

Optimization of interatomic angle geometry in torsion angle space

R. Ajay Kumar

Chemical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India

Received 3 February 1992

Accepted 19 May 1992

Keywords: Interatomic angle; Torsion angles; Distance geometry; TANDY

SUMMARY

Interatomic angle constraints are usually specified as distance constraints in torsion angle distance geometry. Such an approach is inaccurate and often inadequate. We provide a direct definition of the interatomic angle constraint term, which can be incorporated in the target function. The first derivative of this term with respect to the torsion angle has been described for all possible cases. This feature has been implemented in the nucleic acid distance geometry program TANDY [Ajay Kumar et al. (1991) *J. Biomol NMR*, **1**, 363–378], and has been tested on base pairing in the DNA fragment, d(AT)₂. The results clearly indicate the need and adequacy of such angle constraints. Other applications that would also benefit from this technique have been identified.

INTRODUCTION

There are several algorithms and approaches for optimization of interatomic distance-based functions, for applications such as distance geometry (Havel and Wüthrich, 1984; Braun and Gö, 1985; Pardi et al., 1988; Ajay Kumar et al., 1991; Güntert et al., 1991; Havel, 1991) and NOE simulations (Bonvin et al., 1991; Mertz et al., 1991). Some of these algorithms work in Cartesian coordinate space, while others work in torsion angle space. In the torsion angle approach, a target function consisting of terms for interatomic distance constraints, steric check and torsion constraint is minimized by varying the movable torsion angles. This process is greatly aided by the efficient computation of analytical first derivatives of the target function with respect to each movable torsion angle.

For applications of distance geometry which require constraining of interatomic angles, such as in hydrogen bonds, disulphide bridges, cyclic molecules, etc., the specification of interatomic angle constraint is done by fixing the relevant interatomic distances (Williamson et al., 1985; Güntert et al., 1991). Such transformation of angle constraint to distance constraints is often inadequate. For instance, if one were to specify the NH-O hydrogen bond in terms of distance con-

straints (Saenger, 1984; Williamson et al., 1985) alone, then there is a significant degree of inherent flexibility in the NH-O hydrogen bond angle. Figure 1 clearly shows that only a fraction of the distance-allowed region for the oxygen acceptor atom actually provides an acceptable NH-O angle. Moreover, these constraints are not adequate since they do not constrain the placement of the atom covalently bonded to the oxygen acceptor. Similar errors occur with the NH-N, OH-O and NH-S hydrogen bond systems. The requirement for angle constraints has been investigated by others in the context of molecular dynamics (van Gunsteren and Karplus, 1982) and energy minimization of ring molecules (Lavery et al., 1986).

We provide a formalism for direct angle constraining in torsion angle space, which is readily applicable to distance geometry programs. We demonstrate the validity of the proposed method by testing it on a dinucleotide model, where the angle constraints are employed to bring about acceptable hydrogen bond geometry in the base pair.

ANGLE GEOMETRY METHODOLOGY

We introduce here an additional term that can be incorporated into the target function to bring about interatomic angle geometry optimization in a direct manner. This term is written as $T(\varphi_{\alpha\beta\gamma})$, where $\varphi_{\alpha\beta\gamma}$ is the angle between the atoms α , β and γ . The function, $T(\varphi_{\alpha\beta\gamma})$, and consequently, the angle, $\varphi_{\alpha\beta\gamma}$, can be optimized in torsion angle space provided there exist one or more movable torsion angles separating the atoms α , β and γ .

Target function

The target function term, $T(\varphi_{\alpha\beta\gamma})$, is defined as

$$T(\varphi_{\alpha\beta\gamma}) = \begin{cases} 0 & \text{if } |\tilde{\Delta\varphi}| \geq 1 \\ (\tilde{\Delta\varphi}^2 - 1)^2 & \text{otherwise} \end{cases} \quad (1)$$

where

$$\tilde{\Delta\varphi} = \frac{\Delta\varphi - \pi}{\pi - w} \quad (2)$$

$$\Delta\varphi = \varphi_{\alpha\beta\gamma} - m \quad (3)$$

$$m = \frac{1}{2} (\varphi_l + \varphi_u) \quad (4)$$

$$w = \frac{1}{2} (\varphi_u - \varphi_l) \quad (5)$$

and where φ_l and φ_u are the lower and upper extreme values defining the allowed range of values for the interatomic angle, $\varphi_{\alpha\beta\gamma}$. The form of this function is similar to the one used by Braun

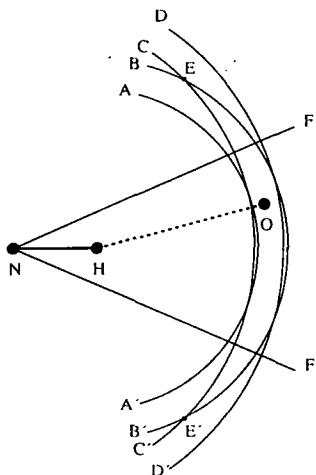


Fig. 1. View of the plane of an NH-O hydrogen bond system. The points, A, B, A' and B', define the annular locus of the oxygen atom, keeping a distance of $1.83 \text{ \AA} \leq \overline{HO} \leq 2.17 \text{ \AA}$ from the hydrogen atom. The points, C, D, C' and D', define the annular locus of the oxygen atom keeping a distance of $2.74 \text{ \AA} \leq \overline{NO} \leq 3.07 \text{ \AA}$ from the nitrogen atom. The intersection of these two regions defines the locus for the oxygen atom, yielding a 'satisfactory' hydrogen bond with distance-constraint criterion alone. This region is bounded by the points, E and E', at the extremes. Lines NF and NF' have been drawn for the extreme allowed hydrogen bond angle, $LHNF = LHNF' = 23^\circ$. All data have been taken from Saenger (1984).

(1987) for the torsion potential. $\Delta\varphi$ is the deviation of the angle, $\varphi_{\alpha\beta\gamma}$, from the desired mean value, m . $\tilde{\Delta\varphi}$ is simply a transformation of the disallowed region of φ to a normalised scale of $(-1, 1)$, to enable easy implementation of the potential described in Eq. 1.

Gradient

Efficient minimization of this target function requires the knowledge of its first derivative with respect to each intervening movable torsion angle. We provide below the theoretical foundation for the computation of such gradients. There are six possibilities for the occurrence of the atoms, α , β and γ , across a movable torsion bond, a , as outlined in Table 1. We define R as the set of atoms on the root (static) side of the bond, a , and T as the set of atoms on its tree (moving) side. For the simple case, $\alpha, \beta \in R$ and $\gamma \in T$, the angle, $\varphi_{\alpha\beta\gamma}$, changes as a function of the torsion angle, θ_a since the γ atom is on the tree side of a .

$$\frac{\partial T(\varphi)}{\partial \theta_a} = \frac{dT(\varphi)}{d\varphi} \cdot \frac{d\varphi}{dr_\gamma} \cdot \frac{\partial r_\gamma}{\partial \theta_a} \quad (6)$$

where r_γ is the position vector of the atom, γ . Differentiating Eq. 1, we have

$$\frac{dT(\varphi)}{d\varphi} = \frac{4}{\pi - w} \cdot (\tilde{\Delta\varphi}^3 - \tilde{\Delta\varphi}) \quad (7)$$

For evaluating $\frac{d\varphi}{dr_\gamma}$, we write

TABLE I
TERMS IN THE GRADIENT (Eq. 14) FOR ALL α , β and γ ROOT-TREE POSSIBILITIES

Type	R	T	P*	C	D	A	B
I	$\alpha\beta$	γ	\mathbf{r}_a	$D \times \mathbf{r}_\gamma$	B-A	$\frac{\hat{\mathbf{r}}_{\alpha\beta}}{r_{\gamma\beta}}$	$\frac{\mathbf{r}_{\alpha\beta} \cdot \mathbf{r}_{\gamma\beta}}{r_{\gamma\beta}^2} \cdot \frac{\hat{\mathbf{r}}_{\gamma\beta}}{r_{\alpha\beta}}$
II	$\alpha\gamma$	β	\mathbf{r}_a	$D \times \mathbf{r}_\beta$	B-A	$\frac{1}{r_{\alpha\beta}^3 r_{\gamma\beta}} \{ \mathbf{r}_{\alpha\beta} \cdot (\mathbf{r}_{\alpha\beta} - \mathbf{r}_{\gamma\beta}) \} \cdot \mathbf{r}_{\alpha\beta}$	$\frac{1}{r_{\alpha\beta} r_{\gamma\beta}^3} \{ \mathbf{r}_{\gamma\beta} \cdot (\mathbf{r}_{\alpha\beta} - \mathbf{r}_{\gamma\beta}) \} \cdot \mathbf{r}_{\gamma\beta}$
III	$\beta\gamma$	α	\mathbf{r}_a	$D \times \mathbf{r}_\alpha$	B-A	$\frac{\hat{\mathbf{r}}_{\gamma\beta}}{r_{\alpha\beta}}$	$\frac{\mathbf{r}_{\gamma\beta} \cdot \mathbf{r}_{\alpha\beta}}{r_{\alpha\beta}^2} \cdot \frac{\hat{\mathbf{r}}_{\alpha\beta}}{r_{\gamma\beta}}$
IV	α	$\beta\gamma$	$\mathbf{r}_{a'}$	$D \times \mathbf{r}_\alpha$	A-B	$\frac{\hat{\mathbf{r}}_{\gamma\beta}}{r_{\alpha\beta}}$	$\frac{\mathbf{r}_{\gamma\beta} \cdot \mathbf{r}_{\alpha\beta}}{r_{\alpha\beta}^2} \cdot \frac{\hat{\mathbf{r}}_{\alpha\beta}}{r_{\gamma\beta}}$
V	β	$\alpha\gamma$	$\mathbf{r}_{a'}$	$D \times \mathbf{r}_\beta$	A-B	same as in type II	
VI	γ	$\alpha\beta$	$\mathbf{r}_{a'}$	$D \times \mathbf{r}_\gamma$	B-A	same as in type I	

* $\mathbf{r}_{a'}$ is the position vector of the beginning atom along bond a .

$$\cos \varphi = \frac{(\mathbf{r}_\alpha - \mathbf{r}_\beta) \cdot (\mathbf{r}_\gamma - \mathbf{r}_\beta)}{r_{\alpha\beta} r_{\gamma\beta}} \quad (8)$$

where \mathbf{r}_α and \mathbf{r}_β are the position vectors of α and β atoms, respectively.

$$r_{\alpha\beta} = |\mathbf{r}_\alpha - \mathbf{r}_\beta| \quad (9)$$

$$\frac{d\varphi}{dr_\gamma} = -(1 - \cos^2 \varphi)^{-\frac{1}{2}} \cdot \left\{ \frac{\hat{\mathbf{r}}_{\alpha\beta}}{r_{\gamma\beta}} - \left(\frac{\mathbf{r}_{\alpha\beta} \cdot \mathbf{r}_{\gamma\beta}}{r_{\gamma\beta}^2} \cdot \frac{\hat{\mathbf{r}}_{\gamma\beta}}{r_{\alpha\beta}} \right) \right\} \quad (10)$$

where

$$\mathbf{r}_{\alpha\beta} = \mathbf{r}_\alpha - \mathbf{r}_\beta \quad (11)$$

$$\hat{\mathbf{r}}_{\alpha\beta} = \mathbf{r}_{\alpha\beta} / r_{\alpha\beta} \quad (12)$$

and similarly for terms involving β and γ atoms. The potential singularity arising at $\varphi = 0$ can be easily checked for and regularized in the implementation. From Eq. 9 in the paper by Noguti and Gō (1983),

$$\frac{\partial \mathbf{r}_\gamma}{\partial \theta_\alpha} = \hat{\mathbf{e}}_\alpha \times (\mathbf{r}_\gamma - \mathbf{r}_\alpha) \quad (13)$$

where $\hat{\mathbf{e}}_a$ is the unit vector along the bond, a , pointing to the tree side and \mathbf{r}_a is the position vector of the end atom along bond, a . Combination of Eqs. 7, 10 and 13 gives us the gradient for this case. The general form of the derivative may be written as:

$$\frac{\partial T(\varphi)}{\partial \theta_a} = c\varphi \cdot (\hat{\mathbf{e}}_a \cdot \hat{\mathbf{e}}_a \times \mathbf{P}) \cdot \begin{pmatrix} C \\ D \end{pmatrix} \quad (14)$$

where

$$c\varphi = \frac{-4(\tilde{\Delta}\varphi^3 - \tilde{\Delta}\varphi)}{(\pi - w)(1 - \cos^2\varphi)^2} \quad (15)$$

The terms C, D and P for this and other cases are defined in Table 1.

For the case where $\alpha, \gamma \in R$ and $\beta \in T$, Eq. 6 may be rewritten as

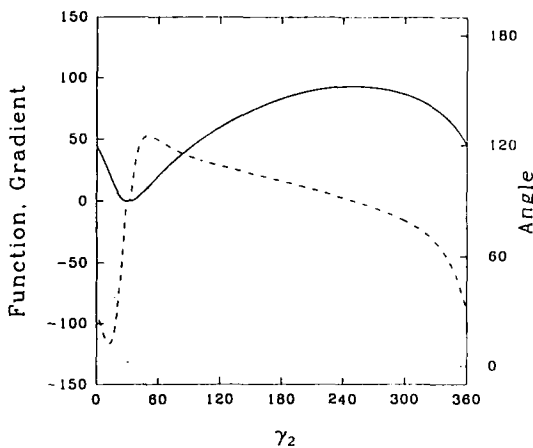
$$\frac{\partial T(\varphi)}{\partial \theta_a} = \frac{dT(\varphi)}{d\varphi} \cdot \frac{d\varphi}{d\mathbf{r}_\beta} \cdot \frac{\partial \mathbf{r}_\beta}{\partial \theta_a} \quad (16)$$

where

$$\frac{d\varphi}{d\mathbf{r}_\beta} = \frac{1}{r_{\alpha\beta} r_{\gamma\beta}^3} \cdot \{\mathbf{r}_{\gamma\beta} \cdot (\mathbf{r}_{\alpha\beta} - \mathbf{r}_{\gamma\beta})\} \cdot \mathbf{r}_{\gamma\beta} - \frac{1}{r_{\alpha\beta}^3 r_{\gamma\beta}} \cdot \{\mathbf{r}_{\alpha\beta} \cdot (\mathbf{r}_{\alpha\beta} - \mathbf{r}_{\gamma\beta})\} \cdot \mathbf{r}_{\alpha\beta} \quad (17)$$

The other four cases for the occurrence of α, β and γ atoms across the bond, a , can be understood as trivial modifications of either of the above two cases, as is shown in Table 1.

A



B

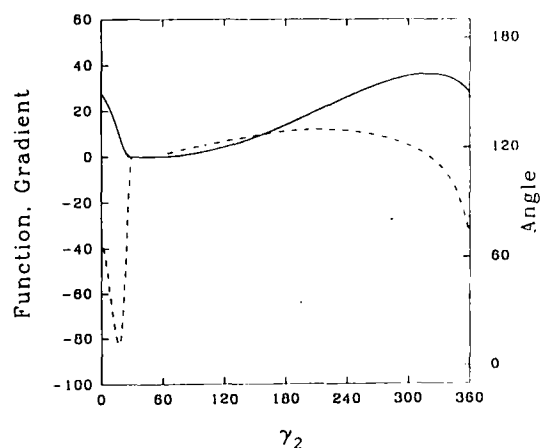


Fig. 2. Variation in the angle (A) T2H3-T2N3-A4N1 and (B) T2H3-A4N1-A4C2 as a function of the backbone torsion angle γ_2 of the T2 residue on first strand (.....). The target function term (—) and its gradient with respect to γ_2 (----) has been evaluated at an 1° interval for γ_2 . The θ_1 and θ_u used for these calculations are listed in Table 2.

TABLE 2
HYDROGEN BOND GEOMETRY, WITH AND WITHOUT ANGLE CONSTRAINTS*

	Distance (Å)				Angle (degree)						CPU time (min:sec)
	d ₁	d ₂	d ₃	d ₄	a ₁	a ₂	a ₃	a ₄	a ₅	a ₆	
<i>D Alone test</i>											
lower	2.75	1.78	2.74	1.83	unconstrained						
upper	3.15	2.02	3.07	2.17							
1	2.90	2.03	2.97	2.13	29.2	137.2	157.4	27.2	139.4	81.3	5:39
2	2.94	2.03	2.73	1.83	24.6	142.0	122.9	23.2	145.3	106.2	70:25
3	2.77	1.90	2.74	2.10	51.7	106.8	81.8	26.2	140.0	125.0	24:32
4	2.83	2.00	3.07	2.18	25.7	142.5	122.8	29.7	135.4	109.1	2:56
5	2.74	2.02	3.06	2.16	23.9	144.9	126.2	37.2	124.9	106.0	17:30
<i>D&A test</i>											
lower					0.0	175.0	120.0	0.0	175.0	120.0	
upper	same as above				5.0	180.0	180.0	5.0	180.0	180.0	
1	2.97	1.94	2.86	1.83	0.3	179.5	125.4	2.1	176.8	120.2	18:12
2	3.01	1.98	2.87	1.84	2.3	176.4	122.2	2.2	176.6	122.2	39:27
3	2.83	1.80	2.85	1.83	2.3	176.4	123.4	2.0	176.8	120.2	75:37
4	3.05	2.02	2.93	1.90	1.5	177.7	123.0	2.0	176.9	122.2	36:36
5	3.06	2.03	3.07	2.06	1.8	177.3	123.0	1.0	178.7	120.3	74:09

*d₁: T2N3-A4N1; d₂: T2H3-A4N1; d₃: T2O4-A4N6; d₄: T2O4-A4H62; a₁: T2O4-A4N6-A4H62; a₂: T2O4-A4H62-A4N6; a₃: T2C4-T2O4-A4H62; a₄: T2H3-T2N3-A4N1; a₅: T2N3-T2H3-A4N1; a₆: T2H3-A4N1-A4C2.

The angle geometry terms have been incorporated in the target function, and gradient computation routines of TANDY and several other tests have been performed.

TESTS AND RESULTS

The validity of the expressions for gradient computation was verified by the following test. In a model dinucleotide, d(AT)₂, in B-form, only the torsion angle, γ_2 , on the first strand was varied in steps of 1°. For each value of γ_2 , the target function and its gradient were computed, using the angle geometry routines of TANDY, for two interatomic angle constraints: T2H3-A4N1-A4C2 (type I) and T2H3-T2N3-A4N1 (type VI). These are shown in Fig. 2. Each function and gradient evaluation for this test dinucleotide takes 0.12 seconds on an Iris 4D/70G computer.

The necessity of angle geometry optimization and the performance of its implementation in the Quasi-Newton minimization procedure of TANDY was studied by the following tests.

Distance constraints alone (D alone)

In the dinucleotide, d(AT)₂, the relative orientation of the two bases forming the bottom base

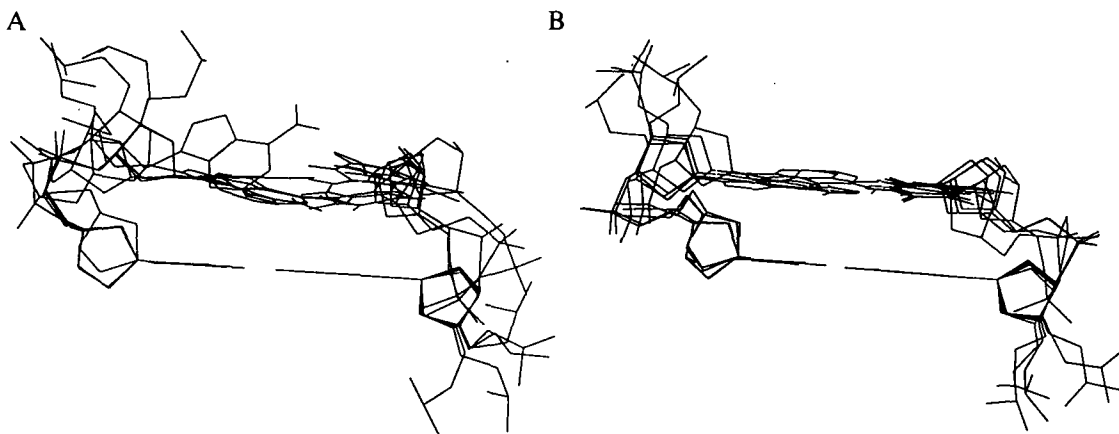


Fig. 3. Superimposed view of the structures obtained from (A) D Alone and (B) D&A tests. The structures are juxtaposed to preserve the orientation of the bottom base pair. Some hydrogen atoms have been omitted for greater clarity.

pair was held rigid, while all the torsion angles, α , β , γ , χ , ϵ , ζ , on both strands were allowed to flex freely. Typical intrastrand distance constraints, described as the BCNOE set in the paper by Ajay Kumar et al. (1991) were used with lower and upper bounds as ± 0.3 Å of the actual distance in B-DNA. In addition, interstrand distance constraints for typical Watson–Crick hydrogen bonds were also specified. The weights on all terms were uniform, except for the torsion term, whose weight was kept at 0 throughout. Five structures were generated, using the option for initial randomization of all movable torsion angles, in TANDY. For all the outputted structures, the final distances for the constrained atom pairs were analysed. To assess the possible adequacy of distance constraints alone in bringing about satisfactory interatomic angles between hydrogen bond forming atoms, all these angles were calculated and analysed in all the final structures.

Distance and angle constraints (D&A)

In addition to all the distance constraints used in the D Alone test, angle constraints between the hydrogen bond forming atoms were introduced. We deliberately employed extremely stringent angle constraints for the following reasons: (1) we wanted a rigorous test for the angle geometry optimization procedure. Such, or even more stringent constraints would be necessary in applications such as disulphide bridges and ring closure; (2) the bounds used are more stringent than the comparatively relaxed bounds used in energy minimization and molecular dynamics programs, for hydrogen bond systems. In the latter programs, the relaxed bounds are used for screening of likely candidates for hydrogen bonding. For geometries within the bounds, the hydrogen bond potentials drive the angles to mean values. Whereas, in the present test, we have deliberately used stringent constraints to get proper base geometry without having to follow it with energy minimization. TANDY was used, with distance and angle geometry provisions, to generate five structures after initial randomization. For each generated structure, the distances and angles for the constrained atoms were analysed.

There were no residual violations in the intrastrand distance constraints, or overall steric constraints, beyond ± 0.01 Å in either of the tests. Table 2 shows an analysis of the final interstrand

distances and angles in the two tests. For the D Alone test, the hydrogen bond angles are clearly in violation of the desired range. Most angles have values well beyond the maximum values in DNA base pairs observed experimentally (Saenger, 1984). While, in the D&A test, both the hydrogen bond forming distances, as well as angles, are simultaneously satisfied. Table 2 also lists the CPU time for these calculations, performed on an Iris 4D/70G computer. The inclusion of angle constraint terms causes a marginal increase in the computation time. Figure 3 shows a graphic comparison of the sets of structures obtained in the two tests.

The angle constraints are clearly necessary to provide more realistic hydrogen bonding in torsion space distance geometry calculations. The formalism presented here has potential utility in other applications as well, e.g., disulphide bridges and ring closure in cyclic structures such as ribose, proline and cyclic peptides.

ACKNOWLEDGEMENTS

The facilities provided by the 500 MHz FTNMR National Facility, supported by the Department of Science and Technology, Government of India, are gratefully acknowledged.

REFERENCES

- Ajay Kumar, R., Hosur, R.V. and Govil, G. (1991) *J. Biomol. NMR*, **1**, 363–378.
Bonvin, A.M.J.J., Boelens, R. and Kaptein, R. (1991) *J. Biomol. NMR*, **1**, 305–309.
Braun, W. (1987) *Quart. Rev. Biophys.*, **19**, 115–157.
Braun, W. and Gō, N. (1985) *J. Mol. Biol.*, **186**, 611–626.
Güntert, P., Braun, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.
Havel, T.F. (1991) *Prog. Biophys. Mol. Biol.*, **56**, 43–78.
Havel, T.F. and Wüthrich, K. (1984) *Bull. Math. Biol.*, **46**, 673–698.
Lavery, R., Parker, I. and Kendrick, J. (1986) *J. Biomol. Struct. Dyn.*, **4**, 443–462.
Mertz, J.E., Güntert, P., Wüthrich, K. and Braun, W. (1991) *J. Biomol. NMR*, **1**, 257–269.
Noguti, T. and Gō, N. (1983) *J. Phys. Soc. Jpn.*, **52**, 3685–3690.
Pardi, A., Hare, D.R. and Wang, C. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 8785–8789.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, p. 118.
Van Gunsteren, W.F. and Karplus, M. (1982) *Macromolecules*, **15**, 1528–1544.
Williamsen, M.P., Havel, T.F. and Wüthrich, K. (1982) *J. Mol. Biol.*, **182**, 295–315.