Journal of Peptide Science
J. Peptide Sci. **7**: 121–127 (2001)

Influence of Thermal Motion on ^{1}H Chemical Shifts in Proteins: The Case of Bovine Pancreatic Trypsin Inhibitor

BERNARD BUSETTA, PHILIPPE PICARD* and GILLES PRECIGOUX

Unité de Biophysique Structurale (UMR 5471-CNRS), Université Bordeaux I, Talence, France

Received 28 May 2000
Accepted 30 June 2000

Abstract: The possible influence of thermal motion on ^{1}H chemical shifts is discussed for a small stable protein, the bovine pancreatic Kunitz trypsin inhibitor (BPTI). The thermal effects on the aromatic side chains and on the backbone are treated separately. The thermal motion of the aromatic side chains is accounted for in terms of their rotation around the C_α–C_β bond and the motion of each individual proton is interpreted as a ratio between the amount of ordered and quite disordered states. The influence of hydrogen bonds is introduced as an extra contribution to the chemical shifts of the bonded proton. Their contribution to the chemical shifts resulting from the polarization of the peptide bond is investigated, as is their influence on local flexibility. Finally, the relative importance of each contribution to the chemical shift information is compared. Copyright © 2001 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: ^{1}H chemical shifts; thermal motion; bovine pancreatic trypsin inhibitor

INTRODUCTION

During the past few years, NMR spectroscopy has been shown to be a powerful technique for investigating the structure of proteins and peptides [1]. Three-dimensional (3D) structures are deduced from distance information derived from nOe and dihedral angle constraints derived from coupling constants. Presently, this type of information is sufficient to determine NMR-derived protein 3D structure with the same accuracy as X-ray crystallography [2].

On the contrary, for small peptides in water, thermal motion leads to poor nOe information and extra information is strongly required to determine the peptide structure. Hydrogen bonds were previously proposed as efficient extra constraints to improve the determination of peptide structure [3]. However, if a resolved and fully assigned NMR spectrum can be obtained, considerable spectroscopic information in the form of isotropic chemical shifts is usually available and, in principle, further structural information should follow.

To date, there has been little progress in the use of chemical shifts in NMR structure determination. Due to the very large number of factors that can influence them (including, but not limited to, the presence of aromatic, charged or polar groups, the local flexibility, etc. [4]), chemical shifts remain one of the most difficult NMR parameters to interpret in structural terms. However, as they represent a mixture of coexisting conformers, they provide an experimental snapshot of the dynamics of a protein or peptide structure and remain an essential tool for investigating molecular structure.

A direct evaluation of ^{1}H chemical shifts from the protein 3D crystal structure was proposed some years ago [5], but no objective description of the thermal motion takes into account the strong influence that the flexibility of a polypeptide chain may have on their estimation [6]. For large proteins, the crystal structure appears to be a good approximation of the averaged structure of the solution state [5,7]. In the crystal state, the effects of thermal

Abbreviations: BPTI, bovine pancreatic Kunitz trypsin inhibitor.

* Correspondence to: Unité de Biophysique Structurale (UMR 5471-CNRS), Bât. B8 Avenue des Facultés, Université Bordeaux I, 33405 Talence, France; e-mail: P.Picard@ubs.u-bordeaux.fr

motion and the flexibility are less pronounced than in the solution state and, thus, can be more easily analysed. The differences between the solution and crystal states should be more pronounced for small proteins and peptides where crystal lattice packing forces and solvation play a more important role.

In this paper, we analyse the influence of thermal motion on chemical shifts, taking as an example a stable protein (bovine pancreatic Kunitz trypsin inhibitor; BPTI) that is large enough to have a well-defined structure but, at the same time, is small enough to allow the partition and differentiation of simultaneous effects. To initiate this analysis, any 3D crystal structure was presumed to provide an approximation of the BTPI averaged solution structure that was good enough to be used without any structural modification. As any structural refinement could involve an overfitting of the estimated chemical shifts and might alter the conclusions of the present analysis, we used directly all of the X-ray coordinates throughout this study. We then tried to define structure-dependent factors that may explain the thermal motion derived from the observed chemical shifts.

MATERIALS AND METHODS

X-ray coordinates used in this study correspond to the file *5pti* in the Protein Data Bank [8]. The observed chemical shifts are taken from the determination of the 3D polypeptide fold in solution [9].

^{1}H Chemical Shift Estimation

Following the process developed by Osapay and Case [5], ^{1}H chemical shifts (δ) are estimated from the protein 3D structure as the sum of elementary contributions:

$$\delta = \delta_{\text{local}} + \delta_{\text{rc}} + \delta_{\text{m}} + \delta_{\text{el}} \qquad (1)$$

where

- δ_{local} is the local contribution, approximated to the protein chemical shifts observed at 309 K for short peptides in random conformation [9];
- δ_{rc} is the aromatic ring contribution (i.e. the ring current effect), computed with Haigh–Mallion formalism ([5] and references herein);
- δ_{m} is the magnetic contribution of the peptide group, according to McConnell's estimation. In the case of axial symmetry, this contribution is reduced to the trigonometric expression $3K\cos^2\Theta$, where Θ is the angle between the vector connecting the amide group to the proton in question and the vector normal to the amide plane, and K is Flygare's constant [5]; and
- δ_{el} is the electrostatic contribution, restricted to the contribution of the main chain as suggested in Osapay and Case [5].

The solvent contribution is limited to the extra corrections $\Delta\delta$, which are necessary when taking hydrogen bonding into account (see 'Results and Discussion' section). In this paper we suppose that, besides the thermal motion and hydrogen bonding effects, no other contribution is necessary to explain the variations of the chemical shifts. For instance, the relative solvent exposure of HN protons not engaged in a hydrogen bond is not taken into account, as an evident explanation for its influence on the chemical shifts is lacking.

The efficiency of the chemical shift estimation differs greatly according to the nature of the proton being considered. For H_X side chain and H_α protons, an error within a 0.23–0.28 ppm range is retrieved as described in Williamson *et al.* [7]. The HN protons that are carried by an N atom behave differently depending on whether they are engaged in an intramolecular hydrogen bond (in the following, they are referred to as H_N) or remain exposed to a solvent (H_n). Because the BPTI chemical shifts were collected at the same 309 K room temperature, no temperature-dependent correction was applied to the HN protons.

To appreciate the overall estimation of chemical shifts, a function combining the average absolute errors on the estimated chemical shifts obtained for each proton family has been used:

$$\mathsf{F} = \langle \Delta\delta H_n \rangle + \langle \Delta\delta H_N \rangle + 2(\langle \Delta\delta H_\alpha \rangle + \langle \Delta\delta H_X \rangle), \qquad (2)$$

where $\langle \Delta\delta H_i \rangle = (\Sigma|\delta H_{i\text{obs}} - \delta H_{i\text{calc}}|)/(\text{number of assigned } H_i)$.

Simulation of Thermal Motion

In the basic estimation [5], thermal motion is not taken into account. In the following, thermal motion is introduced in the estimation of the chemical shifts as the superposition of two different effects:

(i) a rotation of aromatic side chains around the C_α–C_β bond (χ^1 angle). Different j positions, defined by a χ^1_j torsion angle associated with a weight p_j, are considered. They are initially determined by a 10° scan over χ^1 monitored by

the F value from Equation (2), and then the set of the initial locations (p_j, χ_j^1) is refined with a simplex refinement process; and

(ii) a local vibration of each proton around its average position which modifies its location relative to the whole polypeptide chain and mainly affects the magnetic contribution by averaging the trigonometric factor. The full δ_m contribution, which affects a given proton of the reference 3D structure in the Osapay and Case basic estimation, should be modified into $r * \delta_m$ by the adjunction of a randomization r factor. Subsequently, following its location in the protein structure, each proton displays a specific motion and its randomization factor r should be expressed as a product:

$$r = r_o * r_i,$$

where r_o represents the overall flexibility of the conformer and r_i represents the relative randomization of this proton. The overall randomization r_o factor is refined together with the aromatic cycle locations and weights, starting from an initial ($r_o = 1$) value. For each residue type or specific location, the variations of the error function F were followed during a scan around the $r_i = 1$ value with a 0.05 step.

RESULTS AND DISCUSSION

Rotation of Aromatic Side Chains

Because ring current effect is the most important contribution to the chemical shifts, we first investigated the influence of the rotation of aromatic side chains. Their positions were determined according to the fingerprint left behind when they are freely rotating around the χ^1 torsion axis. Up to three different positions were considered with a suitable weight p to account for the rotation of these chains. As reported in Table 1, the possibility of rotation is dependent on the position of the aromatic side chain in the structure. When they are quite buried, no rotation is possible and a single apparent position close to the actual one is obtained with a relative weight around 1 (this is the case for Phe22, Phe33, Tyr35 and Phe45 in BPTI). On the contrary, when they are partly exposed to solvent, rotation is possible and a minimum of two positions with roughly similar weights is required to explain the observed fingerprint (Tyr21 and Tyr23).

It should be noted that refinement of the ring positions, as used in this paper, involves an overfitting of the estimated chemical shifts on the observed ones and may conceal or modify other weaker contributions. Owing to the importance of the contributions due to ring current effects, the use of chemical shift information in a structure refinement process cannot be straightforward. It is necessary to define the exact positions of the rotating side chains before using chemical shift information to improve an initial conformation.

Table 1 Final Positions for each Aromatic Ring in BPTI provided by the Simplex Minimization of the Differences between the Observed and the Estimated ^{1}H Chemical Shifts

Residue	p_1	χ_1^1 (°)	p_2	χ_2^1 (°)
Phe4	0.52	45	0.22	−42
Tyr11	0.76	0		
Tyr21	0.74	33	0.41	−89
Phe22	1.30	−27		
Tyr23	0.57	0	0.54	−135
Phe33	1.05	−14		
Tyr35	0.90	0		
Phe45	1.17	−15		

The position(s) of each ring is defined by a torsion angle χ_1^j associated with a weight p_j. Two locations are used during the refinement when they are suggested by the scanning process.

Influence of Hydrogen Bonding on the Chemical Shift of HN Protons

When the ring current effects are thoroughly corrected, large discrepancies subsist on the estimation of the chemical shift of HN protons and extra corrections are needed. The chemical shifts of hydrogen-bonded H_N protons were already shown to be strongly dependent on the length d_h of the hydrogen bond and a $1/d_h^3$ or $1/d_h$ linear dependency was previously suggested [10,11]. The use of a correction such as

$$\Delta\delta(d_h) = k[19.0(1/d_h^3) - 2.3], \quad (3)$$

where k is equal to -0.85 and d_h is defined for a given HN proton as the distance to the closest electron donor in the polypeptide chain, provides the most suitable estimations. However, during this analysis, it was not possible to rule out completely a $1/d_h$ linear dependency.

No correction ($\Delta\delta_h = 0$) of the estimated H_N chemical shifts is necessary if the hydrogen bond length is equal to 2.03 Å, a value that should be considered as the average bond length with the solvent in

random coiled peptides. Such a correction accounts for the large up-field deviations of the HN chemical shifts involved by large d_h lengths (2.5–3.0 Å), which are not usually considered as possible hydrogen bond lengths. The influence of thermal motion at this level is difficult to define; it should correspond to a variation of the d_h length. An accurate analysis of the hydrogen bond contribution on more flexible peptides is then required.

On the another hand, the estimation of both H_n and H_N chemical shifts is strongly improved if we consider that a polarization effect occurs through the peptide bond when the carbonyl oxygen is hydrogen bonded. Similar effects have already been observed by Llinas and Klein [12]. In this case, a coulombic $1/d_h^2$ correction should be applied on both H_N and H_n protons as soon as an effective hydrogen bond that is shorter than the average bond with the solvent is observed:

$$\Delta\delta_{co}(d_h) = 2.60[4.0(1/d_h^2) - 1].$$

In the case of BTPI, Equation (3) still holds for d_h distances up to 2.9 Å. Use of the HN chemical shifts in a structure refinement requires analysis of the cut-off distance that should be used to distinguish the bonded H_N protons from the non-bonded H_n ones. Obviously, a weak or an important motion of the protein or peptide may affect this cut-off distance.

Remaining errors on the estimation of the H_n chemical shifts could be related to solvent accessibility. Unfortunately, it seems to be very difficult to derive objective rules to realize a convenient correction for this phenomenon.

Individual Randomization Factor

When ring current and hydrogen bond effects are corrected, the remaining discrepancy between the observed and the estimated chemical shifts may be attributed to the relative motion of the protons within the protein structure, which involves a modification of the δ_m magnetic contribution due to the different peptide groups. At this stage, the difference between observed and estimated chemical shifts provides information on the relative motion of the different protons. For this study, we considered only two factors to explain the relative thermal motion of a given proton:

- the hydrogen bond pattern; and
- the nature of the residue that harbours the considered proton or the nature of the adjacent residue.

Those protons that were prevented from any particular motion by the formation of a nearby hydrogen bond were used to define the overall randomization of the protein and were arbitrarily associated with an individual randomization factor $r_i = 1$ (i.e. the H_N and H_n protons adjacent to a hydrogen-bonded carbonyl group, the H_α protons adjacent to a hydrogen-bonded HN or carbonyl group). For the prolyl residue, estimation of the chemical shifts of H_n protons is considered to be optimal when the same thermal motion ($r_i = 1$) is applied.

The relative agitation of the other protons is then determined in comparison with the overall agitation of the protein. A modification of the relative motion of a group of residues is assumed to be satisfactory when the resulting estimation of the chemical shifts is improved. When a disulphide bridge is formed, a better estimation of the chemical shifts of the non-HN protons of the cysteines is obtained when they are less agitated and when a greater value of the individual randomization factor is considered ($r_i =$ 1.25). In the absence of any adjacent hydrogen bonds, the backbone is not stabilized and should be more agitated ($r_i = 0.7$). The glycyl residue is usually considered to be a disturbing element. The H_α protons of a glycyl appear to be more agitated than the protein ($r_i = 0.7$) if this or the adjacent residues are not hydrogen-bonded; the same strong agitation ($r_i = 0.7$) is used for the protons of the amino acid that follows a glycine if it is not hydrogen-bonded.

For all amino acids, the motion of the side chain protons is usually greater than the motion of H_α atoms ($r_{H_X} = 0.7 * r_{H_\alpha}$). In the inner part of BPTI (residues 30–38), the side chains are closely packed together and form a strong, poorly agitated hydrophobic core. For Tyr35, all of the ring protons may be distinguished in NMR spectroscopy [2], suggesting no possible motion at the level of this aromatic ring in solution. The absence of side chain motion is communicated to the corresponding H_α of the 30–38 portion of the polypeptide chain and a suitable estimation of their chemical shifts required the weakest agitation ($r_i = 1.30$).

Influence of Thermal Motion on the Chemical Shift Estimation

To illustrate the importance that the different points discussed above have on the estimation of chemical shifts, we report in Table 2 the average error on the estimation of ^1H protons of BPTI solution structure after the following calculation steps were successively performed:

Table 2 Averaged Errors (in Å) on the Estimation of the ^{1}H Chemical Shift for BPTI

	F	$\langle \Delta\delta H_n \rangle$	$\langle \Delta\delta H_N \rangle$	$\langle \Delta\delta H_\alpha \rangle$	$\langle \Delta\delta H_X \rangle$
(a) Osapay and Case estimation	1.770	0.462	0.477	0.229	0.186
(b) Ring current optimization	1.514	0.359	0.470	0.186	0.155
(c) Hydrogen bond contribution	1.193	0.359	0.148	0.187	0.155
(d) Polarization of peptide group	1.067	0.237	0.145	0.187	0.155
(e) Individual randomization	0.879	0.215	0.135	0.141	0.124

(a) basic estimation corresponding to Osapay and Case's [5] approach with an overall randomization factor $r_o = 1$. All aromatic rings are located as they occur in the X-ray structure and all of the ring currents have the same weight ($p_j = 1$);
(b) optimization (location and weight) of the ring current effects and refinement of the overall randomization factor (final value $r_o = 1.14$);
(c) influence of the hydrogen bond (direct length-dependent effect on the H_N protons);
(d) influence of the hydrogen bond (polarization effect through the peptide bond on both H_N and H_n protons); and
(e) improvement finally brought about by the use of individual randomization factors.

The errors on the hydrogen-bonded H_N protons are comparable with the corresponding errors on the H_α and H_X protons. Taking into account the inaccuracies of the estimation, which have the same effect on all of the different protons, it may be thought that the X-ray hydrogen bond lengths are, in the main, compatible with the average bond lengths which appear in the solution state.

Another illustration of the improvement provided by the present analysis is given in Figure 1, where our final estimations are compared for each H_α and HN proton with the basic Osapay and Case estimation. Figure 1 shows clearly the poor estimations of the chemical shifts obtained for the non-bonded H_n protons when local fluctuations of the polypeptide chain occur, as is the case for the Cys_{14}–Cys_{38} disulphide bridge. Only a mixture of different conformations of these flexible loops could explain the observed chemical shifts.

CONCLUSION

In this study, some structure-dependent rules are derived to define the relative flexibility of the different protein portions and to estimate the relative thermal motion of each proton of BTPI in solution. Despite the fact that an X-ray structure, which surely differs from the solution structure, was used without any coordinate refinement throughout this study, the current estimation of the chemical shifts is strongly improved when thermal motion is taken into account. This relative flexibility in the protein is mainly dependent on the location of the hydrogen bonds and the nature of the local residues.

On the another hand, we showed that the HN proton chemical shifts are strongly dependent on the formation and the strength of hydrogen bonds. A careful analysis of their contribution should be performed in order to make available the distance information that is present in the HN chemical shifts. However, large errors persist in the estimation of the chemical shifts of non-bonded H_n protons (i.e. at the level of a flexible loop). If the H_n protons are likely to be dependent on the local conformation of the very amino acid that harbours them (i.e. on the ϕ angle) and on their solvent accessibility, these contributions will be difficult to quantify. For these H_n protons, it is very difficult to separate and analyse with accuracy the different contributions (ring current effects, local magnetic contribution, hydrogen bonding, temperature effects, etc.) that are superimposed in a large protein.

If the ^{1}H chemical shifts can be correctly interpreted, they are convenient parameters to estimate the validity of any 3D structure: H_α chemical shifts rather contain information on the validity of the overall geometry of the protein, whereas the HN ones contain information on solvent accessibility.

If the influence of thermal motion on chemical shift variations is important for large, stable protein structures, where the formation of a hydrophobic core prevents the free rotation of side chains, it must be expected that this influence is more pronounced in a peptide. In this case, can we always use a single conformation to describe the solution state? For a peptide, solvent accessibility is more important. Thus, at the hydrogen bond level, competition with the surrounding solvent is greater and

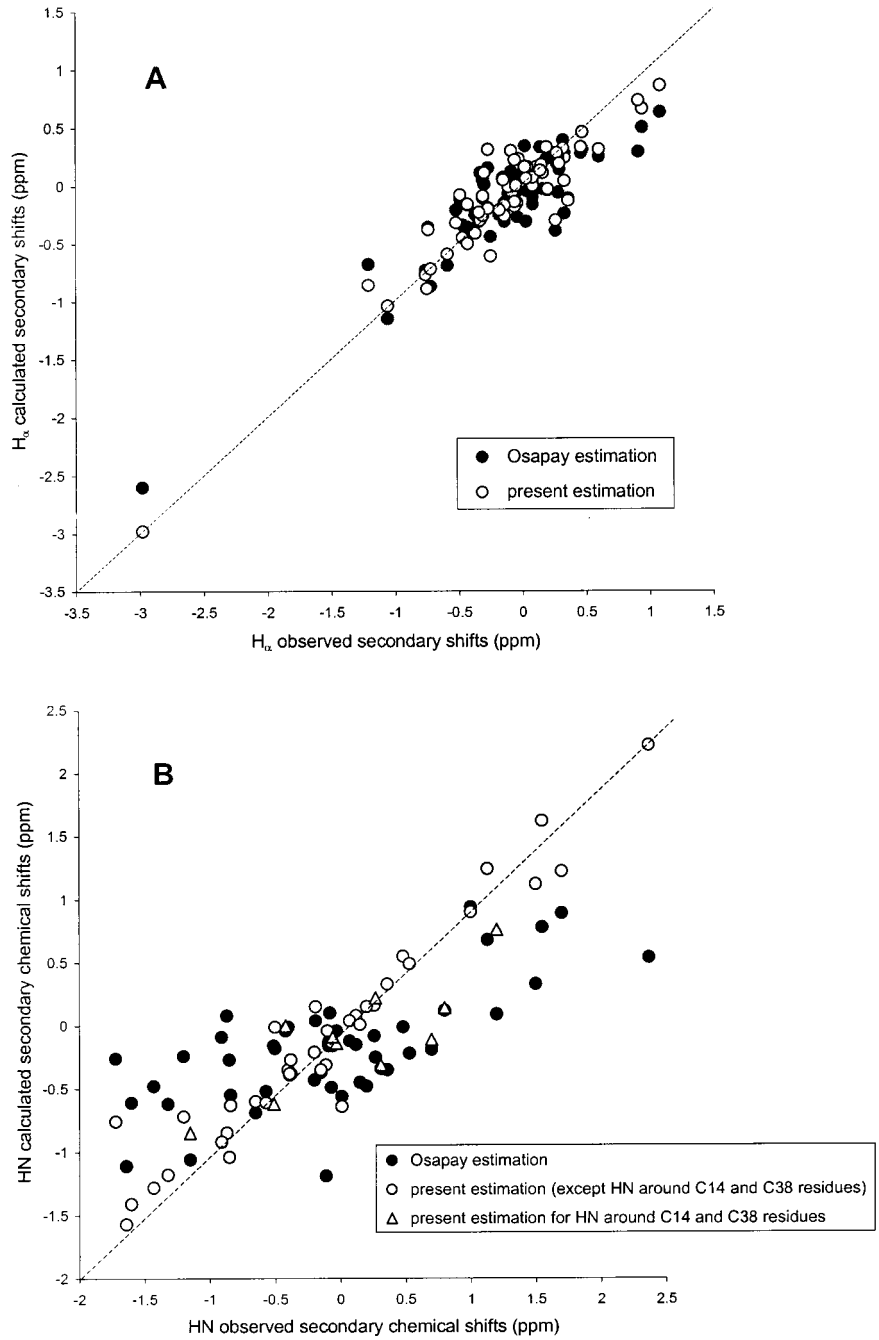

Figure 1 A comparison of the estimated secondary chemical shifts ($\delta_{calc} - \delta_{local}$) *vs.* the observed ones ($\delta_{obs} - \delta_{local}$) in the Osapay and Case estimation (full circle) and with our improved estimation (open circle) for the (A) H_α and (B) HN protons. Estimations of the HN protons around the flexible Cys_{14}–Cys_{38} disulphide bridge are represented by an open triangle.

it may be expected that the cut-off distance to accept an intramolecular bond would be shorter than in proteins and would be strongly dependent on the relative flexibility of the peptide structure. The weight of the current rings, when they are freely rotating, may be different and their effects in peptides may be reduced. Thus, in this case, thermal effects should be carefully analysed before any conformational information is derived from the chemical shifts. In a follow-up paper, we intend to

determine a possible use for chemical shift information when the solution structure of a small flexible peptide is studied.

Nevertheless, it is very apparent that conformational information provided by chemical shifts cannot be directly introduced in a structure refinement process as extra constraints, together with the other NMR-derived constraints. They are too dependent on the thermal motion. Following the definition of this motion, quite different refined 3D structures could be obtained. Thus, a different approach must be found for use of chemical shift information.

REFERENCES

1. Wüthrich K. *NMR of Proteins and Nucleic Acids*. Wiley: New York, 1986.
2. Berndt KD, Güntert P, Orbons LPM, Wüthrich K. Determination of a high quality nuclear magnetic resonance solution structure of the bovine pancreatic trypsin inhibitor and comparison with three crystal structures. *J. Mol. Biol.* 1982; **227**: 757–775.
3. Mierke DF, Geyer A, Kessler H. Coupling constants and hydrogen bonds as experimental restraints in a distance geometry refinement protocol. *Int. J. Peptide Protein Res.* 1994; **44**: 325–331.
4. Wishart DS, Sykes RD, Richards FM. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* 1991; **222**: 311–333.
5. Osapay K, Case DA. A new analysis of proton chemical shifts in proteins. *J. Am. Chem. Soc.* 1991; **113**: 9436–9444.
6. Wishart DS, Sykes BD. Chemical shifts as a tool for structure determinations. In *Methods in Enzymology*, ***239***, Oppenheimer NJ, James TL (eds), 1994; 363–392.
7. Williamson MP, Asakura T, Nakamura E, Demura M. A method for the calculation of protein α-CH chemical shifts. *J. Biomol. NMR* 1992; **2**: 83–98.
8. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers R, Kennard O, Schimanouchi T, Tasumi M. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 1977; **112**: 535–542.
9. Wagner D, Braun W, Havel TF, Schaumann T, Go N, Wüthrich K. Protein structures in solution by nuclear magnetic resonance and distance geometry: the polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol.* 1987; **196**: 611–639.
10. Bundi A, Wüthrich K. ^{1}H-NMR parameters of the common amino-acid residues measured in aqueous solutions of the linear tetrapeptides H-Gly-Gly-X-Ala-OH. *Biopolymers* 1979; **18**: 285–297.
11. Pardi A, Wagner G, Wüthrich K. Protein conformation and proton nuclear magnetic resonance chemical shifts. *Eur. J. Biochem.* 1983; **137**: 445–454.
12. Llinas M, Klein MP. Charge relay at the peptide bond—resonance study of solvation effects on the amide electron density distribution. *J. Am. Chem. Soc.* 1975; **9**: 4731–4737.

