Journal of Biomolecular NMR, 3 (1993) 675–700
ESCOM

J-Bio NMR 143

Statistical strategy for stereospecific hydrogen NMR assignments: Validation procedures for the floating prochirality method

Robert A. Beckman[a,*], Samuel Litwin[a] and A. Joshua Wand[a,b,**]

[a]*Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, U.S.A.*
[b]*Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.*

Received 26 April 1993
Accepted 9 June 1993

Keywords: Stereospecific assignments; Floating prochirality; Structure refinement; Distance geometry; Simulated annealing

SUMMARY

We examine the statistical and other considerations which determine the validity and reproducibility of stereospecific hydrogen NMR assignments obtained by the floating prochirality method. In this method, the assignment of a prochiral configuration of hydrogens at selected centers is allowed to 'float' during the structure refinement, and the distribution of prochiral orientations in highly refined structures is subjected to statistical analysis. The underlying statistical basis for this approach is examined and potential limitations of current approaches are identified. As an example, approximately 1300 distance constraints obtained from NOESY spectra of oxidized horse cytochrome c have been used to examine several computational strategies. Repeated calculations were done by several different methods on both the whole molecule (104 residues plus heme) and on a 23-residue fragment containing two helices, a turn, and flanking residues. The results show that, even with NOE constraints alone, one third of the centers may be reproducibly assigned, provided appropriate precautions are taken. These precautions include adjustments for multiple statistical comparisons and characterization of statistical interactions between prochiral centers. The analysis demonstrates that inadequately constrained systems, such as fragments from a larger molecule, may produce misleading results, raising concerns about methods which rely solely on intraresidue and sequential interresidue constraints. A mathematical model describing interactions among prochiral centers is described and validated, and protocols for assignment and statistical validation are presented.

INTRODUCTION

The precision of solution structures of proteins determined by NMR can be greatly enhanced by obtaining stereospecific assignments for prochiral beta methylene hydrogens and prochiral

*Present address: Biophysics Research Division, University of Michigan, Ann Arbor, MI 48109-1055, U.S.A.
**To whom correspondence should be addressed at: Dept. of Biochemistry, 600 S. Mathew, University of Illinois, Urbana, IL 61801, U.S.A.

methyl substituents of leucine and valine residues. This is particularly true for helical and loop regions of proteins, where the addition of stereospecific assignments can halve all-atom r.m.s.d and reduce main chain r.m.s.d by a factor of four (Güntert et al., 1989). The improvement in precision due to stereospecific assignments is a consequence of the elimination of the need to broaden NMR distance constraints with a 'pseudoatom correction', which replaces the two diastereotopic substituents with one pseudoatom at their mean position (Wüthrich et al., 1983). Recent improvements in the pseudoatom concept (Güntert et al., 1991) only partially ameliorate the loss of information when stereospecific assignments are not available.

Many approaches have been proposed and used for determining stereospecific hydrogen assignments from NMR data. While direct experimental determination of prochirality by use of isotopic labeling has been used (Neri et al., 1989; Sattler et al., 1992), by far the most popular approach has been the search for preference or bias in a given set of distance and angular constraints for a particular set of prochiral assignments. These approaches can be classified on the basis of the scope of the NMR data analyzed, the mode and extent of conformational sampling, the final criteria for a preferred stereospecific assignment, the settings under which the method has been tested, and the level of precision of NMR constraints required for the method to work.

Several methods, including HABAS (Güntert et al., 1989), STEREOSEARCH (both its systematic data base and crystallographic data base components) (Nilges et al., 1990), and approaches involving qualitative or quantitative back calculation (Zuiderweg et al., 1985; Hyberts et al., 1987) consider only intraresidue or sequential interresidue NMR data. While these restrictions allow the rapid evaluation of many conformations, they also ignore medium- and long-range constraints which may affect the distribution of feasible conformations. Medium-range periodic constraints are important in defining helical secondary structure, and long-range constraints will limit possible conformations in both helices and loops.

In contrast, GLOMSA (Güntert et al., 1991) and the floating prochirality method (Weber et al., 1988) both evaluate results of full structure refinements. These methods therefore utilize the full complement of available NMR data, and the resultant population of conformers may thus be a more accurate reflection of this data. However, it is more difficult to sample conformational space in a complete and unbiased manner when the full molecule is considered.

The method of conformational sampling and extent of sampling also varies among the different methods. HABAS, STEREOSEARCH, and the back-calculation approaches typically step through a grid search of a limited number of conformational variables. The sampling is unbiased, and typically 10^4 to 10^5 conformers are sampled. Nonideal bond lengths and angles are not considered. The limited dimensionality of the conformational space allows for finer sampling of this space with this number of conformers.

Another method of conformational searching is used with GLOMSA. This program works in concert with the program DIANA, which utilizes a variable target function algorithm (Braun and Gō, 1985). This first considers sequentially local constraints only, eventually adding increasingly sequentially remote constraints. The sampling is not guaranteed to be uniform, and there is the possibility that local constraints will have undue weight in the overall process. Since the algorithm operates in torsion angle space, nonideal geometries are not considered. Typically, 10^1 to 10^2 final conformations are available for analysis; however, since each refinement may reject many conformations on the basis of the target function before producing a final conformation, one may

estimate that perhaps 10^5 to 10^6 conformations receive some consideration. Since the conformational space available to the whole molecule has a very high dimensionality, this still corresponds to a relatively sparse sampling compared to the purely local methods.

The floating prochirality method (Weber et al., 1988) has typically been used in conjunction with a metric matrix distance geometry algorithm (Crippen, 1978; Crippen and Havel, 1978; Havel et al., 1983; Havel, 1991) followed by simulated annealing, as in the program Dspace (Nerdal et al., 1988). This approach considers both sequentially local and remote constraints in every phase of the calculation. The number of conformers sampled initially, and then available for final analysis, is comparable to that available through the variable target function approach, although the distance geometry/simulated annealing protocols may be more computationally intensive. Nonideal geometries are automatically considered as the program operates in Cartesian coordinate space. Again, the high dimensionality of the whole molecule conformational space precludes fine and uniform sampling.

Finally, STEREOSEARCH also samples conformers from a crystallographic database. Typically, another 10^3 to 10^4 conformers are available (10^1 to 10^2 crystal structures with 10^2 centers evaluable per structure). This enables the program to consider nonideal geometries, but the sampling for this component will be biased towards previously observed geometries, perhaps precluding the discovery of new geometries.

With regard to the criteria which must be met before a preferred assignment is declared, the methods also vary. HABAS and the quantitative back-calculation approach (Hyberts et al., 1987) both demand that the preferred assignment be the *only* one which satisfies the constraints. This has the advantage of being very conservative, and may even overcome the problem of irrelevant conformers which are considered when medium- and long-range constraints are ignored, since it is hard to argue with an assignment if no single acceptable conformer can be found for the alternative assignment. However, this method critically depends on the completeness and precision of the data set, if the alternative assignment is really to be completely excluded (Güntert et al., 1989). A very limited number of stereospecific assignments will be available, and these will be the ones which are less helpful for further refinement if the alternative was indeed already inaccessible. Finally, since in practice it is difficult to generate conformers which are in agreement with all constraints, HABAS provides an option by which the user can accept conformers which have small violations. Once this option is invoked for practical reasons, the advantages of rigor and conservatism are lost, and one is again dealing with a *statistically based* assignment criterion.

STEREOSEARCH requires that the preferred assignment be highly predominant, i.e., a 10:1 preference over at least 20 conformers. This cutoff is relatively conservative, but arbitrary. Such an arbitrary cutoff will either allow too many wrong assignments or be too insensitive in detecting assignments (depending on the situation), when compared to a criterion which is more firmly grounded in statistical theory.

GLOMSA analyzes relative distances between the two diastereotopic substituents and a third atom in the ensemble of structures generated by DIANA. The sign of the average relative distance and the sign of the relative upper distance bounds must agree for the input assignment to be judged correct. Moreover, the relative distances in individual structures must also agree with the relative upper distance bounds for a majority of structures. The degree to which these quantities must agree is left to the user's discretion, again resulting in an assignment criterion that lacks a firm statistical basis.

The floating prochirality method has been used in conjunction with an assignment criterion which, for the first time, recognized that binomial statistics are applicable to the analysis of the problem of distinguishing between two discrete alternatives. The original work (Weber et al., 1988) was done with the binomial distribution, a one-tailed comparison, and $p \leq 0.05$ as the assignment cutoff. This approach has a firmer statistical basis than the others, and in principle allows for higher sensitivity in detecting stereospecific biases if a large number of conformers are evaluated. However, the original statistical treatment has several pitfalls which must be surmounted if it is to be truly valid and generally applicable (see below).

The various methods have been evaluated and tested in a variety of contexts. GLOMSA, STEREOSEARCH, and HABAS have been tested with both real and simulated data, and limited back calculation (Hyberts et al., 1987) and floating prochirality have been tested with real data only. The precision and completeness of the data used to evaluate these methods has, in our opinion, been high to ideal. All had extensive NOE data, and the simulations typically had very narrow NOE distance bounds. Most had extensive χ_1 and ϕ torsion angle constraints. In reality, side-chain coupling constants may not always be available for larger proteins. Main chain coupling constants are small in helical regions and may be unresolvable in larger proteins. Moreover, the relationship between scalar coupling constants and torsion angles is not one to one (Karplus, 1959). This is often surmounted by rejecting unlikely solutions, although in at least one published case such a conformation was observed (Billeter et al., 1990).

The tests of GLOMSA and the floating prochirality method to date have included hydrogen bond constraints in the refinements. These are sometimes justified; however, we feel that the justification is questionable in helices and loops. Although often used in this context, hydrogen exchange data cannot unambiguously identify an acceptor. Main chain coupling constants are nearly identical for 3_{10}- and α-helices despite the different hydrogen bonding patterns. Qualitative NOE patterns have been shown to be diagnostic of helices, but no pattern has been demonstrated to discriminate reliably between 3_{10}- and α-helices. Hydrogen bond constraints are very powerful in limiting the regions of allowable conformational space, and their use should be based only on very firm evidence.

HABAS, STEREOSEARCH, and limited back calculation make extensive use of 'relative distances', that is, it is stated or assumed that the relative magnitude of two inaccurately and imprecisely determined distances may somehow be determined with greater accuracy and precision. While this is perhaps true in selected extreme cases (Zuiderweg et al., 1985), a rigorous justification for the general case remains to be presented.

Here, we have used methods similar to those originally employed by Weber et al. (1988). In particular, we have used the floating prochirality method in conjunction with a distance geometry approach (Crippen, 1978; Crippen and Havel, 1978; Havel et al., 1983) coupled with a simulated annealing (Nerdal et al., 1988) refinement protocol, and have exploited binomial statistics to define assignment thresholds. However, there are several significant differences between our study and earlier ones. We have used a nonideal data set with no coupling constant constraints, hydrogen bond constraints, or relative distance constraints. We seek to determine whether stereospecific assignments can be reliably obtained without these ancillary constraints. If so, a significant improvement in structural information would be expected, since these assignments would represent new, not redundant, information. One might, for example, be able to identify distortions in helices, or to distinguish 3_{10}- from α-helices without imposing a preconceived hydrogen

bonding pattern early in the refinement, and without discarding valid solutions of the Karplus equation.

We also examine in great detail the statistical assumptions of the floating prochirality method, and define more rigorously the conditions required for reproducible, statistically valid stereospecific assignments. A number of statistical and procedural issues are examined, including: (1) the adequacy of a one-tailed statistical comparison for this application; (2) the possible need for adjustment of the assignment threshold for the fact that multiple statistical comparisons are being made; (3) whether or not stereospecific assignments can be obtained iteratively; (4) whether or not the commonly employed assumption of the absence of interaction between the states of different prochiral centers during the refinement is valid; (5) whether or not it is advisable to undertake local stereochemical determinations; and (6) whether or not bias is introduced by initial, arbitrary assignment of prochiral labels and, if present, whether or not it can be adequately compensated for.

MATERIALS AND METHODS

NMR spectroscopy

Spectroscopic methods and assignments will be described in more detail in a subsequent paper on the solution structure of oxidized cytochrome c (R.A. Beckman et al., unpublished results). Briefly, ^{1}H spectra of 5–10 mM oxidized horse cytochrome c (Sigma Type VI) in 50 mM sodium phosphate buffer, pH 5.75, were obtained with Bruker AM-600 and AM-300 NMR spectrometers at 20 °C. The bulk of NOE-based distance constraints were obtained from NOESY spectra (Macura and Ernst, 1980) recorded at 600 MHz with 30, 50, 75, and 110 ms mixing times. NOESY spectra with short mixing times (10 and 30 ms) at 300 MHz were used to obtain supplemental constraints involving hyperfine-shifted resonances. All data processing, assignment work, and NOE cross-peak identification and integration were done by using a modified version of the program FTNMR (Hare Research, Bothell, Washington, USA). NOESY cross peaks were identified on the basis of previously reported resonance assignments (Feng et al., 1989) and a small number of additional assignments (R.A. Beckman et al., unpublished results).

Constraint generation and structure refinement

NOEs were identified and integrated in FTNMR, as were samples of the baseline surrounding each NOE cross peak. Baseline correction of each NOESY cross-peak volume, using an average of the neighboring baseline samples, as well as subsequent statistical analysis of NOE buildup curves, was performed using an in-house program (A.J. Wand, unpublished results).

The distance estimates were obtained from the slope of the best-fit linear regression lines through the NOE buildup curves, after calibration based on distances within known helical segments which are known to be nearly identical in 3_{10}- and α-helices. Cross peaks used for calibration all had linear regression correlation coefficients of 0.95 or better, were free of overlap or assignment ambiguity, and were at least five residues from the end of a known helix, to minimize effects of conformational flexibility at the end of the helices. For the 300-MHz data, calibration of peaks near the heme used known helical distances supplemented by selected NOEs between geminal hydrogens and between methyl groups in the vicinity of the heme. The breadth of the distance constraints was determined by using the variances of the calibration constant

estimate and of the buildup curve slope estimate. A supplementary set of looser constraints was generated with the lower bound at the van der Waals contact limit (1.7 Å) and the upper bound at the NOE detection limit (4.5 Å), corresponding to NOEs which could not be reliably integrated due to overlap or for which the linear correlation coefficient for the fitted buildup curve was less than 0.7. Corrections were also applied for methyl groups, flipping aromatic rings, and spin diffusion in a limited subset of the constraints. All atom pairs with hydrogen bonding potential were allowed to approach each other as closely as half the van der Waals contact limit, and the heme iron was allowed to approach its ligand atoms up to 0.2 times the van der Waals limit. Constraints of 1 Å breadth were also applied between the heme iron and its two ligands. These constraints were centered about the distances determined by solution state EXAFS (Korszun et al., 1982, 1989). The final constraint set used here comprised approximately 13 NOE distance bounds per residue.

Structure calculation and refinement was done by using the programs Dspace 3.55 and 4.20 (Hare Research, Bothell, Washington, USA) with amino acid geometry templates modified to correspond to an average of ECEPP geometry (Momany et al., 1975; Nèmethy et al., 1983) and that utilized in the program AMBER (Singh et al., 1986). Structures were embedded from a metric matrix and were checked for correct handedness by using a signed vector product. This was followed by least-squares minimization in Cartesian space, using a conjugate gradients procedure. The penalty function minimized consisted of a sum of squares of violations of ideal geometry terms (bond lengths, bond angles, peptide bond planarity, improper dihedral angle terms, etc.) and of NOE distance constraints, all given equal weight. No improper chirality constraints were applied to prochiral centers which lacked stereospecific assignments for the two attached hydrogens. The labeling of prochiral pairs was allowed to 'float' during the refinement to the prochiral orientation providing the lowest penalty. Half of the embeds used one initial prochiral orientation ('forward embeds') and the other the reverse ('reverse embeds'). Extensive simulated annealing (Nerdal et al., 1988) was then applied in an attempt to further minimize the penalty function. The simulated annealing, using the sum of the squares of the violations as a pseudo-temperature variable, was generally carried out at penalty values two to three orders of magnitude above the final penalty achieved. Those embedded structures that could not be refined below a certain penalty threshold were rejected (first penalty selection). Accepted structures then underwent further simulated annealing subject to ideal geometry constraints according to the SHAKE algorithm (Ryckaert et al., 1977). From this population, only structures which could be refined below a second penalty threshold were accepted (second penalty selection). Generally, 10–20 embeds were required to generate one structure which had appropriate handedness and met the two penalty thresholds. Thus, the final population consisted only of a minority of the very best structures, as judged by the penalty function.

For the fragment studies described below, 1192 fully refined structures were analyzed in the 'aggressive' approach, and 1000 fully refined structures in the 'conservative' approach. For the whole molecule studies described below, 264 fully refined structures were independently generated, refined, and analyzed. In all cases, none of the final structures had violations of the input constraints greater than 0.25 Å.

The results of structure generation runs, in terms of the orientation of selected prochiral centers, were tabulated, and binomial statistics were calculated by using a family of Fortran 77 programs. These programs are available from R.A.B.

THEORETICAL CONSIDERATIONS

Prior to embarking on a statistical discussion of stereospecific NMR assignments, we wish to clarify the meaning of the statement $p \leq 0.05$. In particular, it does not mean that the stereospecific assignment suggested has a less than 5% chance of being incorrect. Rather, it means that the constraints are more consistent with, or prefer, the given assignment, and the probability that this preference is merely due to chance is less than 5%. Thus, the probability statement really refers to the likelihood that the statistical preference for one assignment manifested in the constraints has not been correctly identified. Confidence that a given stereospecific assignment is correct or incorrect must also depend on the validity, accuracy, and precision of the constraints, and in general cannot be specified.

The strength of the preference for the assignment given may vary at the same confidence level. By evaluating more structures, a weaker preference in the constraints can be detected at the $p \leq 0.05$ level of confidence. In general, the rejected assignment will also have some conformations which are consistent with the constraints. Thus, the statistically based procedure for stereospecific assignments is like any other part of the refinement, such as the simulated annealing, in that it makes decisions which result in an incremental optimization of the agreement between the structure and the constraints. The agreement between the structures and the constraints can be ensured at a high level of statistical confidence, but the agreement between the structures and the unknown true structure can never be completely guaranteed.

To avoid circumlocution, we will occasionally use the terms 'correct' and 'incorrect' to describe the stereospecific assignments and their validity, but it should be understood that the meanings of the terms 'correct' and 'incorrect' are qualified by the points raised above.

One-tailed versus two-tailed binomial probabilities

These terms refer to the 'tails' of the binomial distribution. For a given result, with n conformers of which the majority n_r are in the pro-*R* orientation, p (one-tailed) is the probability that a result at least this biased *in favor of the R orientation* would occur by chance with no inherent preference in the constraints. It is represented by the area of one tail of the distribution from n_r to n. P (two-tailed) is the probability that a result at least this lopsided *in either direction* would occur by chance with no inherent preference in the constraints. It is the sum of the areas of the two tails of the binomial probability distribution from n_r to n and from $(n - n_r)$ to zero. Since the binomial distribution, with $p = 1/2$, is symmetric, p (one-tailed) is equal to half of p (two-tailed).

Previous studies (Weber et al., 1988) have used one-tailed binomial probabilities in their analysis. In this study we analyze separate groups of refinements of the fragment by an 'aggressive' approach and by a 'conservative' approach, and compare the results. The 'aggressive' approach differs from the 'conservative' approach in a number of ways, one of which is that the former uses an assignment cutoff based on the one-tailed, rather than the two-tailed, binomial probability. A rigorous statistical approach demands the use of two-tailed binomial probabilities, as we cannot know in advance of the analysis which of the stereospecific orientations will be preferred.

The issue of multiple comparisons

While the probability of any one 'preference' being spurious (i.e., due to chance) may be held to less than 0.05, if multiple prochiral centers are examined, the chance that at least one of the

preferences is spurious increases. For example, if 20 'preferences' are determined with $p \leq 0.05$, we expect one of the 20 preferences to be spurious. Therefore, the more prochiral centers that are investigated, the more stringent the statistical criteria required to avoid spurious results. We assume that the goal is to obtain sufficient statistical stringency to make even a single spurious result unlikely, even if this means fewer centers assigned.

The expected number of spurious assignments for n stereospecific centers for which the *i*th center is assigned at significance level p_i is simply the sum of all the individual probabilities. The probability of exactly m spurious results (p_m) in n independent stereospecific assignments made with confidence level $p = 0.05$ is:

$$p_m = \frac{(0.05)^m(1-0.05)^{n-m}n!}{m!(n-m)!} \tag{1}$$

The probability of m or more spurious assignments out of n assignments made with confidence level $p = 0.05$ can be obtained by summing terms like Eq. 1 from m to n inclusive.

The total probability, p_t, of at least one spurious result, given n assignments, each assigned at the same individual confidence level p_i, is $p_t = 1-(1-p_i)^n$ since $(1-p_i)$ is the probability that an individual preference is real, $(1-p_i)^n$ is the probability that all n individual preferences are real, and therefore the expression given is the probability that one or more of the preferences is spurious.

The true statistical criterion should be based on reducing p_t below a certain value (we choose 0.05 arbitrarily, in agreement with statistical convention), rather than on reducing p_i below a certain value. This is because the true goal is to make it very unlikely that there are *any* wrong stereospecific assignments entering into the structure refinement. It is not adequate to simply ensure that *most* of the stereospecific assignments are correct. This is often overlooked in studies applying these kinds of arguments.

We obtain the following expression, which gives the required confidence level p_i for each of the individual attempted assignments, if, for n attempted assignments, we are to keep the total probability of *at least* one assignment occurring by chance at p_t or less.

$$p_i = 1-(1-p_t)^{\frac{1}{n}} \tag{2}$$

Equation 2 can be approximated by a Taylor series about $p_t = 0$, obtaining:

$$p_i = \frac{p_t}{n} + \frac{n-1}{2n^2}p_t^2 + \cdots \tag{3}$$

Since we choose conditions where p_t is small, the Taylor series may be truncated at the linear term, obtaining the approximate criterion:

$$p_i \cong \frac{p_t}{n} \tag{4}$$

so that the required individual confidence level is the required total confidence level divided by the number of statistical comparisons. This approximation is often termed the Bonferroni upper bound on error, or Bonferroni criterion. We note that the requirement placed on the confidence level for each individual assignment becomes much more stringent for increasing numbers of

assignments. We further note that the stringency required for an individual assignment now depends on how many assignments are sought. This raises the strategic issue of how many assignments to seek, since seeking many low-priority assignments may unduly increase the difficulty of meeting the statistical criteria for the higher priority assignments.

The derivation of the Bonferroni criterion assumes that each statistical comparison is independent of the next. This is not strictly true in the case where the orientations of stereospecific centers interact. This will be illustrated below. Nevertheless, the Bonferroni criterion effectively addresses the issue of multiple comparisons in all cases. In the case of interactions, the various statistical comparisons are not independent, so that the effective number of total comparisons is less. Thus, the Bonferroni criterion may be too conservative in these instances. But since statistical theory does not offer a way to determine the exact level of stringency required in the face of interacting comparisons, we are forced to take this conservative option. The Bonferroni criterion does not, however, address the issue of interactions per se; this issue is far less straightforward and requires a separate treatment (see below).

Validity of iterative stereospecific assignments

One can envision refinement protocols in which, after obtaining a set of stereospecific assignments at a given level of statistical confidence, these assignments could be used to tighten the distance constraints on which they depend, and further structures generated with the augmented constraints. Structures resulting from use of the augmented constraints could be used separately or could be pooled with the structures generated by refinement against previous constraint sets in an attempt to assign more centers. If the centers were independent, the assigned centers would not affect the statistical properties of the others, and would only speed the refinement. On the other hand, if the centers interact, then the runs with some new centers assigned would really represent a whole new series of statistical comparisons. The increased number of statistical comparisons would increase the overall stringency required by the Bonferroni criterion. Moreover, pooling data obtained with and without an initial set of assignments would be valid only if the remaining floating centers oriented independently of the assigned set.

Model for interaction among prochiral centers

In this section, we develop two related mathematical models to quantify the interactions between prochiral centers. The models serve two purposes. First, they provide a means to evaluate the mutual consistency of a set of individual stereospecific assignments. Thus, if prochiral centers do interact, the orientation which appeared to be favored for center i in the context of the total available constraint set may no longer be favored in the context of other centers assigned to their preferred orientations. In such a case, the assignment for center i is inconsistent with other assignments. Clearly, to be acceptable, a putative assignment must be sufficiently robust to be preferred both within the total constraint set and in the context provided by other prochiral assignments. This will become an important issue and is elaborated on below. The second purpose served by the two models developed below is to provide a means to evaluate the importance of higher order interactions (dependencies) above those that can be represented by pairs and triplets of interacting prochiral centers.

Here the individual probabilities for a center i, P(i) and P(i'), are defined as the probability that the orientation of a given conformer corresponds to the preferred orientation ('correct') or with

its reverse ('incorrect'), respectively. These probabilities are allowed to depend on the context provided by the orientation of other centers j, k, l, etc. The orientational probabilities, P, should not be confused with the statistical *p* values. The center i will have pairwise interactions with all other centers j, k, and l, which refers to the direct effect of the orientation of j, k, and l on the probability distribution of the orientation of i. These pairwise interactions are characterized by pairwise conditional probabilities. For example, P(i|j) is the probability that the orientation of center i is correct given j is correct; P(i|j') is the probability that the orientation of center i is correct given j incorrect; P(i'|j) is the probability that i is incorrect given j correct; and P(i'|j') is the probability that i is incorrect given j is incorrect.

The center i may also have triplet interactions whereby the centers j and k influence center i indirectly. In this triplet, center k may have two types of indirect interactions with center i: (1) center k can have a pairwise interaction with center j, which in turn has a pairwise interaction with center i. This mechanism for center k affecting center i is not a true triplet interaction, but rather a chain of pairwise interactions. (2) Center k may actually affect the pairwise conditional probability connecting centers i and j. Thus, for example, the value of P(i|j) might depend on the orientation of center k. To describe this effect, a triplet conditional probability is introduced. For example, P(i|j|k) = P(i|j,k) = P(i|k,j) denotes the probability that center i is correct given that j and k are correct. Similarly, P(i|j'|k) = P(i|j',k) = P(i|k,j') denotes the probability that center i is correct given that center j is incorrect and center k correct.

The models can generate exact expressions for conditional probabilities for any arbitrary number of centers. These expressions contain terms of increasingly higher-order conditional probabilities. The probability P(i) of center i being correct can be expressed in terms of the state of center j as follows:

$$P(i) = P(i|j)P(j) + P(i|j')P(j') \tag{5}$$

i.e., a conditional probability expansion. Note that when j is assigned correct, P(j) = 1, and P(j') = 0 which, upon substitution into Eq. 5 yields,

$$P(i\{j\}) = P(i|j) \tag{6}$$

where P(i{j}) is the value of P(i) when j is assigned correct. Subtracting Eq. 5 from Eq. 6 gives [{ΔP(i)}|j], the change in the probability that i will be correct due to the assignment of j as correct:

$$[\{\Delta P(i)\}|j] = P(j')[P(i|j) - P(i|j')] \tag{7}$$

where in deriving Eq. 7 we have also used the identity 1 − P(j) = P(j'). To maintain compact expressions, we introduce a new notation:

$$D(i,j) = P(i|j) - P(i|j') \tag{8}$$

and

$$D(i',j) = P(i'|j) - P(i'|j') \tag{9}$$

noting also that D(i,j) = −D(i',j), since P(i|j) = 1 − P(i'|j) and P(i|j') = 1 − P(i'|j'). Thus, in this notation,

$$[\{\Delta P(i)\}|j] = P(j')D(i,j) \tag{10}$$

or

$$P(i|j) = P(i) + P(j')D(i,j) \tag{11}$$

Equations 10 and 11 summarize a primary pairwise interaction.

Let us now consider a triplet interaction. Equation 11 can be made conditional on center k and center k assigned correct such that P(k), P(k|j), etc., equal one and P(k'), P(k'|j) etc. equal zero, giving:

$$P(i|j,k) = P(i|k) + P(j'|k)D(i,j,k) \tag{12}$$

where D(i,j,k) = P(i|j,k) − P(i|j',k). This expression is not obviously symmetric in j and k as its derivation corresponds to separating the simultaneous assignment of j and k into two sequential steps with a specified order. The reverse order yields:

$$P(i|k) = P(i) + P(k')D(i,k) \tag{13}$$

leading, upon assignment of j, to:

$$P(i|k,j) = P(i|j) + P(k'|j)D(i,k,j) \tag{14}$$

Now P(i|k,j) = P(i|j,k), and the expressions on the right hand side of Eqs. 12 and 14 can also be shown to be equivalent. Thus one can obtain Eq. 15 which is symmetrical in j and k:

$$P(i|j,k) = P(i|k,j) = \frac{P(i|k) + P(i|j)}{2} + \frac{P(j'|k)D(i,j,k) + P(k'|j)D(i,k,j)}{2} \tag{15}$$

and leads directly to Eq. 16 which involves P(i) and 'correction terms' of increasingly higher order:

$$P(i|j,k) = P(i) + \frac{P(j')D(i,j) + P(k')D(i,k)}{2} + \frac{P(j'|k)D(i,j,k) + P(k'|j)D(i,k,j)}{2} \tag{16}$$

Similar arguments lead, after averaging the six possible orders of assigning centers j, k, and l, to an exact expression for the quartet conditional probability P(i|j,k,l):

$$P(i|j,k,l) = P(i) + \frac{P(j')D(i,j) + P(k')D(i,k) + P(l')D(i,l)}{3} +$$

$$\left[\frac{P(j'|k)D(i,j,k) + P(k'|j)D(i,k,j) + P(j'|l)D(i,j,l)}{6}\right] + \left[\frac{P(k'|l)D(i,k,l) + P(l'|j)D(i,l,j) + P(l'|k)D(i,l,k)}{6}\right]$$

$$+ \left[\frac{P(j'|k,l)D(i,j,k,l) + P(k'|j,l)D(i,k,j,l) + P(j'|l,k)D(i,j,l,k)}{6}\right]$$

$$+ \left[\frac{P(k'|l,j)D(i,k,l,j) + P(l'|j,k)D(i,l,j,k) + P(l'|k,j)D(i,l,k,j)}{6}\right] \tag{17}$$

This probability was modeled using only lower-order terms with a third-order approximation to Eq. 17 in which D(i,j,k,l) is equated with D(i,j,k) and so on for all fourth-order conditional probabilities. Such a treatment will exactly model the first- and second-order terms for quartets, will approximate the third-order terms, and neglects fourth-order terms. As will be shown below, this treatment is a consistently good estimator of the behavior of quartets.

Following this line of approach, the behavior of larger systems can be approximated by considering the average behavior of all the quartets comprising the system. Thus P(i|j,k,l,m,n...) would be approximated as the value of P(i|j,k,l) averaged over all possible combinations j, k, and l. Using a third-order approximation and averaging over all j, k, and l gives:

$$P(i|j,k,l,m,...) \cong P(i) + \frac{\sum_{j\neq i} P(j')D(i,j)}{N-1} + \frac{\sum_{k\neq i,j}\sum_{j\neq l} P(j'|k)D(i,j,k)}{(N-1)(N-2)} + \frac{\sum_{l\neq i,j,k}\sum_{k\neq i,j}\sum_{j\neq i} P(j'|k,l)D(i,j,k)}{(N-1)(N-2)(N-3)} \quad (18)$$

where N is the number of tentatively assigned prochiral centers in the system, as distinguished from n, the number of prochiral assignments originally sought. In what follows, we do not attempt to model the effect of the (n − N) centers sought but not assigned, nor the effect of other prochiral centers for which no assignment was sought. We are primarily interested in the mutual consistency or inconsistency of the proposed assignments.

We term this approximation of the larger system the third-order average quartet model. This model will prove useful not only for approximating the behavior of higher-order systems, but also for developing an insight into the range of effects of different contexts as it also yields a variance of the population of quartets involving center i.

Another approach to the approximation of larger systems is to write down the exact expression by using a series of conditional probability expansions as outlined above, and then to approximate all fourth- and higher-order terms by their third-order counterparts as also described above. It can be shown that this leads to:

$$P(i|j,k,l,m,...) \cong P(i) + \frac{\sum_{j\neq i} P(j')D(i,j)}{N-1} + \frac{\sum_{k\neq i,j}\sum_{j\neq l} P(j'|k)D(i,j,k)}{(N-1)(N-2)} + \frac{\sum_{l\neq i,j,k}\sum_{k\neq i,j}\sum_{j\neq i} P(j'|k,l)D(i,j,k)}{(N-1)(N-2)} \quad (19)$$

which is identical to Eq. 18 except for the denominator of the third-order term. This treatment we shall refer to as the third-order complete model.

While exact expressions may be written for any arbitrary number of centers, it becomes increasingly difficult to estimate accurately higher-order conditional probabilities from a finite population of conformers. Moreover, the number of terms and types of terms to be estimated increases dramatically with the number of centers considered. Therefore, we will use the models described above as approximations to the total system. These approximations are supported in many cases by the excellent agreement of Eqs. 18 and 19 with the exact conditional probabilities determined directly from the data.

RESULTS

The computational complexity of structure generation and refinement of molecular models of

proteins on the basis of NMR-derived structural constraints increases in a nonlinear manner with the number of atoms in the system. Hence, it would be of obvious utility to carry out preliminary, local refinements to determine prochiral assignments. Indeed, as discussed above, this approach is taken by many algorithms currently in use. However, it is not at all clear that this is statistically sound in general or appropriate for the floating chirality method in particular. In order to assess the appropriateness of applying the floating chirality method in a local, piecemeal fashion, refinements were performed on the whole oxidized cytochrome c molecule and an isolated fragment of the protein.

Fragment refinements

Refinements were performed using a fragment of the oxidized cytochrome c molecule comprising residues 57–79. This region includes two helices, a connecting turn, and flanking residues. The refinements used covalent and NOE constraints exclusively involving these residues. The refinements were undertaken using two different protocols in an effort to evaluate the adequacy of a one-tailed statistical comparison, the need for adjustment for multiple comparisons, and to determine whether iteration of the procedure to obtain more assignments is valid. One protocol, the 'aggressive approach', is similar to what is commonly in use, and takes the less rigorous option in all of these cases. The statistical cutoff is set at $p \leq 0.05$ for a *one-tailed* binomial distribution, without adjustment for multiple comparisons. When a stereospecific assignment is indicated by this criterion, the assignment is made, the bounds matrix is appropriately modified, and new structural refinements proceed, in an attempt to iteratively obtain more assignments. The results of all structure refinement runs are continuously pooled.

Statistical theory suggests that the one-tailed binomial cutoff and the lack of adjustment for multiple comparisons may be misleading. The error distribution expected due solely to the neglect of these points can be calculated by using Eq. 1. However, since the 'aggressive' approach is in common usage, and since it is in principle more efficient than the 'conservative' approach outlined below, we wished to determine whether these issues really affect the validity of the result in a practical case.

The second protocol used, termed the 'conservative' approach, involves the analysis of a family of structures without iteration, consideration of a predefined and limited number of prochiral centers, application of the Bonferroni criterion, and a binomial probability of $p \leq 0.05$ for a *two-tailed* comparison.

The application of the 'aggressive' approach involved 1192 refined structures and yielded stereospecific assignments for 12 prochiral β- or α-centers, two thirds of those queried. Since multiple comparisons were not considered, there was no penalty for also querying centers further down the side chains. This resulted in nine additional stereospecific assignments, for a total of 21 of the available 26 prochiral centers, or 80.8%. The number of conformers required to obtain a stereospecific assignment ranged from 24 to 1192, corresponding to a total of 12 cycles of iteration. The orientational probabilities for the preferred orientations ranged from 0.53 to 0.75. The corresponding one-tailed binomial p values ranged from 0.0076 to 0.0488. In no case was the p value low enough to meet the Bonferroni criterion for either one-tailed or two-tailed comparisons. An appealing feature of the 'aggressive' approach is its rapid convergence; nearly two thirds of assignable centers met the criteria within 700 structures. Application of the 'conservative' approach to the analysis of fragment 57–79 involved 1000 independently refined structures.

Stereospecific assignments were sought for the 12 prochiral centers assigned by the 'aggressive' approach. Assignments for other prochiral centers were not sought. This establishes the Bonferroni criterion as $p = 0.00416$, for a two-tailed comparison. Of the 12 prochiral centers sought, seven were assigned by this method. The orientational probabilities for the seven centers are listed in Table 1, and range from 0.550 to 0.729.

Lack of agreement between the aggressive and conservative approaches

Comparison of the results of the two prochiral assignment protocols reveals that only three of the seven stereospecific assignments generated by the conservative approach agreed with the assignments generated by the aggressive approach. The probability that four or more spurious results would occur in a sample of seven centers assigned by the aggressive approach can be computed using Eq. 1 to be 0.27% ($p = 0.0027$). Furthermore, the expected number of errors in seven centers should have been on the order of one. Thus, the lack of validity of the aggressive approach is too dramatic to be explained merely by the lack of adjustment for multiple comparisons or the failure to use a two-tailed test. Thus an additional problem is indicated.

Interaction model for fragment refinements

In an effort to determine the origin of these discrepancies, pairwise interactions among the 12 prochiral β-centers sought in the conservative approach were evaluated in the ensemble of 1000 structures using Fisher's exact test for 2×2 contingency tables (Matthews and Farewell, 1988). This is a test of the independence of centers i and j, and a statistically significant failure of this test was taken as evidence for interaction between the centers. Of the 66 possible pairwise interactions, 41 were significant at the 0.05 level, far more than the approximately three expected by chance. Of these 41 interactions, 17 met the more stringent Bonferonni significance criterion ($p \leq 0.00075$). These latter 17 interactions are shown in Fig. 1.

The orientational behavior of the seven assigned individual centers, pairs of centers, and triplets of centers was used in attempts to estimate the behavior of quartets. For each group of

TABLE 1
OBSERVED AND MODELED PROCHIRAL ORIENTATIONAL PROBABILITIES FOR THE CONSERVATIVE REFINEMENT OF RESIDUES 57–79[a]

Center	Individual probability[a]	Quartet conditional probabilities[b]	
		Modeled[c]	Observed
cβ[60]	0.563	0.550 (0.071)	0.558 (0.088)
cβ[62]	0.590	0.614 (0.065)	0.619 (0.079)
cβ[66]	0.729	0.732 (0.049)	0.725 (0.065)
cβ[67]	0.554	0.519 (0.076)	0.518 (0.100)
cβ[69]	0.550	0.562 (0.043)	0.565 (0.056)
cβ[72]	0.630	0.594 (0.079)	0.597 (0.100)
cβ[73]	0.627	0.645 (0.058)	0.632 (0.077)

[a] P(i) for the given center i determined directly from the data for the conservative refinement of this fragment.
[b] P(i|j,k,l) averaged over all j, k and l centers listed for the given center i. Standard deviations in parentheses.
[c] Modeled probabilities generated with the third-order complete model for quartets.

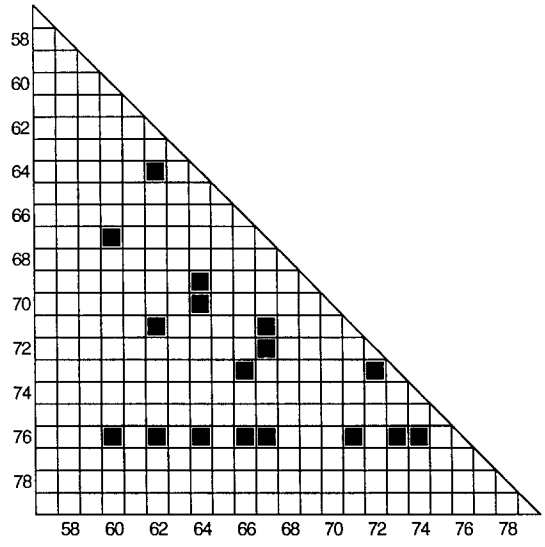

Fig. 1. Summary of strong interactions found among queried prochiral β-carbon centers in the refinement of residues 57–79 of oxidized horse cytochrome c by the conservative approach (see text). Pairwise interactions that are significant at the Bonferroni level $p \leq 0.00075$ are shown by filled squares. There are 17 such interactions. For the sake of clarity, the 24 pairwise interactions that were significant at $0.00075 < p \leq 0.05$ are omitted. The secondary structure in this region is: 57–60, loop/turn; 61–69, mixed 3_{10}/α-helix; 69–72, turn; 72–75, helix; 76–79, loop/turn.

four centers i, j, k, and l, all individual orientational probabilities and pairwise and triplet conditional probabilities were estimated from the family of 1000 structures refined without prochiral constraints. The third-order complete model was then used to estimate the single quartet conditional probability P(i|j,k,l) for all i, j, k, and l. These quartet conditional probabilities were also estimated from the observed distribution in the structure ensemble. As shown in Table 2, the correlation between the modeled and the observed quartet conditional probabilities is excellent, and represents a considerable improvement over the crude model which ignores interactions.

Each modeled quartet conditional probability was also individually compared to its corresponding observed value, and the discrepancies evaluated for significance using a two-tailed binomial statistic. These comparisons indicate that, while the correlation between modeled and observed quartetwise conditional probabilities is generally quite good, some of the small discrepancies observed may be due to real quartetwise interactions, rather than merely to chance. The presence of small but measurable higher order interactions may explain some of the difficulties encountered in extending the model to larger systems (see below).

The size of the interaction effects can be seen by comparing, for a given center i, the average value and standard deviation of P(i|j,k,l) over all j, k, and l (modeled and observed) with the values of P(i) (see Table 1). Such a comparison shows that the interaction effects are significant. For six of the centers, the 95% confidence interval for the quartet conditional probability overlaps 0.5, suggesting that the influence of context can easily alter the preferred orientation.

An attempt was made, using several models, to estimate the septet behavior of the seven centers assigned by the conservative approach, i.e., to estimate P(i|j,k,l,m,n,o) for all i. Using the 1000 conformer data set from the conservative approach refinements, estimates for these seventh-order

conditional probabilities were generated using the crude model, the third-order average quartet model, and the third-order complete model, and compared to the actual seventh-order conditional probabilities determined directly from the data. As seen in Table 3, all models failed to adequately estimate the behavior of septets in the fragment system. While third-order models were adequate to describe the behavior of quartets, their failure to describe the behavior of septets demonstrates the importance of higher-order interactions in the fragment system, and suggests that statistical analysis of this system to determine stereospecific assignments may be inadvisable.

Whole molecule refinement

A family of 264 structures was generated, 114 structures obtained using one arbitrary initial set of prochiral labels and 150 structures obtained with opposite initial prochiral labeling. These two sets are further referred to as 'forward' and 'reverse' runs. Stereospecific assignments were sought for hydrogens from 88 centers, using the 'conservative approach', representing all prochiral α- and β-methylene centers and all prochiral dimethyl centers. The Bonferroni assignment criterion is therefore $p \leq 0.00057$. Stereospecific assignments were obtained for 35 of the 88 centers sought in oxidized cytochrome c. This included five of the 12 glycine α carbons, two of six leucine γ carbons, and 28 of 70 β carbons. Separate analysis of forward and reverse runs shows that the slight imbalance does not affect the results.

Of six centers assigned by both the aggressive fragment approach and the whole molecule conservative approach, only three are in agreement. The probability of this degree of disagreement occurring by chance can be estimated by using Eq. 1 to be less than 1.5%. Of the seven centers assigned in the fragment studies by the conservative approach, three were also assigned in the whole molecule studies. All three agreed between the two approaches. Though not a stringent test, this result clearly indicates that the conservative approach is a significant improvement over the aggressive approach and indicates that additional problems other than those quantifiable by

TABLE 2
CORRELATIONS BETWEEN MODELED AND ACTUAL QUARTET CONDITIONAL PROBABILITIES[a]

System	Model[b]	Correlation coefficient[c]
Fragment (7 centers)[d]	Crude	0.570 (0.498)
	Complete	0.953 (0.947)
Whole molecule (7 centers)[e]	Crude	0.360 (0.341)
	Complete	0.994 (0.985)
Whole molecule (35 centers)[f]	Crude	0.962 (n.c.)
	Complete	0.975 (n.c.)

[a] The quartet conditional probabilities are P(i|j,k,l) for all i, j, k, l.

[b] Crude model neglects all interactions; complete model refers to the third-order complete model defined in the text.

[c] Standard correlation coefficient. Spearman's non-parametric, distribution-free rank correlation coefficient is listed in parentheses.

[d] Seven centers assigned by the conservative approach (see Table 1).

[e] Five centers assigned in residues 61–69 inclusive (cβ[61], cβ[64], cγ[64], cβ[66], cβ[67]) and two centers (cγ[35] and cα[41]) found to have pairwise interactions with 60s centers at the Bonferroni level.

[f] All assigned centers (see Table 4).

n.c. not calculated; the system is too large.

TABLE 3
CORRELATIONS BETWEEN MODELED AND ACTUAL SEPTET CONDITIONAL PROBABILITIES[a]

System	Model[b]	Correlation coefficient[c]
Fragment (7 centers)[d]	Crude	0.132 (−0.182)
	Average quartet	0.005 (−0.036)
	Complete	0.218 (−0.346)
Whole molecule[e] (7 centers)	Crude	0.223 (0.280)
	Average quartet	0.982 (0.944)
	Complete	0.970 (0.964)
Whole molecule[f] (seven 7-center sets)	Crude	0.975 (0.837)
	Average quartet	0.994 (0.948)
	Complete	0.992 (0.924)

[a] The septet conditional probabilities are P(i|j,k,l,m,n,o) for all i, j, k, l, m, n, o.

[b] Crude refers to the crude model that neglects all interactions; average quartet refers to the third-order average quartet model (Eq. 18, N = 7) and complete refers to the third-order complete model (Eq. 19, N = 7) defined in the text.

[c] Standard correlation coefficient. Spearman's non-parametric, distribution-free rank correlation coefficient is listed in parentheses.

[d] Seven centers assigned by the conservative approach (see Table 1).

[e] Five centers assigned in residues 61–69 inclusive (cβ[61], cβ[64], cγ[64], cβ[66], cβ[67]) and two centers (cγ[35] and cα[41]) found to have pairwise interactions with 60s centers at the Bonferroni level.

[f] Average value of the correlation coefficients for seven sets of seven centers in the whole molecule involving centers which are indicated to interact by Fisher's exact test ($p \leq 0.05$).

using Eq. 1 (use of a one-tailed test and the failure to adjust for multiple comparisons) affect the validity of the aggressive fragment approach.

Individual orientational probabilities

The individual orientational probabilities for all 35 centers assigned in the whole molecule study are given in Table 4. It can be seen that, when refining the molecule as a whole, the preferred orientations are favored with higher orientational probability than in the fragment. The observed orientational probabilities for assigned centers range from 0.610 to 0.985, with a mean value of 0.793, itself higher than any single orientational probability observed in the fragment studies.

Reproducibility

In order to determine whether an assignment determined in the whole molecule study was reproducible the family of refined structures was divided into two sets. The first data set contained 91 forward runs and 65 reverse runs (156 structures total); the second data set contained 23 forward and 85 reverse runs (108 structures total). Each data set was analyzed for stereospecific assignments as described for the total whole molecule study.

Twenty-nine centers were assignable by the first data set alone, and 29 by the second data set alone. In no case were opposite assignments obtained by the two data sets. Twenty-six of the 29 centers assigned by data set 1 were confirmed by data set 2; the other three centers all had the same favored orientation in data set 2, but had not yet achieved the Bonferroni level of statistical significance (they were significant at $p \leq 0.05$). Twenty-five of the 29 centers assigned by data set 2 were confirmed by data set 1; the others all had the same favored orientation in data set 1, but

had not yet achieved the Bonferroni level of statistical significance (two of these four were significant at $p \leq 0.05$). These data indicate that conclusions drawn by the whole molecule conservative approach are highly reproducible and, more importantly, are predictable on the basis of simple statistical projections.

Effect of initial orientation

To test the effect of the initial orientation on the final conclusions drawn about the prochiral centers, the results obtained from the two initial orientations were analyzed separately by the methods detailed above. Twenty-three centers were assigned by the 'forward' runs; of these 19 were confirmed by the 'reverse' runs, and the other four had the same favored orientation but had not achieved statistical significance at the Bonferroni level (two of these were significant at $p \leq 0.05$). Twenty-two centers were assigned by the 'reverse' runs; of these 19 were confirmed by the 'forward' runs, and two of the other three had the same favored orientation but had not achieved statistical significance at the Bonferroni level (these were significant at $p \leq 0.05$). The final center was assigned in the 'forward' orientation by the 'reverse' runs, but had only 45 of 91 structures in the 'forward' orientation in the 'forward' runs. Pooling all the data from both data sets, this center (cβ[82]) was not assigned at the Bonferroni level, but did have a tendency towards the 'forward' orientation at $p \leq 0.05$. These results indicate that the effect of the initial orientation

TABLE 4
OBSERVED AND MODELED PROCHIRAL ORIENTATIONAL PROBABILITIES FOR THE CONSERVATIVE WHOLE MOLECULE REFINEMENT

Center	Individual probability[a]	Quartet conditional probabilities[b]		Center	Individual probability[a]	Quartet conditional probabilities[b]	
		Modeled[c]	Observed			Modeled[c]	Observed
cα[1]	0.614	0.615 (0.033)	0.619 (0.032)	cβ[48]	0.833	0.835 (0.024)	0.840 (0.025)
cα[6]	0.648	0.642 (0.033)	0.645 (0.026)	cβ[52]	0.610	0.618 (0.037)	0.621 (0.034)
cβ[8]	0.693	0.693 (0.034)	0.698 (0.030)	cβ[53]	0.701	0.705 (0.035)	0.711 (0.037)
cβ[10]	0.977	0.987 (0.021)	0.992 (0.004)	cβ[59]	0.871	0.881 (0.026)	0.886 (0.024)
cβ[12]	0.962	0.970 (0.022)	0.976 (0.014)	cβ[60]	0.667	0.663 (0.033)	0.667 (0.029)
cβ[14]	0.610	0.632 (0.031)	0.636 (0.031)	cβ[61]	0.962	0.979 (0.022)	0.984 (0.010)
cβ[17]	0.651	0.652 (0.030)	0.657 (0.031)	cβ[64]	0.947	0.956 (0.024)	0.963 (0.016)
cβ[18]	0.670	0.704 (0.034)	0.708 (0.032)	cγ[64]	0.977	0.975 (0.024)	0.982 (0.009)
cβ[20]	0.644	0.638 (0.027)	0.644 (0.028)	cβ[66]	0.951	0.973 (0.025)	0.978 (0.015)
cα[24]	0.939	0.966 (0.026)	0.971 (0.019)	cβ[67]	0.970	0.976 (0.022)	0.983 (0.012)
cβ[31]	0.750	0.750 (0.031)	0.754 (0.024)	cβ[74]	0.727	0.721 (0.034)	0.550 (0.037)
cβ[32]	0.833	0.850 (0.024)	0.855 (0.022)	cβ[76]	0.610	0.623 (0.048)	0.626 (0.043)
cβ[35]	0.735	0.756 (0.034)	0.759 (0.032)	cβ[87]	0.670	0.687 (0.034)	0.691 (0.030)
cγ[35]	0.951	0.977 (0.023)	0.981 (0.013)	cβ[91]	0.674	0.676 (0.029)	0.681 (0.033)
cβ[36]	0.943	0.949 (0.026)	0.954 (0.016)	cβ[94]	0.977	0.983 (0.024)	0.988 (0.008)
cα[37]	0.689	0.694 (0.039)	0.697 (0.032)	cβ[97]	0.682	0.699 (0.034)	0.703 (0.030)
cα[41]	0.966	0.977 (0.023)	0.983 (0.012)	cβ[98]	0.682	0.705 (0.036)	0.709 (0.035)
cβ[46]	0.974	0.987 (0.022)	0.993 (0.003)				

[a] P(i) for the given center i.

[b] P(i|j,k,l) averaged over all j, k and l centers listed for the given center i determined directly from the data. Standard deviations in parentheses.

[c] Modeled probabilities generated using the third-order complete model for quartets (Eq. 19 with N = 4).

on the final conclusion drawn is either small or nonexistent indicating that the simulated annealing employed was sufficient to overcome any potential bias introduced by the initial arbitrary prochiral labeling employed.

Interaction model for whole molecule refinements

Pairwise interactions among the 88 prochiral centers considered in the whole oxidized cytochrome c molecule were evaluated within the ensemble of 264 whole molecule structures by Fisher's exact test for 2×2 contingency tables. Of the 3828 possible pairwise interactions, only 266 were significant at the $p \leq 0.05$ level, slightly more than the 191 expected by chance. This is far fewer than in the fragment study, where nearly two thirds of the interactions were significant at the $p \leq 0.05$ level. Only five of the 3828 potential pairwise interactions were significant at the Bonferroni level for 3828 comparisons. This is slightly more than the expected number (zero), but far less than in the fragment case, in which 17 of 66 interactions were significant at the Bonferroni level. The five significant interactions at the Bonferroni level were: cβ[35] and cγ[35], cγ[35] and cα[41], cγ[35] and cβ[66], cγ[35] and cβ[67], and cα[41] and cβ[66]. It is important to note that only one of these interactions is an intraresidue or nearest sequential neighbor interaction.

Comparing the 12 centers analyzed in the fragment studies in that context and in the whole molecule, we find that only four of the 41 significant ($p \leq 0.05$) interactions present in the fragment studies are also present in the whole molecule. Conversely, of six significant ($p \leq 0.05$) interactions among these 12 centers in the whole molecule, only four are present in the fragment studies. Thus, the greater prevalence of interactions in the fragment studies is not due to the particular centers chosen for analysis in these studies. Moreover, the average p value in the whole molecule studies for the 37 interactions absent in the whole molecule studies but present in the fragment studies is 0.61, with a standard deviation of 0.30. This suggests that the failure to observe these interactions in the whole molecule studies is not due to insufficient statistical power (i.e., too small a sampling).

Table 5 lists the number of observed significant ($p \leq 0.05$) interactions in the whole molecule studies, compared to the expected number on the basis of chance, as a function of various contingencies. The overall ratio of expected to observed interactions is 0.718. Centers separated by less than five residues in the linear sequence are not more likely to interact with each other (expected/observed ratio 0.692). Centers within helices separated by an integral number of α- or 3_{10}- turns from each other are only slightly more likely to interact. Interactions appear to occur more frequently among those centers which have strong orientational preferences as individuals. Thus, when both centers of interacting pairs are unassigned, the expected/observed ratio is 0.986, when one center of interacting centers is assigned, the expected/observed ratio is 0.930, and if both centers of interacting pairs are assigned, the expected/observed ratio is 0.313. Thus, it would appear that centers which are strongly constrained by the whole molecule into one orientation may have broader effects on the entire molecular conformation when their individual orientation is reversed. It should also be noted that the five centers participating in interactions that were significant at the Bonferroni level had an average individual orientational probability of 0.915, somewhat higher than the overall average for assigned centers (0.793).

Detailed interaction model for whole molecule refinements

The interaction model using pairwise and triplet conditional probabilities was used to model

the behavior of quartets among the 35 centers assigned in the whole molecule by the conservative approach. As can be seen in Table 2, the correlation between predicted and observed behavior of quartets is still excellent, in agreement with the fragment studies. However, in this case, we now see that the 'crude' model, which ignores interactions, assuming that P(i|j,k,l) = P(i), correlates nearly as well with the actual behavior of quartets. This result follows from the diminished significance of interactions in the whole molecule in general (see above). We have a very clear distinction between this case and that of the fragment studies, where consideration of pairwise and triplet interactions markedly improved the ability to predict the behavior of quartets.

Note that this analysis is capable of detecting regions within the whole molecule which still do have significant isolated interactions. For example, Fisher's exact test showed several centers in the 61–69 helix which had pairwise interactions which were significant at the Bonferroni level of confidence, especially with cγ[35] and cα[41]. When the five assigned centers in this helix were considered together with cγ[35] and cα[41] in the interaction model, the 'crude' model was again a poor predictor of quartetwise behavior, and the correlation was greatly improved using the third-order complete model (Table 2). The models were also used to estimate the seventh-order conditional probabilities governing the behavior of this group of seven strongly interacting centers, where the actual conditional probabilities could be directly determined from the data. As seen in Table 3, the crude model fails badly to estimate the behavior of this septet. However, both the third-order average quartet model and the third-order complete model estimate the behavior of the septet very well and give rise to similar estimates of P(i|j,k,l,m,n,o) (Table 3). These results strongly imply that higher-order interaction terms are of far less importance in the whole molecule system than they are in the fragment case. In this highly interacting group of centers, inclusion of second- and third-order terms is sufficient to model seventh-order behavior. This further suggests that it may be possible to model the entire system using a third-order approximation. Thus, even in the context of the whole molecule refinements, the interaction models are

TABLE 5
SUMMARY OF SIGNIFICANT PAIRWISE INTERACTIONS BETWEEN PROCHIRAL CENTERS OBSERVED DURING WHOLE MOLECULE REFINEMENTS[a]

Criterion	Number of interactions		Expected/observed
	Observed	Expected[b]	
Total	266	191	0.718
Bonferroni[c]	5	0	0.000
Both centers unassigned	70	69	0.986
One center assigned	100	93	0.930
Both centers assigned	96	30	0.313
Helical centers[d]	9	5	0.555
Close in sequence[e]	26	18	0.692

[a] All pairwise interactions reported are significant at $p \leq 0.05$ by Fisher's exact test.

[b] Expected on the basis of chance, the number of pairs multiplied by 0.05.

[c] Strong pairwise interactions which meet the Bonferroni significance criterion which, in this case, is $p \leq 0.00001$.

[d] Defined as separation by 3m + 4n residues in the primary sequence (m,n = 0,1,2; m + n = 1,2) within the helical regions of the molecule defined as residues 4–18, 61–69 and 88–104 inclusive.

[e] Separated by five or fewer residues in the primary sequence.

TABLE 6
OBSERVED AND MODELED PROCHIRAL ORIENTATIONAL PROBABILITIES FOR A REINFORCING SUBSET OF CENTERS DURING WHOLE MOLECULE REFINEMENT[a]

Center	Individual probability[b]	Quartet conditional probabilities[c]	
		Modeled[d]	Observed
cβ[35]	0.951	0.982 (0.004)	0.982 (0.004)
cα[41]	0.966	0.993 (0.005)	0.993 (0.005)
cβ[61]	0.962	0.990 (0.006)	0.990 (0.006)
cβ[64]	0.947	0.967 (0.004)	0.967 (0.004)
cγ[64]	0.977	0.979 (0.001)	0.979 (0.001)
cβ[66]	0.951	0.990 (0.006)	0.990 (0.006)
cβ[67]	0.970	0.993 (0.004)	0.993 (0.004)

[a] All assigned centers in the sequence spanning residues 61 through 69 inclusive and two centers which had pairwise interactions with one or more of these centers at the Bonferroni level of significance (in this case, $p \leq 0.00001$).
[b] P(i) for the given center i.
[c] P(i|j,k,l) averaged over all j, k and l centers listed for the given center i. Standard deviations in parentheses.
[d] Modeled probabilities generated using the third-order complete model for quartets (Eq. 19 with N = 4). For averages in this table, the modeled and observed probabilities are exactly identical. The correlation between individual modeled and observed quartet probabilities is 0.994 (Table 2).

useful for analyzing the behavior of interaction 'hot spots', revealing in this case a highly mutually consistent interaction among seven centers, two of which were distant from the others in the primary structure (Table 6).

The size of the interaction effects at the level of quartets, using either estimated or actual quartetwise conditional probabilities, is smaller than in the fragment case (Table 4). This, together with the stronger individual orientational preferences in the whole molecule case, means that the quartet conditional probabilities P(i|j,k,l) for each given i, averaged over all j, k, and l are, for the most part, more than two standard deviations above 0.5. In fact, there is only one case in which the 95% confidence interval for these quartet conditional probabilities, averaged over all contexts j, k, and l, overlaps 0.5 (Table 4). This center, cβ[74], is identified from the mean and variance of the actual observed quartet conditional probabilities. Because this assignment may be more dependent on the context provided by other assignments, it was deleted from the list of assigned centers, leaving a total of 34 centers assigned, or roughly 40% of those originally sought.

DISCUSSION

We have demonstrated that reliable, reproducible stereospecific assignments are achievable by the floating prochirality method, with appropriate modifications, even for nonideal NMR data sets lacking in hydrogen bond constraints, relative distance constraints, and coupling constant information. We have used a protein comprising loops and atypical helices, which are the most difficult to assign unless the constraint set is extraordinarily complete and precise (Güntert et al., 1989). The percentage of centers assigned is comparable to that assigned within helices by most other methods, with the exception of STEREOSEARCH, which, however, is highly dependent on relative distance constraints (Nilges et al., 1990), the use of which has never been rigorously justified.

The results are reproducible and reliable if, and only if, appropriate statistical precautions are taken and if the system is sufficiently well constrained by NMR and packing constraints to preclude excessive compensatory interactions between floating prochiral centers during the refinement. The resulting procedures will identify centers for which a given orientation of the prochiral center is preferred due to greater consistency with the input constraint set than the alternative orientation. The probability that there is even one assignment preference identified anywhere within the molecule that is spurious (i.e., due to chance) may be reduced to 5% or less. However, the ultimate validity of the assignments clearly must also depend on the validity of the input constraint set.

Established statistical theory suggests that the current procedure for analysis of floating prochirality data (Weber et al., 1988) must be modified to use a *two-tailed* binomial probability, and to adjust the assignment threshold for the effect of multiple statistical comparisons. The comparison between the results from the 'aggressive' and 'conservative' approaches for the 57–79 fragment further reinforces these points in that it confirms that the former approach does not produce valid assignments. The 'aggressive' approach is essentially modeled on current applications of the floating prochirality method. The very high error frequency of the 'aggressive' approach in this study further strongly suggests that iteration of the procedure to obtain more assignments is not valid, in that the total error rate observed was higher than would be expected even when the effects of using a one-tailed binomial probability and of failing to adjust for multiple comparisons are taken into account.

Previous studies have implicitly assumed that individual prochiral centers 'float' independently during refinement, but this work demonstrates the presence of extensive and important interactions between these floating centers. The analysis of interactions in the fragment studies revealed a surprising number of pairwise interactions between the prochiral centers. Accordingly, an explicit consideration of pairwise and triplet interactions dramatically improved the correlation between predicted and observed behavior of quartets of centers, when compared with a crude model which ignored interactions (Table 2). Many of these interacting pairs included individuals separated by a considerable distance in the primary sequence (Fig. 1). As judged by the mean and variance of the quartet conditional probabilities, most of the assignments, even those obtained by the 'conservative' fragment approach, appear to be highly dependent on the context provided by the orientation of the other assigned centers. Thus, it appears that a change in the orientation of one prochiral center may trigger conformational adjustments which will affect the status of other prochiral centers at a distance. This is a highly significant result as it indicates that *an error in one assignment can be compensated for by errors in subsequent assignments such that the refinement criteria are fully met.*

When interactions in the whole molecule are considered, we find that they are still present, but are greatly reduced both in numbers and in overall importance, except for a few 'hot spots' with strong interactions (which meet the Bonferroni criterion for significance), and for certain centers whose individual orientational preference is sufficiently weak to potentially be dependent on context. Thus, for the whole molecule system, consideration of pairwise and triplet interactions does not improve the correlation between actual and predicted behavior of quartets, when compared with a crude model ignoring interactions (Table 2).

The interactions observed in the whole molecule studies are not necessarily near each other in the primary sequence. However, interactions are more common among assigned centers, suggest-

ing that an orientation reversal at one center is more likely to have distant effects if the original orientation was already highly preferred by the input constraint set.

One must conclude that the decreased interactions in the whole molecule system are due to the fact that it is much more highly constrained. Hence, many conformations which represent an incorrect prochiral orientation at one center with compensation involving an incorrect orientation at another center are more probable within the isolated fragment. In the whole molecule, such conformers containing compensatory adjustments appear to be ruled out by the many additional packing constraints and NMR constraints provided.

We note that conformers containing an incorrectly oriented center and a compensatory change elsewhere may tend to confuse any conformational search method which seeks to evaluate a statistical distribution of conformers for obtaining stereospecific assignments. Hence, it is essential that the system be sufficiently well constrained so that these irrelevant conformers will not dominate. The methods described in this paper allow one to judge whether the system is sufficiently well constrained. In particular, for a well-constrained system, it appears that the consideration of pairwise and triplet interactions will not lead to a markedly better ability to predict the behavior of quartets than provided by a crude model ignoring interactions.

Though complete agreement was found, the comparison between the assignments obtained by the 'conservative' fragment approach and the whole molecule studies is not in itself sufficient to either establish or refute the validity of the fragment approach. Three assignments were obtained by both methods, and they all agree. Thus, although no discrepancy was found, the test of the fragment approach was not sufficiently demanding. However, the demonstration of many significant pairwise interactions which could confuse the fragment approach, and of a change in the strength and pattern of these interactions in the whole molecule, raises significant concerns about the fragment approach.

Therefore, our recommended procedure for obtaining stereospecific assignments by the floating prochirality method involves using all the available constraints from the whole system, rather than from a small fragment or from solely intraresidue and sequential interresidue constraints. The assignments must reflect a two-tailed binomial probability, with assignment threshold further adjusted for multiple comparisons by application of the Bonferroni criterion. Since the Bonferroni threshold becomes more difficult to achieve as more assignments are sought, one must set priorities rather than seeking assignments for every prochiral center within a large molecule. Iteration of the procedure to obtain additional assignments appears to be a flawed approach and is not recommended.

The system must then be further evaluated for interaction. This evaluation may lead one to conclude that the system was not sufficiently well constrained to offer reliable stereospecific assignments by a conformational searching technique. In particular, one ought to demand that the correlation between predicted and observed quartet conditional probabilities be nearly as good for a model ignoring interactions as for one considering pairwise and triplet interactions. This would then suggest that the effect of interactions is not predominant.

Even when the overall system is well constrained, some individual centers may still have their orientational preferences influenced to an unacceptable degree by the context provided by the assignment of other centers. In our study, we found one center for which the two-tailed 95% confidence interval for the given orientational probability, when considered in the averaged context of all assigned quartets containing that center (P(i|j,k,l) averaged over all j, k, and l for a

given i) overlapped 0.5, either for the predicted or actual quartetwise conditional probabilities. We believe that prudent practice must also involve deletion of these assignments, which may be too dependent on context. Both the actual and predicted quartetwise conditional probabilities should be used for this evaluation. The former represents an experimental determination, but is nonetheless based on a limited sample size, so that the experimental estimates of quartetwise conditional probabilities may be less accurate than those of pairwise and triplet conditional probabilities which are used to infer quartetwise conditional probabilities in the latter approach. Finally, overall estimates of conditional probabilities for each preferred orientation given, in the context of the other proposed assignments, may be obtained from the third-order complete model. In the whole molecule system, these contextual probabilities were all greater than 0.5. We believe that it is also prudent to require that all proposed assignments have contextual probabilities greater than 0.5 as values less than this indicate inconsistent assignments.

The use of third-order models and of average behavior of quartetwise conditional probabilities is an imperfect way to look at the total effect of context in the overall system; however, it is a significant improvement over simply ignoring interactions. In all eight whole molecule cases tested, these third-order approximations adequately modeled seventh-order systems (Table 3). We have not found a statistical method for rigorously characterizing the behavior of an entire large system without approximations.

It is gratifying to note that in a well-constrained system, the rigorous standards recommended do not compromise the overall ability of the method to obtain assignments, but rather that the percentage of assignments obtained is comparable to other methods. These rigorous standards provide the investigator with the important assurance that no spurious orientational preferences are included in the subsequent structure refinement. The probability that even a single such error should occur anywhere in the whole molecule is held to less than 5%. If care has been taken to ensure the validity of the input constraint set, the investigator can be further assured that it is unlikely that any of the assignments used in subsequent structure refinements are wrong.

Concerns have been raised herein about the use of conformational searching in an inadequately constrained conformational space that may be dominated by misleading orientational interactions. While our work has utilized the floating prochirality method, these concerns may also apply to methods like HABAS (Güntert et al., 1989) and STEREOSEARCH (Nilges et al., 1990) which utilize uniform conformational searching on a very local level. Clearly, this concern does not apply in those cases where no conformation can be found which is consistent with the rejected prochiral assignment after an exhaustive, systematic search, since in that instance the basis of the assignment is not statistical. Nevertheless, to be exempt from these statistical considerations the method ought not use variable tolerances for an accepted conformer.

The methods discussed herein provide for the correct identification of a statistical preference within a given data set for a given prochiral orientation. As discussed above, one can never completely guarantee that the statistically preferred orientation is the correct one. This caveat applies to other analytic methods of determining prochiral assignments as well, except in the case where an exhaustive conformational search yields only one prochiral orientation which has perfect agreement with the constraint set. In general, similar qualifications apply to any optimization of a penalty function for refinement wherein structures are rejected on the basis of small differences in the value of a somewhat arbitrary function.

In an attempt to examine some of the issues raised here, Havel (1991) has used simulated NMR

data sets for the protein bovine pancreatic trypsin inhibitor. Though the sources of errors in real data which may contribute to errors in stereospecific assignments are incompletely characterized and therefore cannot be adequately simulated, the use of simulated data sets does provide the advantage that one knows which answer is truly correct and thereby allows one to determine an absolute error rate.

Using a small structure set (five structures each for 10 different simulation conditions), Havel (1991) found an apparently worrisome error rate which led to the observation that the floating chirality method was unreliable. However, it should be pointed out that this study used a one-tailed binomial cutoff without adjustment for multiple comparisons. Thus, unanimity among five structures was taken as an adequate statistical criterion when in fact unanimity among 12 structures would have been required to satisfy a two-tailed Bonferroni criterion. Given the loose assignment criterion used, one calculates that 6.25% of the assignments should have been incorrect simply due to chance. This is in sharp contrast to the actual error rate of 8 in 830 tries (83 centers, 10 simulated data sets) or 0.96%. To have such a low error rate despite this loose assignment criterion, one calculates that the correct orientation must have been favored in an average of 65.6% of the conformers, a percentage similar to that observed here with real data. Therefore, there is no evidence that incorrect orientations can actually be statistically favored and this provides further support for the assumption that statistical preference for a given prochiral assignment faithfully corresponds to the correct prochiral assignment.

Finally, we note that assignment of hydrogen bonding on the basis of a statistical analysis of hydrogen bonding geometries observed in a family of independently refined structures will, in principle, suffer from the same problems outlined here for the assignment of prochirality. This is currently being investigated and the results will be reported elsewhere.

CONCLUSIONS

Using a set of constraints for oxidized horse cytochrome c, it has been shown that reproducible, valid stereospecific assignments are possible even without coupling constant information, hydrogen bond constraints, or relative distance constraints, provided appropriate statistical precautions are taken, including evaluation of the suitability of the proposed system for this type of study. A number of statistical and procedural issues have been examined and resolved. It has been shown (1) that the bias introduced by initial, arbitrary assignment of prochiral labels is both measurable and small; (2) that the one-tailed statistical comparison is inappropriate in the context of prochiral assignments; (3) that the assignment threshold needs to be adjusted to accommodate the fact that multiple statistical comparisons are being made, with the Bonferroni criterion shown to be an adequately conservative threshold; (4) that interactions between centers do occur, the error in prochiral assignments can be compensatory and, therefore, stereospecific assignments should not be obtained iteratively; and (5) that it is not advisable to undertake local or global stereochemical determinations without analysis of these interactions.

ACKNOWLEDGEMENTS

We thank Xiurong Qi for assistance with the whole molecule refinements. We are also indebted to Ms. Perry Watts for providing several subroutines employed in the software used for these

studies. We also thank Dr. Dirk Moore for stimulating discussions and for an attempt to model prochiral interactions with log-linear models. This work was supported by NIH Research Grants GM-35940 and DK-39806 to AJW, by NIH Grants CA-06927 and RR-05539, by an appropriation from the Commonwealth of Pennsylvania, and by a grant from the F. Ripple Foundation to the Fox Chase Cancer Center. RAB is the recipient of an NIH Physician Scientist Award (CA-01456).

REFERENCES

Billeter, M., Qian, Y., Otting, G., Muller, M., Gehring, W.J. and Wüthrich, K. (1990) *J. Mol. Biol.*, **214**, 183–197.
Braun, W. and Gō, N. (1985) *J. Mol. Biol.*, **186**, 611–626.
Crippen, G.M. (1978) *J. Comp. Phys.*, **26**, 449–452.
Crippen, G.M. and Havel, T.F. (1978) *Acta Crystallogr.*, **A34**, 282–284.
Feng, Y., Roder, H., Englander, S.W., Wand, A.J. and Di Stefano, D.L. (1989) *Biochemistry*, **28**, 195–203.
Güntert, P., Braun, W., Billeter, M. and Wüthrich, K. (1989) *J. Am. Chem. Soc.*, **111**, 3997–4004.
Güntert, P., Braun, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.
Havel, T.F. (1991) *Prog. Biophys. Molec. Biol.*, **56**, 43–78.
Havel, T.F., Kuntz, I.D. and Crippen, G.M. (1983) *Bull. Math. Biol.*, **45**, 665–720.
Hyberts, S.G., Märki, W. and Wagner, G. (1987) *Eur. J. Biochem.*, **164**, 625–635.
Karplus, M. (1959) *J. Chem. Phys.*, **30**, 11–15.
Korszun, Z.R., Bunker, G., Khalid, S., Scheidt, W.R., Cusanovich, M.A. and Meyer, T.E. (1989) *Biochemistry*, **28**, 1513–1517.
Korszun, Z.R., Moffat, K., Frank, K. and Cusanovich, M.A. (1982) *Biochemistry*, **21**, 2253–2258.
Macura, S. and Ernst, R.R. (1980) *Mol. Phys.*, **41**, 95–117.
Mathews, D.E. and Farewell, V.T. (1988) *Using and Understanding Medical Statistics*, 2nd edition, S. Karger AG, Basel, pp. 20–38.
Momany, F.A., McGuire, R.F., Burgess, A.W. and Scheraga, H.A. (1975) *J. Phys. Chem.*, **79**, 2361–2381.
Nèmethy, G., Pottle, M.S. and Scheraga, H.A. (1983) *J. Phys. Chem.*, **87**, 1883–1887.
Nerdal, W., Hare, D.R. and Reid, B.R. (1988) *J. Mol. Biol.*, **201**, 717–721.
Neri, D., Szperski, T., Otting, G., Senn, H. and Wüthrich, K. (1989) *Biochemistry*, **28**, 7510–7516.
Nilges, M., Clore, G.M. and Gronenborn, A.M. (1990) *Biopolymers*, **29**, 813–822.
Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C. (1977) *J. Comp. Phys.*, **23**, 327–341.
Sattler, M., Schwalbe, H. and Griesinger, C. (1992) *J. Am. Chem. Soc.*, **114**, 1126–1127.
Singh, U.C., Weiner, P.K., Caldwell, J.W. and Kollman, P.A. (1986) Amber 3.0 University of California, San Fransisco.
Weber, P.L., Morrison, R. and Hare, D. (1988) *J. Mol. Biol.*, **204**, 483–487.
Wüthrich, K., Billeter, M. and Braun, W. (1983) *J. Mol. Biol.*, **169**, 949–961.
Zuiderweg, E.R.P., Boelens, R. and Kaptein, R. (1985) *Biopolymers*, **24**, 601–611.