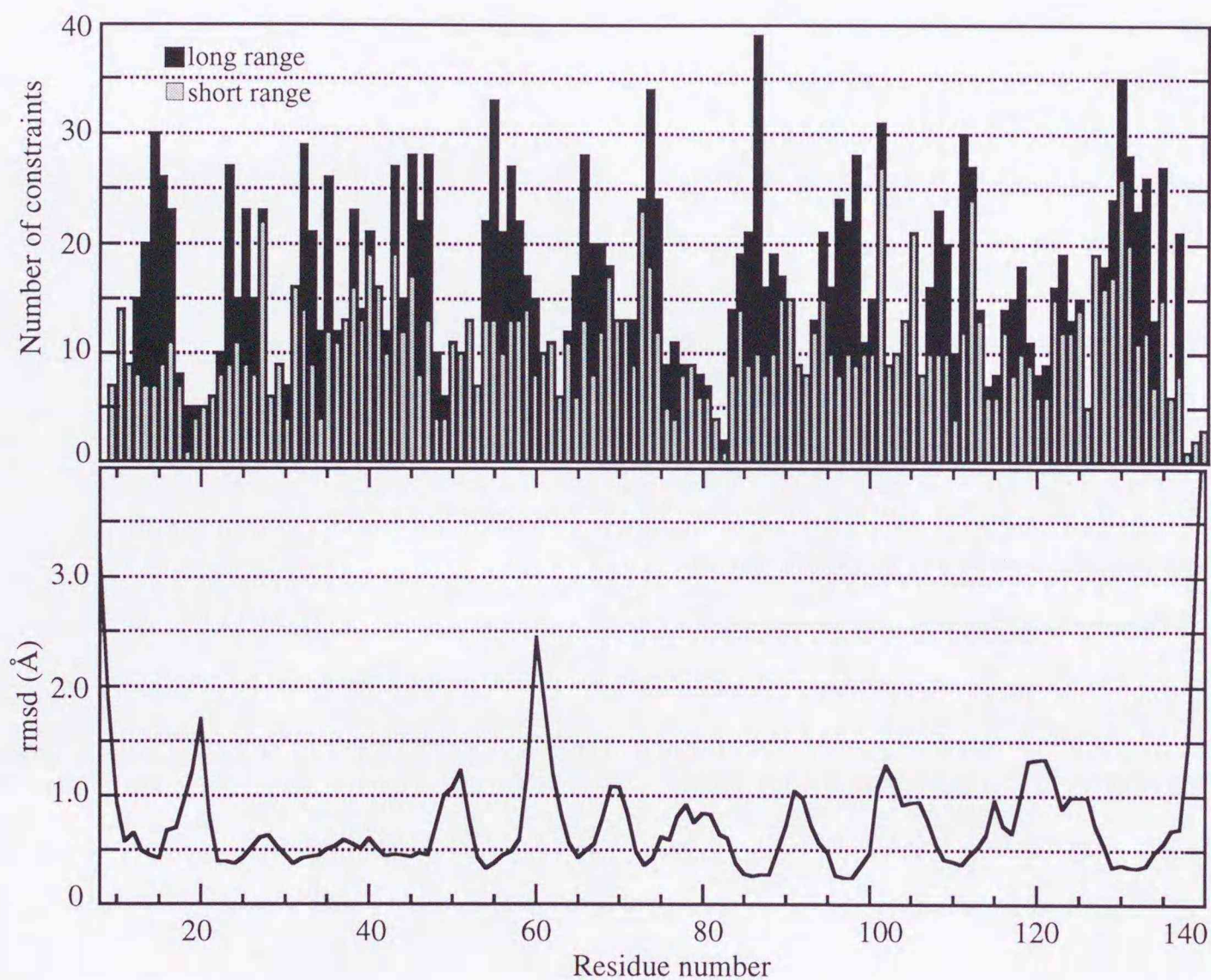


Figure 6. Number of NOE constraints (upper) and rmsd values (lower) for each residue. The number of long range ($|i-j| > 5$) and short range ($|i-j| \leq 5$) distance constraints were plotted by black and gray bars respectively as a function of residue number. The average value of the 20 main chain (N, C α , C') rmsd from the minimized mean structure were also plotted as a function of residue number.

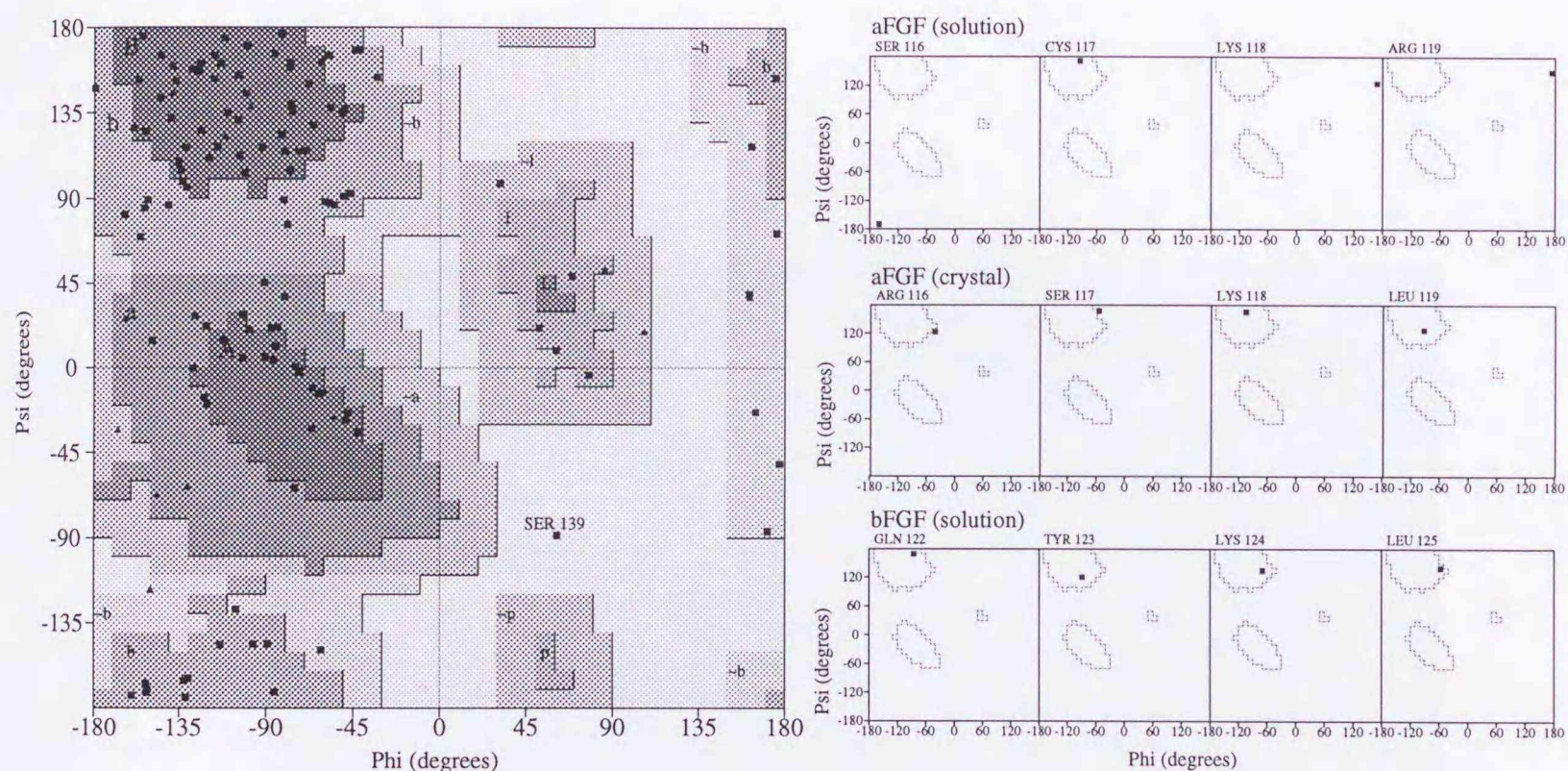


Figure 7 (Left) Ramachandran plot for all residues of the minimized averaged structure of aFGF. (Right) Ramachandran plot for four residues in the β_{XI} region of three FGF molecules (solution and X-ray structure of aFGF and solution structure of bFGF). The figures were produced with the program PROCHECK-NMR (Laskowski et al., 1996).

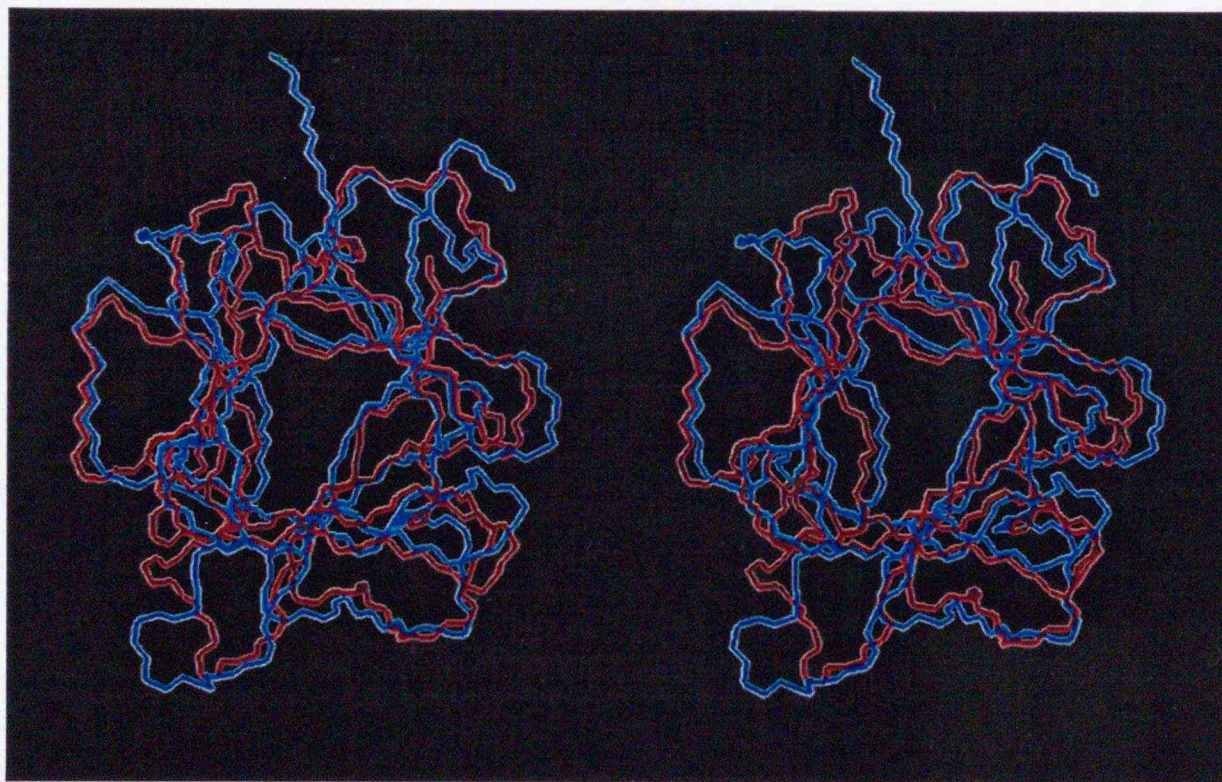


Figure 8. Stereo view of an overlay plot for the X-ray (red) and the solution structure (blue) of aFGF. The main-chain (H^N , C^α , and C') atoms are displayed. The figure was produced with QUANTA (Molecular Simulations, San Diego, CA).

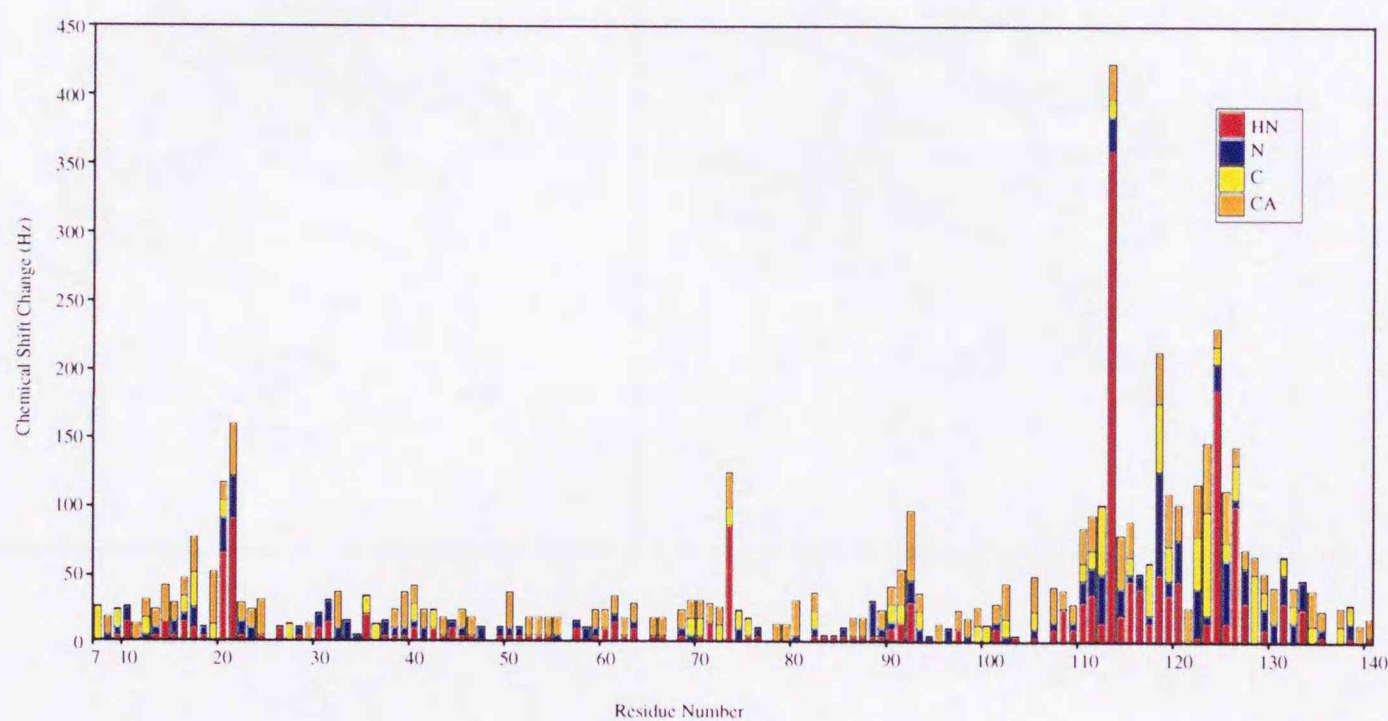


Figure 9. Hep6-induced chemical shift changes of the backbone amide hydrogen ($^1\text{H}^{\text{N}}$), nitrogen (^{15}N), α -carbon ($^{13}\text{C}^{\alpha}$), and carbonyl carbon ($^{13}\text{C}'$) atoms of aFGF.

The absolute values of the chemical shift differences for H^{N} (red), N (blue), C^{α} (yellow) and C' (orange) atoms between the complex of $^{13}\text{C}/^{15}\text{N}$ -labeled aFGF:Hep6 (1:1.2 ratio) and free $^{13}\text{C}/^{15}\text{N}$ -labeled aFGF are plotted in Hz unit.

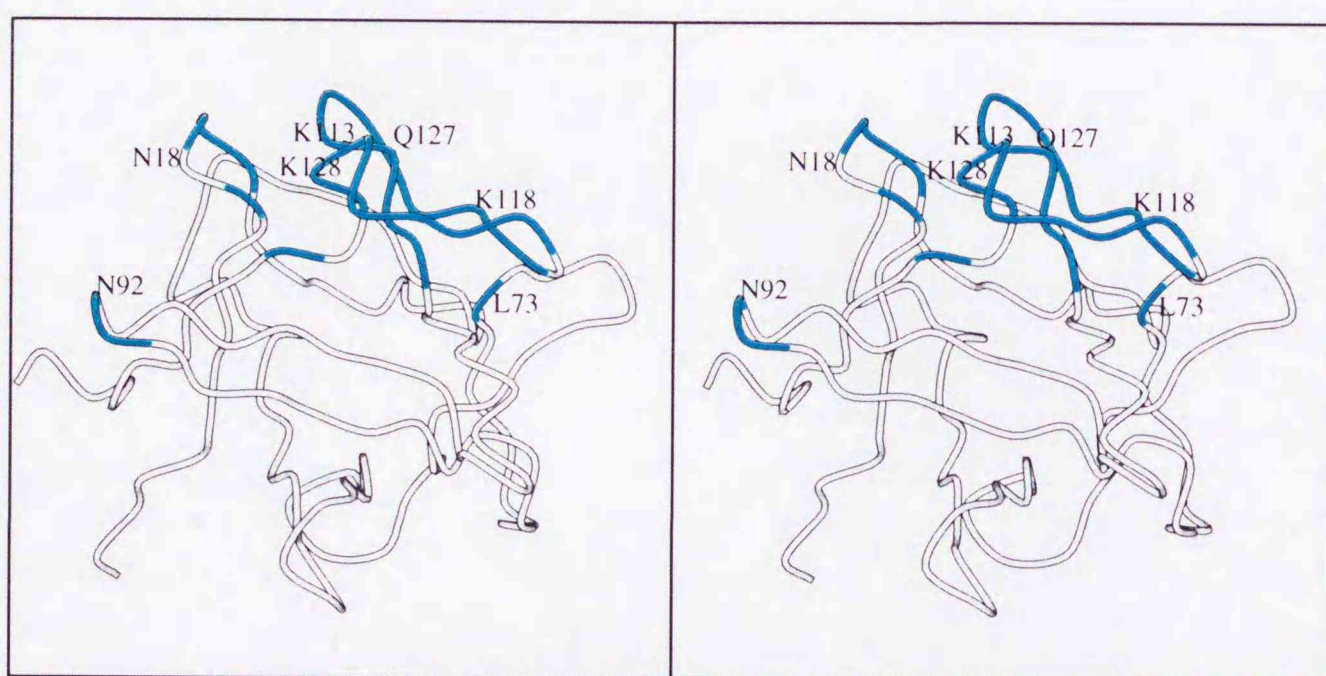


Figure 10. The location of the amino acid residues with chemical shift perturbation more than 50 Hz is plotted in stereo on the main-chain of the mean aFGF structure, where the side chain residues responsible for the heparin-hexasaccharide binding are also shown; Asn18, Asn92, Lys113, Lys118, Gln127, Lys128. These residues correspond to the counterparts which directly interact with the heparin-hexasaccharide in bFGF. The figure is prepared in quite similar orientation to Figure 3(b) of Fahm et al., (1996), and was produced with the program MOLSCRIPT (Kraulis, 1991)

Chapter 4

Solution structure of the SH2 domain of Grb2 complexed with the Shc-derived phosphotyrosine-containing peptide

Abstract

Src homology 2 (SH2) domains are protein modules found within a wide variety of cytoplasmic signaling molecules that bind with high affinity to phosphotyrosine (pTyr)-containing protein sequences. The solution structure of Growth Factor Receptor-Bound Protein 2 (Grb2) SH2 complexed with a Shc-derived pTyr-containing peptide was determined by nuclear magnetic resonance (NMR) spectroscopy. ^1H , ^{13}C , and ^{15}N resonances of the SH2 domain of Grb2 protein were assigned by analysis of multi-dimensional, double- and triple-resonance NMR experiments. Twenty structures of the complex were refined with the program X-PLOR using a total of 1649 experimentally derived conformational restraints containing 89 intermolecular NOEs. The structures converged to a root-mean-squared distance deviation (rmsd) of 0.66 and 1.11 Å for the backbone atoms and for the non-hydrogen heavy atoms, respectively. The pTyr binding site of Grb2 SH2 was similar to that of other SH2 domains. In contrast, amino acid residues C-terminal to pTyr did not form an extended structure because of steric hinderance by a bulky side-chain of Trp121 (EF1) of Grb2 SH2. As a result, the peptide formed a turn-structure on the surface of Grb2 SH2. The asparagine residue at the pTyr+2 position of the Shc peptide, which is the specific residue for protein-binding from library screening experiments, interacted with the main chain carbonyls of Lys109 and Leu120. From a comparison with the peptide-free structure of Grb2 SH2, the conformation of the BC loop region changes upon peptide binding. The present solution structure was similar to the crystal structure reported for Grb2 SH2 complexed with a BCR-Abl derived phosphotyrosine-containing peptide. The author compared the structures of the complex of Src, phospholipase C- γ (PLC- γ 1) and Grb2 SH2 with the cognate peptides.

Introduction

Three-dimensional structures of Src homology 2 (SH2) domains complexed with phosphotyrosine-containing peptides have been determined by X-ray crystallography (Waksman et al., 1992; Eck et al., 1993; Waksman et al., 1993; Lee et al., 1994; Mikol et al., 1995; Rahuel et al., 1996) and NMR spectroscopy (Pascal et al., 1994; Xu et al., 1995; Narula et al., 1995; Zho et al., 1995; Breeze et al., 1996, Mulhern et al., 1997). These studies show that the phosphotyrosine-containing peptides bind to SH2 at several sites. The first binding site is a conserved pocket lined by consensus basic residues which accommodate the phosphotyrosine residue. The second binding site is more variant in several SH2 domains and recognizes the specific amino acid residue(s) on the C-terminal side following to the phosphotyrosine residue. For example, *Src* family SH2 domains (Waksman et al., 1992; Eck et al., 1993; Waksman et al., 1993; Mikol et al., 1995; Xu et al., 1995, Mulhern et al., 1997) have a specific hydrophobic pocket which is served as a binding site for a hydrophobic residue at the pTyr+3 position, thus preferring an isoleucine residue at this position. In contrast, the SH2 domains of PLC- γ 1 (Pascal et al., 1994) and the Syp tyrosine phosphatase (Lee et al., 1994) recognize at least five, primarily hydrophobic residues, which fit into an extended hydrophobic groove running over the ligand-binding surface. These suggest that the surface topography of the SH2 domains can be altered with resulting changes in specificity, while conserving the structure of the central core of the domain.

Growth factor receptor bound protein 2 (Grb2) (Lowenstein et al., 1992) is an adaptor protein with a domain structure of SH3-SH2-SH3. The two SH3 domains of Grb2 bind to proline-rich sequences in the carboxyl terminal region of the nucleotide exchange factor, Son of sevenless (Sos) in cytosol (Buday and Downward, 1993).

Upon EGF stimulation, Grb2 SH2 binds to the EGF receptor directly or indirectly through proteins such as Shc, FAK, Syp and IRS-1 by recognizing the phosphotyrosine-containing sequences and relocates Sos to interact with Ras (Rozakis-Adcock et al., 1992; Skolnik et al., 1992; Buday and Downward, 1993; Pronk et al., 1994; Schlaepfer et al., 1994). Thus, Grb2 mediates signal transduction from EGF to Ras. The consensus phosphotyrosine-containing sequence for Grb2 SH2 binding has been extensively studied by peptide library screening (Songyang et al., 1994) and is now accepted as pTyr-(Leu/Val)-Asn-(Val/Pro), where Asn at pTyr+2 is essential, implying that Grb2 SH2 recognizes phosphotyrosine-containing sequence in a specific manner.

Recently, solution structure of free Grb2 SH2 was reported by NMR (Thornton et al., 1996). The phosphotyrosine-binding pocket of Grb2 SH2 was similar to that of other SH2 domains. However, the side-chain of Trp121 (EF1) in Grb2 occupied the pTyr+3 binding pocket in Src SH2, so that phosphotyrosine-containing peptide did not seem to bind in an extended conformation. X-ray crystal structure of Grb2 SH2 complexed with the ligand peptide derived from BCR-Abl (Lys-Pro-Phe-pTyr-Val-Asn-Val) demonstrated that the peptide formed a β -turn at pTyr, pTyr+1, pTyr+2 and pTyr+3 positions (Rahuel et al., 1996). These studies prompted the author to investigate the three dimensional structure of the Grb2 SH2 domain complexed with a phosphotyrosine-containing peptide derived from Shc (Asp-Asp-Pro-Ser-pTyr-Val-Asn-Val-Gln-Asn-Leu-Asp-Lys, abbreviated as Shc-peptide). Shc is a physiologically relevant target protein for Grb2 SH2 and the signal transduction pathway mediated by the Grb2-Shc complex was proved to be a major mitogenic signaling pathway, so that elucidation of the structural basis for molecular recognition between Grb2 and the Shc-peptide is much more demanding. In the present study, the author determined the structure of the complex of Grb2 SH2 and

the Shc-peptide. Based on the structural alignment, the author compared the recognition mode of the cognate peptides by the Grb2, Src and PLC- γ 1 SH2 domains.

Materials and methods

Expression and Purification of the Grb2 SH2 domain

The plasmid to express the Grb2 SH2 domain (residues His58 to Thr159 and the extrinsic N-terminal two residues, Gly56 and Ser57) was constructed by using a pGEX-4T-2 vector. For preparation of $^{13}\text{C}/^{15}\text{N}$ labeled Grb2 SH2, transformed *Escherichia coli* BL21 (DE3) cells were grown in M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$ (1 g/l), [^{13}C] glucose (2 g/l) and CELTONE-CN (Martek, USA) (1 g/l) at 37 °C and induced at $\text{OD}_{600} = 1.0$ with 0.4 mM isopropyl-thio- β -D-galactoside. The cells were collected after 4 hours, suspended into 20 mM potassium phosphate buffer at pH 8.0 containing 5 mM DTT, 1 mM PMSF and 1 mM EDTA, and then sonicated and centrifuged. The supernatants were incubated with Glutathione Sepharose 4B beads (Pharmacia, Sweden) for 1 h at 4 °C. After extensive washing in potassium phosphate buffer, the GST fusion proteins were eluted with 20 mM reduced glutathione, 50 mM Tris-HCl and 2 mM DTT at pH 8.2, and fractions were collected. The proteins were cleaved with trypsin (0.01 % wt/wt fusion protein) for 1 h at 23 °C. The buffer was exchanged to 20 mM HEPES, 2 mM DTT and pH 7.5 using a desalting column. The protein was applied to a Mono S 10/10 HPLC column (Pharmacia) and eluted with 0 - 600 mM NaCl gradient in 20 mM HEPES at pH 7.5. Finally, the SH2 domain was purified using a Superose 12 column (Pharmacia) with 20 mM potassium phosphate buffer and 150 mM NaCl at pH 6.3.

Peptide Synthesis

The phosphotyrosine-containing peptide (Asp-Asp-Pro-Ser-pTyr-Val-Asn-Val-Gln-Asn-Leu-Asp-Lys) corresponding to the sequence of Shc (313-325) was synthesized by the solid-phase Fmoc strategy using Na-Fmoc-O-(O, O-

dimethoxyphosphoryl)-L-tyrosine (Fmoc-Tyr(OP(OMe)₂) (Watanabe Chem., Japan).

The synthesized peptide was purified by reverse-phase HPLC on a Resource RPC column (Pharmacia). The purified peptide was characterized by mass spectroscopy and amino acid analysis.

NMR Sample Preparation

Protein samples for NMR experiments were prepared in buffers containing 20 mM potassium phosphate and 150 mM NaCl at pH 6.3. Buffer exchange and sample concentration were made by a 10 kDa cutoff Centricon concentrator (Amicon, USA) at 4 °C. Final protein concentration was 1.5 mM in 230 µl sample volume.

Approximately 1.2 mol equiv of the phosphotyrosine-containing peptide was added to the sample solution. After a final concentration step with 10 kDa cutoff concentrator, each sample was adjusted to contain 10 % ²H₂O or exchanged into the sample buffer system prepared with 99 % ²H₂O and adjusted to contain 5 mM DTT-d₁₀ and 0.5 % (w/v) NaN₃ from a fresh concentrated stock solution. All the samples were put into restricted volume 5 mm NMR tubes (Shigemi, Inc., Japan).

NMR Spectroscopy

All NMR spectra were recorded at a sample temperature of 28 °C on Varian Unity 500 and Unity-plus 600 spectrometers equipped with an UltraShim unit (Resonance Research, Inc., USA), three rf channels, and a pulsed field gradient triple-resonance probe with an actively shielded z gradient coil. Proton chemical shifts were referenced to internal DSS (0 ppm). ¹³C and ¹⁵N chemical shifts were referenced indirectly to liquid NH₃ and DSS according to Bax and Subramanian (1986). States-TPPI (Marion et al., 1989) quadrature detection was employed in the

indirect dimensions. All heteronuclear experiments and isotope-filtered experiments were adapted to incorporate z-gradient pulses for suppression of artifacts and elimination of the water resonance. Gradients were also used for coherence selection in the experiments for detection of NH protons. The 2D homonuclear correlated spectra were processed with the standard VNMR (Varian Instruments, USA) software, while other spectra were processed by the NMR Pipe program (Delaglio et al., 1995) on a Sun Ultra-1 workstation, respectively. Time domain convolution was used to remove the solvent signal (Marion et al., 1989), and the indirect ^{15}N and/or ^{13}C dimensions were extended as needed with linear prediction. Multidimensional data peak picking was employed using a home-written software. NMR data visualization and analysis were performed with the Felix 95.0 program and a home-written software (Hatanaka et al., 1994) on a Sun SPARC Station computer.

Chemical Shift Assignment

The backbone resonance assignments of the Grb2 SH2 domain were obtained by a sequential walk in the following triple resonance spectra using a 1.5 mM solution of uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled protein complexed with 1.2 mol equiv. of Shc-peptide in the H_2O solution. All of the following spectra were collected on a Unity 500 spectrometer with a ^1H (F3) carrier frequency of 4.75 ppm with a spectral width of 7000 Hz, a ^{15}N (F2) carrier frequency of 117.7 ppm with a spectral width of 1800 Hz.

To assign the protein-backbone sequentially, the following four 3D experiments were carried out. HNCA and HN(CO)CA experiments (Cavanagh et al., 1996) were collected with a ^{13}C (F1) carrier frequency of 55.3 ppm, a spectral width of 3600 Hz,

and eight transients for each increment. CBCANH and CBCA(CO)NH experiments (Cavanagh et al., 1996) were collected with a ^{13}C (F₁) carrier frequency of 41.0 ppm, a spectral width of 8800 Hz, and 16 transients for each transient. The walk for sequential assignments started at several spin systems assigned to specific amino acid types on the basis of characteristic $^{13}\text{C}\alpha/^{13}\text{C}\beta$ chemical shifts. The peptide fragments connected by the walk were unambiguously assigned to specific sequences on Grb2 SH2 considering the alignments of the amino acid types.

The assignments of other resonances of the Grb2 SH2 domain were subsequently made by the following five experiments. 3D C(CO)NH (Grzesiek et al., 1993) experiment was performed with a ^{13}C (F₁) carrier frequency of 41.0 ppm, with a ^{13}C spectral width of 8800 Hz, and with 32 scans for each transient in 90 % H₂O solution. 3D HBHA(CO)NH (Grzesiek and Bax, 1992) experiment was performed with ^1H (F₁) spectral width of 3500 Hz, and with 16 scans for each transient in 90 % H₂O solution. 3D HNCO (Cavanagh et al., 1996) was collected with a ^{13}C (F₁) spectral width of 1800 Hz centered on 177 ppm, and eight transients. 3D HN(CA)HA (Clubb et al., 1992; Seip et al., 1992) was collected with 16 transients, and a ^1H (F₁) spectral width of 2000 Hz. 4D HCCH-TOCSY (Olejniczak et al., 1992) spectrum was obtained from a 1.5 mM solution of uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled protein with 1.2 mol equiv. of the Shc-peptide in $^2\text{H}_2\text{O}$. This spectrum was collected with eight transient, a ^1H (F₁) spectral width of 3500 Hz, ^{13}C (F₂ and F₃) spectral widths of 2933 Hz centered on 41.0 ppm, and a mixing time of 15 ms.

The assignments of the aromatic resonances of the Grb2 SH2 domain were

achieved with the conventional 2D COSY, TOCSY and NOESY spectra using ^{15}N -labeled sample in 90 % H_2O solution.

The sequential and side-chain assignments of the Shc-peptide were completed by the following five isotope-filtered experiments using the sample of $^{13}\text{C}/^{15}\text{N}$ -labeled Grb2 SH2 complexed with the unlabeled Shc-peptide. The 2D ^{13}C -filtered COSY, ^{13}C -filtered TOCSY, [F1, F2] ^{13}C -filtered NOESY (Ogura et al., 1996) experiments were carried out on the 500 MHz spectrometer with the spectral widths of 7000 Hz, 64 scans for each transient and the mixing time of 75 ms for TOCSY and 150 ms for NOESY experiments. These measurements were performed in $^2\text{H}_2\text{O}$ solution.

The [F1,F2] $^{13}\text{C}/^{15}\text{N}$ -filtered 2D TOCSY and NOESY (Ogura et al., 1996) experiments were performed in H_2O solution, where the S-SS spin-echo pulse sequence was used for an observing pulse. In the S-SS spin-echo sequence, a 313.8 μs S pulse (3.08 kHz strength) and a 517.2 μs SS pulse (1.23 kHz strength) were applied as 180° and 90° pulses, respectively. The spectral widths were 7000 Hz, and the number of scans was 128. The TOCSY and NOESY mixing time of 50 ms and 150 ms was used, respectively.

Distance Restraints

The NOE distance restraints were derived from the following five experiments.

3D ^{15}N -edited NOESY (Cavanagh et al., 1996) experiment was carried out on the 500 MHz spectrometer using the ^{15}N -labeled protein complexed with the peptide in the H_2O solution. The spectrum was collected with eight transients for each

increment, with ^1H (F_1 and F_3) spectral widths of 7000 Hz, with a ^{15}N spectral width of 1800 Hz, and with a mixing time of 60 ms.

3D ^{13}C -edited NOESY (Cavanagh et al., 1996) experiments was performed on the 500 MHz spectrometer using the $^{13}\text{C}/^{15}\text{N}$ -labeled protein complexed with the peptide in the $^2\text{H}_2\text{O}$ solution. The spectrum was collected with eight transient, with ^1H spectral widths of 7000 Hz, with a ^{13}C (F_2) spectral width of 6000 Hz centered on 75.2 ppm, and with a mixing time of 100 ms.

3D [F_1] ^{13}C -edited, [F_3] ^{13}C -filtered NOESY (Ogura et al., 1996) experiment was recorded with the $^{13}\text{C}/^{15}\text{N}$ -labeled sample in the $^2\text{H}_2\text{O}$ solution to collect the intermolecular NOE restraints on the 600 MHz spectrometer. The spectrum was obtained with ^1H (F_2 and F_3) spectral widths of 8500 Hz, with a ^{13}C spectral width of 7334 Hz centered on 72.4 ppm, with 16 scans per increment, and with a mixing time of 150 ms.

2D [F_1 , F_2] ^{13}C -filtered NOESY (Ogura et al., 1996) spectrum in the $^2\text{H}_2\text{O}$ solution described above was also utilized to determine the intramolecular NOE restraints for the unlabeled Shc-peptide.

2D [F_1 , F_2] ^{15}N -filtered NOESY (Ogura, unpublished) experiment was recorded with the ^{15}N -labeled sample in the H_2O solution on the 600 MHz spectrometer. In this NOESY experiment, spectra cause NOE correlations between the amide proton of the unlabeled peptide and all the aliphatic/aromatic proton since directly ^{15}N -attached proton signals are eliminated by the filtering pulse scheme. The spectrum was obtained with spectral widths of 8000 Hz, with 128 scans per

increment, and with a mixing time of 100 ms. The NOE restraints for H δ of Asn (pTyr +2), which is a key residue for Grb2 SH2-Shc specific binding, was derived from the 2D [F1, F2] ^{15}N -filtered NOESY spectrum.

Dihedral angle and hydrogen bond restraints

A 3D HNHA (Vuister and Bax, 1993) spectrum was used to measure H $^{\text{N}}$ -H $^{\alpha}$ coupling constants ($^3\text{J}_{\text{HN-H}\alpha}$). The spectrum was acquired using the 600 MHz spectrometer with uniformly ^{15}N -labeled protein complexed with the peptide in H $_2\text{O}$ solution with the spectral widths of 5000 Hz in ^1H (F $_1$), 2100 Hz in ^{15}N (F $_2$), and 8000 Hz in ^1H (F $_3$). The calculated $^3\text{J}_{\text{HN-H}\alpha}$ values were corrected by simple multiplying by 1.11 as described by Vuister and Bax (1993). Dihedral restraints for the ϕ angles were set to $(-120 \pm 40^\circ)$ in the β -strand region when $^3\text{J}_{\text{HN-H}\alpha} > 9$ Hz, or $(-65 \pm 25^\circ)$ in the α -helix regions only when < 6 Hz.

74 distance restraints were incorporated for 37 hydrogen bonds on the basis of qualitative assessment of NH exchange rates estimated from the signal intensities of the solvent-exchanging peak on the ^{15}N -edited NOESY spectrum. The hydrogen-bonding restraints were incorporated after identification of the acceptor from structures calculated in the absence of hydrogenbond restraints.

Structure Calculations

The NOESY peak assignments were obtained using the laboratory-made programs for semi-automated NOESY cycles (Hatanaka et al., 1994). Upper limits of distance restraints were calculated as $kI^{-1/6}$, where I is the peak intensity and k is a

constant adjusted in each NOESY spectrum and relaxed by 0.5 Å, considering mobility. Lower limits of distance restraints were all 1.8 Å. The structures were calculated using the X-PLOR version 3.1 (Brünger, 1993). Initial coordinates were generated using random ϕ and ψ angles, while peptide bonds and side-chains took extended conformations. The *sa.inp* in X-PLOR 2.1 was used to carry out simulated annealing calculation. The target function minimized during simulated annealing comprised only potential terms for covalent geometry, experimental distance restraints, and van der Waals nonbonded repulsion. No electrostatic and 6-12 Lennard-Jones potential terms were included in the target function. The final structure calculations were based on 1649 interproton distance restraints. 1106 restraints were assigned to single pairs of resonances (432 intra-, 206 sequential [$|\text{li-j}| = 1$], 126 middle range [$2 \leq |\text{li-j}| \leq 5$], and 342 long-range [$|\text{li-j}| > 5$] NOEs). In addition, 543 restraints were assigned to multiple pairs using X-PLOR "or" operators and dynamically assigned to each pair during structural calculations. A final set of 20 converged structures was selected from 200 calculations on the basis of agreement with experimental data and van der Waals energy. A mean structure was obtained by averaging coordinates of the structures that had been superimposed in advance to the best converged structure and then minimized under the restraints.

Results and discussion

Resonance assignments of the Grb2 SH2 domain

The ^1H , ^{13}C and ^{15}N chemical shifts of the Grb2 SH2 domain complexed with the Shc-peptide were assigned from an array of heteronuclear multidimensional NMR experiments. Backbone sequential assignments for Grb2 SH2 were obtained by a conventional strategy using a combination of triple resonance 3D HN(CO)CA, 3D HNCA, 3D CBCA(CO)NH and 3D CBCANH spectra. Assignments for C^α and C^β chemical shifts were complete except for the extrinsic N-terminal two residues and His58. The N and H^N chemical shifts were identified in the 2D ^1H - ^{15}N HSQC spectrum (Figure 1) with the exception of Leu120, Trp121, Val132, His135 and Ser139, possibly due to local exchange broadening. The nitrogen chemical shifts of proline residues were not assigned. Assignments of other aliphatic resonances were complete except for Arg86 C^γ and H^γ , Lys124 C^γ , H^γ , C^δ , H^δ , C^ϵ and H^ϵ by 3D HNCO, 3D HBHA(CO)NH, 3D HN(CA)HA, 3D C(CO)NH and 4D HCCH-TOCSY spectra.

Resonance assignments of the Shc-peptide

The NMR resonances of the unlabeled Shc-peptide complexed with labeled Grb2 SH2 were assigned using a suite of PFG z-filtered pulse sequences (Ogura et al., 1996). Figure 2 shows the 1D NMR spectra of the complex with and without isotope-filter. Using the PFG z-filter together with the wide-band ^{13}C inversion pulse, both aromatic and aliphatic proton resonances directly bonded to ^{13}C were completely filtered out and the proton resonances derived from the unlabeled peptide were observed as shown in Figure 2(b), demonstrating the high filtering efficiency

and the wide-band nature of the present pulse sequence. For a sample in $^2\text{H}_2\text{O}$ solution, ^{13}C -filtered COSY, ^{13}C -filtered TOCSY and [F1, F2] ^{13}C -filtered 2D NOESY spectra were applied. While, $^{13}\text{C}/^{15}\text{N}$ -filtered TOCSY and [F1, F2] $^{13}\text{C}/^{15}\text{N}$ -filtered NOESY were applied for H_2O sample (90 % H_2O / 10% $^2\text{H}_2\text{O}$), where the S-SS spin-echo pulse sequence (Smallcombe, 1993) was employed to eliminate the solvent-derived signal. The spin systems of individual amino acid residues in the Shc-peptide were identified using ^{13}C -filtered COSY, ^{13}C -filtered TOCSY in $^2\text{H}_2\text{O}$ and $^{13}\text{C}/^{15}\text{N}$ -filtered TOCSY in H_2O . The sequential resonance assignments were subsequently obtained using NOE connectivities observed from [F1, F2] $^{13}\text{C}/^{15}\text{N}$ -filtered 2D NOESY in H_2O . In addition to the high suppression efficiency of the filtered experiments for $^{13}\text{C}/^{15}\text{N}$ -bound proton resonances, the solvent suppression by S-SS spin echo pulse sequence was sufficient to observe the NMR resonances of the unlabeled Shc-peptide even in H_2O solution. Figure 3 shows the C^αH -NH cross peak region of the [F1, F2] $^{13}\text{C}/^{15}\text{N}$ -filtered 2D NOESY spectrum recorded on the unlabeled Shc-peptide complexed with labeled Grb2 SH2 in H_2O solution, where the author observes intramolecular NOE correlations in the Shc-peptide and establish its sequential connectivity. The resonance assignments of both Grb2 SH2 and the Shc-peptide are given in Table 1.

Analysis for NOE signals originated from the Shc-peptide

On the basis of the resonance assignments, the author can now investigate in detail the conformation of the Shc-peptide bound to Grb2 SH2 both from

intramolecular and intermolecular NOE correlations.

Figure 4 shows NOE correlations from pTyr5 C δ , ϵ H resonances (overlapped) in the [F1, F2] ^{13}C -filtered 2D NOESY spectrum. The author found several intramolecular NOE correlations from pTyr5 C δ , ϵ Hs to Pro3 C γ H, Ser4 C α , β H, Val8 C α , γ H and Gln9 C β H resonances, showing that the aromatic ring of pTyr5 was surrounded by the several residues located on both N- and C-terminal sides. In contrast to the absence of NOE correlations between pTyr5 C δ , ϵ Hs and Val6 C γ H (+1), strong NOE peaks observed between pTyr5 C δ , ϵ Hs and Val8 C γ H (+3) suggested that the Shc-peptide formed a turn at pTyr5, Val6 (+1), Asn7 (+2) and Val8 (+3).

Intermolecular NOE correlations between the Shc-peptide and Grb2 SH2 were subsequently analyzed. The [F1] ^{13}C -edited, [F3] ^{13}C -filtered 3D NOESY experiment was quite helpful for this purpose, where intermolecular NOEs were observed between ^{13}C directly bonded protons and ^{12}C directly bonded protons. In this experiment, the starting proton on the labeled protein specified by the chemical shifts of a ^{13}C (F2)- ^1H (F1) pair and the destinating proton on the unlabeled peptide specified by a ^1H chemical shift (F3) are correlated with each other (Figure 5). For example, pTyr5 C δ , ϵ Hs of the Shc-peptide (6.91 ppm) showed strong intermolecular NOE correlations with Ser96 ($\beta\text{C}3$), Lys109 ($\beta\text{D}6$), Phe108 ($\beta\text{D}5$) and His107 ($\beta\text{D}4$) of Grb2 SH2. Considering the conserved structure of the phosphotyrosine-binding pocket and the observed intermolecular NOEs, pTyr5 was expected to bind to a similar pocket as elucidated for other SH2 domains. Intermolecular NOESY correlations were identified between the residues on the βD strand and the Shc-

peptide. Gln106 (β D3) had strong NOE correlations with Val6 (+1). Phe108 (β D5) showed NOEs to pTyr5, Asn7 (+2) and Val8 (+3), as well as Val6 (+1). These NOE correlations confirmed that the Shc-peptide formed a turn at pTyr5, Val6 (+1), Asn7 (+2) and Val8 (+3). The aromatic side-chain of Trp121 (EF1) had strong intermolecular NOE correlations with Val6 (+1) and Asn7 (+2) but not with Val8 (+3) which seemed to block the binding of the peptide in an extended conformation. In fact, Val8 (+3) had NOE correlations with Lys109 (β D6), indicating that the Shc-peptide does not form an extended conformation. The intermolecular NOE correlations between the peptide and Grb2 SH2 are summarized on a schematic drawing in Figure 6.

Determination of the solution structure

Solution structures of the Grb2 SH2 domain complexed with the Shc-peptide were determined utilizing a simulated annealing protocol using the program X-PLOR (Brünger, 1993). The calculations included 1427 protein-protein NOE-derived distance restraints, 30 ϕ dihedral angle restraints, and restraints for 37 hydrogen bonds. A total of 89 NOEs between Grb2 SH2 and the Shc-peptide and 133 NOEs within the Shc-peptide were also included. A summary of the final structural statistics for a set of the final structures and for the mean structure is presented in Table 2. There were no NOE violations greater than 0.4 Å or dihedral angle violations greater than 5° in any structure.

Description of the Grb2 SH2 domain complexed with the Shc-peptide

A superposition of the backbone of the Grb2 SH2 domain and the Shc-peptide complex from the 20 final structures is shown in Figure 7. The portion of the Shc-peptide in contact with the protein (residues 5-8; pYVNV) was well defined by the

NMR data. However, the remaining peptide residues (1-4, 9-13) were not determined due to a limited number of NMR-derived constraints involving these residues. The atomic rmsd about the mean coordinate positions for the protein residues 6-95 and the Shc-peptide residues 5-8 of the complex was 0.66 ± 0.09 Å for the backbone atoms and 1.11 ± 0.07 Å for all heavy atoms.

The topology of the Grb2 SH2 domain was very similar to that of other SH2 domains so far reported. Figure 8 shows a schematic ribbon diagram of the NMR-derived minimized mean structure of Grb2 SH2 domain complexed with the Shc-peptide in stereo. The structure consists of a three-stranded antiparallel β -sheet composed of strand β B, β C and β D that is flanked on each side by two α -helices, α A and α B. The central β -sheet is extended by two short parallel strands (β A and β G) located on the opposite side of the peptide-binding site. The protein also contains a second, smaller β -sheet-like structure composed of strands β D', β E and β F. The connections between strands β C and β D, strands β D' and β E, and strands β E and β F are tight β turns. The N- and C-termini of the Grb2 SH2 domain are located on the opposite side of the protein from the peptide binding surface. Residues defining the regions of secondary structure are 61-62 (β A), 67-77(α A), 82-87 (β B), 95-100 (β C), 105-110 (β D), 111-112 (β D'), 118-119 (β E), 124-125 (β F), 128-138 (α B), and 149-150 (β G), which are the same as those reported for the peptide-free structure (Thornton et al., 1996). As in the crystal structure of the Grb2 SH2 domain complexed with the peptide derived from BCR-Abl (Rahuel et al., 1996), the Shc-peptide lies in a turn-conformation on the surface of the protein in the NMR structure. The pTyr binding site is similar to that of other SH2 domains, and is formed by residues from α A, β B, β C, BC loop, and β D. However, the bulky side-chain of Trp121 located on the EF loop blocks the Shc-peptide to form an extended conformation.

Chemical shift perturbations

The chemical shift is sensitive to small perturbations in local geometry and conformation, and has therefore been proved to be a useful tool for monitoring the effects of ligand binding and conformational changes that occur within biological macromolecules. ^1H - ^{15}N HSQC spectra were measured for both free ^{15}N -labeled Grb2 SH2 and ^{15}N -labeled Grb2 SH2 complexed with the nonlabeled Shc-peptide. Chemical shift assignments for the backbone amide ^1H and ^{15}N resonances of free Grb2 SH2 were referenced to the recent study by Thornton et al. (1996). The backbone N and H^{N} chemical shift differences for each residue in Grb2 SH2 upon complex formation are summarized in Figure 9. Residues with large chemical shift changes of their backbone amide groups were found in several regions of the Grb2 SH2 sequences, Trp60 - Phe62, Ile65 - Ala68, Ser88 - Ser96, Phe101, and His107 - Leu111. Figure 10 shows a ribbon drawing of the solution structure of Grb2 SH2 domain. The residues containing the highly shifted amide groups are colored red, where the author classified the chemical shift change with more than 0.8 ppm in ^{15}N or 0.25 ppm in ^1H as highly shifted residues. In contrast to other SH2 domains, the BG loop region of the Grb2 SH2 domain does not actually participate in the recognition of the Shc-peptide as is shown in Figure 8.

Comparison between the Peptide-Bound and Peptide-Free Grb2 SH2 Structures

Figure 11 shows the superposition of the protein-backbone atoms of peptide bound and peptide-free structures (Thornton et al., 1996) of Grb2 SH2. The rmsd of all structurally defined backbone atoms between the two proteins is 1.00 Å and drop to 0.82 Å if the AB, BC, CD, and DE loops are excluded. There is no remarkable

conformational change in the EF loop where Trp121 is located. The striking differences between the peptide-bound and the peptide-free structures were observed on the position of the β B strand and the BC loop (residues 87-93). The conformational change in the BC loop of Grb2 SH2 upon peptide binding have led to the suggestion of a gated peptide binding site which is similar to the p56^{lck} SH2 domain (Mikol et al., 1995) and the p55^{blk} SH2 domain (Metzler et al., 1996). In addition, the α A helix of the peptide-bound Grb2 SH2 domain slightly changes its orientation. This oritational change is probably caused by binding of the phosphate group of pTyr residue to the side-chain of Arg67 (α A1) residue.

Specific Interactions between Grb2 SH2 and Shc-peptide (1)

Interaction between Grb2 SH2 and phosphotyrosine residue

Figure 12(a) shows the residues with appreciable NOE interactions and therefore have strong contact with the Shc-peptide. The phosphotyrosine binding pocket is formed by positively charged residues; Arg67 (α A1), Arg86 (β B5), His107 (β D3) nad Lys109 (β D5). In addition, Ser88 (BC1), Ser90 (BC3) and Ser96 (β C2) also participate in the pocket (Figure 12b). The aromatic ring of pTyr lies close to the β C and β D strands and make numerous NOEs with residues Ser96 (β C2), His107 (β D3), Phe108 (β D4), and Lys109 (β D5). Although the NH2 protons of Lys109 were reported to form a hydrogen bond with the side-chain of Asp94 of the protein in the crystal structur (Rahuel et al., 1996), a number of protons in the Lys109 have strong NOE interactions with the aromatic ring protons of pTyr residue rather than with the side-chain of Asp94 residue in solution. Therefore, the author could presume that the NH2 group of Lys109 contacts with the aromatic and phosphate groups of pTyr residue.

The guanidium groups of the arginine residues in the phosphotyrosine binding

site could not be assigned under the present solution condition. Although it was not feasible to obtain direct NOEs between the guanidium groups and the phosphotyrosine group, the solution structure clearly located Arg86 (β B5) at the bottom of the pTyr-binding pocket. On the other hands, the author could observe weak NOE interactions between Arg67 H^δ (α A1) and the aromatic ring proton of pTyr. This NOE located the side-chain of Arg67 (α A1) in the vicinity of the aromatic ring and the phosphate group of the pTyr residue. The cluster of positively charged side-chains creates an intense positive field at the bottom of the phosphate-binding pocket as shown in Figure 12(b).

Specific Interactions between Grb2 SH2 and Shc-peptide (2)

Recognition between Grb2 SH2 and the regulatory site of the Shc-peptide

Valine (pTyr+1) residue has intermolecular NOE with Gln106, Phe108 and Trp121 of the Grb2 SH2 domain (Figure 13). The peptide screening experiments showed that a hydrophobic amino acid was required at pTyr+1 for Grb2 SH2-binding. Binding of Val (pTyr+1) to the Grb2 SH2 domain is not an electrostatic interaction but van der Waals interaction. Thus, the hydrophobic residue is required at pTyr+1 for specific binding.

Asparagine (pTyr+2) is known as an essential residue of the ligand peptide for specific binding to the Grb2 SH2 domain. Actually, the aliphatic protons of Asn (pTyr+2) had strong NOE interactions with the side-chain of Trp121. The author also observed two intermolecular NOE interactions: (1) between H^δ protons of Asn (pTyr+2) and the amide proton of Lys109 (β D6) and (2) between H^δ protons of Asn (pTyr+2) and H^α of Leu120 (β E4). Furthermore, NMR signals of the H^δ protons of Asn (pTyr+2) were extremely low-field shifted to 9.08 and 8.30 ppm. This low-field

shift of Asn (pTyr+2) H δ s suggests that these protons form the hydrogen bonds with the Grb2 SH2 domain. The present NMR evidence is consistent with the X-ray structure revealed by Grb2 SH2-BCR Abl peptide, where three hydrogen bonds are formed between the side-chain of Asn (pTyr+2) and the main chain of the protein: the first hydrogen bond is formed between Asn (pTyr+2) O δ and Lys109 NH, the second hydrogen bond, between Asn (pTyr+2) H δ and the carbonyl group of Leu120, and the third one, between Asn (pTyr+2) H δ and the carbonyl group of Lys109.

Val (pTyr+3) residue is well defined which weakly contacts with Lys109 of the protein. This also interacts with the aromatic protons of pTyr, confirming again that the Shc-peptide forms a turn on Grb2 SH2. The residue at pTyr+3 is not conserved or type-conserved but hydrophobic residues are preferred for specific binding with the Grb2 SH2 domain. Actually, this position is Ala or Pro in PDGFR and EGFR, respectively.

Specific Interactions between Grb2 SH2 and Shc-peptide (3)

Interaction at N-termini to pTyr and beyond Val (pTyr+3)

There are no NOE interactions between N-terminal three residues, Asp (pTyr-4), Asp (pTyr-3), and Ser (pTyr-1) and the Grb2 SH2 domain. However, side-chain protons of Pro (pTyr-2) have some NOE interactions with those of Arg67 (α A1). These NOE interactions possibly reflect the formation of the hydrogenbond between phosphate group of pTyr and the guanidium group of Arg67. The C-terminal five residues, Gln (pTyr+4), Asn (pTyr+5), Leu (pTyr+6), Asp (pTyr+7), and Lys (pTyr+8) of the Shc-peptide have no NOE interactions with the Grb2 SH2 domain. In the present solution structure, therefore, residues beyond Val (pTyr+3) have no restraints with the peptide, and as a result, C-terminal five residues of the Shc-peptide

are highly flexible.

Comparison of the structures of the SH2 domains complexed with phosphopeptides

Studies using phosphopeptide libraries to probe sequence specificity have shown that there are two types of SH2 domains. Type-1 involves the non-receptor tyrosine kinases, Src, Lck, Fyn and Abl which displays a preference for the motif of pTyr-hydrophilic-hydrophilic-Ile/Pro (Songyang et al., 1994). Another type comprises phospholipase C- γ (PLC- γ 1), Syp tyrosine phosphatase, Shc and Grb2 which prefers the motif of pTyr-hydrophobic-X-hydrophobic (Songyang et al., 1994). However, Grb2 exclusively selects Asn at pTyr+2, suggesting that Grb2 recognizes a target phosphopeptide in a different manner from other SH2 domains. Here the author compared the three dimensional structure of the Grb2-Shc peptide complex with those of Src and PLC- γ 1-cognate peptide complexes (Figure 14). All of the SH2 domains display deep positively charged pockets lined with consensus basic residues of Arg(α A2), Arg(β B5), His(β D4) and Lys/Arg (β D6). The geometry of the pTyr binding pocket is similar with each other, confirming that pTyr is a major binding site for SH2 and the sequence C-terminal to pTyr forms a regulatory site for the specific binding. Actually, the recognition of the cognate phosphopeptide by SH2 is different from each other, although the three dimensional structure of SH2 domains are quite similar. In the case of Src, the pTyr containing peptide binds on the groove of the SH2 surface in an extended conformation. Ile at pTyr+3 binds to the hydrophobic pocket and hydrophilic residues at pTyr+2 and pTyr+3 are relatively exposed. Thus, the specific recognition of the peptide is determined by the residue at pTyr+3 (Figure 14c and 14f). In the case of PLC- γ 1, extensive hydrophobic contacts are observed between the residues, pTyr+1-pTyr+6 and the shallow groove on the SH2 surface (Figure 14b and 14e). The side view of both Src and PLC- γ 1 peptide complex,

display the extensive interactions between the binding peptides and the SH2 BG and EF loops (Figure 14c and 14f). In contrast, the Shc-peptide takes a turn conformation on Grb2 SH2. The bulky side-chain of Trp121 (EF1) takes over the pTyr+3 binding sites, thus blocks the peptide to take an extended conformation but to take a turn conformation (Figure 14a and 14d). This turn conformation is further stabilized by the hydrogen bond formation between the side-chain amide protons of Asn (pTyr+2) and main chain carbonyl groups of SH2 Lys109 (β D6) and Leu120 (β E4). Based on the present NMR study and the X-ray crystal structure of the complex between Grb2 SH2 and the BCR-Abl peptide, Grb2 SH2 binds the peptide with a turn conformation at pTyr, pTyr+1, pTyr+2 and pTyr+3. Interestingly, Grb2 SH2 binds a cyclic peptide containing pTyr-Val-Asn-Val as a core sequence (Gay et al., 1997). Grb2 is unique among the known SH2 domains in a sense that this is the only SH2 domain which has a tryptophan residue at position EF1 and requires Asn at pTyr+2 for the binding peptide. Since EF1 is either threonine or glycine in Src or PLC- γ 1, the residue at EF1 plays an important role in determining the both conformation and specificity of the binding peptides. The mutation of the threonine residue at position EF1 in Src SH2 to a tryptophan residue changed the peptide specificity of Src to Grb2-type (Marengere et al., 1994). Thus, the author can conclude that it is the topography of the binding surface which is responsible for determining the specificity of the peptides. The author can also change the signal transduction pathway by mutation of one residue at EF1 on the SH2 domains.

References

- Bax, A. and Subramanian, S. (1986) *J. Magn. Reson.* 67, 565-569.
- Breeze, A.L., Kara, B.V., Barratt, D.G., Anderson, M., Smith, J.C., Luke, R.W., Best, J.R. and Cartlidge, S.A. (1996) *EMBO J.* 15, 3579-3589.
- Brünger, A.T. (1993) *X-PLOR Version 3.1: A System for X-Ray Crystallography and NMR*, Yale University Press, New Haven
- Buday, L. and Downward, J. (1993) *Cell* 73, 611-620.
- Cavanagh, J., Fairbrother, W.J., Palmer III, A.G. and Skelton, N.J. (1996) *Protein NMR Spectroscopy, Principles and Practice*, Academic Press, New York, NY, USA.
- Clubb, R.T., Thanabal, V. and Wagner, G. (1992) *J. Biomol. NMR* 2, 203-210.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR* 6, 277-293.
- Eck, M.J., Shoelson, S.E. and Harrison, S.C. (1993) *Nature* 362, 87-91.
- Gay, B., Furet, P., Garcia-Echeverria, C., Rahuel, J., Chene, P., Fretz, H., Schoepfer, J., and Caravatti, G. (1997) *Biochemistry* 36, 5712-5718.
- Grzesiek, S. and Bax, A. (1992) *J. Am. Chem. Soc.* 114, 6291-6293.
- Grzesiek, S., Anglister, J. and Bax, A. (1993) *J. Magn. Reson. B* 101, 114-119.
- Hatanaka, H., Oka, M., Kohda, D., Tate, S., Suda, A., Tamiya, N. and Inagaki, F. (1994) *J. Mol. Biol.* 240, 155-166.
- Kraulis, P.J. (1991) *J. Appl. Cryst.* 24, 946-950.
- Lee, C.-H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S.E. and Kuriyan, J. (1994) *Structure* 2, 423-438.
- Lowenstein, E.J., Daly, R.J., Batzer, A.G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E.Y., Bar-Sagi, D. and Schlessinger, J. (1992) *Cell* 70,

431-442.

- Marengere, L.E.M., Songyang, Z., Gish, G.D., Schaller, M.D., Parsons, J.T., Stern, M.J., Cantley, L.C. and Pawson, T. (1994) *Nature* 369, 502-505.
- Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989) *Biochemistry* 28, 6150-6156.
- Metzler, W.J., Leiting, B., Pryor, K., Müller, L., and Farmer, B.T., II (1996) *Biochemistry* 35, 6201-6211.
- Mikol, V., Baumann, G., Keller, T.H., Manning, U. and Zurini, M.G.M. (1995) *J. Mol. Biol.* 246, 344-355.
- Mulhern, T.D., Shaw, G.L., Morton, C.J., Day, A.J., and Campbell, I.D. (1997) *Structure* 5, 1313-1323.
- Narula, S.S., Yuan, R.W., Adams, S.E., Green, O.M., Green, J., Philips, T.B., Zydowsky, L.D., Botfield, M.C., Hatada, M., Laird, E.R., Zoller, M.J., Karas, J.L. and Dalgarno, D.C. (1995) *Structure* 3, 1061-1073.
- Nicholls, A., Sharp, K. and Honig, B. (1991) *Protein Struct. Funct. Genet.*, 11, 281-296.
- Ogura, K., Terasawa, H. and Inagaki, F. (1996) *J. Biomol. NMR*, 8, 492-498.
- Olejniczak, E.T., Xu, R.X., and Fesik, S.W. (1992) *J. Biomol. NMR* 2, 655-659.
- Pascal, S.M., Singer, A.U., Gish, G., Yamazaki, T., Shoelson, S.E., Pawson, T., Kay, L.E. and Forman-Kay, J.D. (1994) *Cell* 77, 461-472.
- Pronk, G.J., de Vries-Smits, A.M.M., Buday, L., Downward, J., Maassen, J.A., Medema, R.H., and Bos, J.L. (1994) *Mol. Cell. Biol.* 14, 1575-1581.
- Rahuel, J., Gay, B., Erdmann, D., Strauss, A., García-Echeverría, C., Furet, P., Caravatti, G., Fretz, H., Schoepfer, J. and Grütter, M.G., (1996) *Nature Struct. Biol.* 3, 586-589.
- Rozakis-Adcock, M., McGlade, J., Mbamalu, G., Pelicci, G., Daly, R., Li, W.,

- Batzer,A., Thomas,S., Brugge,J., Pelicci,P.G., Schlessinger,J. and Pawson,T.
(1992) *Nature* 260, 689-692.
- Schlaepfer,D.D., Hanks,S.K., Hunter,T. and van der Geer,P. (1994) *Nature* 372,
786-791.
- Seip,S., Balbach,J. and Kessler,H. (1992) *J. Magn. Reson.* 100, 406-410.
- Skolnik,E.Y., Batzer,A., Li,N., Lowenstein,L.E., Mohammadi,M., Margolis,B. and
Schlessinger,J. (1992) *Nature* 260, 689-692.
- Smallcombe,S.H. (1993) *J. Am. Chem. Soc.* 115, 4776-4785.
- Songyang,Z., Shoelson,S.E., McGlade,J., Olivier,P., Pawson,T., Bustelo,X.R.,
Barbacid,M., Hanafusa,H., Yi,T., Ren,R., Baltimore,D., Ratnofsky,S.,
Feldman,R.A. and Cantley,L.C. (1994) *Mol. Cell. Biol.* 14, 2777-2785.
- Thornton,K.H., Mueller,W.T., McConnell,P., Zhu,G., Saltiel,A.R. and
Thanabal,V. (1996) *Biochemistry* 35, 11852-11864.
- Vuister,G.W. and Bax,A. (1993) *J. Am. Chem. Soc.* 115, 7772-7777.
- Waksman,G., Kominos,D., Robertson,S.C., Pant,N., Baltimore,D., Birge,R.B.,
Cowburn,D., Hanafusa,H., Mayer,B.J., Overduin,M., Resh,M.D., Rios,C.B.,
Silverman,L. and Kuriyan,J. (1992) *Nature* 358, 646-653.
- Waksman,G., Shoelson,S.E., Pant,N., Cowburn,D. and Kuriyan,J. (1993) *Cell* 72,
779-790.
- Xu,R.X., Word,J.M., Davis,D.G., Rink,M.J., Willard Jr.,D.H. and Gampe Jr.,R.T.
(1995) *Biochemistry* 34, 2107-2121.
- Zho,M.-M., Meadows,R.T., Logan,T.M., Yoon,H.S., Wade,W.S.,
Ravichandran,K.S., Burakoff,S.J. and Fesik,S.W. (1995) *Proc. Natl. Acad.
Sci. USA* 92, 7784-7788.

Table 1. ¹H, ¹³C and ¹⁵N resonance assignments for the SH2 domain of Grb2 complexed with Shc-derived peptide

Residue		HN	N	C'	Cα	Hα	Cβ	Hβ	Hβ'	Cγ	Hγ	Hγ'	Cδ	Hδ	other
56	G														
57	S														
58	H					4.51									
59	P			176.3	64.4	4.28	32.0	2.30	2.03	27.2	2.00	1.91	50.8	3.55	Hδ' 3.01
60	W	7.52	111.9	175.0	55.7	4.80	32.1	3.33	3.07				129.7	7.52	Ne1 130.8 Cζ2 124.2 Cη2 115.1 Cζ3 121.6 Ce3 121.6 He1 10.65 Hζ2 6.75 Hη2 7.42 Hζ3 7.42 He3 6.76
61	F	8.62	125.2	175.2	57.9	5.36	38.0	2.60	2.43				131.9	7.17	He 7.43
62	F	8.82	127.6	175.3	57.6	4.41	40.7	3.33	2.60				131.4	7.26	He 7.35 Hζ 7.72
63	G	5.59	103.8	174.5	47.3	3.83 3.62									
64	K	8.78	127.6	175.5	56.4	4.52	31.3	1.97	1.71	25.1	1.43	1.31	29.8	1.72	Cε 42.3 He 3.05
65	I	7.36	121.7		57.6	4.82	39.9	1.83		26.4	1.57	0.57	14.3	0.74	Cγ' 17.7 Cγ' 1.02
66	P			176.6	62.7	4.40	32.9	2.64	1.90	27.9	2.13	2.07	51.3	4.04	Hδ' 3.56
67	R	8.81	125.4	178.4	60.9	3.54	30.9	1.87	1.67	28.1	1.12	0.57	43.1	2.81	
68	A	9.36	118.2	180.8	55.1	4.18	18.4	1.43							
69	K	6.93	117.6	178.6	57.1	4.30	31.5	1.97	1.97	24.5	1.55	1.55	28.1	1.77	Cε 42.1 He 2.98
70	A	8.15	122.4		55.1	3.98	18.2	1.54							
71	E	8.17	115.1	178.4	60.3	3.91	28.6	2.35	2.17	37.1	2.83	2.15			
72	E	7.92	121.0	179.2	59.7	4.00	29.9	2.24	2.24	36.0	2.44	2.15			
73	M	8.17	116.8	179.7	59.1	4.07	33.9	2.08	1.91	31.3	2.57	2.36			Cε 16.6 He 1.72
74	L	8.29	117.8	179.1	57.3	3.99	42.3	1.89	1.15	27.3	1.71		27.3	0.71	Cδ' 22.5 Hδ' 0.61
75	S	8.33	114.2	175.3	61.5	4.10	63.0	3.97	3.97						
76	K	6.89	117.1	177.6	56.0	4.35	33.0	2.10	1.81	25.1	1.63	1.50	29.0	2.08	Hδ' 2.03 Cε 42.4 He 2.99
77	Q	7.43	117.8	176.0	54.9	4.20	28.9	2.29	2.29	35.0	2.71	2.17			Ne2 119.8 He21 8.26 He22 7.83
78	R	8.72	118.3	176.9	57.0	4.26	31.1	1.68	1.57	26.9	1.48	1.48	43.3	3.14	
79	H	7.82	116.3	174.3	54.9	4.92	31.9	3.08	3.08				120.5	7.27	
80	D	8.69	123.0	177.7	55.9	4.57	41.0	2.86	2.75						
81	G	9.52	111.5	174.6	45.5	5.10 3.24									
82	A	7.60	124.6	177.5	53.2	5.08	18.5	1.47							
83	F	7.46	118.5	171.5	55.5	6.14	44.2	3.25	3.16						
84	L	8.85	111.9	175.8	54.4	4.71	45.6	1.74	1.74	25.6	1.71		25.6	0.78	Cδ' 27.4 Hδ' 0.35

85	I	9.46	119.8	173.0	60.0	5.47	39.3	2.34		29.5	1.91	1.39	12.7	0.72	C γ ' 18.2	H γ ' 1.29
86	R	9.36	123.2	175.4	51.9	5.30	33.9	2.22	1.36				44.1	3.52	H δ ' 2.51	
87	E	8.60	123.0	176.4	57.0	4.50	30.1	2.02	1.82	37.1	2.47	2.28				
88	S	8.75	121.0	175.2	58.2	4.21	63.1	3.87	3.60							
89	E	8.52	128.8	177.8	58.2	4.12	30.6	2.18	2.18	36.3	2.41	2.18				
90	S	8.99	114.3	175.0	59.8	4.25	63.4	3.96	3.86							
91	A	7.55	126.7		49.1	4.83	19.3	1.18								
92	P			178.2	64.1	4.35	31.4	2.27	1.82	27.8	2.05	1.91	49.9	3.52	H δ ' 3.43	
93	G	9.00	114.0	173.6	45.2	4.28 3.47										
94	D	8.11	120.6	175.5	52.9	4.93	42.6	2.93	2.79							
95	F	9.33	118.5	174.8	57.2	5.46	41.9	3.20	2.83				131.6	7.36	C ϵ 131.5	He 7.55 H ζ 7.40
96	S	9.56	116.5	171.0	57.2	5.44	65.9	4.13	3.56							
97	L	9.71	127.6	175.3	53.9	5.42	45.8	1.80	1.23	27.9	1.62		25.7	0.71	C δ ' 25.7	H δ ' 0.41
98	S	9.31	124.8	172.2	58.6	5.74	66.1	3.63	3.46							
99	V	9.27	122.3	173.4	59.7	5.25	36.5	1.90		20.5	0.90				C γ 22.3	H γ 0.90
100	K	9.27	129.0	175.7	56.6	4.59	35.2	2.12	1.74	25.7	1.43	1.43	29.1	1.54	C ϵ 43.3	He 2.98
101	F	8.75	128.3	174.9	59.0	4.67	40.9	2.88	2.88				131.0	7.13	He 7.36	H ζ 7.44
102	G	9.40	120.6	174.7	46.6	3.71 3.56										
103	N	8.82	123.8	174.1	53.1	4.83	38.6	3.00	2.69						N δ 2 113.0	H δ 21 7.63 H δ 22 6.93
104	D	7.98	119.1	174.6	53.1	5.02	44.8	2.76	2.76							
105	V	8.36	120.2	174.7	61.5	4.56	33.9	1.79		22.1	0.90				C γ 22.1	H γ 0.43
106	Q	9.06	125.5	174.5	54.4	4.37	31.2	1.94	1.74	33.4	2.23	2.23			Ne2 105.9	He21 7.29 He22 6.22
107	H	8.53	118.8	173.9	54.7	5.90	34.3	2.89	2.78				121.4	7.12		
108	F	10.42	121.6	175.3	56.2	4.84	42.2	3.07	2.60				131.3	6.92	He 7.35	H ζ 7.18
109	K	9.52	125.7	175.7	55.1	4.31	31.8	1.30	0.71	22.6	0.81	0.81	27.2	1.39	H δ ' 1.24	C ϵ 40.4 He 2.50 He' 2.39
110	V	8.44	125.4	174.8	63.1	3.90	29.8	2.39		20.0	0.71				C γ 23.0	H γ 1.05
111	L	8.38	132.8	174.9	54.3	4.11	41.0	0.57	-0.07	26.9	1.24		26.2	0.56	C δ ' 23.2	H δ ' 0.57
112	R	7.85	115.3	177.7	52.4	5.67	32.8	1.80	1.60	26.6	1.57	1.33	42.6	3.01		
113	D	7.97	121.3	178.6	51.7	5.08	42.4	3.29	2.47							
114	G	8.63	106.3	174.6	46.5	3.97 3.89										
115	A	7.96	122.3	177.6	51.2	4.58	19.3	1.47								
116	G	8.12	107.4	174.2	45.3	4.23 3.63										

117	K	8.72	121.5	175.6	55.8	4.54	32.5	2.01	1.82	26.1		28.3		Cε 42.6	
118	Y	8.94	117.9	176.7	56.7	5.61	41.8	2.78	2.65			133.0	6.94	Cε 117.7	He 6.85
119	F	9.25	116.5		57.5	5.12	41.8	3.47	2.81			132.8	6.76	Cε 131.1	He 7.03 Hζ 7.20
120	L				55.5	4.53	45.8	1.96	0.85	27.1	1.57	24.1	0.67	Cδ' 25.7	Hδ' 0.37
121	W			174.5	56.9	5.28	31.2	3.63	3.54			128.8	7.54	Ne1 127.6 Cζ2 114.0 Cη2 124.7 Cζ3 122.0 Cε3 122.0	He1 10.54 Hζ2 7.86 Hη2 7.31 Hζ3 7.30 He3 7.94
122	V	8.56	118.9	175.4	64.0	4.21	35.0	2.13		21.2	1.02			Cγ 21.2	Hγ 0.93
123	V	8.27	121.8	174.4	64.0	3.58	31.8	1.73		21.8	0.97			Cγ 21.8	Hγ 0.31
124	K	7.74	121.8	176.2	54.4	4.78	35.6	1.61	1.42						
125	F	9.33	117.8	176.2	56.5	4.93	43.5	3.43	2.79			132.0	7.24		
126	N	9.48	118.2	175.2	54.8	4.90	39.4	3.11	3.02					Nδ2 115.3	Hδ21 8.22 Hδ22 7.12
127	S	7.61	109.3	173.8	56.7	4.73	66.7	4.19	4.04						
128	L	8.62	120.4	177.9	56.9	3.58	41.8	1.54	1.15	26.1	1.02	24.5	0.24	Cδ' 23.4	Hδ' 0.38
129	N	8.31	114.9	176.8	56.8	4.29	38.3	2.84	2.79					Nδ2 113.7	Hδ21 7.84 Hδ22 7.04
130	E	7.68	117.8	177.9	59.2	4.02	31.2	2.35	2.19	37.7	2.45	2.45			
131	L	6.88	122.6		58.5	2.12	42.4	1.68	1.14	27.4	1/36	28.0	0.83	Cδ' 23.7	Hδ' 0.48
132	V			177.0	66.4	2.75	31.1	1.45		22.3	-0.26			Cγ 21.2	Hγ 0.14
133	D	7.84	115.7	180.2	57.6	4.21	40.2	2.63	2.63						
134	Y	7.96	121.2		61.5	4.10	39.9	3.04	2.60			133.7	6.33	Cε 117.9	He 6.73
135	H			174.0	57.9	5.30	27.3	3.40	2.68			120.2	6.98		
136	R	7.46	116.2	178.3	58.0	4.61	30.2	1.87	1.87	26.8	2.00	2.00	43.2	3.26	
137	S	7.20	107.3	172.8	57.9	4.29	64.6	3.84	3.62						
138	T	7.31	120.1		63.0	4.45	70.4	3.02		21.4	1.28				
139	S			176.8	57.1	4.18	63.5	3.94	3.61						
140	V	8.39	120.1	173.7	62.7	3.81	31.9	1.37		20.3	0.31			Cγ 20.3	Hγ 0.31
141	S	7.78	111.1	174.8	55.1	5.03	64.6	3.77	3.69						
142	R	8.65	124.3		57.3	4.08	30.4	1.55	1.27	26.7	0.99	0.76	43.0	2.49	
143	N	8.17	115.2	174.1	53.5	4.78	40.0	2.78	2.68					Nδ2 112.1	Hδ21 7.56 Hδ22 6.89
144	Q	7.61	117.6	175.8	54.2	4.46	32.4	1.62	1.48	33.9	2.19	1.99		Ne2 111.7	He21 7.50 He22 6.96
145	Q	8.68	121.5	173.0	55.0	3.99	28.4	2.02	1.95	33.9	2.21	2.02		Ne2 112.8	He21 7.55 He22 6.90

146	I	8.06	124.3	172.8	60.7	3.91	40.1	1.31		25.3	0.64	0.01	13.1	0.25	C γ' 17.4	H γ' 0.53
147	F	8.50	126.2	175.5	54.6	4.94	40.4	3.12	2.68				131.0	6.86	He 7.14	H ζ 7.23
148	L	8.22	119.1	177.4	55.4	4.42	39.3	0.91	0.48	25.4	1.34		21.4	0.29	C δ' 25.4	H δ' -0.10
149	R	8.34	123.5	173.9	54.2	4.65	34.1	1.88	1.69	26.5	1.74	1.74	43.8	3.24		
150	D	8.68	121.8	178.3	55.3	4.92	40.7	2.83	2.60							
151	I	9.38	120.9	175.9	63.1	3.88	39.4	1.88		28.9	1.67	1.06	14.6	0.76	C γ' 17.4	H γ' 0.95
152	E	8.90	125.1	176.5	55.5	4.60	31.9	2.14	2.10	36.8	2.50	2.34				
153	Q	8.77	123.4	175.7	56.0	4.40	29.4	2.12	2.03	34.0	2.40	2.40			Ne2 112.1	He21 7.66 He22 6.87
154	V	8.34	123.1		59.9	4.46	32.7	2.12		21.2	1.01				C γ 20.5	H γ 0.96
155	P			176.8	63.2	4.46	32.2	2.34	1.93	27.4	2.06	2.01	51.0	3.89	H δ' 3.71	
156	Q	8.48	120.7	178.4	55.7	4.31	29.7	2.10	2.00	33.9	2.41	2.41			Ne2 112.6	He21 7.59 He22 6.89
157	Q	8.41	122.6		53.7	4.67	29.2	2.13	1.97	33.6	2.41	2.41			Ne2 112.3	He21 7.59 He22 6.89
158	P			176.4	63.5	4.53	32.0	2.34	2.03	27.3	2.05	2.05	50.8	3.81	H δ' 3.73	
159	T	7.83	119.5		63.3	4.15	70.8	4.25		22.1	1.21					
1	D					4.27		2.82	2.66							
2	D					4.31		2.07	1.75							
3	P					4.96		2.77	2.49		2.07	1.75		3.84	H δ' 3.72	
4	S	7.96				4.47		3.76	3.72							
5	pY	8.91				5.02		3.08	2.64					6.91	He 6.91	
6	V	8.87				3.58		0.93			1.03	0.66				
7	N	8.64				4.62		2.04	1.38						H δ 21 9.08	H δ 22 8.30
8	V	7.23				4.31		2.15			0.96	0.82				
9	Q	8.50				4.52		2.19	2.09		2.46	2.46				
10	N	8.64				4.83		2.93	2.83							
11	L	8.41				4.47		1.69	1.69		1.69			0.97	H δ' 0.92	
12	D	7.96				4.43		2.73	2.63							
13	K	7.76				4.21		1.74	1.43		1.87	1.43		1.74	He 3.05	

Table 2. Structural Statistics

Parameter	<SA>	<SA> _r
rmsd from experimental distance constraints (Å) (1649)	0.038 ± 0.001	0.031
Number of distance constraints violations greater than 0.5 Å	0 (max. 0.38)	0 (max. 0.28)
rmsd from experimental dihedral constraints (deg) (30)	0.17 ± 0.30	0
Number of dihedral constraints violations greater than 5.0 deg	0 (max. 4.4)	0 (max. 0.5)
F _{NOE} (kcal/mol)	117.1 ± 5.3	79.6
F _{cdih} (kcal/mol)	0.04 ± 0.10	0
F _{repel} (kcal/mol)	67.0 ± 5.3	51.0
E _{L-J} (kcal/mol)	-101.8 ± 36.1	36.4
rmsd from idealized geometry		
Bonds (Å) (1910)	0.005 ± 0.000	0.005
Angle (deg) (3438)	0.953 ± 0.015	0.873
Improp(deg) (1035)	1.658 ± 0.014	1.611

<SA> refers to the final set of simulated annealing structures, and <SA>_r is the mean structure. The number of terms is given in parenthesis.

The value of the square-well NOE potential, F_{NOE}, is calculated with a force constant of 50 kcal/mol per Å². The value of F_{cdih} is calculated with a force constant of 50 kcal/mol per rad². The value of F_{repel} is calculated with a force constant of 4 kcal/mol per Å² with the van der Waals radii scaled by a factor of 0.8 of the standard values used in the CHARMM empirical function.

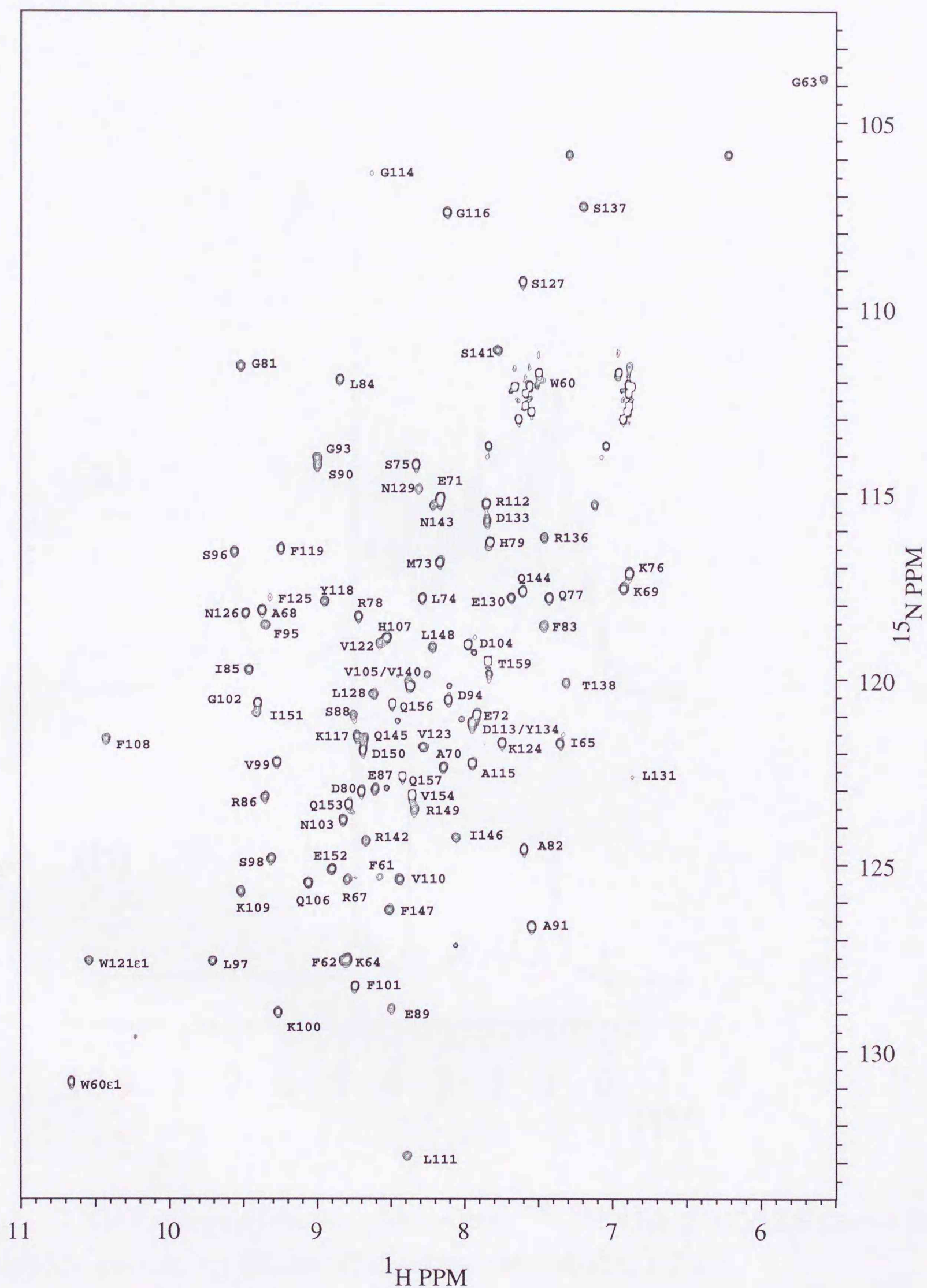


Figure 1. Two-dimensional ^1H - ^{15}N HSQC spectrum of the 1.2 mM ^{15}N -labeled Grb2 SH2 complexed with the unlabeled Shc-peptide. The peaks are labeled with their assignments.

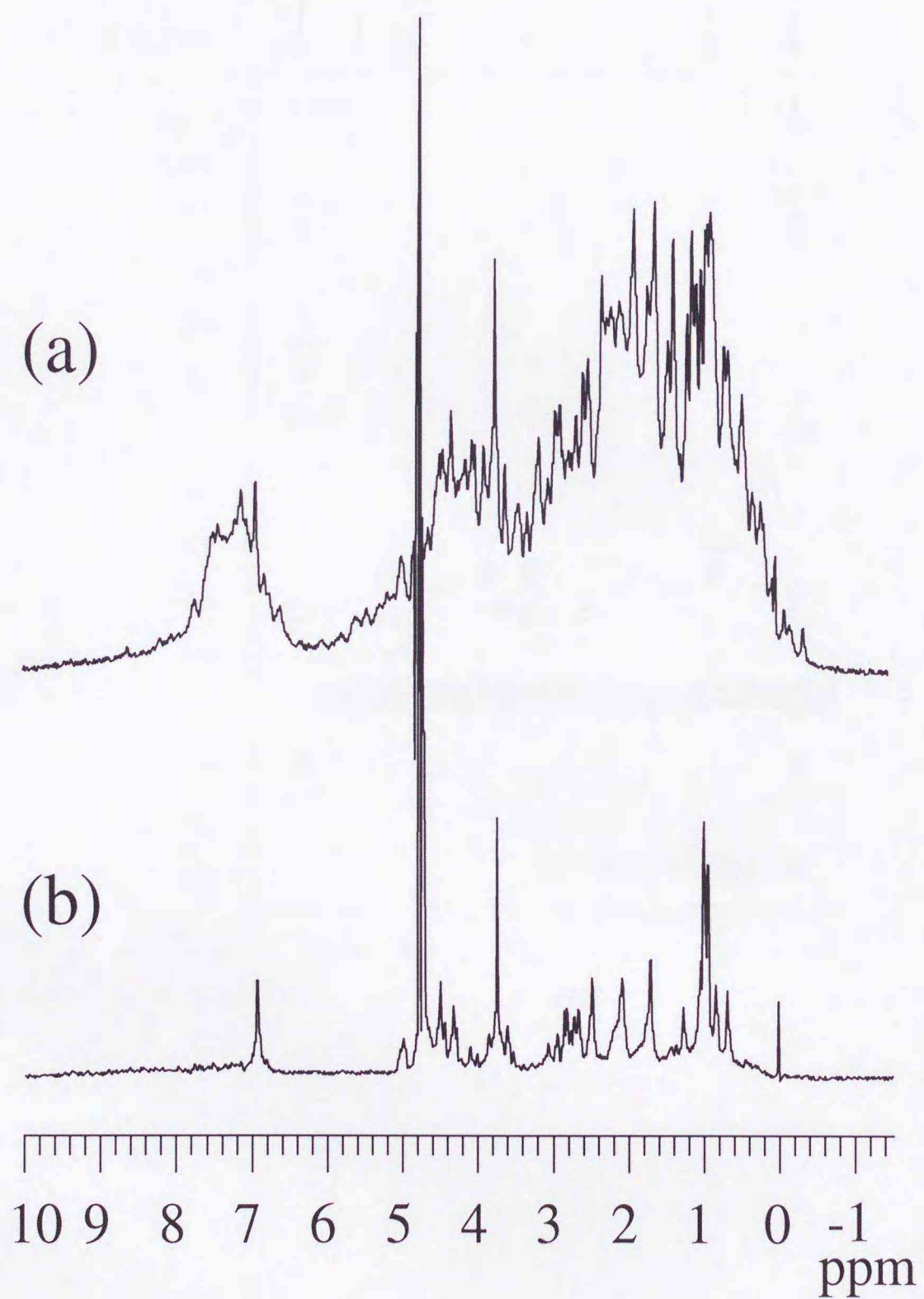


Figure 2. 1D NMR spectra of the complex between $^{13}\text{C}/^{15}\text{N}$ labeled Grb2 SH2 and the unlabeled Shc-peptide. (a) Without PFG z-filter; (b) with PFG z-filter.

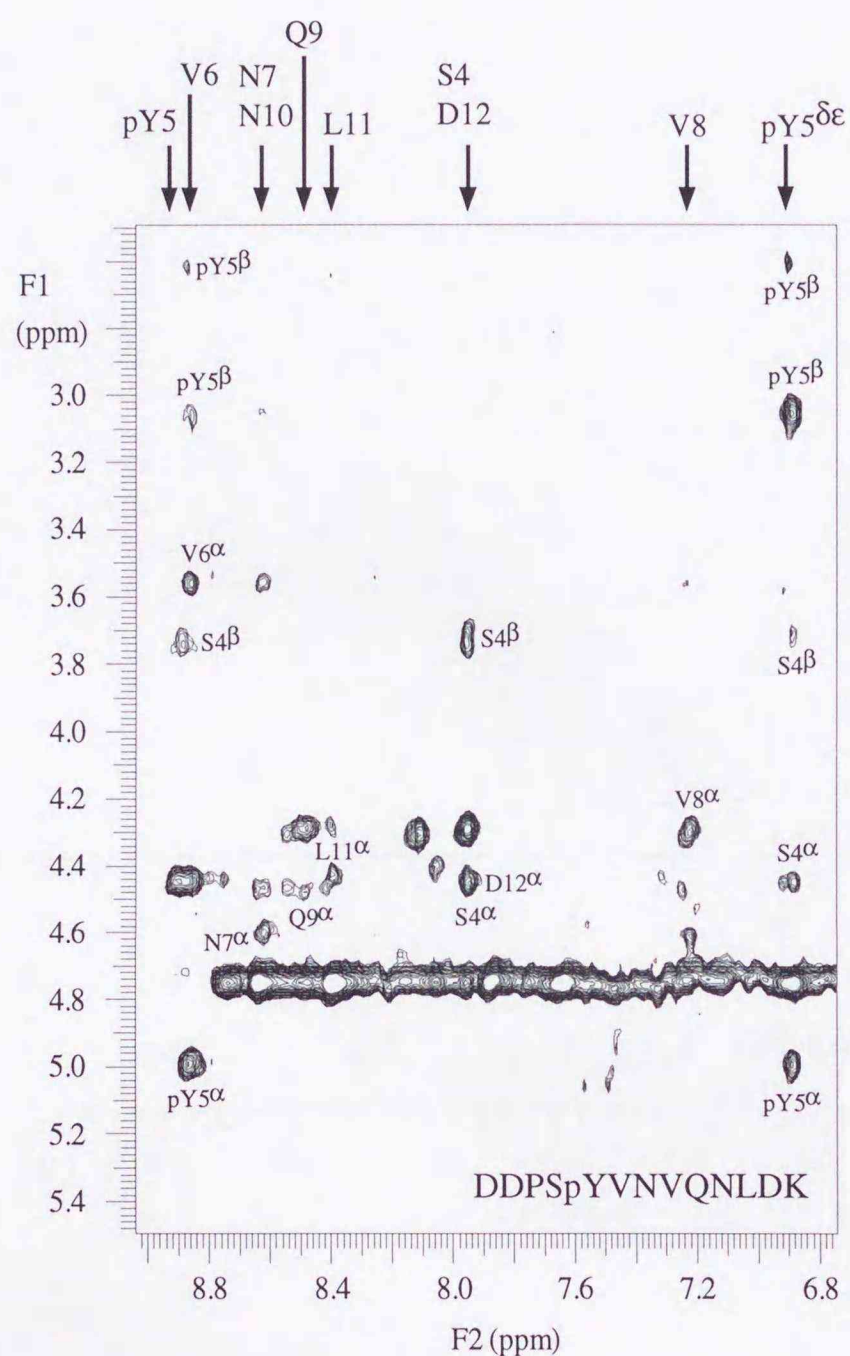


Figure 3. $C^{\alpha}H$ -NH region of the [F1, F2] $^{13}C/^{15}N$ -filtered 2D NOESY spectrum recorded with the unlabeled Shc-peptide complexed with the $^{13}C/^{15}N$ -labeled Grb2 SH2 domain in 90 % H_2O solution at pH 6.3 and 28 °C. Inter- and intraresidue NOE signals corresponding to the unlabeled Shc-peptide are shown.

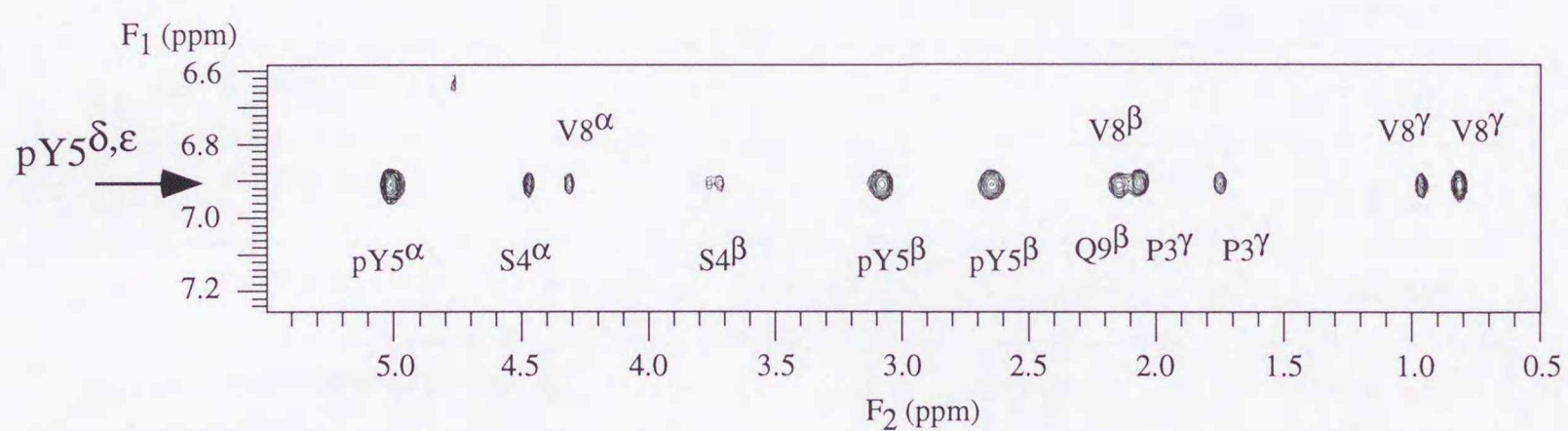


Figure 4. Aromatic-aliphatic region of [F1, F2] ^{13}C -filtered 2D NOESY spectrum of the unlabeled Shc- peptide complexed with the $^{13}\text{C}/^{15}\text{N}$ -labeled Grb2 SH2 domain recorded in $^2\text{H}_2\text{O}$ at pH 6.3 and 28 $^{\circ}\text{C}$. Intramolecular NOE signals between aliphatic protons and the protons at 6.91 ppm are shown. These protons are assigned to H^{δ} and H^{ϵ} of pTyr5.

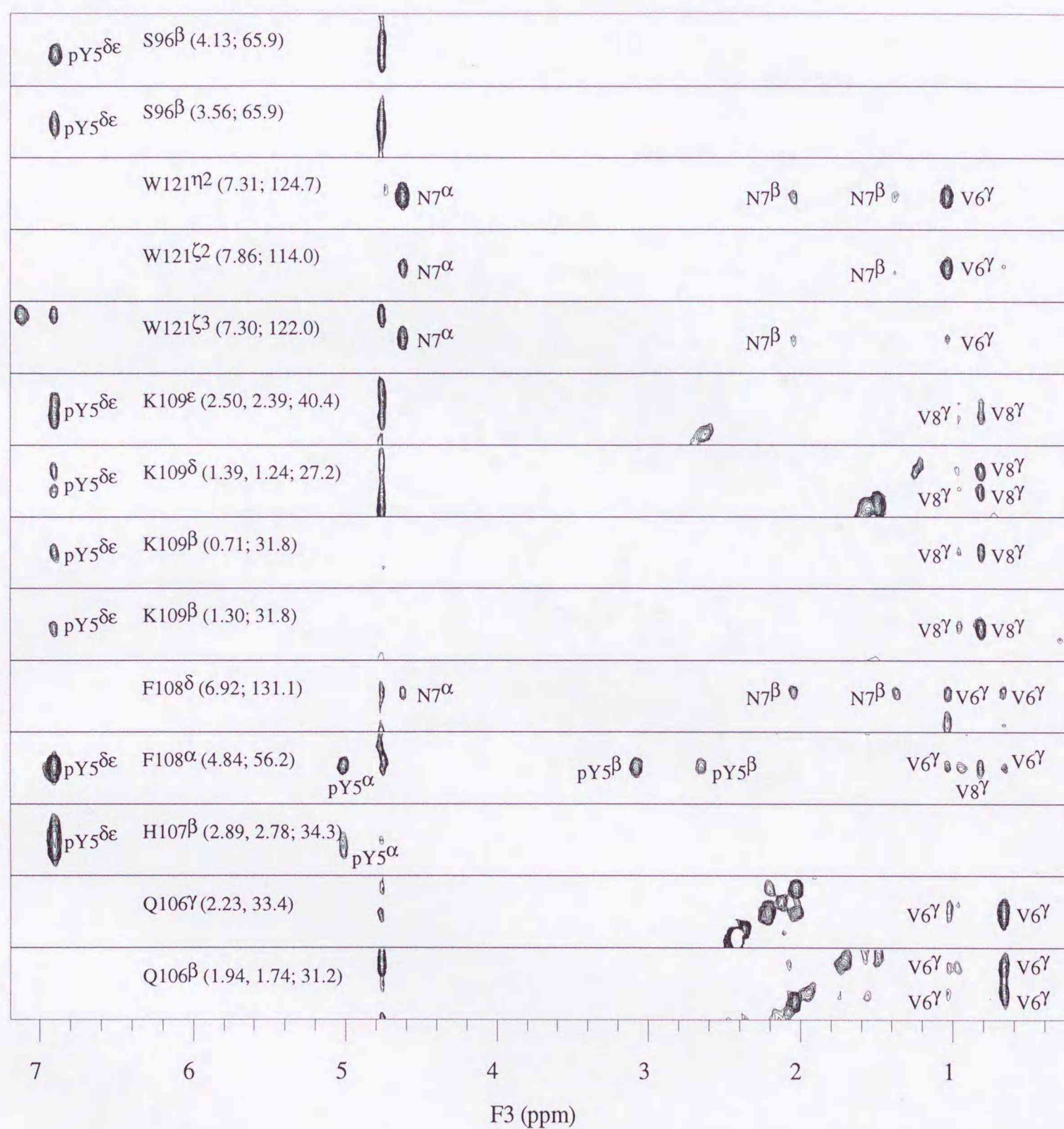


Figure 5. Strip plots of ^1H - ^1H planes extracted from [F1] ^{13}C -edited, [F3] ^{13}C -filtered 3D NOESY spectrum of the unlabeled Shc-peptide complexed with $^{13}\text{C}/^{15}\text{N}$ -labeled Grb2 SH2 recorded in $^2\text{H}_2\text{O}$ at pH 6.3 and 28 $^\circ\text{C}$. Signals at 4.7 ppm on F3 arise from the solvent. The resonance assignments, and corresponding ^1H and ^{13}C chemical shifts are given in each strip. Intermolecular NOE correlations between the Shc-peptide and Grb2 SH2 are annotated.

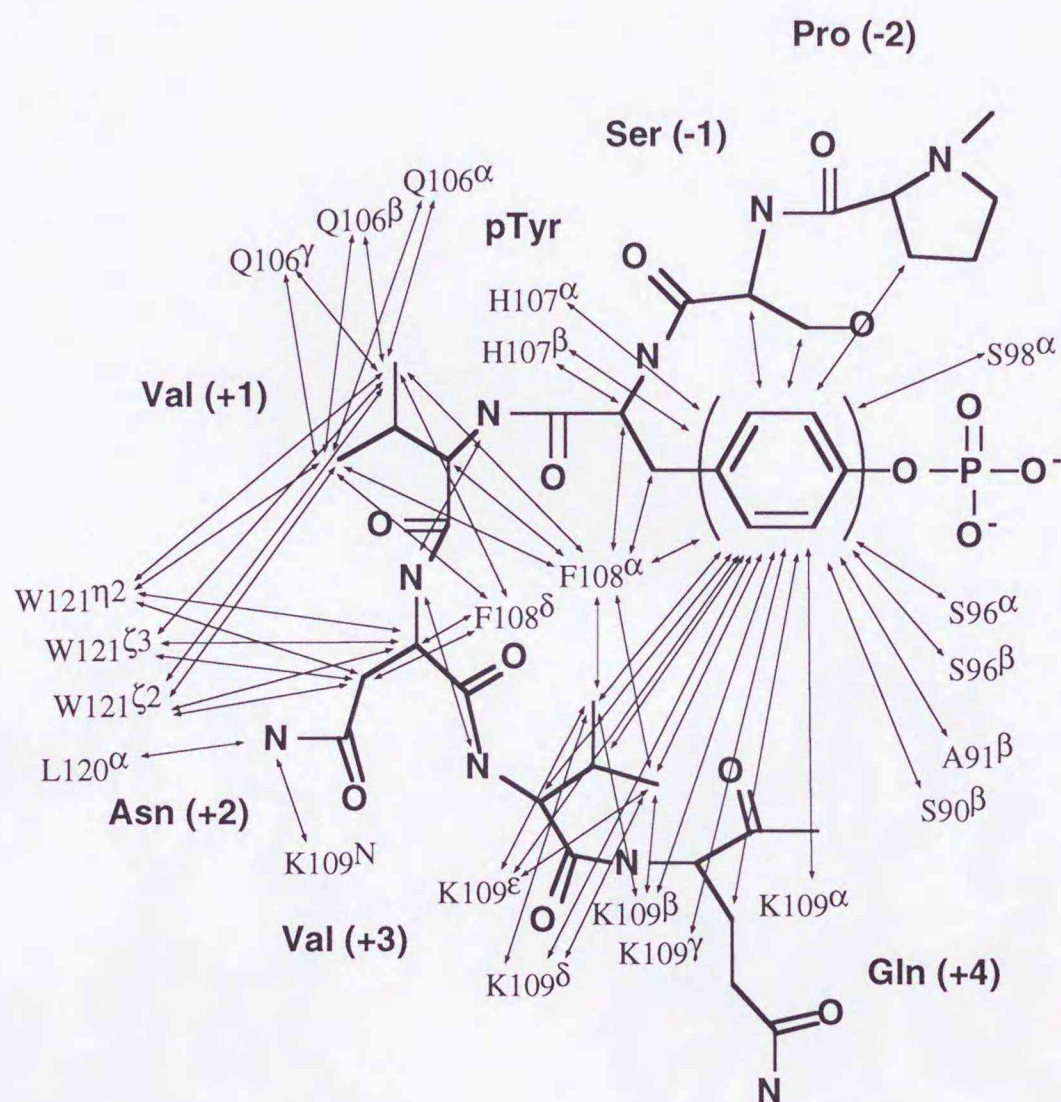


Figure 6. Schematic drawing of the intermolecular NOE interactions observed between Grb2 SH2 and the Shc-peptide. The peptide residues from Pro 3 (-2) to Gln 9 (+4) are shown. A bracket on the aromatic ring of pTyr 5 means that the chemical shifts of pTyr 5 δ and pTyr 5 ϵ are degenerate. The observed NOEs are shown as arrows.

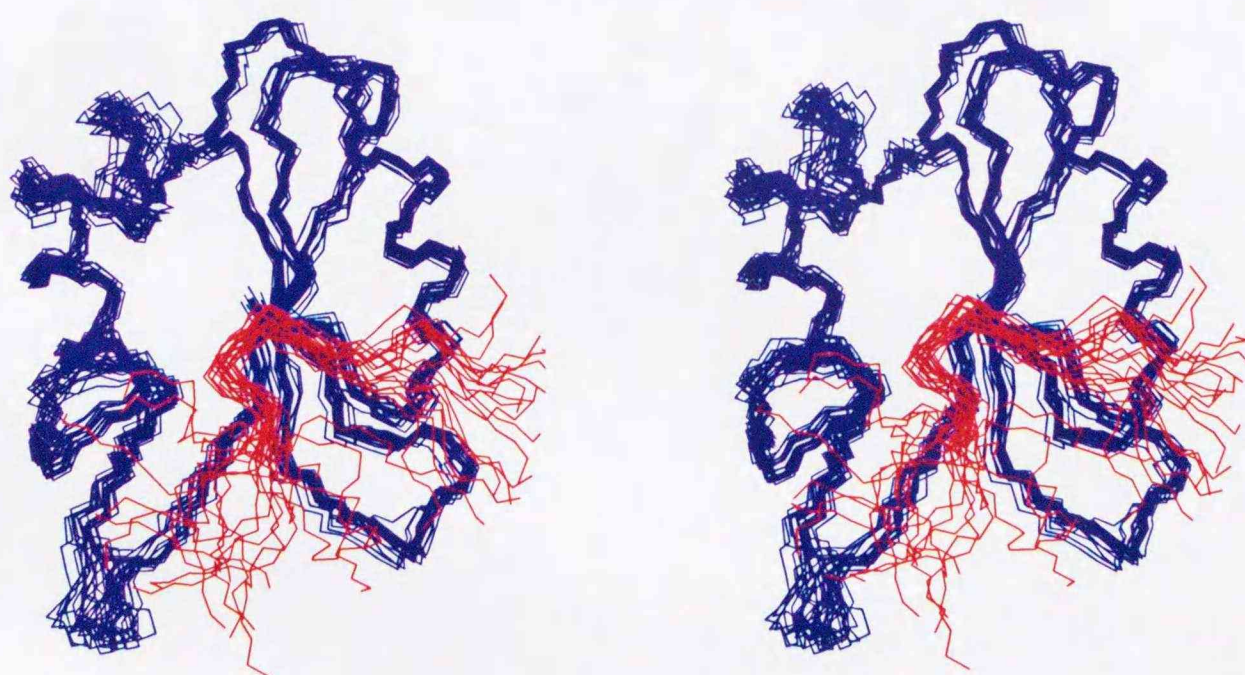


Figure 7. Stereo view of overlaid 20 NMR-derived structures of the Grb2 SH2 domain complexed with the Shc-peptide. Residues 60-152 of the SH2 are colored blue. Residues 1-13 of the peptide are colored red. Structures were superimposed by least square fit to the average N, C α , and C' coordinates of residues 60-152 of the Grb2 SH2 domain.

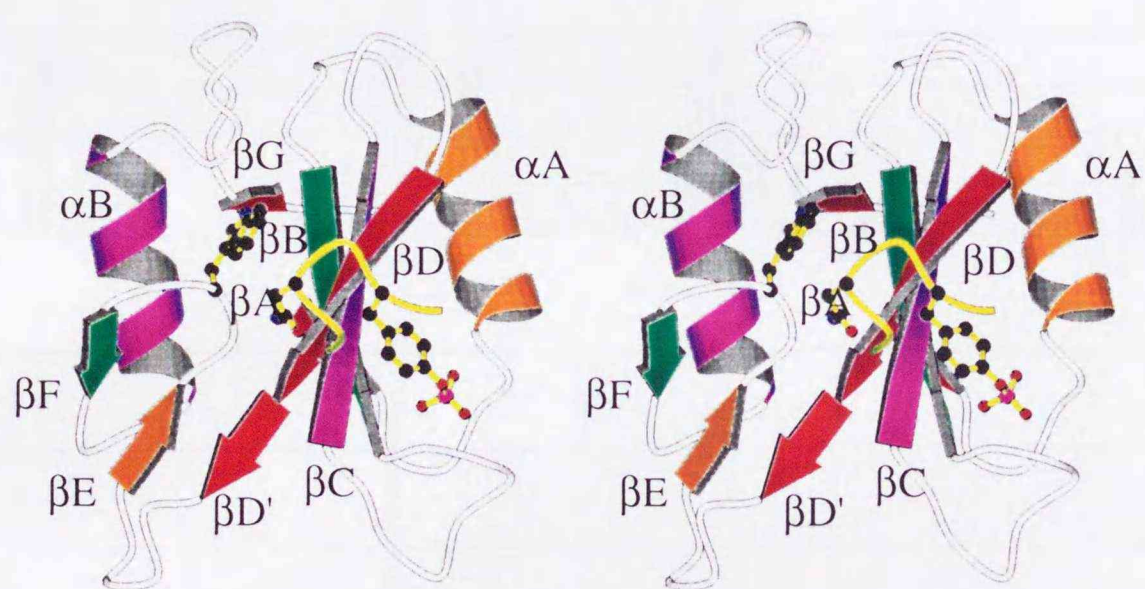


Figure 8. Stereoview of backbone and ribbon diagram of the Grb2 SH2 domain complexed with the Shc-peptide. Residues 60-152 of the SH2 and residues 4-9 of the peptide are shown. Sidechains are also shown for the following residues: Trp121 of the SH2, pTyr and Asn (pTyr+2) of the Shc-peptide. The figure was made using the program MOLSCRIPT (Kraulis, 1991). The regular secondary structure elements are labeled.

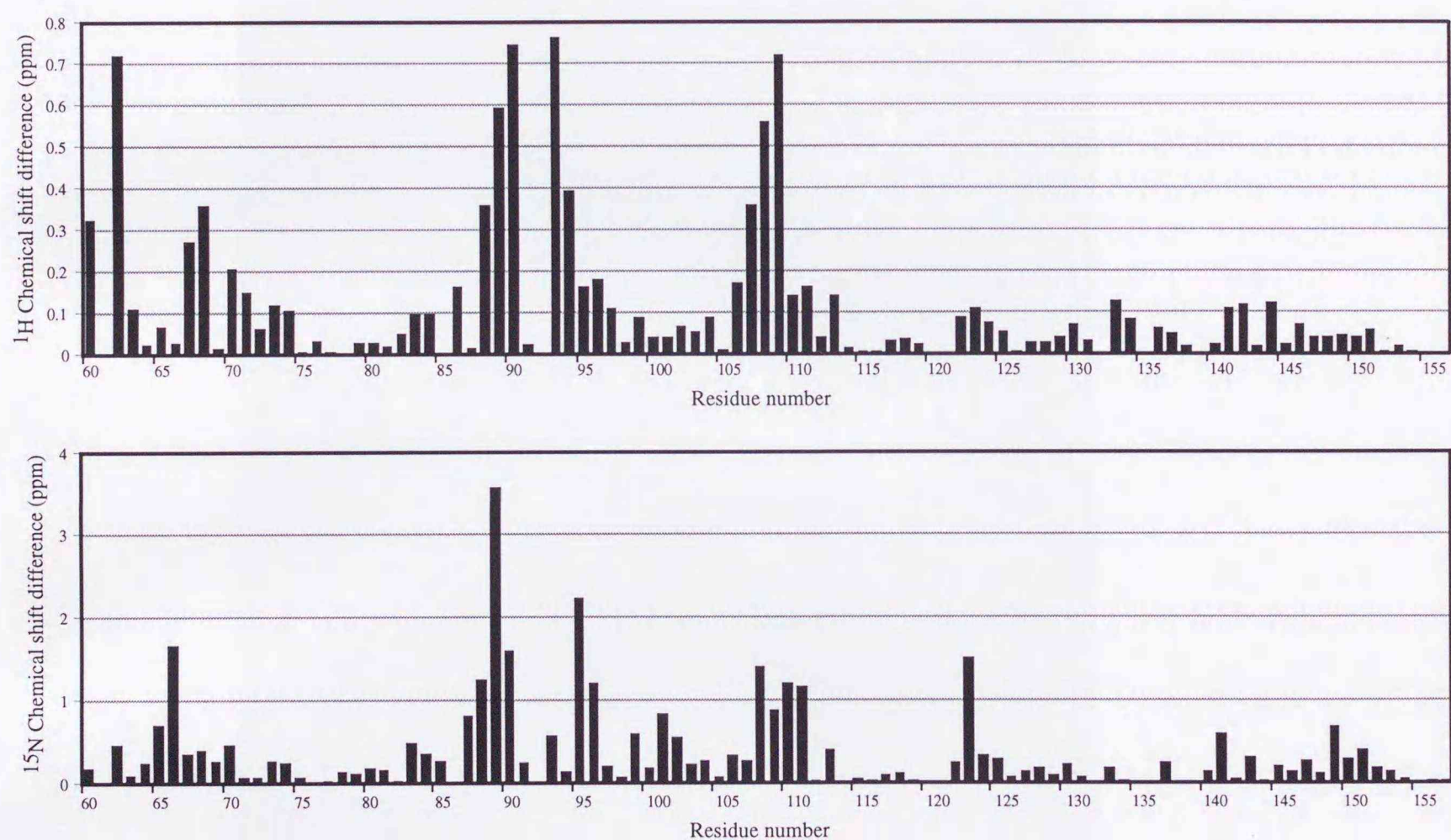


Figure 9. Chemical shift changes for each backbone amide group of Grb2 SH2 on complex formation with the Shc-peptide, where the values are presented in the absolute values. The chemical shift change was measured from the ^1H - ^{15}N HSQC spectra. Upper and lower panel includes backbone ^{15}N and ^1H shift perturbations, respectively, where the chemical shift difference is calculated by subtracting shift values for the free form of the protein from those of the peptide-bound form of the protein.

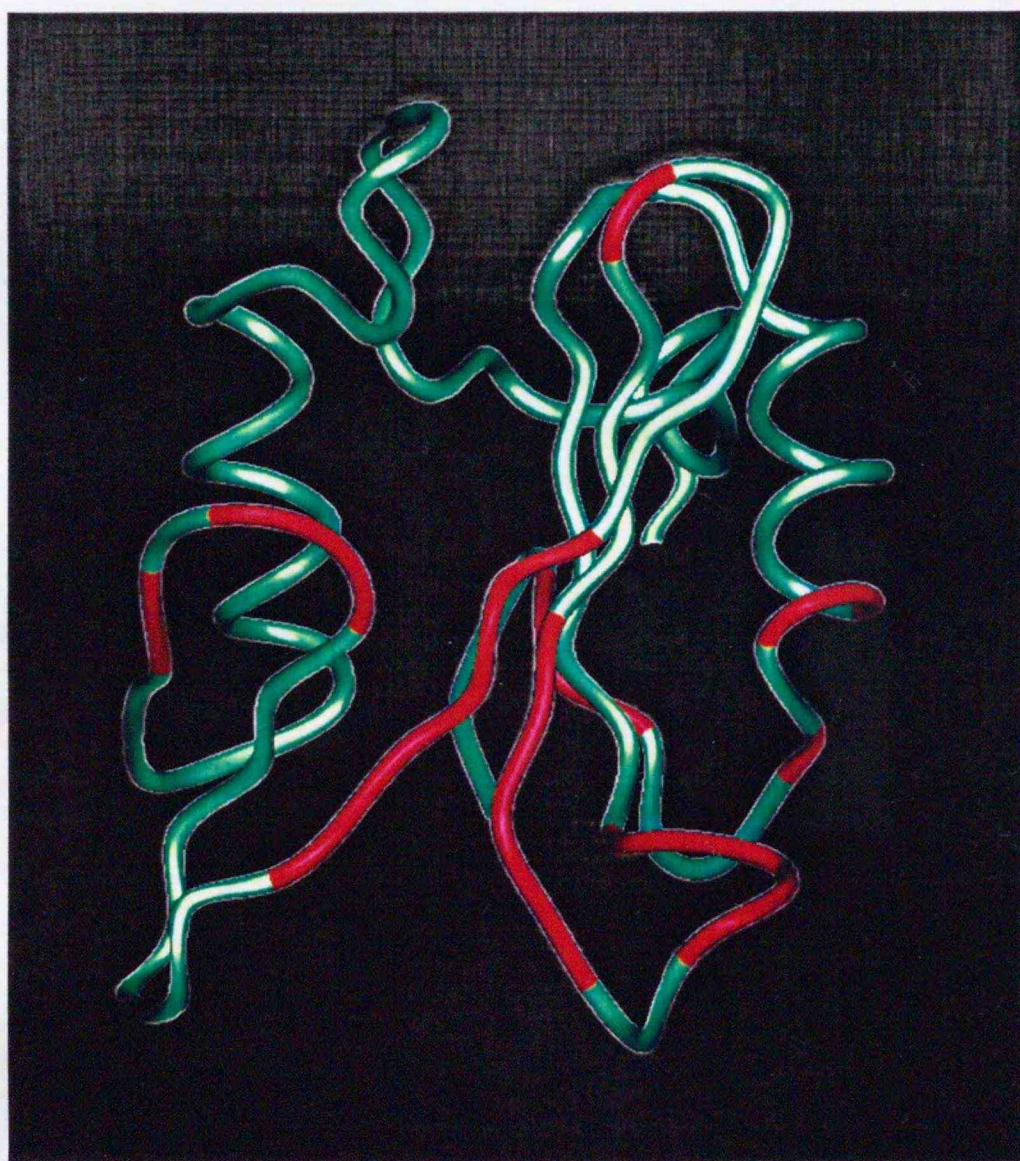


Figure 10. Correspondence of chemical shift perturbations that occur upon the Shc-peptide for the Grb2 SH2. The residues containing amide ^1H and ^{15}N in large chemical shift difference are colored red.

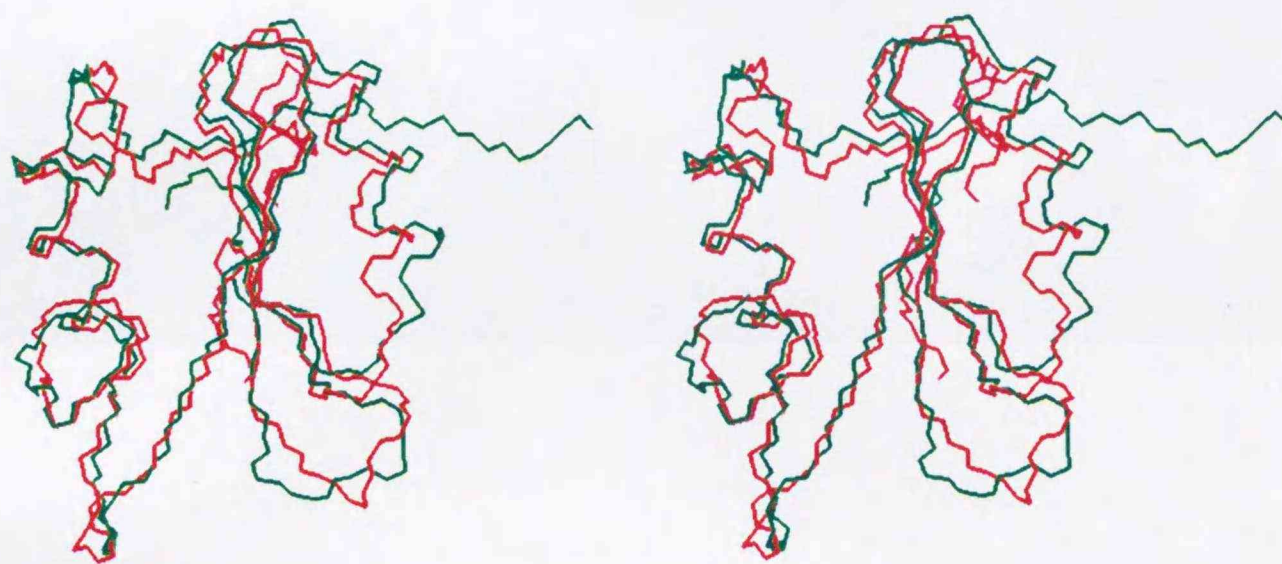


Figure 11. Stereoview of the Grb2 SH2 domain for peptide-bound (red) and peptide-free (green) structures superimposed for the backbone (N, C α , C') atoms of residues 65-145. The Shc-peptide is omitted from the present structure.

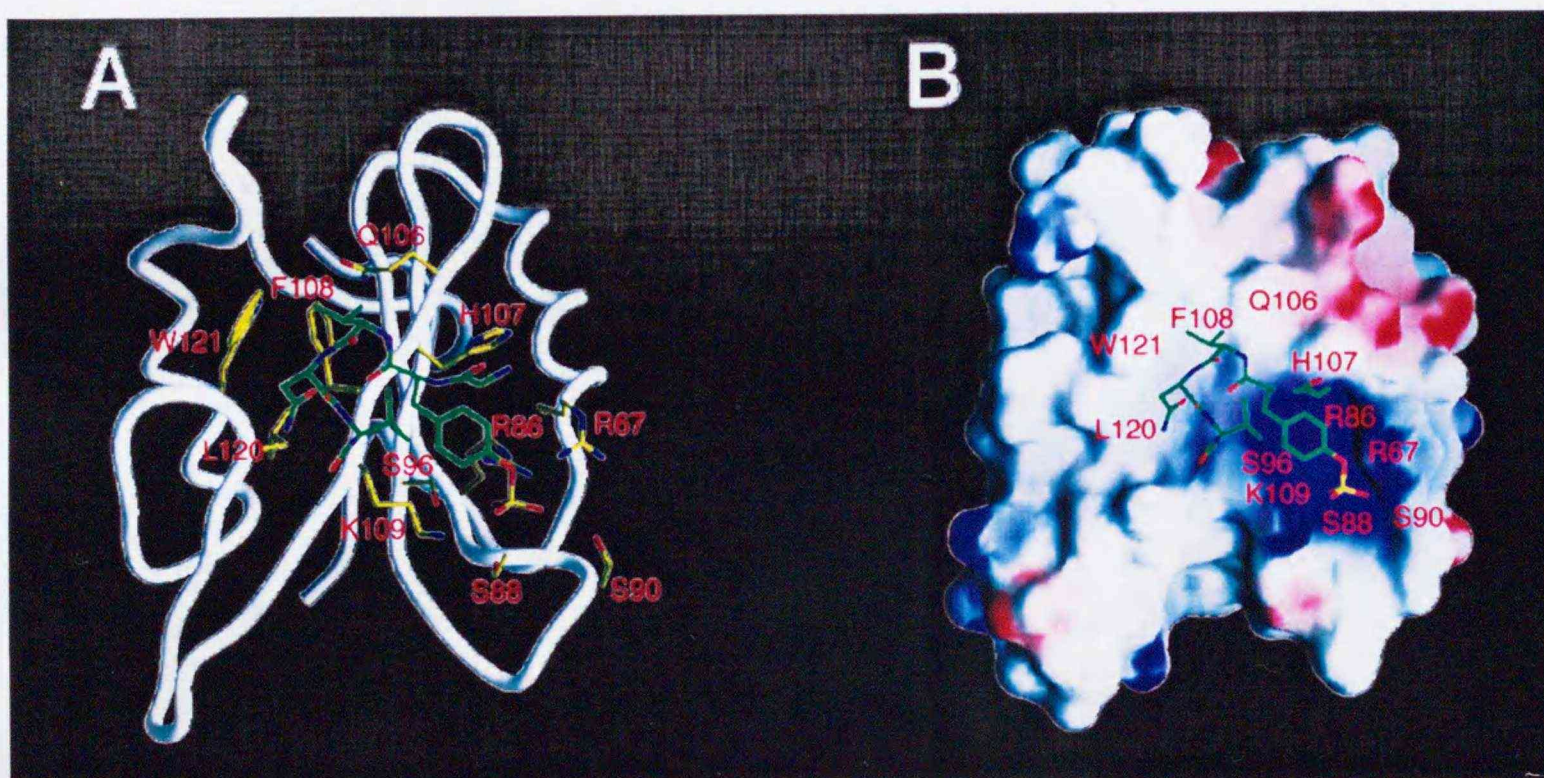


Figure 12. (A) View of the restrained, minimized averaged structure of the Grb2 SH2 domain-peptide complex. SH2 residues involved in the major binding interactions (labeled) are shown with yellow carbon atoms, with the peptide having green carbon atoms. The backbone (residues 56-159) of the SH2 domain is shown as a white tube. (B) Same view as in (A), but with the SH2 domain depicted as a molecular surface. Surface electrostatic potential is shown in red (negative) and blue (positive), with white representing neutral potential. Figure was generated using GRASP (Nicholls, et al., 1991).

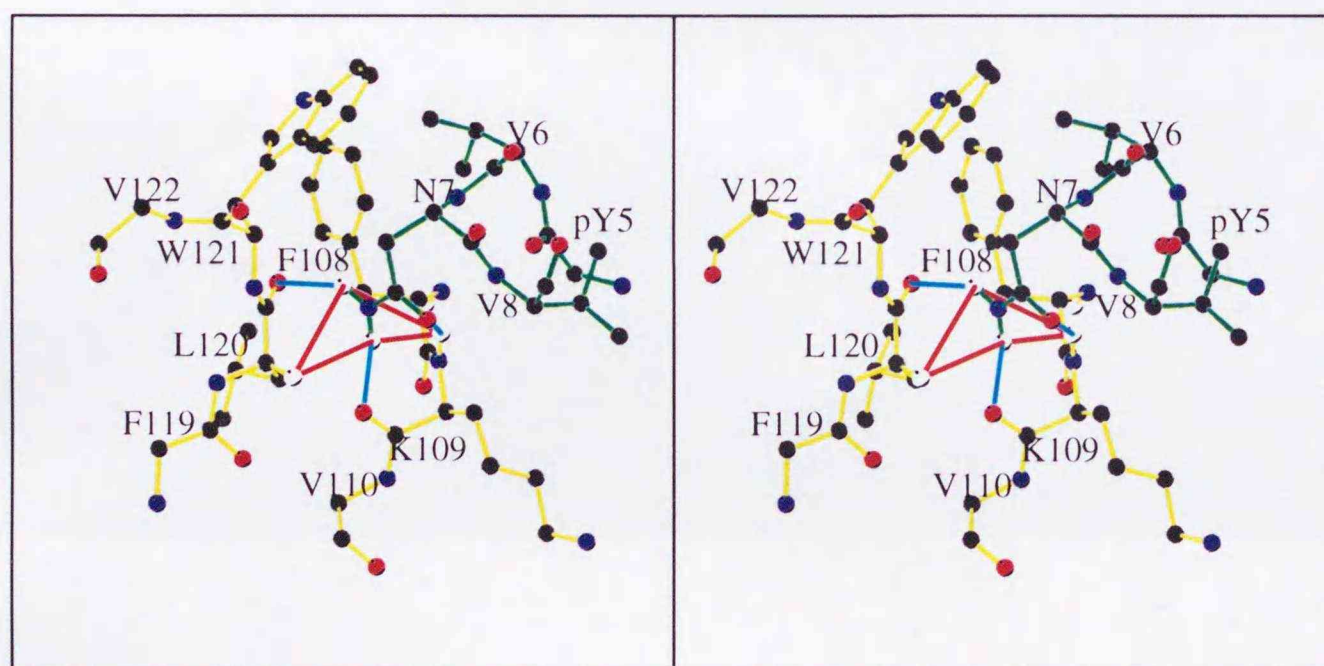


Figure 13. Stereo view of the solution structure around Asn (pTyr +2) of the Shc-peptide bound on the Grb2 SH2 domain. Residues from Val (pTyr +1) to Val (pTyr +3) of the Shc-peptide are shown in ball-and-stick mode (green). Residues of Phe108, Lys109 and Trp121 of the protein are also shown in ball-and-stick mode (yellow). Red lines indicate actually detected intermolecular NOE correlations. Blue lines indicate the hydrogen bonds between $H\delta$ of Asn (pTyr +2) and the protein, have been shown by Rafael et al. (1996). The figure was generated using the program MOLSCRIPT.

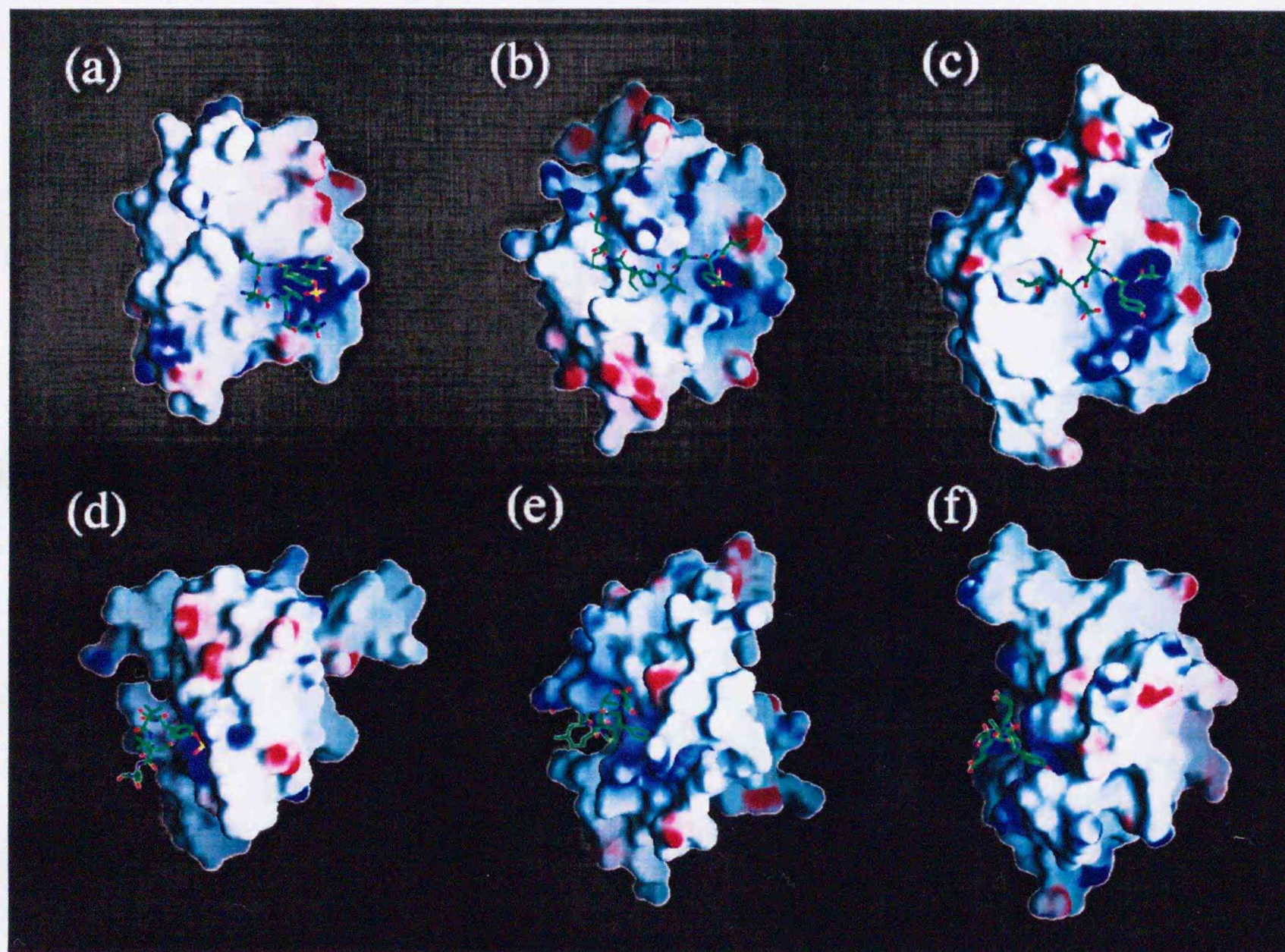


Figure 14. Molecular surfaces of the Grb2 (a,d), PLC- γ 1(b,e) (Pascal et al., 1994, PDB Entry 2PLD), and Src (c,f) SH2 (Waksman et al., 1993, PDB Entry 1SPS) domains, with bound SpYVNVQ (-1 to +4), DpYIIPLPD (-1 to +6) and QpYEEI (-1 to +3) peptides respectively, shown in green. The surface is colored by electrostatic potential (red = negative, blue = positive, white = uncharged). Orientation directly facing the peptide binding surface with the positively charged pTyr binding site on the right (a - c) and 90° rotated (d - f) are shown. Figure generated using GRASP.

Acknowledgment

I want to thank Professor Dr. K. Hikichi of Hokkaido University for accepting me as an undergraduate student of his research group in April 1987, and later for allowing me to submit the work as a dissertation.

I am grateful to Dr. F. Inagaki of Tokyo Metropolitan Institute of Medical Science for inviting me as a research scientist of his laboratory in October 1993, and for his valuable discussion and advice about my whole work.

This dissertation could not have been realized without the substantial help I received, in the forms of instrumental support or stimulating discussions, from the following persons:

Dr. H. Hatanaka, Dr. K. Nagata, H. Terasawa, S. Tsuchiya and S. Yuzawa of Tokyo Metropolitan Institute of Medical Science, Professor Dr. J. Schlessinger and Dr. V. Mandiyan of New York University, Dr. S. Tate of Tokyo Metropolitan University, Dr. H. Habuchi and Professor Dr. K. Kimata of Aichi Medical School, Dr. M. W. Ravera and Dr. M. Jaye of Rhone-Poulenc Rorer Central Research, Dr. E. Kupce and P. Sandor of Varian Instruments.

October, 1998

