Journal of Biomolecular NMR, 8 (1996) 292–310
ESCOM

J-Bio NMR 376

Treatment of NOE constraints involving equivalent or nonstereoassigned protons in calculations of biomacromolecular structures

C. Mark Fletcher[a], David N.M. Jones[b], Robert Diamond[c] and David Neuhaus[c,*]

[a]*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, U.S.A.*
[b]*Department of Biochemistry and Molecular Biology, The University of Chicago, 920 East 58th Street, Chicago, IL 60637, U.S.A.*
[c]*MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.*

Received 10 January 1996
Accepted 24 June 1996

Keywords: Structure determination; Pseudoatom corrections; Multiplicity corrections; Stereoassignments; Distance-geometry calculations; Simulated-annealing calculations

Summary

Two modifications to the commonly used protocols for calculating NMR structures are developed, relating to the treatment of NOE constraints involving groups of equivalent protons or nonstereoassigned diastereotopic protons. Firstly, a modified method is investigated for correcting for multiplicity, which is applicable whenever all NOE intensities are calibrated as a single set and categorised in broad intensity ranges. Secondly, a new set of values for 'pseudoatom corrections' is proposed for use with calculations employing 'centre-averaging'. The effect of these protocols on structure calculations is demonstrated using two proteins, one of which is well defined by the NOE data, the other less so. It is shown that failure to correct for multiplicity when using 'r^{-6} averaging' results in overly precise structures, higher NOE energies and deviations from geometric ideality, while failure to correct for multiplicity when using 'r^{-6} summation' can cause an avoidable degradation of precision if the NOE data are sparse. Conversely, when multiplicities are treated correctly, r^{-6} averaging, r^{-6} summation and centre averaging all give closely comparable results when the structure is well defined by the data. When the NOE data contain less information, r^{-6} averaging or r^{-6} summation offer a significant advantage over centre averaging, both in terms of precision and in terms of the proportion of calculations that converge on a consistent result.

Introduction

The determination of biomacromolecular structures in the solution state using NMR generally depends on the measurement and interpretation of a large number of NOE interactions between protons (Wüthrich, 1986). Each assigned NOE interaction is used to define a semiquantitative distance constraint, and the three-dimensional structure is calculated by computing conformations of the known covalent structure that satisfy, to a good approximation, all constraints simultaneously. In almost all cases, a proportion of the NOE-derived distance constraints involve one or more groups of equivalent spins, or spins in diastereotopic groups for which no stereoassignments are available. Examples of the former include methyl groups or rapidly flipping symmetrical aromatic rings, and examples of the latter include nonstereoassigned methylene groups or isopropyl groups (e.g. of valine and leucine residues in proteins). In such cases, there are two particular issues of interpretation that must be addressed. Firstly, when translating a measured NOE intensity into a distance constraint, a multiplicity correction may be needed if one or both interacting groups contain more than one proton, depending on how the intensities are calibrated. Secondly, it must be decided how to test whether an NOE-derived distance constraint corresponding to a group of equivalent protons is satisfied in the evolving model structure.

These are fundamental issues that were addressed early

*To whom correspondence should be addressed.

Abbreviations: HMG, high mobility group; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; rmsd, root-mean-square deviation; YASAP, yet another simulated-annealing protocol.

in the development of the field, but recent experience has prompted us to re-examine and compare some aspects of the methods currently in use. In this paper we demonstrate two modifications to the way in which NOE-based distance constraints are derived for groups of equivalent and nonstereoassigned protons. Firstly, we investigate a modified method of correcting for multiplicity (Constantine et al., 1992). This method is applicable when all NOE intensities are calibrated as a single set and are estimated only semiquantitatively, resulting in broad intensity ranges (typically labelled 'strong', 'medium', 'weak', etc.); such a calibration is generally used in conjunction with a manual approach to intensity estimation. Secondly, we propose a new set of values for the 'pseudoatom corrections' that may be used in conjunction with calculations employing 'centre-averaging'. We then examine the application of these new protocols to the determination of protein structure using simulated annealing, both for a case in which the structure is well defined by the constraints (CD59) (Fletcher et al., 1994), and for a case in which the data constrain the structure less completely (HMG-D) (Jones et al., 1994).

Theory

Effective distances sensed by the NOE

Simple interpretations of NOE intensities are usually based on the initial rate approximation. This states that, for short values of the NOE mixing time, the growth rate of an NOE between two spins or equivalent spin groups I and S remains substantially linear and independent of other spins (Neuhaus and Williamson, 1989). When applied to NOE cross peaks in NOESY experiments, the formal expression of the initial rate approximation is (Macura and Ernst, 1980):

$$\left.\frac{d(a_{IS})}{d\tau_m}\right|_{\tau_m=0} = -n_I n_S \sigma_{IS} \qquad (1)$$

where a_{IS} is the intensity of the IS cross peak, n_I is the number of equivalent spins in group I, n_S is the number of equivalent spins in group S, σ_{IS} is the cross-relaxation rate constant between I and S and τ_m is the NOESY mixing time.* In all that follows, we assume that the two symmetry-related IS cross peaks are identical, and we refer to the intensity of either of them as the IS NOE interaction intensity. The cross-relaxation rate σ_{IS} is in turn given by (Neuhaus and Williamson, 1989):

$$\sigma_{IS} = \frac{1}{10}\left(\frac{\mu_0}{4\pi}\right)\gamma_I^2\gamma_S^2\hbar^2\langle r_{IS}^{-6}\rangle[6J(\omega_I+\omega_S) - J(\omega_I-\omega_S)] \qquad (2)$$

where γ_I and γ_S are the gyromagnetic ratios of spins I and S, respectively, $J(\omega_X)$ is the spectral density function at frequency ω_X, and μ_0 and $\hbar$ have their usual meanings.

Since Eq. 2 is not scaled by $n_I n_S$, this definition of σ_{IS} corresponds to the cross-relaxation rate constant appropriate for a *single* spin I interacting with a *single* spin S (which is also why σ_{IS} appears multiplied by $n_I n_S$ in Eq. 1). The appropriate distance-related term in the definition of σ_{IS}, written here as $\langle r_{IS}^{-6}\rangle$, is some form of *average* over all the $n_I n_S$ different individual values of $r_{I_i S_j}^{-6}$, where I_i represents an individual spin in the I group and S_j represents an individual spin in the S group. It is convenient to rewrite this term so that it has the dimensions of distance, thereby defining an 'effective distance' as sensed by the NOE:

$$r_{eff} = \langle r_{IS}^{-6}\rangle^{-1/6} \qquad (3)$$

The way in which the individual contributions to r_{eff} are averaged during evolution of an NOE interaction depends on how, if at all, the different spins within each equivalent group interchange with one another over time. For cases where the spins in both groups interchange more slowly than the whole molecule tumbles in solution ($k_{exch} < \tau_c^{-1}$), or do not interchange at all, the so-called 'r^{-6} average' is appropriate (Brünger et al., 1986), defined by:

$$r_{eff} = \left(\frac{1}{n_I n_S}\sum_{i,j} r_{I_i S_j}^{-6}\right)^{-1/6} \qquad (4)$$

Examples include flipping of symmetric aromatic rings and all cases of nonstereoassigned diastereotopic groups. Of course, the latter are not strictly cases of equivalence, but they are commonly treated as such in order that constraints can be placed upon the overall position of a diastereotopic group without having to differentiate between the nonstereoassigned signals.

For cases where spins interchange with one another faster than the whole molecule tumbles in solution ($k_{exch} > \tau_c^{-1}$), the averaging of NOE interactions is more complicated, since the anisotropy of the fast internal motion relative to the molecular frame must be taken into account. Tropp has derived general expressions for spectral densities that treat internal motions of any rate in combination with either isotropic or anisotropic overall tumbling (Tropp, 1980). From these he obtains simplified equations that treat methyl group rotations faster than overall tumbling using a three-site jump model. This model has the advantage that it is independent of the rate at which the methyl group jumps between rotamers (provided this is much faster than overall molecular tumbling), and can therefore be applied directly to a set of coordinates describing an otherwise static structure without needing to specify any additional parameters. More complex treatments are required when the internal motion is only some-

*On page 127 of the first printing of the book by Neuhaus and Williamson (1989), it is incorrectly stated that NOESY cross-peak intensity is proportional to $n_I n_S/(n_I+n_S)$. This error was corrected by Yip (1990).

what faster than overall tumbling, but these depend also on the value of the jump rate between methyl rotamers (see, for instance, Kalk and Berendsen, 1976).

Although many internal motions may be significantly faster than overall tumbling, methyl group rotations are the only common example where a group of spins *are made equivalent* as a result of such an internal motion, so the treatment of fast internal motions that follows is limited to those of methyl groups. Rearrangement of Tropp's equations for the case of isotropic overall tumbling leads to an expression for r_{eff} for the interaction of three methyl protons I_i with a fourth proton S:

$$r_{eff} = \left[\frac{1}{5} \sum_{n=-2}^{2} \left| \frac{1}{N} \sum_{i=1}^{3} \frac{Y_{2n}(\Phi_i^{mol})}{r_{I_iS}^3} \right|^2 \right]^{-1/6} \qquad (5)$$

Here, the summation over i runs over the three vectors $\vec{r}_{I_iS}$ linking the methyl protons to S. The terms $Y_{2n}(\Phi_i^{mol})$ are second-order spherical harmonic functions of the polar angles θ_i and ϕ_i that describe the orientation of each vector $\vec{r}_{I_iS}$ in (any) fixed molecular frame, and the summation over n runs over the second index of the spherical harmonic functions. For convenience in what follows, we shall refer to this value of r_{eff} for the interaction of a methyl group with a single proton as the 'Tropp' averaged distance, abbreviated r_{Tropp}.

The formulation of r_{Tropp} given in Eq. 5 emphasises the separate contributions from radial and angular terms. Yip and Case (1991) have given a much simpler formulation of Tropp's equations that leads to the following expression:

$$r_{Tropp} = \left[\frac{1}{2(3)^2} \sum_{i=1}^{3} \sum_{j=1}^{3} \left(\frac{3(\vec{r}_{I_iS} \cdot \vec{r}_{I_jS})^2 - r_{I_iS}^2 r_{I_jS}^2}{r_{I_iS}^5 r_{I_jS}^5} \right) \right]^{-1/6} \qquad (6)$$

where the indices i and j correspond to two independent summations, each running separately over all three methyl protons.

Since the maximum value of $(\vec{r}_{I_iS} \cdot \vec{r}_{I_jS})^2$ is $r_{I_iS}^2 r_{I_jS}^2$, it can be shown that:

$$r_{Tropp} \geq \left(\frac{1}{3} \sum_{i=1}^{3} \frac{1}{r_{I_iS}^3} \right)^{-1/3} \qquad (7)$$

Thus, it may be seen from Eqs. 5–7 that the radial dependence leads to an averaging over r^{-3}, while the angular dependence leads to a geometry-dependent reduction in the NOE interaction strength relative to that expected based on r^{-3} averaging alone. This reduction due to the angular term is greatest near the axis of methyl group rotation and at relatively short distances, while near the plane of the methyl group protons it is relatively small.

For cases where one group undergoes fast interchange and the other slow, or where two types of averaging occur within one group, appropriate mixed averages result. For example, for the interaction of a single proton S with a nonstereoassigned isopropyl group in a valine or leucine residue of a protein, the appropriate average would be:

$$\begin{aligned} r_{eff} &= \left[\frac{1}{2} \left(r_{Tropp}^{Me_A-S} \right)^{-6} + \frac{1}{2} \left(r_{Tropp}^{Me_B-S} \right)^{-6} \right]^{-1/6} \\ &= \left[\frac{1}{4(3)^2} \sum_{i=1}^{3} \sum_{j=1}^{3} \left(\frac{3(\vec{r}_{A_iS} \cdot \vec{r}_{A_jS})^2 - r_{A_iS}^2 r_{A_jS}^2}{r_{A_iS}^5 r_{A_jS}^5} \right) \right. \\ &\quad \left. + \frac{1}{4(3)^2} \sum_{i=1}^{3} \sum_{j=1}^{3} \left(\frac{3(\vec{r}_{B_iS} \cdot \vec{r}_{B_jS})^2 - r_{B_iS}^2 r_{B_jS}^2}{r_{B_iS}^5 r_{B_jS}^5} \right) \right]^{-1/6} \end{aligned} \qquad (8)$$

where the terms $r_{Tropp}^{Me_A-S}$, r_{A_iS}, r_{A_jS}, $\vec{r}_{A_iS}$ and $\vec{r}_{A_jS}$ all relate to the three vectors from the external spin S to the protons of one methyl group (A), while the terms $r_{Tropp}^{Me_B-S}$, r_{B_iS}, r_{B_jS}, $\vec{r}_{B_iS}$ and $\vec{r}_{B_jS}$ relate to the three vectors from the external spin S to the protons of the other methyl group (B).

Testing for constraint violations

During calculations of biomolecular structures based on NMR data, violations of distance constraints are assessed by comparing the length of specified distances in the evolving or completed structural model against NOE-based constraints on those distances. The upper bound of each distance constraint is set using the intensity of the corresponding NOE interaction, so that, ideally, each upper bound would correspond to the distance r_{eff} that is 'sensed' by the NOE. In practice, cross-peak intensities are often treated semiquantitatively by grouping them into a small number of broad categories, for each of which a particular upper bound value is specified.

In cases where either or both of the signals involved in an NOE interaction correspond to a group of equivalent spins, there will be several relevant distances in the structural model (i.e., the $n_I n_S$ different values of $r_{I_iS_j}$ referred to in, for instance, Eq. 4). All of these must be combined in some suitable fashion to yield a single distance that may be compared with the upper bound in the constraint list. There are two approaches to this in common use at present, usually called 'centre averaging' and 'r^{-6} averaging'. In addition, there is a subtle variant of r^{-6} averaging, called 'r^{-6} summation'. Although r^{-6} averaging and r^{-6} summation appear to differ somewhat conceptually (see below), in reality they are virtually identical, and most of the conclusions in the theory section concerning the effects of r^{-6} averaging apply equally to those of r^{-6} summation. For simplicity in what follows, we shall discuss these various approaches in terms of a constraint between one equivalent group of protons I and a single further proton S (i.e., $n_S = 1$).

In the 'centre averaging' approach, originally introduced by Wüthrich et al. (1983), such a constraint is referred to a single pseudoatom at the mean position of

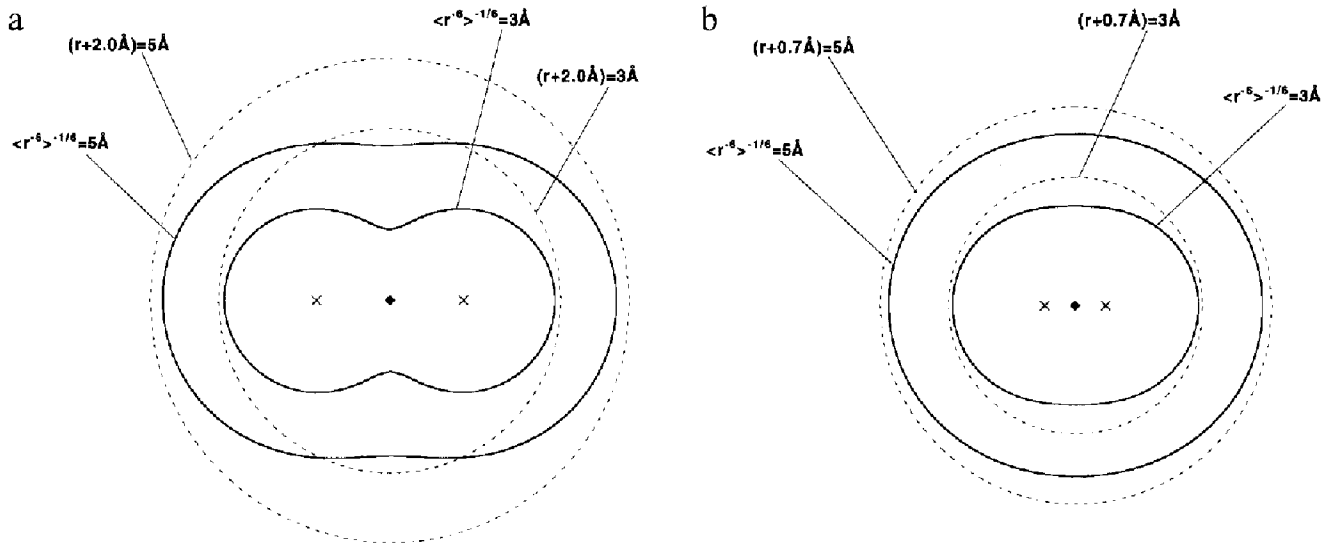

Fig. 1. Validity boundaries for constraint upper bounds of 3 Å and 5 Å calculated using r^{-6} averaging (solid lines) and centre averaging (dashed lines) for (a) a pair of symmetry-related aromatic protons (separation 4.3 Å), and (b) a CH_2 group (interproton separation 1.76 Å). In each case, the proton positions (I_1 and I_2) are indicated by crosses, and the pseudoatom position (Q) by a filled diamond. When the constraint upper bound is short relative to the interproton separation $r_{I_1I_2}$, the volume bounded by the r^{-6}-averaged constraint is significantly smaller than that bounded by the centre-averaged constraint. The difference between the two constraints is at its least close to the line connecting the two protons, and at its largest close to the perpendicular plane through the pseudoatom. (In the limit $r_{upper\text{-}bound} \ll r_{I_1I_2}$, the r^{-6}-averaged constraint would define two separate spherical volumes, each centred on one of the two protons and with radius $r_{upper\text{-}bound}$, whereas the centre-averaged constraint would define one large sphere centred on the pseudoatom and with radius $r_{upper\text{-}bound} + r_{IQ}$.) When the constraint upper bound is long relative to the interproton separation, the r^{-6}-averaged constraint is shorter than the centre-averaged constraint, because the pseudoatom correction is not included in the former.

the atoms in the equivalent group. This immediately yields a single distance in the model (the distance between the pseudoatom and the spin S) that may be compared with the upper bound in the constraint list, but at the cost that the distance measured in the model is necessarily different from that 'sensed' by the NOE. Consequently, a 'pseudoatom correction' must be added to the constraint upper bound. In the more recently proposed 'r^{-6} averaging' approach, all the individual inter-spin distances r_{IS} in the model are measured separately, and a single 'r^{-6}-averaged' distance calculated from them according to Eq. 4 (Brünger et al., 1986; Levy et al., 1989). This approach seeks to match the averaged distance determined from the model directly with that 'sensed' by the NOE, and therefore does not require addition of a pseudoatom correction to the constraint upper bound. However, this approach cannot be used with methods such as metric matrix distance geometry, where the structural model is not defined in three dimensions until near the end of the calculation. Figure 1 shows the validity boundary for various constraints calculated using both centre averaging and r^{-6} averaging.

As generally used at present, the r^{-6} averaging and summation methods ignore the fact that NOE intensities involving methyl groups are actually determined by values of r_{Tropp}. However, in practice this probably makes little difference. It is true that r_{Tropp} can be significantly longer than the r^{-6} average distance for locations at short distances close to the methyl axis, but elsewhere the difference between the functions is much smaller. Furthermore, what error there is makes a constraint calculated using r^{-6} averaging over-conservative. An r^{-6}-averaged distance is more strongly dominated by the shortest of the individual distances, and is thus somewhat shorter than the corresponding value of r_{Tropp}. Thus, when the model structure is identical to the true structure, the model-derived r^{-6}-averaged distance that is used to check constraint validity will still be shorter than the distance 'sensed' by the NOE (r_{Tropp}) and used to set the constraint upper bound, implying that the methyl group could move further away from its NOE partner without violating the constraint.

Correcting for multiplicity in equivalent groups

If all constraints are based upon a common calibration of the intensity-to-distance relationship, then it follows straightforwardly from Eq. 1 that NOE intensities involving equivalent groups must be divided by the total number of spins involved ($n_I n_S$) before being translated into an upper bound on the corresponding distance for use in conjunction with centre averaging or r^{-6} averaging (no correction is needed when r^{-6} summation is used, as discussed later). If numerical values for NOE intensities are available, for instance as a result of volume integration of cross peaks, division by $n_I n_S$ is of course trivial. In still probably the majority of cases, however, NOE intensities are estimated semiquantitatively (e.g. by counting contour levels in an evenly contoured plot of the NOESY spectrum) and then classified into categories such as 'strong',

'medium' and 'weak'. In such cases, division of the intensity by $n_I n_S$ is not possible directly.

One way around this difficulty, often adopted, is to calibrate different classes of constraints separately; thus, all methyl–single-proton constraints might be calibrated using $d_{\alpha\beta}(i,i)$ in alanine, methyl–methyl constraints using $d_{\delta 1\delta 2}(i,i)$ in leucine, and so on (e.g. see Krezel et al., 1994). However, if a common calibration is used for all interactions, it is necessary to correct individual constraints for different multiplicities. Quite commonly, 0.5 Å is added to the upper bound of all constraints involving a methyl group (and 1 Å to those of methyl–methyl constraints; Clore et al., 1987; Wagner et al., 1987) to compensate for multiplicity (although the precise reason for introducing this correction originally is not completely clear). No related corrections appear to have been adopted for constraints involving degenerate aromatic signals or methylene groups. Furthermore, it is clear that any such *addition* to the upper bound cannot in general correspond to a *division* of the NOE intensity, and that what is needed instead is a multiplicative correction to the upper bound. It follows directly from the r^{-6} dependence of the NOE intensity in Eqs. 1–4 that the appropriate multiplicity correction factors Z are given by (Constantine et al., 1992):

$$Z = (n_I n_S)^{1/6} \tag{9}$$

Thus, for a constraint between a single spin and a methyl group ($n_I n_S = 3$), the uncorrected upper bound should be multiplied by $Z = 3^{1/6} \approx 1.20$ to account for multiplicity, while for $n_I n_S = 2$ (e.g. for a constraint between a single spin and a fast-flipping aromatic ring), an uncorrected upper bound should be multiplied by $Z = 2^{1/6} \approx 1.12$. For comparison, addition of 0.5 Å to the upper bound of methyl group constraints gives the same result as Eq. 9 when the uncorrected upper bound is about 2.49 Å; for longer upper bounds, addition of 0.5 Å generally results in too small a correction, the error increasing with the length of the constraint.

Nonstereoassigned diastereotopic groups; One constraint or two?

Application of multiplicity corrections as outlined above is straightforward in all cases of genuine equivalence, but additional complications arise when dealing with nonstereoassigned diastereotopic groups. However, it should be noted that these complications are not a new feature of the present approach; rather, they concern issues present in all previous approaches to deriving such constraints.

Consider a nonstereoassigned methylene group interacting with a single additional proton S. The first issue is whether to treat the methylene protons separately or as a group. In the former case, a separate constraint to S would be specified for each methylene proton, using the weaker of the two NOE interaction intensities to set the upper bound for both constraints (obviously, this method is only applicable if two separate NOE interactions are measured). Alternatively, both constraints can be specified using the actual intensities of the corresponding cross peaks if the 'floating stereoassignment' approach is employed (Weber et al., 1988). This works by allowing the calculation protocol to switch individual stereoassignments dynamically whenever it detects that, by doing so, a lower overall NOE violation penalty would result. However, if the methylene protons are treated as a group only one constraint would be specified, using r^{-6} averaging, r^{-6} summation or centre averaging to account simultaneously for the interaction of S with both methylene protons.

Applying a separate constraint to each methylene proton generally leads to a tighter restriction on the position of the group. However, this approach takes no account of spin diffusion between the methylene protons, which is likely to be efficient and may partially or completely average the intensities of the corresponding NOE interactions. Thus, were it not for spin diffusion, the weaker intensity might have been still weaker or even undetectable. When separate constraints are used, the structure calculation might therefore be forced to place the external proton S closer than it should be to the more distant methylene proton, even if the weaker intensity is used to set both constraints. For this reason, we suggest that use of a single constraint treating the nonstereoassigned methylene protons as an equivalent group is preferable; this approach does not attempt to partition the total NOE intensity between the two methylene protons, and is therefore unaffected by spin diffusion between them (Levy et al., 1989). The same argument applies to nonstereoassigned isopropyl groups, although the extent of intermethyl spin diffusion in an isopropyl group is generally less than that between the two protons of a methylene group. A related argument also applies in cases where methylene or isopropyl groups *are* stereoassigned: if the two NOE interactions with S have different intensities, we suggest that one should use only the stronger interaction to define a constraint, discarding that corresponding to the weaker interaction due to the potential for spin diffusion, while if the two NOE interactions with S have the same intensity, one should set a single averaged constraint as though no stereoassignment had been made (including multiplicity corrections as described below).

Against these arguments, however, there is an advantage of using the 'floating stereoassignment' approach that is lost when using a single averaged constraint. Consider a four-spin system comprising a methylene group and two other protons X and Y, where one methylene proton is close to both X and Y, while the other is relatively distant from both X and Y. If no stereoassignment has been made, it is not known which methylene proton

TABLE 1
UPPER BOUNDS FOR A CONSTRAINT BETWEEN TWO NONSTEREOASSIGNED DIASTEREOTOPIC GROUPS

Case	Individual upper bounds	Combination horizontally	Combination vertically
1	A, A A, A	A A	A
2	A, B A, A	1.12 A A	1.12 A
3	A, B A, B	1.12 A 1.12 A	1.12 A
4	A, B B, B	1.12 A B	$(1.12)^2$ A
5	A, B A, C	1.12 A 1.12 A	1.12 A
6	A, B B, C	1.12 A 1.12 B	$(1.12)^2$ A
7	A, B C, C	1.12 A C	$(1.12)^2$ A
8	A, B C, D	1.12 A 1.12 C	$(1.12)^2$ A

A, B, C and D represent the constraint upper bounds corresponding to the intensities of the four possible inter-group NOE interactions, in decreasing order of corresponding NOE strength (i.e. in increasing order of upper bound, A < B < C < D). In each case, the lowest intensity (longest upper bound) may correspond to an absent, unresolved or unassigned NOE interaction, in which case the corresponding upper bound (B, C or D) would take a notional value of ∞. See text for further explanation.

is the one that gives strong cross peaks to X and Y, but it is known that *the same* methylene proton contacts both X and Y. Information of this sort is utilised during calculations that employ floating stereoassignment, since the consistency of a particular stereoassignment generated by the calculation is enforced across the whole structure, even if that stereoassignment should be incorrect. However, when averaged constraints are employed, no stereoassignment is made and information of this sort is effectively discarded.

Nonstereoassigned diastereotopic groups; Correcting for multiplicity

When an interaction between a nonstereoassigned diastereotopic group and a single proton is treated using a single constraint with centre averaging or r^{-6} averaging, the intensity that should be used to set the upper bound is the sum of both NOE interaction intensities, divided by two to correct for multiplicity; in other words, it is the *average* intensity of the two NOE interactions. (For interactions involving isopropyl groups a further division by three is required; this is assumed in what follows). In practice, one must deal with the three possible situations described below, for each of which the most appropriate method of deriving the constraint upper bound is considered:

(i) The two possible NOE interactions both correspond to resolved and measurable NOESY cross peaks, and both fall into the same intensity category. In this case, no multiplicity correction is needed, because when both cross peaks are in the same intensity category the average must also lie in the same category.

(ii) Both NOE interactions correspond to resolved and measurable cross peaks, but fall into different intensity categories. There are two possible ways to set the upper bound: (a) the shorter upper bound is used with a multiplicity correction (Z = 1.12, corresponding to division of the intensity by two); or (b) the longer upper bound is used with no multiplicity correction. Neither approach would result in an unjustifiably short upper bound, so to make best use of the data one should use the approach yielding the shorter upper bound. Provided that the intensity categories are not divided so finely that upper bounds in successive categories differ by less than 12%, approach (a) will always give the shorter upper bound.

(iii) Only one cross peak is resolved, visible, or assigned. In such cases this cross peak may represent the only intensity resulting from the NOE interaction, since the other cross peak may have zero intensity or the two cross peaks may be degenerate. Therefore, the single observed intensity should be used to set the upper bound with a multiplicity correction corresponding to division of the intensity by two (Z = 1.12).

Identical principles apply for constraints between two nonstereoassigned diastereotopic groups. In such situations there are up to four distinct NOE interactions between individual signals in the different groups, and each must be assessed when determining the appropriate upper bound for the corresponding single averaged constraint. Table 1 shows how an overall upper bound may be deduced for any distribution of the four individual NOE interaction intensities amongst the different intensity categories, using the principles outlined above in (i), (iia) and (iii). For each case shown in Table 1, the four intensities are first grouped into two pairs and a combined upper bound deduced for each pair; these two upper bounds are then further combined to give a single upper bound derived from all four individual intensities. Although there are three possible initial pairings of four intensities (only one of which is illustrated, corresponding to 'combination horizontally'), these all give the same final result in every case.

The simple rule that emerges from Table 1 is as follows. Suppose that the strongest intensity from amongst the four possible NOE interactions linking the signals of the two diastereotopic groups corresponds to an upper bound of length A. If all four NOE intensities fall into this category, then the overall upper bound is also given by A. When either three or two intensities correspond to an upper bound of A, then the overall upper bound is given by 1.12 A, and when only one intensity corresponds

to an upper bound of A, the overall upper bound is given by 1.26 A [that is $4^{1/6}$ A ≈ $(1.12)^2$ A]. Provided, as before, that the constraint boundaries are all separated by at least 12% (i.e., 1.12 A ≤ B, 1.12 B ≤ C and 1.12 C ≤ D), this approach will always result in the strongest constraint consistent with data in cases 1, 2, 3, 5, 6 and 8 of Table 1. In case 4 it will do so, provided that $(1.12)^2$ A ≤ B, while in case 7 the requirement is: $(1.12)^2$ A ≤ C. These conditions hold for most of the constraint upper bound definitions used in practice, so it is likely that this approach will make the best use of the data in most cases where a single upper bound has to be derived for an NOE interaction between two nonstereoassigned diastereotopic groups.

r^{-6} Summation

Recently, a subtle variation on r^{-6} averaging, called r^{-6} summation, has been proposed by Nilges and incorporated into the program XPLOR (Brünger, 1992). In this approach, the distance measured in the evolving structural model in order to test constraint validity is defined, not as in Eq. 4 by the r^{-6} average, but instead by the (shorter) r^{-6} sum:

$$r_{eff} = \left(\sum_{i,j} r_{I_i S_j}^{-6} \right)^{-1/6} \tag{10}$$

This definition necessarily implies a change in the relationship between distance and NOE interaction intensity. For r^{-6} averaging, the relationship between intensity (I) and distance is given by:

$$\frac{I}{n_I n_S} = k \frac{1}{n_I n_S} \sum_{i,j} r_{I_i S_j}^{-6} \tag{11}$$

(where k is a constant of proportionality), whereas for r^{-6} summation, the division by $n_I n_S$ is simply removed from both sides of Eq. 11, giving:

$$I = k \sum_{i,j} r_{I_i S_j}^{-6} \tag{12}$$

Thus, when r^{-6} summation is used instead of r^{-6} averaging, the appropriate value to use for each NOE interaction intensity changes from an average to a sum, analogously to the change in the definition of r_{eff}.

The implication of these equations is that no multiplicity corrections should be made for NOE interactions involving genuinely equivalent spins when using r^{-6} summation (Constantine et al., 1994,1995). Uncorrected upper bounds should be used, simply as though the corresponding cross peaks had connected two single-proton resonances. However, in the case of nonstereoassigned diastereotopic groups, as before the situation is more complex. The intensity that should now be used is the sum over all the relevant NOE cross peaks, rather than their average intensity as would be used in conjunction with r^{-6} averaging calculations. When numerical intensities are available such sums may be readily calculated, but when intensities are only divided into categories a procedure based on corrections to upper bounds is appropriate.

The required corrections may be found by an 'inverse analogy' to the r^{-6} averaging case, since now it is required to *reduce* some upper bounds to account for the *increase* in intensity that occurs if intensities are summed. For a single proton interacting with a nonstereoassigned diastereotopic group, cases (i)–(iii) again apply, as defined in the previous section. In case (i) (both cross peaks in the same intensity category), the upper bound for the combined constraint must be *divided by* 1.12, corresponding to the doubling of intensity that would occur if the two cross peaks were summed. In case (ii) (cross peaks in different intensity categories), one may either (a) take the stronger cross peak and use no correction, or else (b) take the weaker cross peak and divide the corresponding upper bound by 1.12; as before, approach (a) will give a stronger constraint, provided that successive upper bound categories do not differ by less than 12%. In case (iii) (only one cross peak observed), no correction is needed. For interactions between two nonstereoassigned diastereotopic groups, there is an inverse analogy with the results presented in Table 1. When all four intensities correspond to A (where A is the upper bound corresponding to the strongest intensity category amongst the four NOE interactions between the two groups), the overall upper bound is given by A/1.26; when either three or two intensities correspond to A, the overall upper bound is given by A/1.12, and when only one of the four intensities corresponds to A, the overall upper bound is given simply by A.

Thus, overall, it may be seen that r^{-6} summation and r^{-6} averaging are essentially identical in their effects (when appropriate corrections are applied in each case), even though the conceptual link between r_{eff} and the evolving structural model is perhaps slightly more obscure in the case of r^{-6} summation. However, there is a difference between r^{-6} summation and r^{-6} averaging concerning the effect of omitting all multiplicity corrections. As discussed in the previous section, in the case of r^{-6} averaging this would cause constraints involving equivalent groups and most nonstereoassigned diastereotopic groups to be set too short, resulting in significantly overconstrained structures. In contrast, omission of all multiplicity corrections is not an invalid approach for calculations using r^{-6} summation (Constantine et al., 1994,1995). For equivalent groups it is exactly correct, while for nonstereoassigned diastereotopic groups it causes some constraints to be set too long, so that the information content of some cross peaks is not fully exploited. As far as we are aware, all applications of r^{-6} summation so far reported have omitted multiplicity corrections.

TABLE 2
PSEUDOATOM CORRECTIONS FOR GROUPS OF EQUIVALENT OR NONSTEREOASSIGNED PROTONS

Equivalent group	Averaging regime	Pseudoatom correction[a] (Å)	Conventional value[b] (Å)	$(r_{SQ} - r_{eff})_{max}$ in X-PLOR structures of CD59[c] (Å)
CH_3	Tropp	0.4	1.0	0.415 ± 0.003 (N = 1529)
Nonstereoassigned CH_2	r^{-6}	0.7	1.0	0.667 ± 0.004 (N = 4800)
Aromatic H_2/H_6 or H_3/H_5	r^{-6}	2.0	2.0	1.926 ± 0.014 (N = 794)
Nonstereoassigned $CH(Me)_2$	Mixed (see Eq. 8)	1.5	2.4	1.349 ± 0.044 (N = 490)
Nonregioassigned $CONH_2$	r^{-6}	0.7	1.0	0.634 ± 0.002 (N = 642)
NH_3^+ (e.g. of Lys)	Tropp	0.4[d]	1.0	0.386 ± 0.001 (N = 299)
Nonregioassigned $N^{\eta}H_2$ of Arg	r^{-6}	1.8[e]	2.2	1.675 ± 0.001 (N = 100)

[a] Values are calculated using standard X-PLOR bond lengths and angles and are rounded up to the next highest 0.1 Å, except that for methyl groups (see Methods section for details).
[b] These values represent the maximum distance that the pseudoatom can lie beyond the nearest member of the equivalent group, as viewed from an external spin (see text). Values are taken from Wüthrich et al. (1983), choosing the values for 'long-range' constraints as defined in that paper.
[c] For each type of equivalent group, a grid search to find $(r_{SQ} - r_{eff})_{max}$ was carried out independently for every occurrence of the group in the calculated XPLOR structures of CD59 (ensemble III only, chosen since it has the greatest deviations from geometric ideality). The mean, standard deviation and sample size (N) are reported for each type of equivalent group analysed (a very small number of calculations where the grid search failed were discarded from the statistics). (See Methods section for further details.)
[d] For an unprotonated NH_2 group this value would become 0.7 Å, as for a CH_2 group.
[e] This value assumes the guanidinium group to be protonated at N^{η}, so forming a planar symmetrical group $C^{\zeta}(N^{\eta}H_2)_2^+$.

The r^{-6} summation approach was originally mentioned by Levy (1989), and was first applied to dealing with the assignment ambiguity between intra- and inter-unit NOE interactions in symmetric dimers (Nilges, 1993). More recently, it has been applied to deal more generally with limited NOESY cross-peak assignment ambiguity (Nilges, 1995). Although these uses may appear superficially very different from earlier applications of r^{-6} averaging, in fact r^{-6} averaging could also be used in these contexts in an almost identical fashion to r^{-6} summation, provided only that NOE intensities are divided by the number of possible assignments involved for each cross peak before setting upper bounds.

Pseudoatom corrections

During structure calculations employing 'centre-averaging', the atoms of an equivalent group I are represented using a pseudoatom Q at the mean position of the individual atoms I_i. For most geometries, the distance r_{QS} from Q to some other atom S, involved in a constraint with I, is longer than the shortest of the individual distances r_{I_iS}. However, that shortest distance r_{I_iS} is the main determinant of the distance 'sensed' by the NOE, r_{eff}, using which the constraint upper bound is set. Therefore, in the correct structure, the distance r_{QS} to the pseudoatom may be longer than the upper bound unless a pseudoatom correction is added to the upper bound.

Pseudoatom corrections were originally defined as the distance between any one of the equivalent atoms I_i and the pseudoatom Q (Wüthrich et al., 1983). Such a correction must always be sufficiently long, since it represents the maximum distance that the pseudoatom could lie beyond the closest of the individual spins I_i in the equivalent group, as viewed from the external spin S. However, in effect this definition corresponds to an assumption that only the closest of the individual spins I_i contributes to the NOE interaction with S, whereas in reality the other spins in the group I also contribute to the NOE interaction.

If, instead, one defines the pseudoatom correction as the maximum possible value of $r_{QS} - r_{eff}$, this takes all members of the equivalent group into account. Under this definition, the pseudoatom correction is the maximum value by which the distance from S to the pseudoatom can exceed the averaged distance 'sensed' by the NOE and used to set the constraint upper bound. This correction will always be smaller than that based on the original definition, to an extent that depends on how close together the atoms are within the equivalent group, and whether averaging occurs over r^{-6} values or r_{Tropp} values. However, the value of r_{eff} also depends on the geometry of the interaction (i.e. the relative positions of all the interacting spins I_i and S), so defining a pseudoatom correction for general use implies finding the geometry for which $r_{QS} - r_{eff}$ is a maximum. This geometry is necessarily that in which r_{eff} is the most strongly dominated by the shortest of the individual distances r_{I_iS}, and this property makes it trivial to find in several cases. For instance, for a methyl group the geometry corresponding to the maximum value of $r_{QS} - r_{eff}$ has the external spin S lying on a line connecting the nearest methyl proton and the pseudoatom, and in van der Waals contact with the nearest methyl proton. However, in the more complicated case of an isopropyl group interacting with a single spin, a numerical search was required to establish the pseudoatom correction. The observation that smaller pseudoatom corrections could be defined based on this approach is not new; Koning et al. (1990) have proposed that a pseudoatom correction of 0.3 Å is sufficient for a methyl group. As far as we are aware, this idea has not previous-

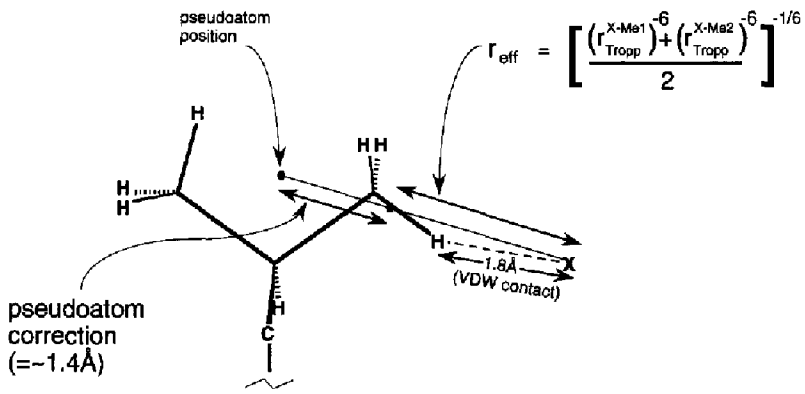

Fig. 2. Geometry used for calculation of the pseudoatom correction appropriate for isopropyl groups of leucine or valine residues. X represents the position of an external seventh spin such that the difference between r_{eff} and the distance from X to the pseudoatom is maximal. See text for further discussion.

ly been generalised, nor has its use in structure calculations been tested. Schemes for reducing the values of pseudoatom corrections for nonstereoassigned diastereotopic groups have also been proposed (Güntert et al., 1991).

Table 2 shows the values obtained for pseudoatom corrections using the revised definition, and compares these to the original values. As expected, the pseudoatom corrections that are most affected by the new definitions are those where the spins in the group are closest together, and where r_{Tropp} rather than r^{-6}-averaged distances are involved in r_{eff}. Thus, the correction for a methyl group differs quite substantially from the conventional value, as does that for an isopropyl group, whereas that for a pair of equivalent aromatic protons differs by less than 0.1 Å. In the original paper on pseudoatom corrections, reduced corrections were defined for sequential and intraresidue distances, based on the additional restrictions of stereochemistry that apply in these cases (Wüthrich et al., 1983). However, none of the pseudoatom corrections proposed here are any longer than the reduced 'short-range' definitions from the original work, so there is no requirement for separate 'short-range' pseudoatom corrections under the present definition.

TABLE 3
RMS DEVIATIONS FROM GEOMETRICAL IDEALITY OF THE VARIOUS ENSEMBLES OF CALCULATED STRUCTURES

Ensemble	Rmsd		
	Bond lengths (Å)	Bond angles (°)	Improper angles (°)
CD59 I	0.00125	0.315	0.188
CD59 II	0.00179	0.426	0.262
CD59 III	0.00492	0.642	0.546
CD59 IV	0.00286	0.469	0.365
CD59 V	0.00272	0.511	0.385
CD59 VI	0.00276	0.515	0.390
HMG-D I	0.00473	0.633	0.544
HMG-D II	0.00288	0.539	0.361
HMG-D III	0.00527	0.714	0.544
HMG-D IV	0.00339	0.560	0.357
HMG-D V	0.00304	0.542	0.331
HMG-D VI	0.00333	0.560	0.385

Deviations from ideality are calculated using XPLOR, and the rmsd calculated in each case for an ensemble of the 30 structures having the lowest total XPLOR energies. For each protein, ensemble I corresponds to conventional centre averaging, ensemble II corresponds to centre averaging with the new pseudoatom and multiplicity corrections, ensemble III corresponds to conventional r^{-6} averaging, ensemble IV corresponds to r^{-6} averaging with the new multiplicity corrections, ensemble V corresponds to conventional r^{-6} summation and ensemble VI corresponds to r^{-6} summation with the new multiplicity corrections.

Methods

Calculation of pseudoatom corrections

Pseudoatom corrections were determined using home-written software. In each case, the coordinates of the equivalent group I were fixed, and the value of $r_{QS} - r_{eff}$ (see Theory section for definitions) was calculated for locations of S at points on a user-defined grid. Locations where S violated the van der Waals volume of protons or heavy atoms of the equivalent group were discarded from the search. The search was localised and the resolution of the grid increased in subsequent rounds of calculation to a final resolution of 0.005 Å, and the highest value of $r_{QS} - r_{eff}$ found then rounded up to the nearest 0.1 Å to give the pseudoatom correction. The only exception is the pseudoatom correction for methyl groups, for which the grid search yielded $(r_{QS} - r_{eff})_{max} = 0.413$ Å. However, this value corresponds to a sterically unfavourable 'eclipsed' geometry in which the external spin S is colinear with the pseudoatom Q and the nearest methyl proton I_1, as opposed to more plausible 'staggered' geometries with S lying in the region between two methyl protons. Very small in-plane movements of the external spin away from the unfavourable 'eclipsed' conformation (< 0.08 Å increase in the radial separation r_{I_1S}, or < 7° axial rotation of

the methyl group) result in $r_{QS} - r_{eff}$ values of less than 0.4 Å, and it was therefore judged unnecessary to increase the methyl pseudoatom correction given in Table 2 to 0.5 Å.

Values for bond lengths and angles were taken from the X-PLOR force-field parameters (file 'parallhdg.pro' for simulated-annealing calculations, see Brünger, 1992). Bond lengths used were C–C (saturated) = 1.53 Å, C–C (aromatic) = 1.40 Å, C–H (all) = 1.08 Å, N–H (Lys NH_3) = 1.04 Å, N–H ($CONH_2$) = 0.98 Å, N–H (Arg N^{η}–H) = 1.00 Å, C–N (Arg C^{ζ}–N^{η}) = 1.305 Å; all tetrahedral bond angles were taken as 109.5°, and all trigonal bond angles were taken as 120°. van der Waals contact distances were taken as 1.8 Å (H–H), 2.2 Å (C–H) and 2.2 Å (N–H). In the case of the isopropyl group, three different limiting conformations were analysed, corresponding to one proton of one methyl group being either *cis*- or *trans*-coplanar with the carbon atom of the other in different combinations (i.e. *cis-cis*, *cis-trans* and *trans-trans*). The geometry shown in Fig. 2 corresponds to the largest value of $r_{QS} - r_{eff}$ found in any of these conformations. A slightly modified version of this software was used to calculate the validity boundaries for r^{-6}-averaged constraints illustrated in Fig. 1.

To test how distortions of equivalent group geometry might affect the required pseudoatom corrections, the same grid search procedure was carried out using coordinates of equivalent groups taken from the calculated structures of CD59; ensemble III was used for this purpose since it includes the greatest deviations from geometric ideality (see Table 3). Note particularly that the coordinates taken from the calculated structures include only the atoms of the equivalent group itself, *not* any external spin S that might interact with that group in the complete molecule. This procedure is intended only to sample distortions of the equivalent group itself, the position of the external spin S being the variable in the grid search procedure used to determine the required pseudoatom corrections. All occurrences of each type of equivalent group in all members of the ensemble were analysed independently, and, again with the exception of methyl groups, no values of $r_{QS} - r_{eff}$ larger than the proposed pseudoatom corrections were found (see Table 2). For methyl groups, values of $(r_{QS} - r_{eff})_{max}$ were mainly in the range 0.410–0.415 Å (see above). The smaller values of $r_{QS} - r_{eff}$ found for isopropyl groups probably reflect the many occurrences of conformations other than that shown in Fig. 2, and therefore do not imply that a value of 1.4 Å could have been defined for the pseudoatom correction in this case.

Structure calculations and NMR constraints

All structures were calculated using a slightly modified version of the YASAP simulated-annealing protocol (Nilges et al., 1988,1991; Brünger, 1992) using the program XPLOR 3.1. Full details are given in the original papers describing the structure determinations (Fletcher et al., 1994; Jones et al., 1994). For CD59, the NMR constraint set comprised 794 NOE-based distance constraints (165 intraresidue, 223 sequential, 119 medium-range ($2 \leq |i-j| \leq 4$) and 287 long-range ($|i-j| \geq 5$)], 3 lower-limit constraints based on missing sequential NH–NH NOE connectivities, 16 dihedral angle constraints (12 for χ_1 and 4 for ϕ), and 20 constraints for 10 hydrogen bonds. The C-terminal linked GPI anchor and the complex glycan linked to Asn^{18} were excluded from the model in these calculations. NOE constraints were classified (prior to pseudoatom and multiplicity corrections) as strong (≤ 2.3 Å),

TABLE 4
RMS DEVIATIONS BETWEEN THE AVERAGE STRUCTURES CALCULATED FOR EACH ENSEMBLE

Protein	Rmsd (Å) and proportion of distance constraints differing					
	I	II	III	IV	V	VI
CD59						
I	–	0.15 (67%)	0.55 (67%)	0.39 (67%)	0.34 (67%)	0.34 (67%)
II	0.15 (67%)	–	0.56 (67%)	0.38 (67%)	0.33 (67%)	0.35 (67%)
III	0.55 (67%)	0.56 (67%)	–	0.33 (61%)	0.41 (31%)	0.36 (41%)
IV	0.39 (67%)	0.38 (67%)	0.33 (61%)	–	0.16 (61%)	0.14 (67%)
V	0.34 (67%)	0.33 (67%)	0.41 (31%)	0.16 (61%)	–	0.11 (11%)
VI	0.34 (67%)	0.35 (67%)	0.36 (41%)	0.14 (67%)	0.11 (11%)	–
HMG-D						
I	–	0.20 (58%)	0.62 (58%)	0.45 (58%)	0.47 (58%)	0.42 (58%)
II	0.20 (58%)	–	0.57 (58%)	0.38 (58%)	0.42 (58%)	0.36 (58%)
III	0.62 (58%)	0.57 (58%)	–	0.33 (57%)	0.36 (22%)	0.32 (29%)
IV	0.45 (58%)	0.38 (58%)	0.33 (57%)	–	0.18 (57%)	0.09 (58%)
V	0.47 (58%)	0.42 (58%)	0.36 (22%)	0.18 (57%)	–	0.17 (9%)
VI	0.42 (58%)	0.36 (58%)	0.32 (29%)	0.09 (58%)	0.17 (9%)	–

The average structure for each ensemble was calculated using the program CLUSTERPOSE (see Methods section), in each case using the 30 ensemble members having the lowest total XPLOR energies and superposing them using the backbone atoms (N, C^{α} and C') of residues 1–75 for CD59, or residues 11–59 for HMG-D. The figures in brackets represent the percentage of NOE-based distance constraints which differ (due to the application of different corrections) between the constraint lists used to calculate the different ensembles.

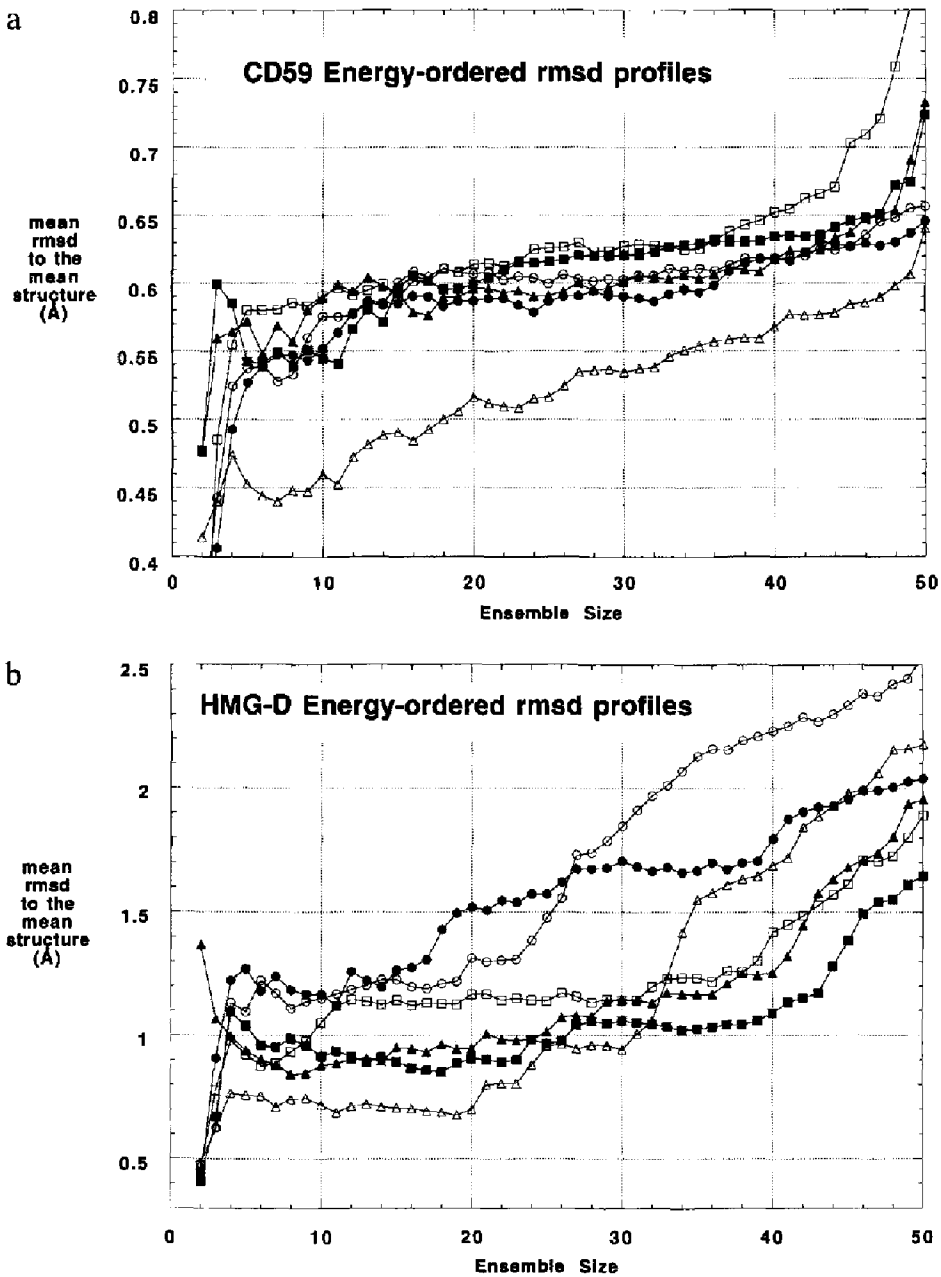

Fig. 3. Energy-ordered rmsd profiles for various ensembles of (a) CD59 and (b) HMG-D structures. For each profile, the mean rmsd to the mean structure is independently calculated for each ensemble size, and successive ensembles are generated by adding structures in order of increasing total XPLOR energy. Thus, the ensemble of size 2 comprises the two lowest energy structures, that of size 3 includes also the third lowest energy structure, and so on. The symbols are defined as follows: conventional centre-averaging (ensemble I), open circles; centre-averaging with new corrections (ensemble II), filled circles; conventional r^{-6} averaging (ensemble III), open triangles; r^{-6} averaging with new corrections (ensemble IV), filled triangles; conventional r^{-6} summation (ensemble V), open squares; r^{-6} summation with new corrections (ensemble VI), filled squares. See text for further discussion. In (c) and (d) are shown the corresponding total XPLOR energies of each structure, arranged in increasing order, i.e. in the same order as structures are added to the ensembles in (a) and (b). Symbols correspond to ensembles I, II, III, IV, V and VI in the same way as in (a) and (b). Note the logarithmic scale of the energy axes.

medium (≤2.9 Å), weak (≤3.6 Å) or very weak (≤5.0 Å). For HMG-D, the NMR constraint set comprised 794 NOE-based distance constraints (360 intraresidue, 203 sequential, 143 medium-range ($2 \leq |i-j| \leq 4$) and 88 long-range ($|i-j| \geq 5$)), 6 lower-limit constraints based on missing sequential NH–NH NOE connectivities, 49 dihedral angle constraints (8 for χ_1 and 41 for ϕ), and 12 constraints for 6 hydrogen bonds. NOE constraints were classified (prior to pseudoatom and multiplicity corrections) as strong (≤ 2.3 Å), medium (≤ 2.9 Å), weak (≤ 3.5 Å) or very weak (≤ 5.0 Å). For each protein, the only differences between the different sets of calculations reported in this paper involve the addition of different pseudoatom corrections and multiplicity corrections to upper limits for

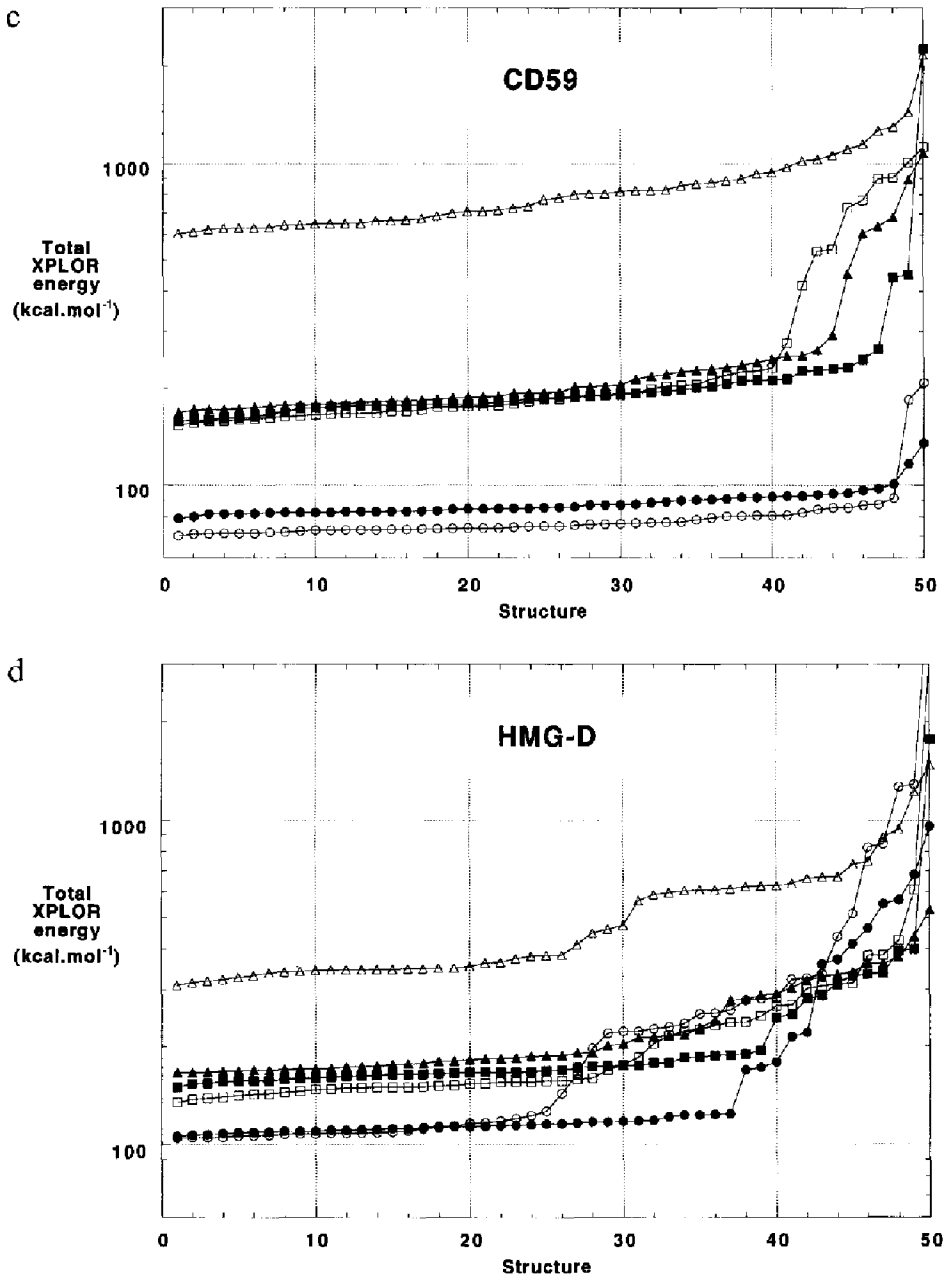

Fig. 3. (continued).

equivalent groups and nonstereoassigned diastereotopic protons; the extent of these differences between the various constraint sets are summarised in Table 4. (Relative to the constraint sets used in the original papers, the only other differences involve the reclassification of constraints involving nonstereoassigned diastereotopic protons according to the protocol given in the Theory section, and a very small number of other changes based on a further round of refinement for each protein.)

Comparison of global precision between ensembles

We have chosen a method of reporting that includes the results of all calculations, similar to recent proposals by Widmer et al. (1993) and Kohda and Inagaki (1992). Ensembles are built up by adding successive structures in increasing order of their total energies, as defined by the force field active at the end of the structure calculation (E(total) = E(bond) + E(angle) + E(improper) + E(van der Waals) + E(NOE constraints) + E(dihedral constraints); force constants as defined in the standard protocol 'sa.inp' in the XPLOR 3.1 release). Thus, the first ensemble is formed from the structure with the lowest total energy and that with the second lowest, the second ensemble comprises the three lowest energy structures, and so on until all fifty structures are included. The mean rmsd to the mean structure is independently calculated at each stage (using a global simultaneous superposition rather than sequential pairwise fits) within the program CLUS-

TERPOSE (Diamond, 1992,1995), and plotted against increasing ensemble size.

We refer to this form of presentation as an 'energy-ordered rmsd profile'. In the literature generally, statistics for NMR structures are often reported for some rather arbitrarily chosen subset of a larger ensemble. In contrast, we suggest that when all of the data are shown in the form of an energy-ordered rmsd profile, together with the corresponding energy profiles themselves, this allows the justification for rejecting structures and the validity of the reported rmsd to be assessed more objectively. Generally, the width of a 'plateau region' in the energy-ordered rmsd profile reflects the ability of the structure calculations to converge consistently on a similar result, which in turn is a property both of the method of calculation and of the information content of the constraint set. In the present case, since an identical calculation protocol was used for both proteins, the fact that plateaux in the energy-ordered rmsd profiles for CD59 (Fig. 3a) are significantly broader and flatter than those for HMG-D (Fig. 3b) must reflect the higher information content of the CD59 constraint set (see Results section).

Comparison of local structure between ensembles

Torsion angles ϕ, ψ, χ_1 and χ_2 were measured where possible for all residues in the 30 structures having the lowest total XPLOR energy in each ensemble, and the corresponding mean angle and angular order parameter S^θ (Hyberts et al., 1992) calculated for each angle in each ensemble. Angular standard deviations $\sigma(\theta)$ were calculated from the angular order parameters using the approximate relationship:

$$\sigma(\theta) = 2\arccos[1 + 0.5\log(S^\theta)] \qquad (13)$$

Results and Discussion

In order to test the impact on protein structure calculations of these revised approaches to multiplicity and pseudoatom corrections, we recalculated the structures of two proteins recently determined in this laboratory, namely CD59 (a 77-residue glycosylated human complement control protein (Fletcher et al., 1994)), and a 74-residue fragment of HMG-D (a DNA-binding protein from *Drosophila melanogaster* (Jones et al., 1994)). These proteins were chosen because they form a contrasting pair. The compact and globular CD59 structure is quite precisely defined by a constraint set that contains many long-range constraints distributed throughout the sequence, and that probably contains much redundant information. The HMG-D structure is significantly less well constrained by the data. Partly this is because the HMG-D molecule is predominantly composed of α-helices and is L-shaped, so that the constraint list contains relatively few long-range distances and the angle between the 'arms' of the structure remains poorly defined. Partly it may also be because the spectra of HMG-D are less well dispersed than those of CD59, making it intrinsically more difficult to assign some NOE connectivities. (Neither protein was available in ^{13}C-labelled form.) For CD59, rmsd values are reported for backbone atoms (N, C^α, C') of residues 1–75. For HMG-D, rmsd values are reported for backbone atoms of residues 11–59 so as to exclude contributions from the N- and C-terminal regions, which although well-defined locally have poorly defined positions relative to the rest of the molecule (Jones et al., 1994).

For each protein, the following six sets of structure calculations were run, in each case generating an ensemble of 50 conformers:

(I) 'Conventional' centre-averaging (pseudoatom corrections as in Table 2, column 4; multiplicity corrections restricted to 0.5 Å per methyl group).

(II) Centre-averaging with new multiplicity corrections (from Eq. 9) and pseudoatom corrections (Table 2, column 3).

(III) 'Conventional' r^{-6} averaging (multiplicity corrections restricted to 0.5 Å per methyl group).

(IV) r^{-6} averaging with new multiplicity corrections (from Eq. 9).

(V) 'Conventional' r^{-6} summation (no multiplicity corrections).

(VI) r^{-6} summation with new multiplicity corrections (division by Z (as defined in Eq. 9) for certain nonstereoassigned diastereotopic groups; see above).

Global comparison of centre-averaged calculations with and without new corrections (ensembles I and II)

Figure 3a shows that, for CD59, introducing the new pseudoatom corrections and multiplicity corrections causes a very small increase in the global precision of the structures calculated using centre averaging (filled circles are generally lower than open circles). This small change results from a combination of opposing factors, since the revised pseudoatom corrections act to tighten constraints, whereas the multiplicity corrections act to loosen them. The balance between these factors varies from constraint to constraint, depending on the type of equivalent group involved and the value of the uncorrected constraint upper bound. For example, for a constraint to a methylene group where the uncorrected upper bound is U Å, conventionally the corrected upper bound would be U+1.0 Å, while under the new definition it would be 1.12 U+0.7 Å. Therefore, the revised definitions will result in a tighter corrected constraint whenever the uncorrected upper bound is shorter than $(1.0-0.7)/(1.12-1) = 2.5$ Å. Similarly, for a methyl group the new definitions result in a tighter constraint whenever the uncorrected upper bound is less than 5.5 Å, while for an isopropyl group the corresponding distance is 5.7 Å (these values assume addition

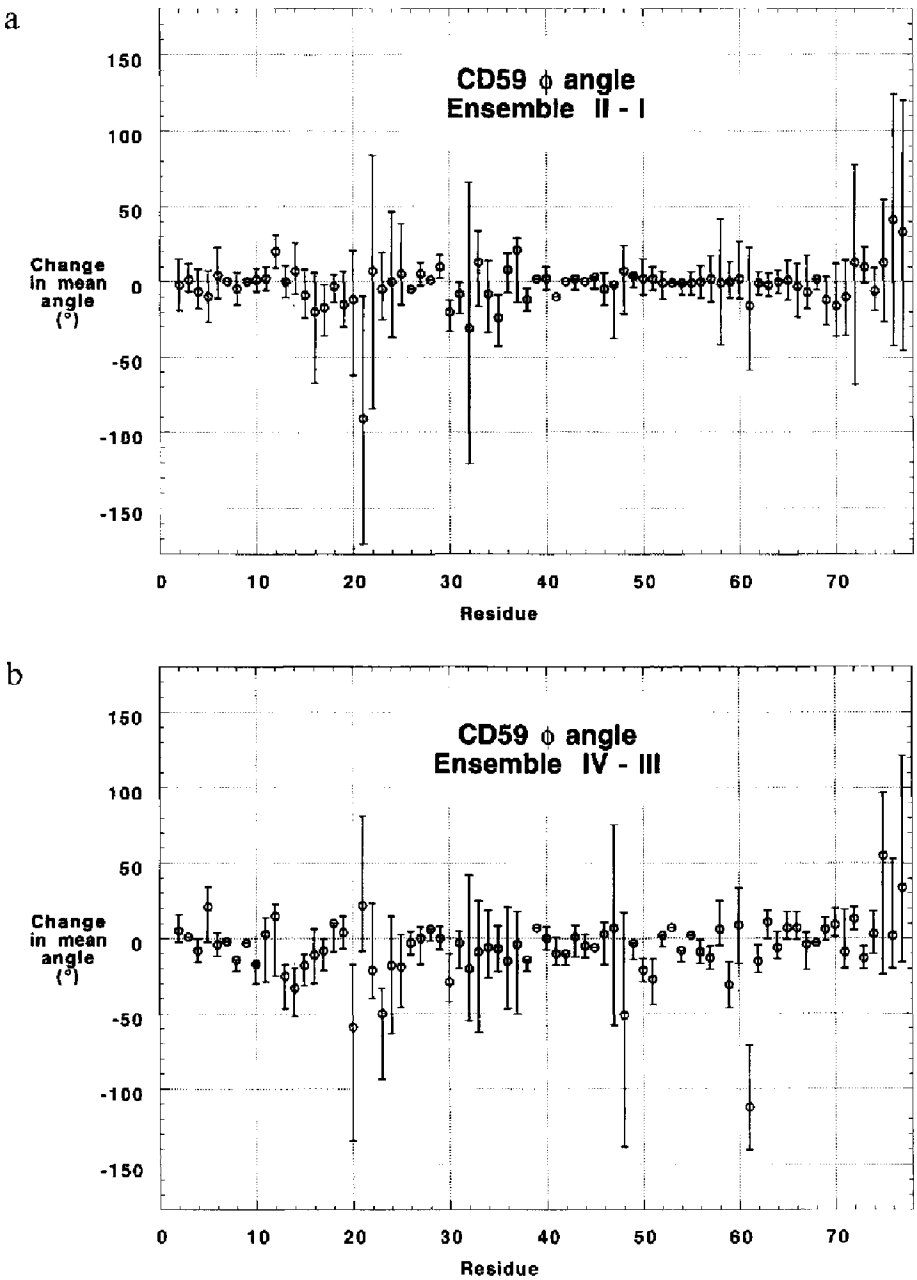

Fig. 4. Changes in the backbone ϕ angles between the various ensembles I–VI calculated for CD59. Each circle represents, for a particular ϕ angle in a particular residue, the shift in the mean value of that angle between the two ensembles being compared. Defining this shift as mean (ensemble A) – mean (ensemble B), the upper error bar on each circle represents the standard deviation of the angle in ensemble A, while the lower error bar represents the standard deviation of the angle in ensemble B. These error bars allow one to assess the relative significance of each shift (see text). Panels (a)–(c) show differences resulting from the introduction of the new corrections proposed in this paper, as follows: (a) ensemble II – ensemble I (new versus conventional centre averaging); (b) ensemble IV – ensemble III (new versus conventional r^{-6} averaging); (c) ensemble VI – ensemble V (new versus conventional r^{-6} summation). Panels (d)–(f) show differences between the various averaging methods, using the new corrections in each case: (d) ensemble IV – ensemble II (r^{-6} averaging versus centre averaging); (e) ensemble VI – ensemble II (r^{-6} summation versus centre averaging); (f) ensemble IV – ensemble VI (r^{-6} averaging versus r^{-6} summation). Similar plots were produced for the ψ, χ_1 and χ_2 angles in CD59, and for the ϕ, ψ, χ_1 and χ_2 angles in HMG-D; these are available as Supplementary Material.

of a 0.5 Å multiplicity correction for methyl groups under the old definitions, as discussed in the Theory section).

In contrast, since the pseudoatom correction for symmetry-related aromatic protons is essentially unchanged, constraints to such protons are always looser under the new definitions, as a result of applying the multiplicity correction. One might argue, however, that multiplicity corrections could almost be neglected in this case. Consider an NOE constraint between a single external proton and a pair of symmetry-related, motionally averaged aromatic protons, and suppose no multiplicity correction is

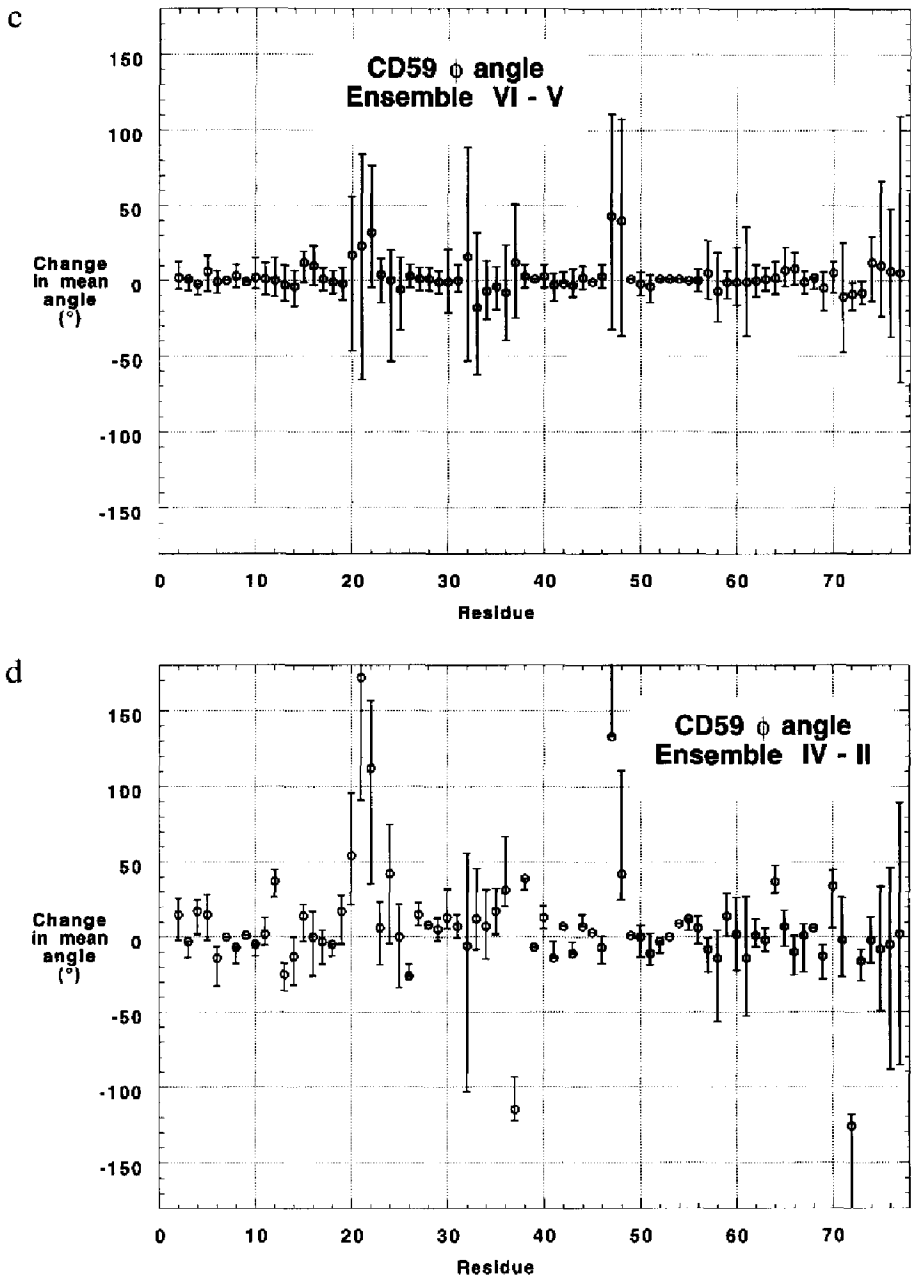

Fig. 4. (continued).

made. If the external proton lies nearly equidistant from the two aromatic protons, then the observed NOE cross peak comprises significant contributions from both aromatic protons, and failure to divide its intensity by 2 would formally be incorrect. However, because the 2 Å pseudoatom correction that is added to the constraint is excessive for this interaction geometry, the constraint is not in fact too short, even if the multiplicity correction is not made. If, instead, the external proton lies close to only one of the aromatic protons, now the NOE cross peak contains no significant contribution from the other aromatic proton. Division of its intensity by 2 is then essentially unnecessary, and again the constraint is not too short, even if the multiplicity correction is not made. This argument may account for the fact that, although an empirical correction to upper bounds for constraints to methyl groups was found to be needed soon after the first NMR protein structures were determined, no corresponding correction seems to have been introduced for constraints to symmetric aromatic groups. The above argument notwithstanding, in this work we have included multiplicity corrections in all cases.

Use of the new pseudoatom corrections and multiplicity corrections results in a small increase in the XPLOR energies of the structures calculated using centre averaging (Fig. 3c). Consistent with this, ensemble II for CD59 shows slightly higher overall deviations from geometric ideality than does ensemble I (calculated in each case for

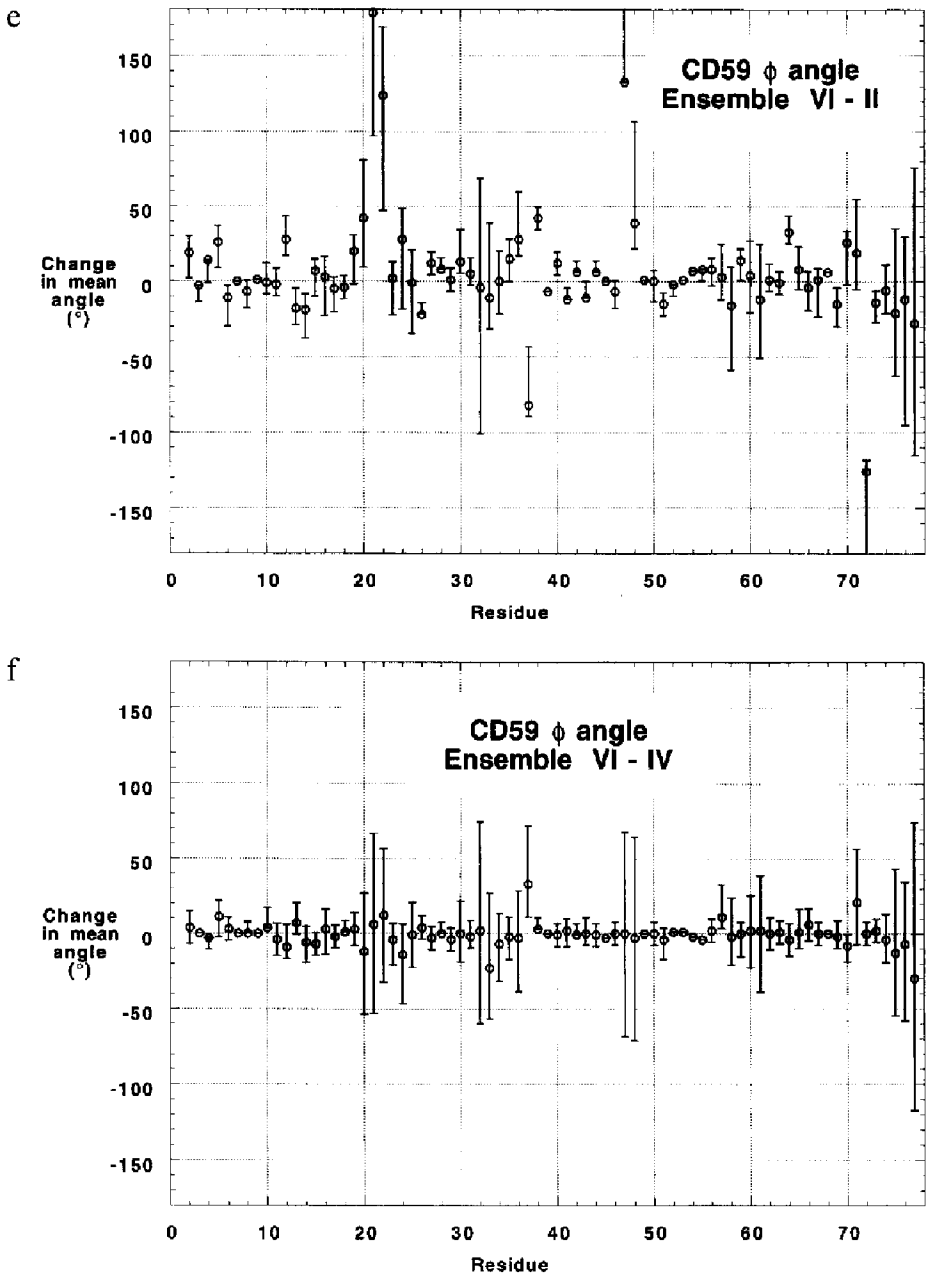

Fig. 4. (continued).

the thirty lowest energy structures; Table 3), showing that the geometric terms and the NMR-derived terms in the force field arrive at a slightly different balance in terms of their effects on the atomic positions in the two cases. However, differences between these ensembles are very small, as emphasised by the small rmsd between their average structures (Table 4).

In the case of HMG-D, the XPLOR energies (Fig. 3d) show that in some ensembles (particularly I and III) only about half of the calculations have converged on a consistent result, making the second half of the energy-ordered rmsd profiles (Fig. 3b) of less significance in these cases. As for CD59, the precision of the two centre-averaged ensembles is again quite similar, although the absence of significant plateau regions in the energy-ordered rmsd profiles for HMG-D makes generalisations more difficult. While the deviations from geometric ideality (Table 3) are lower for ensemble II than for I, this is almost certainly due to the inclusion of several somewhat higher energy structures within the first 30 members of HMG-D ensemble I. Also as for CD59, the rmsd between the average structures of ensembles I and II shows that these are very similar (Table 4).

Global comparison of r^{-6}-averaged calculations with and without new corrections (ensembles III and IV)

In contrast to the centre averaging case, for r^{-6}-averaged structures quite large differences emerge when calculations

with and without the new corrections are compared. For CD59, ensemble III shows a significantly higher precision (smaller rmsd) than ensemble IV throughout the energy-ordered rmsd profile (Fig. 3a), while for HMG-D (Fig. 3b) the same is true for that part of the profile corresponding to converged structures (i.e. until about structure 32). Thus, for both proteins it is clear that failure to include multiplicity corrections leads to unjustifiably overconstrained structures. (We note here that our previously published rmsd figure of 0.85 Å for HMG-D (Jones et al., 1994), obtained from structures calculated without multiplicity corrections, must be revised to 1.1 Å in the light of the present results.)

Not unexpectedly, these differences in precision are also reflected in the energies and geometries of the calculated structures. For CD59, the total XPLOR energy terms for ensemble III (Fig. 3c) are more than threefold higher than those for ensemble IV in the plateau region, while for HMG-D (Fig. 3d) there is also a significant, though somewhat smaller, difference. For CD59, deviations from geometric ideality are considerably higher for ensemble III than for ensemble IV (Table 3), while once more HMG-D shows a rather smaller difference. These differences again reflect the fact that when multiplicity corrections are not included in r^{-6}-averaged calculations, the resulting structures are unjustifiably overconstrained. However, it must be emphasised that these conclusions only emerge when the various ensembles are *compared*. There is no particular feature of the global results for ensemble III alone that would allow the problem to be recognised.

Global comparison of r^{-6}-summed calculations with and without new corrections (ensembles V and VI)

The number of multiplicity corrections required in conjunction with r^{-6} summation calculations is much smaller than for either of the other averaging methods (Table 4), comprising only a small subset of the constraints to nonstereoassigned diastereotopic groups, so it might be expected that their inclusion would have little impact on the results of structure calculations. For CD59, this is indeed the case. The energy-ordered rmsd profiles for ensembles V and VI are almost identical (except that the former has somewhat more outliers; Fig. 3a), as also are the XPLOR total energies (Fig. 3c) and the deviations from geometric ideality (Table 3), while the rmsd between the average structures of ensembles V and VI is very low (Table 4). However, in the case of HMG-D, there is a larger difference between ensembles V and VI. Although the XPLOR total energies and deviations from ideality for ensembles V and VI are again very similar, and the rmsd between their average structures is low, a significantly better precision is obtained for ensemble VI, where multiplicity corrections are included, than for ensemble V, where they are omitted (Fig. 3b).

Global comparisons between different averaging methods

For CD59, the precisions of ensembles I, II, IV, V and VI are all very similar, while that of ensemble III (r^{-6} averaging without multiplicity corrections) differs markedly because it is overconstrained (Fig. 3a). Ensembles I, II and IV are barely distinguishable on grounds of precision, but it does appear that the two r^{-6}-summed ensembles (V and VI) are very slightly less precise than the r^{-6}-averaged ensemble with corrections (IV), at least in the statistically more significant central region of the profiles (ensemble size approximately 20–40). In contrast, when the total XPLOR energies (Fig. 3c) or rmsd values between average structures (Table 3) are compared, it is ensembles IV, V and VI that are essentially identical. The centre-averaged ensembles (I and II) have significantly lower energies and their average structures are somewhat different to those resulting from any of the valid r^{-6} methods (i.e. ensembles IV, V and VI).

For HMG-D, there is much more variation amongst the precisions of the different ensembles. Although comparisons are more difficult than for CD59, since a smaller proportion of all calculations converge on a consistent result in each ensemble (Fig. 3b), two fairly clear conclusions emerge. First, there is a clear advantage in not using centre-averaged calculations when the NOE data are relatively sparse; ensembles I and II show significantly worse precision than any of the others for the case of HMG-D. Second, when multiplicity corrections are made in both cases, r^{-6} averaging and r^{-6} summation give very similar results (although, as previously discussed, they give very different results when multiplicity corrections are omitted). For the total XPLOR energies, deviations from geometrical ideality and differences amongst the average structures (Fig. 3d; Tables 3 and 4), very similar conclusions apply as for CD59.

Analysis of local changes in conformation

Figure 4 shows how the ϕ angles in CD59 vary between the different ensembles I to VI. It is rather difficult to draw concrete conclusions from these data. The main trends as a function of sequence affect the angular standard deviations, which are high in poorly defined loop regions (e.g. residues 20–23 and 31–33) and are generally low in well-ordered regions of regular secondary structure. However, these trends largely reflect the distribution of constraints within the structure, which, since it does not vary from one ensemble to another, probably has little bearing on comparisons between ensembles.

If one defines a significant angular change between two ensembles as one that is greater than the standard deviation of the corresponding angle in either ensemble, then only a relatively small number of ϕ angles change significantly between any of the ensembles (in Fig. 4, these changes correspond to the circles that lie further from zero than the length of either of their attached error

bars). Figures 4a–c show, for each averaging method in turn, the difference between using new and old corrections. The only case where such differences are appreciable is in Fig. 4b (ensemble IV – ensemble III), probably reflecting mainly the overconstrained nature of ensemble III. Figures 4d–f show differences between the three averaging methods, using the new corrections in each case. These figures show the very close similarity between the structures calculated using r^{-6} averaging (ensemble IV) and r^{-6} summation (ensemble VI). Not only are the differences between ensembles IV and VI very small (Fig. 4f), but also the differences between the centre-averaged ensemble (II) and either ensemble IV (Fig. 4d) or ensemble VI (Fig. 4e) are virtually identical to one another.

Similar comparisons (six differences between ensembles, as defined in Fig. 4) were calculated for the ψ, χ_1 and χ_2 angles in CD59, and also for the ϕ, ψ, χ_1 and χ_2 angles in HMG-D (data not shown; plots available as Supplementary Material). Broadly similar conclusions apply for side-chain angles as for main-chain angles. Some side chains apparently change their rotamer distribution appreciably between the different ensembles (i.e. the mean χ_1 angle shifts by approximately 60°), but few of these changes are significant under the definition given earlier. A few significant changes can be found in χ_2 angles. Overall, however, changes in the backbone and side-chain angles appear to be distributed diffusely over the structure, with no obvious correlations between the locations of significant angular changes and the positions of side chains for which relatively large changes in the values of pseudoatom corrections or multiplicity corrections have been introduced.

Conclusions

It would clearly be dangerous to make many generalisations based on a study limited to two proteins and one calculation protocol. Also, our definitions of 'conventional' multiplicity and pseudoatom corrections, against which we compare our proposed new methods, are necessarily rather limited given the tremendous variety of different protocols used for NMR structure determination. Indeed, in some matters it is very difficult to assess what some other workers have done, since details such as how or whether particular categories of signals were corrected for multiplicity are frequently not reported at all. Nonetheless, some conclusions do emerge from the present work that we believe are likely to be of widespread relevance and validity.

Firstly, our results have a bearing on comparisons between structures calculated using centre averaging and those calculated using r^{-6} methods. Although there have not previously been many direct comparisons of these techniques applied to the same protein, the view has evidently gained ground that r^{-6}-averaged or r^{-6}-summed structures are inherently more precise than centre-averaged structures. The results for CD59 show that this is not necessarily true. However, the situation here is complicated by the sensitivity of r^{-6}-averaged calculations to the correct use of multiplicity corrections. The present results show very clearly that failure to apply correct multiplicity corrections results in overconstrained structures whose significantly lower rmsd values are not justified by the data.

It is impossible to tell how widespread this issue may be in the literature generally, since the absence from a paper of any statement that multiplicity corrections were applied may or may not mean that they were actually omitted. Some of the present authors have omitted multiplicity corrections (those for aromatic and methylene protons) in r^{-6}-averaged calculations on two occasions in the past (Schwabe et al., 1993; Jones et al., 1994), but we would be surprised if these should have been the only instances. Cases where volume integration was used to measure NOE interaction intensities are likely to include multiplicity corrections, but in cases where r^{-6}-averaged calculations were used in conjunction with NOE constraints classified into semiquantitative intensity groups such as 'strong', 'medium' and 'weak', very few papers state that multiplicity corrections were made (for possibly the only example, see Constantine et al., 1992). Alternatively, r^{-6} summation may be used instead of r^{-6} averaging, when it is valid to omit all multiplicity corrections (Constantine et al., 1994,1995). The results for CD59 show that, for a structure that is well defined by the data, omission of multiplicity corrections in calculations with r^{-6} summation does not cause any appreciable degradation of precision.

In contrast to the data for CD59, those for HMG-D make it very clear that use of r^{-6} averaging or summation *can* cause a significant increase in precision relative to using centre averaging, and, perhaps more importantly, in the proportion of calculations that converge on a similar result. This difference in the effect of r^{-6} methods between the cases of CD59 and HMG-D is most likely to reflect the much greater number of long-range constraints present in the NOE data for CD59 than for HMG-D. It is clear that the information content of an individual constraint is more effectively harnessed when r^{-6} averaging or summation is used (see, for instance, Fig. 1). However, in cases where a structure is determined by a large number of interlocking constraints, the shortness of individual constraints may be of little importance when the positions of the relevant atoms are also tied down by several other constraints. Thus, in the case of CD59 the difference between r^{-6} methods and centre averaging is relatively slight, because many constraints play an active rôle in determining each atomic position, whereas in HMG-D the difference is more profound, because relatively few constraints are active in determining each atomic position.

These results agree well with those recently published by Gradwell and Feeney (1996), who showed that r^{-6} averaging was much more effective than centre averaging in defining the position and conformation of bound ligands when NOE data are sparse.

However, the data for HMG-D also highlight the importance of making multiplicity corrections for calculations with r^{-6} summation in cases where the NOE data are sparse. The corrections involved are relatively few, but in the context of a relatively poorly defined structure, the impact of allowing even a small proportion of constraints to remain unnecessarily loose is apparently significant. To our knowledge, such corrections have not previously been employed.

As discussed earlier, many of the issues raised in this paper can be avoided by adopting separate calibrations for different classes of constraint, an approach that is limited only by the availability of suitable NOE cross peaks corresponding to known distances for calibration within each class. It is also true that the need for pseudo-atom corrections can be avoided by using r^{-6} averaging or summation rather than centre averaging, and that use of computer-aided volume integration rather than manual 'contour counting' on paper plots largely eliminates ambiguity as to how to correct NOE intensities for multiplicity. We do not suggest that these older, manual approaches are in any way superior to the newer alternatives currently gaining ground. However, manual categorisation of intensities into broad ranges according to one overall distance calibration is even now probably still the commonest approach to quantifying NOE constraints, and calculations employing centre-averaging are widely used. The aim of this paper is to provide a valid framework for the treatment of NOE constraints involving equivalent groups and nonstereoassigned diastereotopic groups in any such circumstances.

Acknowledgements

We thank Martin Billeter, David Case (particularly for comments concerning the Tropp model), Keith Constantine (particularly for comments concerning multiplicity corrections when using r^{-6} summation), Peter Güntert, James Keeler, Michael Nilges (particularly for pointing out the significance of enforcing a consistent assignment when using floating stereoassignment, and for comments concerning r^{-6} averaging and the Tropp model), Michael Summers, Gerhard Wagner, Michael Williamson and Kurt Wüthrich for many helpful discussions, suggestions and critical reading of the manuscript.

References

Brünger, A.T., Clore, G.M., Gronenborn, A.M. and Karplus, M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3801–3805.

Brünger, A.T. (1992) *X-PLOR Version 3.1: A System for Crystallography and NMR*, Yale University, New Haven, CT, U.S.A.

Clore, G.M., Gronenborn, A., Nilges, M. and Ryan, C.A. (1987) *Biochemistry*, **26**, 8012–8023.

Constantine, K.L., Madrid, M., Bányai, L., Trexler, M., Patthy, L. and Llinás, M. (1992) *J. Mol. Biol.*, **223**, 281–298.

Constantine, K.L., Friedrichs, M.S., Metzler, W.J., Wittekind, M., Hensley, P. and Mueller, L. (1994) *J. Mol. Biol.*, **236**, 310–327.

Constantine, K.L., Friedrichs, M.S., Detlefson, D., Nishio, M., Tsunakawa, M., Furumai, T., Ohkuma, H., Oki, T., Hill, S., Bruccoleri, R.E., Lin, P.-F. and Mueller, L. (1995) *J. Biomol. NMR*, **5**, 271–286.

Diamond, R. (1992) *Protein Sci.*, **1**, 1279–1287.

Diamond, R. (1995) *Acta Crystallogr.*, **D51**, 127–135.

Fletcher, C.M., Harrison, R.A., Lachmann, P.J. and Neuhaus, D. (1994) *Structure*, **2**, 185–199.

Gradwell, M.J. and Feeney, J. (1996) *J. Biomol. NMR*, **7**, 48–58.

Güntert, P., Braun, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.

Hyberts, S.G., Goldberg, M.S., Havel, T.F. and Wagner, G. (1992) *Protein Sci.*, **1**, 736–751.

Jones, D.N.M., Searles, M.A., Shaw, G.L., Churchill, M.E.A., Ner, S.S., Keeler, J., Travers, A.A. and Neuhaus, D. (1994) *Structure*, **2**, 609–627.

Kalk, A. and Berendsen, H.J.C. (1976) *J. Magn. Reson.*, **24**, 343–366.

Kohda, D. and Inagaki, F. (1992) *5th International Conference on Magnetic Resonance in Biological Systems*, Jerusalem, Israel, Abstr. P201.

Koning, T.M.G., Boelens, R. and Kaptein, R. (1990) *J. Magn. Reson.*, **90**, 111–123.

Krezel, A.M., Darba, P., Robertson, A.D., Fejzo, J., Macura, S. and Markley, J.L. (1994) *J. Mol. Biol.*, **242**, 203–214.

Levy, R.M., Bassolino, D.A. and Kitchen, D.B. (1989) *Biochemistry*, **28**, 9361–9372.

Macura, S. and Ernst, R.R. (1980) *Mol. Phys.*, **41**, 95–117.

Neuhaus, D. and Williamson, M.P. (1989) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH, New York, NY, U.S.A.

Nilges, M., Gronenborn, A.M., Brünger, A.T. and Clore, G.M. (1988) *Protein Eng.*, **2**, 27–38.

Nilges, M., Kuszewski, J. and Brunger, A.T. (1991) In *Computational Aspects of the Study of Biological Macromolecules by NMR* (Eds, Hoch, J.C., Poulsen, F.M. and Redfield, C.), Plenum Press, New York, NY, U.S.A., pp. 451–455.

Nilges, M. (1993) *Proteins Struct. Funct. Genet.*, **17**, 297–309.

Nilges, M. (1995) *J. Mol. Biol.*, **245**, 645–660.

Schwabe, J.W.R., Chapman, L., Finch, J.T., Rhodes, D. and Neuhaus, D. (1993) *Structure*, **1**, 187–204.

Tropp, J. (1980) *J. Chem. Phys.*, **72**, 6035–6043.

Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Gō, N. and Wüthrich, K. (1987) *J. Mol. Biol.*, **196**, 611–639.

Weber, P.L., Morrison, R. and Hare, D. (1988) *J. Mol. Biol.*, **204**, 483–487.

Widmer, H., Widmer, A. and Braun, W. (1993) *J. Biomol. NMR*, **3**, 307–324.

Wüthrich, K., Billeter, M. and Braun, W. (1983) *J. Mol. Biol.*, **169**, 949–961.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY, U.S.A.

Yip, P.F. (1990) *J. Magn. Reson.*, **90**, 382–383.

Yip, P.F. and Case, D.A. (1991) In *Computational Aspects of the Study of Biological Macromolecules by NMR* (Eds, Hoch, J.C., Poulsen, F.M. and Redfield, C.), Plenum Press, New York, NY, U.S.A., pp. 317–330.