MOLECULAR PHYSICS, 1998, VOL. 95, NO. 6, 1099–1112

NMR structure calculation methods for large proteins Application of torsion angle dynamics and distance geometry/ simulated annealing to the 269-residue protein serine protease PB92

YASMIN KARIMI-NEJAD[1], GREGORY L. WARREN[2], DICK SCHIPPER[3], AXEL T. BRÜNGER[2] and ROLF BOELENS[1]

[1] Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[2] The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA,

[3] Gist Brocades B.V., Research and Development, PO Box 1, 2600 MA Delft, The Netherlands

(*Received 31 March 1998; accepted 10 August 1998*)

Molecular dynamics in torsion angle space together with the well established combination of metric matrix distance geometry and simulated annealing in Cartesian space have been applied to the solution structure determination of serine protease PB92, a 269-residue monomeric protein from *Bacillus alcalophilus*. The input data set comprised distance restraints and a combination of distance and angular restraints derived from NMR data. A number of different modifications of the two calculation strategies were studied with respect to their convergence behaviour. The resulting structural ensembles were evaluated according to several criteria for protein structure quality. Improved protocols for both methods have been developed from these analyses. A comparison of the structures obtained with these protocols demonstrates the superior convergence behaviour of the torsion angle dynamics method. Although both methods are able to fold up the protein correctly, the torsion angle dynamics protocol scores slightly better with respect to measures of structural quality, especially when dihedral angle restraints are available. Torsion angle molecular dynamics presents a simple and robust method for the NMR structure calculations of large proteins.

1. Introduction

Nuclear magnetic resonance (NMR) structural investigations of proteins of increasingly large size have become feasible due to the recent progress in NMR methodology [1] (see Appendix for abbreviations used here). Advanced measurement techniques in combination with sophisticated labelling strategies have made possible studies of systems as large as 64 kDa, as the recently reported resonance assignments for a $^{2}H/^{13}C/^{15}N$ labelled Trp repressor–operator complex demonstrate [2]. However, currently full structure determinations for systems larger than 25 kDa are far from being routine. Until recently, all examples in this range were either oligomers, for which the spectra are simplified due to symmetry, or molecular complexes, the components of which can be analysed separately by isotope-filtering techniques. Although there are several examples of large protein complex structures solved by NMR [3–5], to our knowledge only two NMR structure determinations of monomeric systems larger than 25 kDa molecular weight have been reported thus far: that of the N-terminal domain of enzyme I from *E. coli*, comprising 259 residues [6], and that of the 269-residue protein serine protease PB92 from *Bacillus alcalophilus* from our laboratory [7]. The main reason for the lack of other examples is certainly the sheer complexity of the NOE spectra, which allows for the collection of a sufficiently large number of NOEs only in conjunction with preliminary structures guiding their assignment. This procedure is often carried out in an iterative manner, leading to the so-called 'bootstrap' approach [8, 9]. A drawback of this strategy is that it might artificially induce a self-consistency of the restraint set, enforced by the data collection procedure, eventually leading to an equally artificial precision of the resulting protein structure. Obviously, any progress in the computational area that improves the convergence of structure calculations is highly desirable, especially for this type of NMR structure calculation, not only because the requirement for fewer iterative steps might speed up the procedure, but even more so, because it will eventually minimize the potential bias associated with this approach.

0026–8976/98 $12·00

Molecular dynamics in torsion angle space, a recently introduced new method for protein structure determination [10], has been shown to improve the convergence of NMR structure calculations of proteins and DNA [11]. It was applied to NMR data sets of small to medium size systems, and it was concluded that the method has better sampling properties plus a high success rate compared with conventional simulated annealing in Cartesian space [11]. These advantages should become even more significant with increasing protein size. Therefore we applied the torsion angle dynamics (TAD) protocol to the NMR structure determination of serine protease PB92, the solution structure of which has been determined by us in an eight-step iterative procedure [7] using a combination of metric matrix distance geometry and simuled annealing in Cartesian space known as the DG/SA hybrid method [12–14]. During these iterative calculations 3372 NOE distance restraints were determined. However, in the study presented here, we have used a significantly smaller set of restraints, namely only those that could be derived unambiguously from the spectrum itself, i.e., without a structure-aided assignment of NOE intensities. For comparison, we applied the DG/SA protocol to the same input data set. Both protocols have been modified in several parameters in order to adapt them to the size of the protein serine protease PB92. This system, composed of 3721 atoms, is described by 11 163 degrees of freedom in Cartesian and 1099 degrees of freedom in torsion angle space, respectively. Various measures of structure quality together with an assesment of the sampling and convergence properties of the two methods have been applied to test the modified protocols. For both calculation methods, protocols optimized for performance on large molecules will be presented, and an analysis will be given of the results obtained from data sets comprising distance and also both distance and angular restraints.

2. Materials and methods

2.1. *Sample preparation*

Serine protease PB92 was isolated and purified from *Bacillus alcalophilus* strain PB92 as described previously [15, 16]. Samples for the NMR measurements typically contained 1–2 mM protein in a 25 mM deuterated acetate buffer at pH 5.0 in H_2O/D_2O (95:5).

2.2. *NMR data collection*

NMR spectra were recorded at 42 °C on samples of ^{15}N- or $^{13}C/^{15}N$-labelled protein, inhibited with DFP. A fractionally ^{13}C-labelled sample was used to obtain stereospecific assignments of valine and leucine methyl groups [17] from (^{13}C,^{1}H)-HSQC and/or CT-(^{13}C,^{1}H)-HSQC experiments [18, 19] as described previously [7]. A series of 6 CT-(^{15}N,^{1}H)-HMQC-J spectra [20] with values of 35, 50, 70, 90, 110 and 128 ms for the constant time period was recorded for the determination of $^{3}J_{HNH\alpha}$ values. These data sets were acquired on a Bruker AMXT-600 spectrometer. For details on the acquisition of NOE data, see [7].

2.3. *Generation of conformational restraints*

Stereospecific assignments of isopropyl methyl groups were obtained either from the sign of the methyl cross peaks in CT-(^{13}C,^{1}H)-HSQC spectra or from an inspection of their ω_1 multiplet patterns in (^{13}C,^{1}H)-HSQC spectra.

NOE distance restraints were obtained from three-dimensional (3D) ^{15}N- and ^{13}C-separated NOESY-HSQC and HMQC-NOESY-HSQC spectra by classification of unambiguously assigned NOE crosspeaks, i.e., NOE intensities assigned without the aid of additional structural information. Various upper bound categories of 2.5, 3.0, 3.5, 4.0, 4.5 and 5.0 Å were used, according to the intensities corresponding to short interproton distances observed in elements of regular secondary structure or to covalently fixed distances. All lower bounds were set to the sum of the van der Waals radii, i.e., 1.8 Å. Multiplicative correction factors to the upper bounds were set according to Constantine *et al.* [21]. Pseudo-atom corrections of 0.5 Å were added for methyl groups [22], the corresponding correction for degenerate aromatic protons was set to 2.0 Å [23]. The treatment of distance bounds involving non-stereoassigned diastereotopic protons was accomplished with the algorithm of Güntert *et al.* [24] implemented in the DIANA program, such that the necessary pseudo-atom corrections were generated with a minimum of information loss. DIANA was also used to filter the resulting distance restraint list for redundant, i.e., non-restrictive, restraints.

Distance restraints corresponding to hydrogen bonds were set according to their standard values [25] wherever the acceptors for the slowly exchanging amide protons could be identified from the NOE data [7]. Distances between the proton and its acceptor and between the donor nitrogen and the acceptor oxygen atom were constrained.

$^{3}J_{HNH\alpha}$ coupling were determined by fitting a non-linear model to the decay of experimentally obtained peak intensities in the CT-(^{15}N,^{1}H)-HMQC–J [20, 26]. Note that in these decay curves, the $^{3}J_{HNH\alpha}$ couplings appear as modulation frequencies which deviate from the true J value due to the finite lifetime of the Hα spin states [27]. The apparent reduction in the $^{3}J_{HNH\alpha}$ values caused by this effect increases with the rotational correlation time of the protein, and therefore was taken into account by including the effect of the Hα T_1 relaxa-

tion into the nonlinear model as described by Bax and coworkers [20].

Values < 5 Hz or > 8 Hz were converted into dihedral restraints for the backbone angle ϕ, using allowed intervals of $-80^\circ < \phi - 40^\circ$, if ${}^3J_{HNH\alpha} < 5$ Hz, and $-160^\circ < \phi - 100^\circ$, if ${}^3J_{HNH\alpha} > 8$ Hz [28].

2.4. *Structure calculations*

All DG/SA calculations were performed with the program X-PLOR [29] version 3.8 (available on the Internet, URL http://atb.csb.yale.edu), running on Silicon Graphics R10000 machines. The program for the TAD calculations will be made available in the future (requests should be sent to A. T. Brünger). Force field parameters for both the SA refinement and the TAD calculations were taken from the parallhdg.pro set, where bond angles in proline residues are corrected according to the values of Engh and Huber [30].

2.5. *Conformer selection*

All conformers generated during a calculation run were ranked according to their NOE energy (or their overall energy in case of the Optimized + Phi runs) before computing a matrix containing their pairwise backbone and heavy atom coordinate RMSD values with MOLMOL [31]. A cluster analysis of this RMSD matrix was then performed [32], computing the maximal pairwise coordinate RMSD within an ensemble of successively increased size, i.e., starting from the first two conformers with the lowest energies, then proceeding to the first three lowest energy conformers, and so forth. In order to obtain comparable structural statistics, the first cluster of conformers with the same ensemble size in both the TAD and DGSA runs was then selected for further analysis. Finally, the selected conformers were screened for their average rms deviations from ideal covalent geometry. In all cases, these deviations did not exceed 0.01 Å for bond lengths and 1° for bond angles or improper dihedral angles.

2.6. *Analysis of structures*

Restraint violations were evaluated with X-PLOR [29], MOLMOL [31] and AQUA [33]. The program suite PROCHECK-NMR [33] was used for secondary structure analysis, the computation of Ramachandran maps, and the analysis of deviations from Cα chirality. Backbone angular order parameters and ensemble-averaged ${}^3J$ values were computed from the structure ensembles with home-written software. 3D profile scores were calculated with the program Profile 3D [34]. For each structure ensemble of the protein, the total number of residues involved in consensus secondary structure elements was compared with the corresponding number determined for the crystal structure of PB92 [35], counting only those residues for which the $C\alpha$ B factor is $\leqslant 15$ Å^2 and which participate in secondary structure elements structurally highly conserved among subtilisins [36].

Table 1. Experimental restraints for structure calculations of serine protease PB92.

Total number of NMR restraints	1581
NMR restraints per residue	5.9
Total number of NOE distance restraints[a]	1374
intraresidual $[i = j]$	239
sequential $[\lvert i - j\rvert = 1]$	550
medium range $[1 < \lvert i - j\rvert \leqslant 4]$	98
long range $[\lvert i - j\rvert > 4]$	487
Hydrogen bond restraints[b]	130
ϕ torsion angle restraints[c]	77

[a] Only non-redundant NOEs are counted.
[b] Representing 65 hydrogen bonds.
[c] Used only in the runs denoted Optimized + Phi.

3. Results and discussion

3.1. *Experimental restraints*

Three different types of conformational restraints derived from NMR data have been used during the structure calculations: 1374 non-trivial NOE distance restraints (classified into categories of upper distance limits), 139 H-bond restraints (corresponding to 65 hydrogen bonds), and 77 ϕ angle restraints (derived from 97 ${}^3J_{HNH\alpha}$ values). The latter ones were used only in the last set of calculations (see section 3.7). Stereospecific assignments were made for 42 out of the 44 diastereotopic isopropyl groups in valine and leucine residues. Details of the conversion of NMR data into conformational restraints are given in section 2. A summary of the input data is presented in table 1.

3.2. *Initial structure calculations of serine protease PB92*

The initial torsion angle molecular dynamics (TAD) calculations of serine protease PB92 were carried out with the original TAD NMR structure calculation protocol [11], which has been optimized for small to medium size systems. This TAD protocol comprises four stages (see table 2), the first of which is a high temperature torsion angle dynamics period, followed by slow cooling in torsion angle space, with fixed covalent geometry during both steps. A second molecular dynamics cooling period in Cartesian space allows then for relaxation of bond lengths and angles. Finally, the structures are energy minimized. 50 conformers were generated and refined during the course of the TAD protocol. Details of the computations are given in table 2, with the parameters for these initial calculations listed in brackets. With the exception of the weight on

Table 2. Torsion angle molecular dynamics protocol[a]

	Stage 1 High temperature TAD	Stage 2 Slow cooling TAD	Stage 3 Slow cooling Cartesian MD	Stage 4 1000 steps EM
Temperature/K	60 000 (50 000)	60 000 (50 000) → 1000	1000 → 300	
Duration/ps	30 (15)	30 (15)	6	
Time steps Δt/ps	0.015	0.015	0.003	
w_{NOE}	150	150	150	150
w_{dihedral}[b]	100	100	100	300
w_{vdW}	0.1	0.1 → 1.0	1.0	1.0

[a] Listed are the temperature for each stage of the protocol, the duration of the stage and the weights for the energy terms during the structure calculation. Arrows indicate a linear change over the course of a particular period. A conjugate gradient minimizer was used for the EM. Parameters for the unoptimized Initial run [11] are given in parentheses whenever they differ from the final Optimized set. During all stages, a square-well potential with soft asymptotic behaviour [12, 13, 38] was used for the NOE energy term:

$$E_{\mathrm{NOE}} = \min\begin{cases} \Delta^2 & \text{for } d_{\mathrm{upper}} + r_{\mathrm{sw}} > R, \\ a + b/\Delta + \Delta & \text{for } d_{\mathrm{upper}} + r_{\mathrm{sw}} < R, \end{cases}$$

with $r_{\mathrm{sw}} = 0.5$ Å and Δ defined as

$$\Delta = \begin{cases} (R - d_{\mathrm{upper}}) & \text{for } d_{\mathrm{upper}} < R, \\ 0 & \text{for } d_{\mathrm{lower}} < R < d_{\mathrm{upper}}, \\ (d_{\mathrm{lower}} - R) & \text{for } R < d_{\mathrm{lower}}, \end{cases}$$

where R is the distance between a particular pair of protons in the model, d_{lower} is the lower bound for the distance, d_{upper} is the upper bound and a and b are determined such that E_{NOE} is a differentiable function at the point $R = d_{\mathrm{upper}} + r_{\mathrm{sw}}$. The van der Waals interactions were evaluated according to the quartic repulsive potential given by

$$E_{\mathrm{vdW}} = ((k_{\mathrm{repel}} R_{\mathrm{vdW}})^2 - R^2)^2$$

where k_{repel} was set to 0.8 and R is the distance between two atoms.

[b] Dihedral angle restraints were used only in the Optimized + Phi runs.

the NOE energy term during the final energy minimization, all of these values are identical to the previously published ones [11].

The same input data were used to generate and refine 50 conformers of PB92 with the DG/SA hybrid method [12–14], consisting of three steps (distance geometry submatrix embedding, SA regularization and SA refinement) with several stages each. The protocol was run as implemented in X-PLOR [29], i.e., with the scripts `dg_sub_embed.inp`, `dgsa.inp`, and `refine.inp`. Slight modifications of the recommended parameters [29] were necessary to account for the size of the system: the timestep for the integrator during the SA calculations had to be reduced, and the number of EM steps preceding the final refinement stage was increased. Table 3 lists the parameters used in the DG/SA calculations.

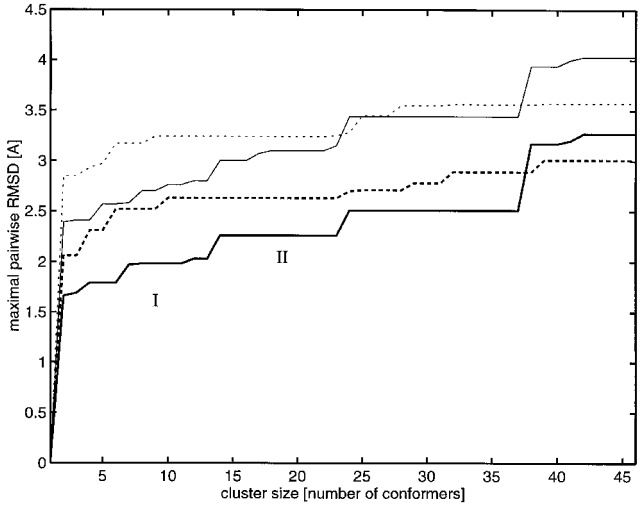
Note that only distance restraints were used at this stage, to avoid any bias in favour of the TAD method caused by the availability of dihedral angle restraints.

The two sets of conformers generated by these initial structure calculation runs were ranked according to their NOE violation energy and then subjected to a cluster analysis [32] in order to select a final ensemble of structures (for details of the conformer selection procedure see section 2). For both runs, these ensembles correspond to the second cluster, comprising 23 conformers each (figure 1(*a*)). They will be referred to as DG/SA Initial and TAD Initial in the following. A closer inspection of the cluster profile for the two initial runs, shown in figure 1(*a*), reveals some interesting differences in the sampling as well as in the convergence properties of the two structure calculation methods. The TAD clusters of PB92 exhibit lower $\mathrm{RMSD}_{\mathrm{max}}$ values for the backbone as well as for all heavy atoms over almost the whole range of generated conformers, indicating that the convergence of the TAD calculations is significantly better compared with the DG/SA calculation run. Furthermore, the backbone cluster profile of the TAD ensemble displays more pronounced steps, thus reflecting the sampling of clearly distinct conformational substates [32]. By contrast, the corresponding profile for the DG/SA ensemble exhibits a relatively monotonic increase towards higher energies without significant step formation. This behaviour is indicative of sampling a single conformational family only.

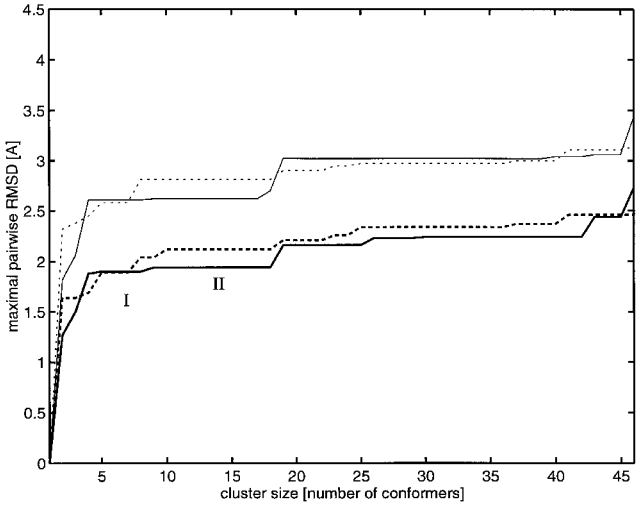
3.3. *Analysis of the initial DG/SA and TAD ensembles*

Table 4 summarizes the results of the structure determination of PB92 carried out with the two initial calcu-

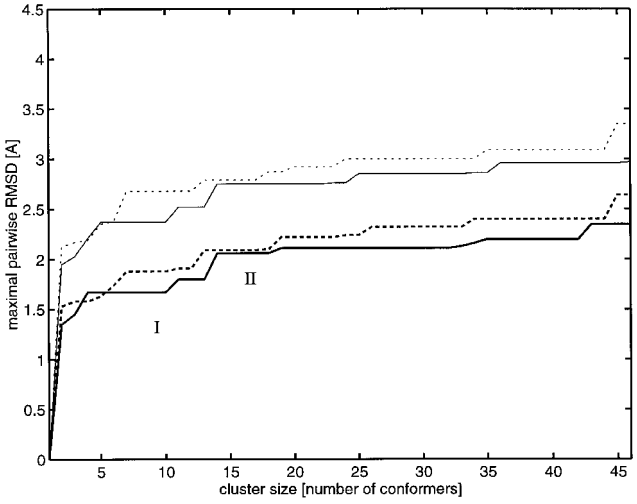
lation protocols described above. Some of the criteria applied for the analysis of the structure ensembles are related directly to the target function minimized during the computations: the agreement with the chemical (i.e. bond lengths, angles, impropers and van der Waals energy) and the experimental restraints as reflected by the energies, the restraint violations and the deviations from idealized covalent geometries. Similarly, the precision of the structures as monitored by the coordinate RMSD is related to the convergence of the calculations. A second category of analysis criteria includes measures of structure quality independent of the calculations: a Ramachandran analysis of the backbone conformations of the PB92 structures, the compatibility of a structural model with the preferred environment of a given residue



(*a*)



(*b*)



(*c*)

Figure 1. Cluster analyses of the structure calculations of serine protease PB92. The RMSD profiles of the TAD ensembles are drawn in solid thick (backbone coordinate RMSD) and thin (all-heavy atom coordinate RMSD) lines, respectively, whereas the corresponding profiles for the DG/SA calculations are rendered in dashed thick (backbone RMSD) and dashed thin (all-heavy atom RMSD) lines, respectively. Backbone coordinate RMSD values were computed by superpositions of N, Cα and C$'$ atoms. The clusters relevant to the analysis are indicated with Roman numbers. (*a*) RMSD profiles of the initial calculations. Cluster II of each Initial run comprises 23 conformers. (*b*) RMSD profiles of the structure calculations performed with the optimized protocols. In both optimized calculational runs, cluster II is constituted by 18 conformers. (*c*) RMSD profiles of the structure determination runs performed with the optimized protocols upon inclusion of constraints for the backbone angle ϕ. Cluster II consists of 18 conformers for both the TAD and the DG/SA run.

in terms of buried surface area, neighbour polarity, and local secondary structure [34], and the agreement of consensus secondary structure elements with the previously reported X-ray structure of the protein [35]. Note that this latter criterion may appear somewhat biased, since it is based on the assumption that the secondary structure of PB92 in solution and in the crystal is identical. On the other hand, the comparison includes only those segments of PB92 which belong to the core secondary structure elements common to all 3D structures of subtilisins known so far [36], thus also excluding the C-terminal α-helix which appears to be absent in solution [7].

From an inspection of table 4, it is evident that the convergence behaviour of the TAD method, reflected in the conformational energy terms, is significantly better than that of the DG/SA protocol. In addition, the coordinate RMSD values of the TAD ensemble are lower

Table 3. Hybrid distance geometry/simulated annealing protocol[a]

	SA refinement	
	Slow cooling Cartesian MD	200 steps EM
Temperature/K	2000 → 100	
Duration/ps	30 (6)	
Timestep Δt /ps	0.003	
w_{NOE}	150	150
w_{dihedral}[b]	100	100
w_{vdW}	0.002 → 1.0	1.0
k_{repel}	0.9 → 0.8	0.8

[a] Listed are only the parameters for the SA refinement stage of the three-step hybrid protocol. The DG substructure embedding and the SA regularization steps were carried out with the standard parameters given in the corresponding X-PLOR scripts, with the following exceptions. The timesteps were reduced to 0.002 ps during the high temperature stage and to 0.003 ps during the cooling stage of the SA regularization, and the final value of k_{repel} for the repulsive potential describing the vdW interactions during this stage was decreased to 0.8 only instead of 0.75. Furthermore, the number of EM steps before the SA refinement stage was increased to 500.

The parameter values listed in parentheses refer to the non-optimized Initial run. Arrows indicate a linear change over the course of a particular period. A conjugate gradient minimizer was used for the EM. The 'soft' potential used for the NOE energy term and the quartic repulsive potential describing the van der Waals interactions were identical to the ones used in the TAD calculations (see table 2, footnote [a]), with the exception of the values for k_{repel} given above.

[b] Dihedral angle restraints were used only in the Optimized + Phi runs.

than the corresponding values of the DG/SA conformers. The average deviations from idealized covalent geometries within the two ensembles are comparable. The stereochemical quality of the two ensembles, judging from an analysis of their Ramachandran maps, is slightly better for the TAD ensemble, which exhibits a smaller percentage of residues in disallowed regions of Ramachandran space. Moreover, the TAD calculations are able to reproduce the known secondary structure elements of PB92 in the crystal structure to a higher degree than the DG/SA method, thus implying a higher degree of accuracy. The 3D profile scores for the two structure families are comparable, indicating that on average, their agreement with the preferred environment of a given residue type is similar.

However, if one considers the restraint violations, the DG/SA protocol produces fewer violations. Both the average number of NOE violations and the value of the maximal NOE violation are lower in the DG/SA ensemble. From these analyses, the need for an optimization of both protocols is evident. We therefore tested a number of variations in both calculational protocols, applying the same criteria for the assessment of the two algorithms as for the initial calculation runs.

3.4. *Optimization of structure calculation protocols*

In order to improve the two calculation methods, numerous parameters were varied in both protocols. All changes were assessed again by generating and refining 50 conformers of serine protease PB92. The resulting ensembles were then subjected to the same conformer selection procedures and analyses as the Initial ensembles (data not shown). The protocol variations tested were aimed at improving the performance of both methods. Only a brief overview of all tested protocol modifications will be given in this section, while the optimized protocols and their results will be analysed in more detail below.

As discussed above,the DG/SA protocol suffers primarily from a lack of convergence. Our attempts to improve the method therefore included a prolongation of the final SA refinement step plus the addition of an extra period of restrained isothermal molecular dynamics (MD) in Cartesian space at 300 K. Accordingly, the first modification of the DG/SA protocol we tested consists of the addition of 6 ps of restrained MD in Cartesian space, performed after the SA refinement stage of the DG/SA protocol, and followed by 1000 steps of restrained EM. This additional simulation period, run at 300 K, turned out to be beneficial in various respects: The average conformational energy terms became significantly lower, and in particular the van der Waals term decreased considerably. Furthermore, the average number of NOE violations decreased, the Ramachandran behaviour improved, and better agreement with the known consensus secondary structure elements of PB92 was achieved. A more dramatic improvement of the performance of the DG/SA method, however, was accomplished simply by prolonging the final SA refinement stage to 30 ps, instead of running a final restrained MD simulation. It produced the best DG/SA structure of PB92 in terms of the various above mentioned criteria applied as measures of structural quality, as will be discussed in more detail below (see section 3.6).

The optimization strategy applied to the TAD method was twofold with respect to its aim: several variations addressed the sample properties of the algorithm, while on the other hand attempts were made to improve the convergence of the method, especially with respect to the residual violations of experimental restraints. The first category of variations included several different procedures which produce randomized extended strands as input for the subsequent TAD calculations instead of the identical extended strand used in the original pro-

Table 4. Structural statistics for the Initial ensembles of serine protease PB92.[a]

Basic statistics

Run	RMSD statistics/Å[b] Mean global backbone[c]	Mean global heavy	⟨Energies⟩/kcal mol^{-1} Bonds	Angles	Impropers	NOE	vdW[d]
DG/SA	2.11 ± 0.17	2.86 ± 0.15	111 ± 14	437 ± 35	67 ± 8	250 ± 23	462 ± 73
TAD	1.84 ± 0.18	2.60 ± 0.21	108 ± 40	345 ± 95	65 ± 25	252 ± 86	320 ± 110

Restraint violations

Run	Average No. of NOE violations > 0.2 Å	> 0.5 Å	> 1.0 Å	No. of consistent NOE violations[e] > 0.2 Å	Maximal NOE violation/Å
DG/SA	9 ± 3	0	0	2	0.47
TAD	10 ± 5	0	0	0	0.93

Average RMS deviations from ideal covalent geometry

Run	Bonds/Å	Angles/deg	Impropers/deg
DG/SA	0.005 ± 0.0003	0.7 ± 0.03	0.5 ± 0.03
TAD	0.005 ± 0.001	0.6 ± 0.08	0.5 ± 0.09

Structure validation

Run	Agreement with X-ray secondary structure[f] % agreement α-Helix	β-Sheet	3D profile score[g] Average total score	Ramachandran behaviour[h] % of residues in[i] Most fav.	Add. all.	Gen. all.	Disallowed
DG/SA	78	65	118 ± 1	56.5	32.1	8.4	3.0
TAD	82	78	120 ± 1	56.1	34.2	7.7	2.0

[a] Ensemble size: 23 conformers for both the TAD and the DG/SA ensemble.
[b] Coordinate RMSD values were computed as average pairwise RMSD values. Only residue 4–269 were included in the computations; the N-terminal residues are disordered.
[c] Backbone defined as: N, Cα, C′ atoms.
[d] Quartic repulsive vdW energy term.
[e] NOE distance restraints violated in all conformers of the ensemble.
[f] Consensus secondary structure elements within the ensemble were determined by PROCHECK-NMR [33] and compared with the secondary structure elements found in the X-ray structure of PB92 [35]. See section 2 for further explanations.
[g] As computed by PROFILE3D [34].
[h] According to PROCHECK-NMR.
[i] Abbreviations refer to regions in the Ramachandran map: Most fav., most favoured; Add. all., additional allowed; Gen. all., generously allowed.

tocol. However, none of these attempts showed significant effects, which is a testament of the sampling power of the TAD method. In order to optimize the convergence of the TAD calculations, increased simulation temperatures and prolonged simulation times in torsion angle space turned out to be very efficient. Raising the simulation temperature to 60 000 K led to a structure ensemble with significantly lower conformational energies and less average NOE violations. The best results, however, were achieved when higher temperatures for the torsion angle dynamics stage were applied during a period of simulation prolonged by a factor of two (see below). By contrast, additional periods of restrained MD in Cartesian space were less efficient in achieving lower NOE violations or conformational energies. Recently we found that complete elimination of the Cartesian stage and extension of the TAD cooling stage to 0 K further improves the convergence [37].

3.5. *Optimized protocols for structure calculations of serine protease PB92*

Tables 2 and 3 list the parameters of the DG/SA and TAD protocols, respectively, that gave the best results in

structure calculations of PB92. Accordingly, the ultimate TAD protocol comprises 30 ps of torsion angle molecular dynamics, followed by another 30 ps of cooling in torsion angle space, and a final period of MD in Cartesian space for 6 ps. The last stage of the protocol consists of an energy minimization. The total CPU time needed to generate and refine one conformer of serine protease PB92 with this TAD protocol is 55 min on an SGI R10000 Indigo, equipped with a 195 MHz processor and a main memory of 128 MB.

The optimized DG/SA protocol includes three steps, the first two of which, namely the generation of DG substructures and their regularization with the SA algorithm, have been described already (see section 3.2). The last optimized step comprises 30 ps of simulated annealing in Cartesian space, followed by a restrained energy minimization. The complete DG/SA protocol takes 35 min of CPU time on the above mentioned workstation, in order to produce one refined conformer of PB92.

Each of the optimized structure determination protocols was used to generate and refine 50 conformers again. Out of these, final ensembles again were selected by cluster analysis and then subjected to the same assessment procedures as before. Figure 1(*b*) shows the RMSD profiles resulting from an analysis of the two optimized runs. From figure 1(*a*,*b*) it is evident that the protocol optimization had the most pronounced effect on the performance of the DG/SA method. An inspection of its backbone RMSD profile shows the considerable improvement in convergence obtained by the prolongation of the SA refinement stage. While the improvement achieved for the TAD protocol appears less pronounced, nevertheless it is obvious that the TAD calculations also benefit substantially from a prolongation of the dynamics in torsion angle space. A persistent feature of the TAD calculations, however, is the occurrence of more pronounced steps in both the backbone and the all-heavy atom RMSD profile, indicating that the differences in the conformational substates sampled by this algorithm are larger than in the conformational families produced by the DG/SA method.

3.6. *Analysis of the optimized DG/SA and TAD calculations*

Statistics for the two PB92 structure ensembles determined with the two protocols are given in table 5. A comparison of these statistics with the values of the Initial ensembles (table 4) reveals that both optimized protocols deliver better structures with respect to *all* applied quality criteria, a result that might not necessarily be taken for granted.

Upon comparison of the two structural ensembles produced by the optimized calculations, it is evident from the RMSD values that the TAD ensemble is defined more precisely than the DG/SA family of conformers. Moreover, the TAD conformers also exhibit lower average conformational energies. The differences are most pronounced in the bond angle and the van der Waals term. The NOE restraint violations within the two ensembles are very similar in both number and magnitude, and the TAD ensemble displays a somewhat higher maximal NOE violation. On average, the agreement with idealized covalent geometries is also very similar for both ensembles, with the only difference occurring in the average deviations from ideal values of improper dihedral angles which are slightly lower in the TAD structure of PB92. In contrast, the *maximal* deviations from the ideal values of Cα chirality (data not shown) are significantly higher for the DG/SA ensemble. Minor differences only are found in the Ramachandran maps of the two ensembles. The DG/SA structure has a slightly higher percentage of residues adopting disallowed backbone conformations. The Ramachandran map of the TAD ensemble, generated by PROCHECK-NMR, is shown in figure 2.

A significant difference between the two structure ensembles occurs with respect to the length of their consensus secondary structure elements. The optimized DG/SA ensemble scores somewhat better in reproducing the secondary structure elements present in the X-ray structure of PB92.

On the other hand, the 3D profile score [34] indicates a higher probability for the TAD structural model of serine protease PB92 (Table 5), considering the statistical preferences of each amino acid residue for a specific environment. On the whole, it appears that the structure family of PB92 obtained with the TAD method has more favourable properties than the DG/SA ensemble.

3.7. *Structure calculations of serine protease PB92 including torsion angle restraints*

Structure determinations of PB92, following the optimized DG/SA and the optimized TAD protocol, were repeated with a modified input data set that contained additional restraints for the backbone torsional angle ϕ (see table 1). As mentioned already, it is important to compare the performance of the two methods with and without torsion angle restraints, given the fact that these restraints might have different impact on the two different methods. Therefore during the early stages of the DG/SA protocol, torsion angle restraints are introduced gradually, i.e., during the SA regularization of the distance geometry structures. In order to avoid the possibility of conformers being trapped in unfavourable local minima, the relative weight of the torsion angle

Table 5. Structural statistics for the Optimized ensembles of serine protease PB92.[a]

Basic statistics

Run	RMSD statistics/Å[b] Mean global backbone[c]	Mean global heavy	⟨Energies⟩/kcal mol^{-1} Bonds	Angles	Impropers	NOE	vdW[d]
DG/SA	1.66 ± 0.18	2.36 ± 0.19	66 ± 6	271 ± 17	40 ± 3	155 ± 13	220 ± 21
TAD	1.52 ± 0.16	2.18 ± 0.19	60 ± 7	210 ± 18	35 ± 3	143 ± 18	171 ± 26

Restraint violations

Run	Average No. of NOE violations > 0.2 Å	> 0.5 Å	> 1.0 Å	No. of consistent NOE violations[e] > 0.2 Å	Maximal NOE violation/Å
DG/SA	5 ± 2	0	0	2	0.36
TAD	5 ± 2	0	0	2	0.39

Average RMS deviations from ideal covalent geometry

Run	Bonds/Å	Angles/deg	Impropers/deg
DG/SA	0.004 ± 0.0002	0.5 ± 0.02	0.4 ± 0.01
TAD	0.004 ± 0.0002	0.5 ± 0.02	0.3 ± 0.02

Structure validation

Run	Agreement with X-ray secondary structure[f] % agreement α-Helix	β-Sheet	3D profile score[g] average Total score	Ramachandran behaviour[h] % of residues in[i] Most fav.	Add. all.	Gen. all.	Disallowed
DG/SA	90	84	122 ± 1	60.4	31.1	6.5	2.1
TAD	84	76	127 ± 1	60.4	31.5	6.4	1.7

[a] Ensemble size: 18 conformers for both the TAD and the DG/SA ensemble.
[b] Coordinate RMSD values were computed as average pairwise RMSD values. Only residues 4–269 were included in the computations; the N-terminal residues are disordered.
[c] Backbone defined as: N, Cα, C′ atoms.
[d] Quartic repulsive vdW energy term.
[e] NOE distance restraints violated in all conformers of the ensemble.
[f] Consensus secondary structure elements within the ensemble were determined by PROCHECK-NMR [33] and compared with the secondary structure elements found in the X-ray structure of PB92 [35]. See section 2 for further explanations.
[g] As computed by PROFILE3D [34].
[h] According to PROCHECK-NMR.
[i] Abbreviations refer to regions in the Ramachandran map: Most fav., most favoured; Add. all., additional allowed; Gen. all., generously allowed.

restraints during the high temperature dynamics of this simulated annealing period is very low, and increased only during the cooling period. Note that the early stages of the DG/SA protocol have not been included in table 3 since, apart from the minor changes mentioned before (section 3.2) they are identical with the default protocols distributed with the program X-PLOR [29].

The same number of conformers was generated and refined as in the other calculation runs. A cluster analysis was again performed on both the resulting conformer ensembles, but in this case, the conformers had been ranked before according to their *overall* energy. This modification of the ranking procedure takes into account that now two different violation energy terms contribute to the outcome of the calculations, corresponding to the two different types of restraint that are now available. Figure 1(*c*) shows the resulting RMSD profiles of both the TAD and the DG/SA conformer families obtained from the extended input restraint set. As for the other optimized runs performed without torsion angle restraints, the finally selected structure clusters comprise 18 conformers each.

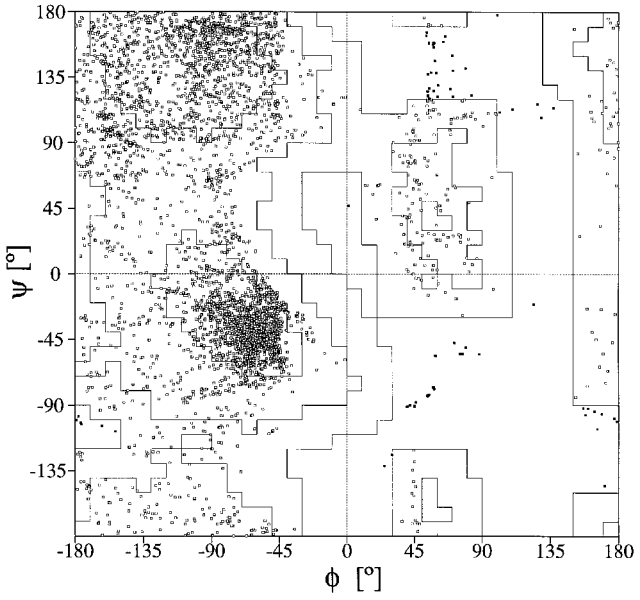


Figure 2. Ramachandran map showing all non-glycine/non-proline residues of the ensemble of 18 conformers of serine protease PB92 determined with the optimized TAD protocol. The map was generated with the program package PROCHECK-NMR [33]. Residues located in allowed ϕ/ψ regions are indicated by open symbols, and those in disallowed regions are indicated by filled symbols.

3.8. *Analysis of the ensembles obtained upon inclusion of torsion angle restraints*

Table 6 lists the statistics of the ensembles computed with both angular and distance restraints. The precision of the two conformer families, termed 'Optimized + Phi' from now on, again differs in favour of the TAD ensemble, an overlay of which is shown in figure 3. An intriguing observation results from a comparison of the given coordinate RMSD values with the corresponding numbers computed for 'Optimized' ensembles which were calculated using distance restraints only (cf. tables 5 and 6): The precision of the TAD structure slightly *increases* with the availability of torsion angle restraints, whereas the opposite is true for the DG/SA ensemble. Apparently, the inherent information content of the input data set comprising distance *and* torsion angle restraints is different for the two structure determination algorithms. On the other hand, an analysis of the corresponding conformational energy terms demonstrates clearly the self-consistency of the restraints. For both calculational methods, the increase in conformational energy resulting from an inclusion of the ϕ angle restraints is only modest (cf. tables 5 and 6), and the van der Waals energies even *decrease* in both 'Optimized + Phi' ensembles. Again, the TAD ensemble scores better with respect to the agreement with the restraints imposed by the force field. This statement also holds for the agreement with the experimental restraints, provided one considers only the *number* of NOE violations. In contrast, the values of the *maximal* violations are slightly lower for the DG/SA ensemble. The agreement with the experimental $^3J_{HNH\alpha}$ data, assessed by the mean J RMSD values, is comparable for the two ensembles; the difference of 0.07 Hz is within the experimental error margin of the values determined for the $^3J_{HNH\alpha}$ couplings of serine protease PB92 from CT-HMQC-J spectra [20, 26]. The average deviations from ideal covalent geometry are also identical for the two structure ensembles, while the maximal deviations from ideal values of Cα chirality (data not shown) are significantly lower in the TAD conformers. The Ramachandran maps of the two structure families are of comparable overall quality, but the number of residues in disallowed regions of the ϕ/ψ map is again lower for the TAD ensemble. In addition, the 3D profile scores reveal a somewhat higher quality for the TAD model of PB92. Furthermore, the consensus secondary structure elements present within the two ensembles differ in length, especially with respect to the β-sheets. The TAD structure reproduces the secondary structure features of the crystal structure of PB92 considerably better than the DG/SA structure.

Figure 4 shows a 'sausage' type backbone representation of the TAD ensemble, obtained upon inclusion of ϕ angle restraints. It is evident that the local precision of this ensemble, which is reflected in the tube width, is quite comparable for β-strands and α-helices. The good definition of the α-helices might be somewhat unexpected, given their inherently lower chemical shift dispersion, but one should keep in mind that the $^3J_{HNH\alpha}$ values encountered in α-helices allow for a tighter definition of ϕ angle restraints, compared with β-sheets. This difference is due to the Karplus equation for the $^3J_{HNH\alpha}$ coupling constant, the steepest part of which corresponds to ϕ angle values typical for α-helices.

Again, an interesting observation results from a comparison of the structure ensembles computed from both distance and angular restraints to their counterparts obtained from distance restraints only (cf. tables 5 and 6). In the latter case, the DG/SA structure family reproduces the secondary structure present in the crystal structure of PB92 better than the TAD structure, implying a higher degree of *accuracy*, whereas upon inclusion of torsion angle restraints, the opposite is true. Similar behaviour was found for the relative *precision* of the two structure calculation methods when correlated with the availability of angular restraints. The precision of the TAD structure ensemble improved slightly upon inclusion of ϕ angle restraints, whereas the precision of the DG/SA ensemble deteriorated somewhat compared with the calculation run where only dis-

Table 6. Structural statistics for the Optimized + Phi ensembles of serine protease PB92.[a]

Basic statistics

Run	RMSD statistics/Å[b] Mean global backbone[c]	Mean global heavy	⟨Energies⟩/kcal mol^{-1} Bonds	Angles	Impropers	NOE	Dih	vdW[d]
DG/SA	170 ± 0.17	2.38 ± 0.19	72 ± 7	286 ± 23	47 ± 3	182 ± 11	9 ± 1	209 ± 23
TAD	1.50 ± 0.16	2.15 ± 0.18	66 ± 5	231 ± 11	43 ± 3	164 ± 14	9 ± 2	165 ± 12

Restraint violations

Run	Average No. of NOE violations > 0.2 Å	> 0.5 Å	> 1.0 Å	No. of consistent NOE violations[e] > 0.2 Å	Maximal NOE violation/Å	Maximal dihedral violation/deg	Global $^3J_{HNH\alpha}$ RMSD/Hz[f]
DG/SA	7 ± 1	0	0	4	0.36	8	1.83
TAD	5 ± 2	0	0	3	0.52	9	1.90

Average RMS deviations from ideal covalent geometry

Run	Bonds/Å	Angles/deg	Impropers/deg
DG/SA	0.004 ± 0.0002	0.5 ± 0.02	0.4 ± 0.01
TAD	0.004 ± 0.0002	0.5 ± 0.01	0.3 ± 0.01

Structure validation

Run	Agreement with X-ray secondary structure[g] % of residues in α-Helix	β-Sheet	3D profile score[h] Average total score	Ramachandran behaviour[i] % of residues in[j] Most fav.	Add. all.	Gen. all.	Disallowed
DG/SA	87	74	121 ± 1	60.2	32.6	5.9	1.2
TAD	89	86	124 ± 1	59.5	34.3	5.3	1.0

[a] Ensemble size: 18 conformers for both the DG/SA and the TAD ensemble.
[b] Coordinate RMSD values were computed as average pairwise RMSD values. Only residues 4–269 were included in the computations; the N-terminal residues are disordered.
[c] Backbone defined as: N, Cα, C′ atoms.
[d] Quartic repulsive vdW energy term.
[e] NOE distance restraints violated in all conformers of the ensemble.
[f] Computed according to ^3J RMSD $= \left(\frac{1}{N}\sum_{i=1}^{N}(J_i^{\exp} - \langle J_i^{DG}\rangle)^2\right)^{1/2}$.
[g] Consensus secondary structure elements within the ensemble were determined by PROCHECK-NMR [33] and compared with the secondary structure elements found in the X-ray structure of PB92 [35]. See section 2 for further explanations.
[h] As computed by PROFILE3D [34].
[i] According to PROCHECK-NMR.
[j] Abbreviations refer to regions in the Ramachandran map: Most fav., most favoured; Add. all., additional allowed; Gen. all., generously allowed.

tance restraints were applied. Thus, it appears from a comparative evaluation of the performance of the two protocols that the TAD algorithm benefits from the additional angular restraints. By contrast, the DG/SA calculation method seems to be unable to exploit the higher information content now present in the input restraints. Although one has to be cautious with general conclusions derived from one data set only, it seems that on the whole the TAD protocol gives better structures from input restraint sets comprising *both* distance and angular restraints than the DG/SA method.

Conclusion

The recently introduced torsion angle dynamics algorithm together with the established DG/SA hybrid method are able to determine correctly the solution structure of the 269-residue protein serine protease PB92 from a relatively sparse input restraint set.

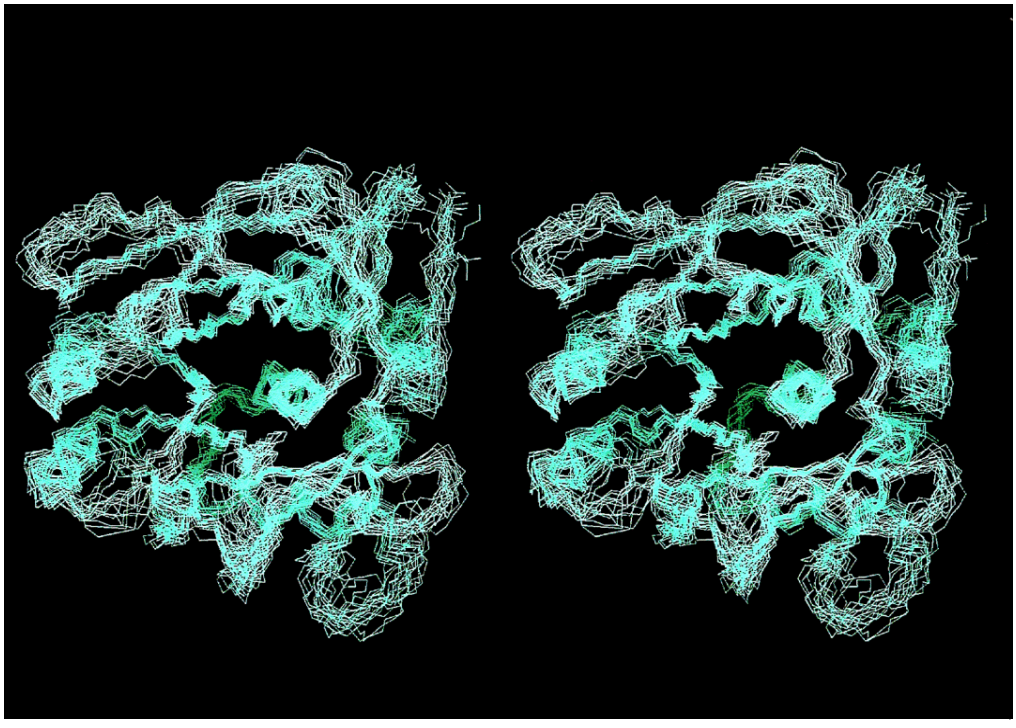


Figure 3. Stereoview of an overlay of the 18 conformers of the TAD solution structure ensemble of PB92 computed from both distance and torsion angle constraints. The conformers are superimposed onto their backbone atoms N, Cα, and C′.

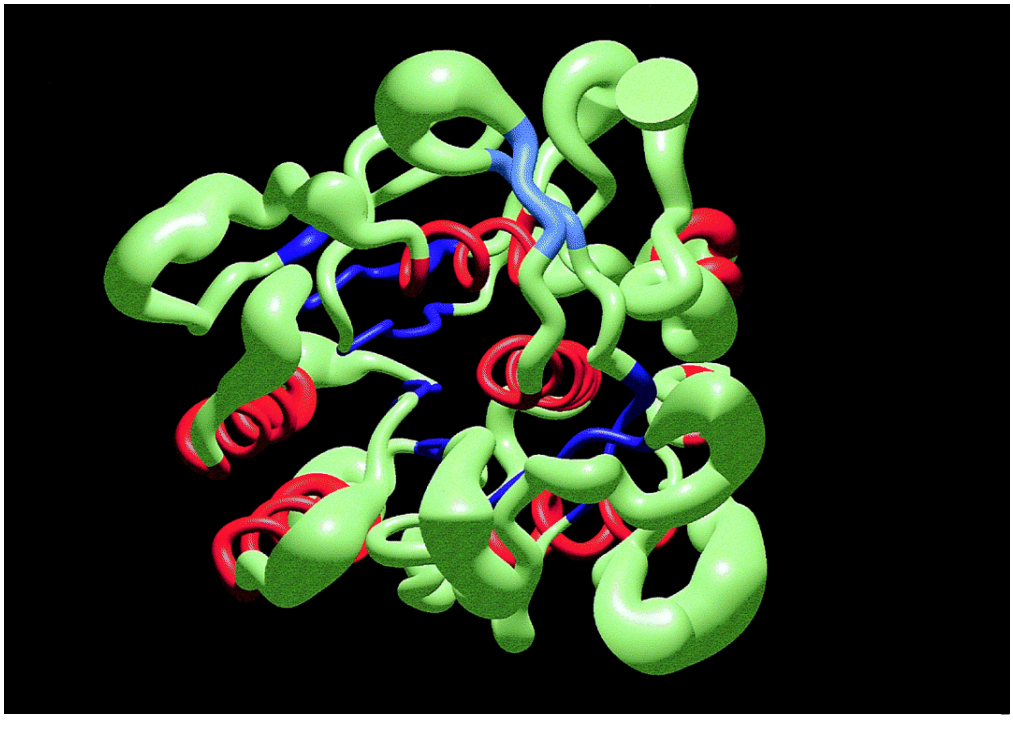


Figure 4. MOLMOL sausage representation of the backbone of serine protease PB92, determined with the TAD protocol including also constraints for the torsion angle ϕ. The width of the tube is related to the backbone RMSD within the 18 conformer ensemble. Consensus secondary structure elements of the protein as determined by PROCHECK-NMR are rendered in different colours: α-helices are depicted in orange-red, the parallel seven-stranded β-sheet of the protein is coloured in dark blue, and the short antiparallel β-sheet is rendered in light blue.

Both calculational methods, however, can be improved substantially in their performance on large molecules by introducing appropriate modifications that adapt the protocols to the size of the protein studied. From the comparative analyses presented in this paper, it is evident that molecular dynamics in torsion angle space is the method of choice for the NMR structure determination of larger biomolecules. The more favourable ratio of parameters to observables intrinsic to the TAD algorithm is reflected in its better performance compared with the DG/SA protocol.

The RMSD profiles of the structure calculations of PB92 presented in this study demonstrate that the TAD calculations sample more conformational substates of the protein than the DG/SA method. At the same time, torsion angle dynamics shows better convergence behaviour, producing structures that exhibit higher coordinate precision and lower conformational energies. When torsion angle restraints are available, the protocol also scores significantly better regarding measures of structure quality independent of the structure generation process. However, the improved quality of the TAD structures comes along with a somewhat higher computational cost in terms of the CPU time needed for the generation of *one* conformer. It should be noted, however, that up to now the TAD method has not yet been optimized in terms of CPU requirements. Furthermore, it can be expected that the iterative refinement of NMR structures using the bootstrap approach will require fewer iteration steps with the TAD method due to its superior convergence behaviour, a feature that will eventually speed up the structure determination process.

The authors wish to thank Alexandre Bonvin, Ton Rullmann and Evan Stein for valuable discussions.

Appendix

Abbreviations used

CRMSD, coordinate root mean square deviation
DFP, diisopropylfluorophosphonate
DIANA, distance geometry algorithm for NMR application
DG, distance geometry
SA, simulated annealing
EM, energy minimization
HMQC, heteronuclear multiple quantum coherence
HSQC, heteronuclear single quantum coherence
NOE, nuclear Overhauser enhancement
NOESY, nuclear Overhauser enhancement and exchange spectroscopy
PB92, serine protease PB92.

References

[1] Bax, A., and Grzesiek, S., 1993, *Accounts chem. Res.*, **26,** 131.
[2] Shan, X., Gardner, K. H., Muhandiram, D. R. Rao, N. S., Arrowsmith, C. H., and Kay, L. E., 1996, *J. Amer. chem. Soc.*, **118,** 6570.
[3] Zhang, H., Zhao, D., Revington, M., Lee, W., Jia, X., Arrowsmith, C., and Jardetzky, O., 1994, *J. molec. Biol.*, **238,** 592.
[4] Matsuo, H., Li, H., McGuire, A. M., Fletcher, C. M., Gingras, A.-C., Sonenberg, N., and Wagner, G., 1997, *Nature struct. Biol.*, **4,** 717.
[5] van den Berg, B., Tessari, M., Boelens, R., Dijkman, R., Kaptein, R., de Haas, G. H., and Verheij, H. M., 1995, *J. biomolec. NMR*, **5,** 110.
[6] Garrett, D. S., Seok, Y.-J., Liao, D.-I., Peterkofsky, A., Gronenborn, A. M., and Clore, G. M., 1997, *Biochemistry*, **36,** 2517.
[7] Martin, J. R., Mulder, F. A. A., Karimi-Nejad, Y., v.d. Zwan, J., Mariani, M., Schipper, D., and Boelens, R., 1997, *Structure*, **5,** 521.
[8] Güntert, P., Berndt, K. D., and Wüthrich, K., 1993, *J. biomolec. NMR*, **3,** 601.
[9] Meadows, R. P., Olejniczak, E. T., and Fesik, S. W., 1994, *J. biomolec. NMR*, **4,** 79.
[10] Rice, L. M., and Brünger, A. T., 1994, *Proteins*, **19,** 277.
[11] Stein, E. G., Rice, L. M., and Brünger, A. T., 1997, *J. magn. Reson.*, **124,** 154.
[12] Nilges, M., Clore, G. M., and Gronenborn, A. M., 1988, *FEBS Lett.*, **229,** 317; Nilges, M., Gronenborn, A. M., Brünger, A. T., and Clore, G. M., 1988, *Protein Eng.*, **2,** 27.
[13] Nilges, M., Kuszewski, J., and Brünger, A. T., 1991, *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy*, edited by J. C. Hoch, F. M. Poulsen and C. Redfield (New York: Plenum Press) p. 451.
[14] Kuszewski, J., Nilges, M., and Brünger, A. T., 1992, *J. biomolec. NMR*, **2,** 33.
[15] Teplyakov, A. V., van der Laan, J. M., Lammers, A. A., Kelders, H., Kalk, K. H., Misset, O., Mulleners, L. J. S. M., and Dijkstra, B. W., 1992, *Protein Eng.*, **5,** 413.
[16] Fogh, R. H., Schipper, D., Boelens, R., and Kaptein, R., 1995, *J. biomolec. NMR*, **5,** 259.
[17] Neri, D., Szyperski, T., Otting, G., Senn, H., and Wüthrich, K., 1989, *Biochemistry*, **28,** 7510.
[18] Santoro, J., and King, G. C., 1992, *J. magn. Reson.*, **97,** 202.
[19] Vuister, G. W., and Bax, A., 1992, *J. magn. Reson.*, **98,** 428.
[20] Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A., 1994, *J. biomolec. NMR*, **4,** 981.
[21] Constantine, K. L., Madrid, M., Bányai, L., Trexler, M., Patthy, L., and Llinás, M., 1992, *J. molec. Biol.*, **223,** 281.
[22] Koning, T. M. G., Boelens, R., and Kaptein, R., 1990, *J. magn. Reson.*, **90,** 111.
[23] Wüthrich, K., Billeter, M., and Braun, W., 1983, *J. molec. Biol.*, **169,** 969.
[24] Güntert, P., Braun, W., and Wüthrich, K., 1991, *J. molec. Biol.*, **217,** 517.

[25] Wüthrich, K., 1986, *NMR of Proteins and Nucleic Acids* (New York: Wiley).
[26] Billeter, M., Neri, D., Otting, G., Qian, Y. Q., and Wüthrich, K., 1992, *J. biomolec. NMR*, **2,** 257.
[27] Harbison, G. S., 1993, *J. Amer. chem. Soc.*, **115,** 3026.
[28] Case, D. A., Dyson, H. J, and Wright, P. E., 1995, *Methods Enzymol.*, **239,** 392.
[29] Brünger, A. T., 1992, *X-PLOR*, Version 3.1, a system for X-ray crystallography and NMR. (New Haven, CT: Yale University Press).
[30] Engh, R. A., and Huber, R., 1991, *Acta Crystallogr.* A, **47,** 392.
[31] Koradi, R., Billeter, M., and Wüthrich, K., 1996, *J. molec. Graph.*, **14,** 51.
[32] Widmer, H., Widmer, A., and Braun, W., 1993, *J. biomolec. NMR*, **3,** 307.
[33] Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J., 1996, *J. biomolec. NMR*, **8,** 477.
[34] Lüthy, R., Bowie, J. U., and Eisenberg, D., 1992, *Nature*, **356,** 83.
[35] van der Laan, J. M., Teplyakov, A. V., Kelders, H., Kalk, K. H., Misset, O., Mulleners, L. J. S. M., and Dijkstra, B. W., 1992, *Protein Eng.*, **5,** 405.
[36] Siezen, R. J., and Leunissen, J. A. M., 1997, *Protein Sci.*, **6,** 501.
[37] Warren, G., and Brünger, A. T., unpublished.
[38] Brünger, A. T., and Nilges, M., 1993, *Quart. Rev. biophys.*, **26,** 49.