The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures

SVEN G. HYBERTS, MATTHEW S. GOLDBERG, TIMOTHY F. HAVEL,
AND GERHARD WAGNER

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School,
240 Longwood Avenue, Boston, Massachusetts 02115

Abstract

A high-precision solution structure of the elastase inhibitor eglin c was determined by NMR and distance geometry calculations. A large set of 947 nuclear Overhauser (NOE) distance constraints was identified, 417 of which were quantified from two-dimensional NOE spectra at short mixing times. In addition, a large number of homonuclear ^{1}H–^{1}H and heteronuclear ^{1}H–^{15}N vicinal coupling constants were used, and constraints on 42 χ^1 and 38 ϕ angles were obtained. Structure calculations were carried out using the distance geometry program DG-II. These calculations had a high convergence rate, in that 66 out of 75 calculations converged with maximum residual NOE violations ranging from 0.17 Å to 0.47 Å. The spread of the structures was characterized with average root mean square deviations (⟨rmsd⟩) between the structures and a mean structure. To calculate the ⟨rmsd⟩ unbiased toward any single structure, a new procedure was used for structure alignment. A canonical structure was calculated from the mean distances, and all structures were aligned relative to that. Furthermore, an angular order parameter S was defined and used to characterize the spread of structures in torsion angle space. To obtain an accurate estimate of the precision of the structure, the number of calculations was increased until the ⟨rmsd⟩ and the angular order parameters stabilized. This was achieved after approximately 40 calculations. The structure consists of a well-defined core whose backbone deviates from the canonical structure ca. 0.4 Å, a disordered N-terminal heptapeptide whose backbone deviates by 0.8–12 Å, and a proteinase-binding loop whose backbone deviates up to 3.0 Å. Analysis of the angular order parameters and inspection of the structures indicates that a hinge-bending motion of the binding loop may occur in solution. Secondary structures were analyzed by comparison of dihedral angle patterns. The high precision of the structure allows one to identify subtle differences with four crystal structures of eglin c determined in complexes with proteinases.

Keywords: comparison X-ray; eglin c; elastase inhibitor; NMR; solution structure

Protein structure and mobility determine biological function. The elastase inhibitor eglin c is a paradigm for a protein containing both a well-defined rigid core and regions of high mobility. Both aspects seem to be important for the function of this protein. Eglin c is a proteinase inhibitor from the leech *Hirudo medicinalis* (Seemüller et al., 1980). It inhibits proteinases, such as elastase, cathepsin G, chymotrypsin, thermistase, and subtilisin.

Eglin c consists of a single polypeptide chain with 70 residues. Although lacking disulfide bonds, it is very stable against denaturation by heat (Seemüller et al., 1980). The crystal structure of the protein has been determined in complexes with subtilisin Carlsberg, subtilisin Novo, and thermistase (McPhalen et al., 1985; Bode et al., 1986, 1987; McPhalen & James, 1987, 1988; Dauter et al., 1988; Gros et al., 1989; Heinz et al., 1991). Some of these structures were determined in different crystal forms (Gros et al., 1989; Heinz et al., 1991), and structures of some mutant eglins have also been investigated (Heinz et al., 1991). The structure of the homologous serine pro-

teinase inhibitor CI-2 from barley seeds has been determined in the complex with proteinases by X-ray crystallography (McPhalen & James, 1988). The structure of uncomplexed CI-2 was determined by X-ray crystallography (McPhalen & James, 1987) and by NMR spectroscopy (Clore at al., 1987a,b).

In this paper we present a high-precision structure determination based on NMR data. Compared to previous studies, we made use of a large number of heteronuclear ^{15}N-$^{1}H^{\beta}$ and homonuclear $^{1}H^{\alpha}$-$^{1}H^{\beta}$ vicinal coupling constants to characterize the dihedral angles χ^{1}. We made efforts to obtain a complete sampling of the conformational space by maximizing the root mean square deviations in a large set of calculations. Only structures that have no significant violations of constraints were accepted. This investigation is part of a larger effort to characterize and separate static and dynamic aspects of protein structure. A preliminary comparison with the crystal structures of eglin c in complexes with subtilisin Carlsberg, subtilisin Novo, and thermistase (Bode et al., 1987; McPhalen & James, 1988; Gros et al., 1989; Heinz et al., 1991) was also made.

Results

NOE distance constraints

The data used for generating the structures of eglin c consist of distance constraints obtained from two-dimensional nuclear Overhauser effect (2D NOE) spectra, and dihedral angle constraints derived from measurements of homonuclear and heteronuclear coupling constants. Four 2D NOE spectroscopy (NOESY) spectra were used for the analysis, two of which were recorded in H_2O, pH 3.0, with mixing times of 200 ms and 50 ms, and two in D_2O

with mixing times of 200 ms and 75 ms. All NOE cross peaks that could be identified in the 200-ms spectra were associated with an upper distance limit of 5.0 Å. All cross peaks that could be observed in the 50-ms spectra and were well resolved were integrated using the integration routines in the software packages FTNMR and FELIX (Hare Research, Inc.). Two reference points were used for calibration of the NOE intensities. First, the sequential $d_{\alpha N}(i - 1, i)$ NOE intensities, measured in a 50-ms NOESY, were plotted versus the amino acid sequence (Fig. 1A). In total, 47 of the possible 63 cross peaks could be integrated. The $d_{\alpha N}(i - 1, i)$ were absent for only four peptide bonds, whereas the peaks for four connectivities were heavily overlapped, and eight $d_{\alpha N}(i - 1, i)$ cross peaks were present but too weak to be integrated quantitatively. Due to the constraints imposed by the covalent structure, the corresponding distances can vary only between 2.2 Å and 3.6 Å (Billeter et al., 1982; Wüthrich, 1986). Many residues of eglin c have strong $d_{\alpha N}(i - 1, i)$ cross peaks. Therefore, we associated the mean value of these points in Figure 1A with a distance of 2.2 Å. Second, the intraresidue $d_{\alpha N}(i, i)$ NOEs were integrated (open circles in Fig. 1A). Due to the efficient elimination of J-cross peaks as described in the Methods (Rance et al., 1985), reliable intensities could be derived. These NOEs have a narrow intensity distribution, which is consistent with the fact that usually only negative ϕ-angles are populated with $d_{\alpha N}(i, i)$ distances between 2.4 Å and 2.9 Å. The mean of these intensities was then associated with a distance of 2.8 Å. For this calibration point we excluded the glycines (residues 4, 15, 40, 59, and 70) and a residue with apparently positive ϕ angles (Asn 61). In Figure 1A the NOE intensities are plotted on a volume$^{-1/6}$ scale, corresponding to a linear distance scale. This scale was then used to calibrate all NOEs.
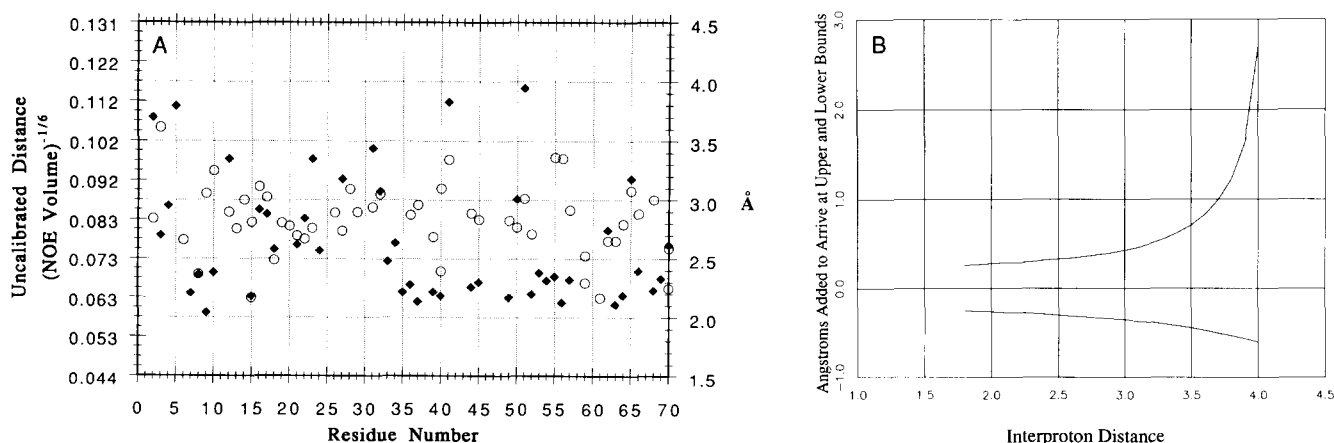

Fig. 1. A: Plot of the integrated NOE cross peak volumes vs. sequence for intraresidue $d_{\alpha N}(i, i)$ NOEs (open circles) and sequential $d_{\alpha N}(i - 1, i)$ NOEs (filled diamonds). The volumes are plotted on a volume$^{-1/6}$ scale (left) and a linear distance scale (right). **B:** Plot of distance limits used for the quantified NOEs. This curve was produced by using the smallest integrated cross-peak intensity as a measure of the error of the integration. The r^{-6} dependence leads to the trumpet-shaped curve (see text).

Error limits for the distance calibration were obtained by the following procedure. The intensity of the smallest cross peak that was well above the noise level and could reliably be integrated was taken as the uncertainty of the measurements. The value of this integral corresponded to a distance of 4.3 Å. The lower and upper distance limits, r_l and r_u, respectively, for the structure calculations were then determined in the following way:

$$r_l = \frac{1}{\left(\dfrac{1.4}{r_o^6} + \dfrac{1}{4.3^6}\right)^6} - 0.15$$

$$r_u = \frac{1}{\left(\dfrac{0.7}{r_o^6} - \dfrac{1}{4.3^6}\right)^6} + 0.15.$$

The factors 1.4 and 0.7 reflect the estimated relative uncertainty of the volume integration, the number 4.3 represents the distance corresponding to the noise level, and 0.15 Å was added and subtracted, respectively, to take account of the uncertainty of the distance calibration. This results in a trumpet-shaped pair of curves when plotted vs. the distance (Fig. 1B). This relation was used to obtain distance constraints on all quantified NOEs. Reinspection of Figure 1A shows that all the NOE intensities used for the calibration are consistent with these error limits, except for the intraresidue $d_{\alpha N}(i, i)$ NOE for Phe 3, which gives too large a lower distance limit. However, it is in the N-terminal strand, which is highly mobile, so that a significantly shorter correlation time might lead to partial quenching of the NOEs.

Whenever an NOE involved a methyl group, a correction of 1.1 Å was added to the upper distance limit, and this limit was related to a pseudoatom M located in the center of the triangle spanned by the three methyl protons instead of the methyl carbon. For methylene groups that were not stereospecifically assigned, the following correction was made: If an NOE was observed to one of the two methylene protons in the 50-ms spectrum, 1.8 Å was added to the upper limit, and the constraint was imposed to both methylene protons. If NOEs were observed to both methylene protons in the 50-ms spectrum, the weaker NOE was used for the constraint of the upper limit, and the stronger NOE was converted to the lower distance limit for both protons. For methyl groups that were not stereoselectively assigned, analogous corrections were also made: If an NOE was observed to one of the two methyl groups in the 50-ms spectrum, 4.2 Å (1.1 Å correction for the methyl group and 3.2 Å for the distance between the two methyl pseudoatoms) was added to the upper limit, and the constraint was imposed to both methyl pseudoatoms M. If NOEs were observed to both methyl groups in the 50-ms spectrum, the weaker NOE was used for the upper limit, and the stronger NOE was

converted to the lower distance limit for both pseudo-atoms M in question. The residues of the proteinase-binding loop, 38–49, and the N-terminal residues 1–7 obviously have a higher mobility than the core of the protein, based on measurements of ^{15}N relaxation parameters. Therefore, for each proton of these regions, a correction of 0.1 Å was added to the upper distance limit. All aromatic side chains of eglin c are rotating rapidly. Whenever an NOE was observed for H^δ or H^ϵ atoms, constraints were imposed to the C^γ and C^ζ atoms, respectively, and 2.1 Å were added to take account of the distance between the hydrogen and carbon positions. For NOEs to methylene protons and isopropyl methyl groups observed only in 200-ms NOESY spectra, corrections of 1.8 Å and 4.2 Å were used always to consider the effect of spin diffusion.

A total of 947 NOEs were identified and used for the structure determination, of which a total of 417 interresidue and intraresidue NOEs were quantified in the 50-ms H_2O NOESY spectrum. Intraresidue NOEs were only used when they were quantified. A complete list of the NOE constraints used for the calculations is available in the Supplementary material (Table S1, Diskette Appendix). All upper distance limits that were larger than the maximum distance possible due to constraints of the covalent structure (mainly intraresidue and sequential NOE constraints) were automatically rejected by the program DG-II and are not counted in the numbers of NOEs given above.

Stereospecific assignments

Stereospecific assignments of the $C^\beta H_2$ groups were made using measurements of heteronuclear vicinal coupling constants between the amide nitrogen and the β-protons together with the homonuclear vicinal coupling constants between α- and β-protons. This method also yields information on the values of the dihedral angles χ^1, as follows. First, one assumes that the side chain is either locked in one of the three staggered rotamer states with $\chi^1 = +60°$, $-60°$, or $180°$, or that it is jumping between the three states. For homonuclear H^α–H^β couplings we expect values of ca. 3.4 Hz and 12.9 Hz if the two protons are gauche or trans, respectively, and a value of ca. 6.6 Hz if the side chain is hopping between the three rotamer states with equal populations. For the heteronuclear N-H^β vicinal couplings we expect values of ca. -0.4 Hz and -5.7 Hz if the two nuclei are gauche or trans, respectively, and a value of -2.2 Hz if the side chain is jumping between the three rotamer states with equal populations (Bystrov, 1976). Measurement of the four vicinal coupling constants shows whether the side chain is locked in one of the three staggered conformations or is jumping between them. In this way, the χ^1 angle can be determined and stereospecific assignments can be derived at the same time. The same analysis can be applied to de-

termine the χ^1 angles of threonines, valines, and isoleucines (although eglin c has none of the latter). The dihedral angles and the stereospecific assignments for the $C^\beta H_2$ protons are listed in the supplementary material (Table S2, Diskette Appendix). We note that the stereospecific assignments for the $C^\beta H_2$ protons derived in this way differ from those achieved by analysis of homonuclear coupling constants and intraresidue NOEs (Hyberts et al., 1987; Hyberts & Wagner, 1990). This seems to be due to an error in the program used for the previous assignments and will be discussed elsewhere together with details of the present analysis (Goldberg, Hyberts, & Wagner, unpubl.).

Dihedral angle constraints

For three of the four threonines the χ^1 rotamer state could be determined uniquely, whereas Thr 1 had homonuclear and heteronuclear coupling constants consistent with rotational averaging of χ^1. For 9 of the 11 valines the rotamer state could also be determined uniquely. Val 13, on the other hand, shows the homonuclear and heteronuclear coupling constants expected for a rotating side chain (although it cannot be excluded with certainty that this residue has an unusual [eclipsed] χ^1 angle). Its side chain is pointing toward the interior of the protein and is in close contact with the side chain of Phe 25, which is rapidly rotating on the NMR time scale. Val 43, which is part of the proteinase-binding loop, also shows coupling constants indicating some rotational mobility. For 29 of the 41 residues with $C^\beta H_2$ groups (except prolines) the χ^1 angle range could be characterized unambiguously. No attempt was made to characterize the χ^1 angles of prolines prior to the structure calculations. Gln 20, Asp 33, and Leu 37 were found to have averaged χ^1 angles, and Ser 9, Lys 16, Glu 39, Ser 41, and Asn 61 could not be analyzed because of overlap in 2D and three-dimensional (3D) spectra. For Lys 8, Phe 10, Glu 12, Asp 19, and Glu 23, the data were inconclusive. If a defined rotamer orientation could be identified, this χ^1 angle ($+60°$, $-60°$, or $180°$) $\pm 30°$ was used in the distance geometry calculations. The Supplementary material (Table S2) contains a list of these constraints along with the constraints on the angle ϕ, which were obtained from measurements of the H^N-H^α coupling constants with a similar procedure as described by Kline et al. (1988).

Hydrogen-bond constraints

Constraints for hydrogen bonds were used in regions of the protein structure where clear indications of a regular secondary structure were obtained. The distance between the carbonyl oxygen and the amide proton was then constrained within the limits of 1.8 Å and 2.5 Å. This range of distances was selected considering the range of hydrogen-bond lengths observed in crystal structures of pro-

teins, which was estimated by comparing the H-bond distances in the crystal structure of the basic pancreatic trypsin inhibitor (Pardi et al., 1983; Wagner et al., 1983). The Supplementary material also contains a list of these H-bond constraints (Table S3, Diskette Appendix).

Distance geometry structure calculations

The complete set of constraints was first subjected to several exploratory distance geometry runs with the DG-II program package to check for inconsistencies in the data. These seven runs of 10–25 structures each were performed prior to the production runs and proved very helpful in tracking down the inevitable typographical errors and a few erroneous NOE assignments. The production runs consisted of a set of 75 calculations. Complete triangle smoothing and a tetrangle smoothing over sequential amino acids was first used. Random distance matrices were chosen using metrization (Havel, 1990) and embedding in four dimensions followed by simulated annealing and conjugant gradient refinement in three dimensions as described in Havel (1991). Out of the 75 calculations, 66 converged with an error function below 0.5, 10 structures had error functions between 0.5 and 0.9, and 3 structures had high error function above 10. The latter 9 structures were considered unacceptable and not included in further analysis. In the set of acceptable structures, the largest violation of an NOE upper bound in each of the 66 structures ranged from 0.17 Å to 0.47 Å. Only 5 of the accepted 66 structures have violations of upper distance limits larger than 0.3 Å. All violations of lower distance bounds are smaller than 0.08 Å. There are only eight upper-bound constraints that are violated on the average by more than 0.1 Å, five of these are intraresidue H^N-H^β constraints that were quantified in a 50-ms NOESY spectrum, apparently underestimating the effect of spin diffusion via the second H^β. The other three are sequential NOE constraints. In total, only 15 NOE constraints are violated in all 66 structures. The average largest upper-bound violation is 0.19 Å, and there are 36 structures with no violation of an NOE upper bound larger than 0.20 Å. There is only one lower bound that is violated in all 66 structures (maximum violation 0.02 Å). All violations of the dihedral angle constraints were smaller than 2°. These results indicate that there are no serious geometric inconsistencies in the constraints.

Alignment of structures

The alignment of structures was achieved without giving preference to any one of them using the following procedure. Each of the squared interatomic distances was averaged over the 66 accepted structures. The corresponding root mean square (rms) distance matrix was used as input for the EMBED program of the DG-II package (Crippen & Havel, 1978). This yields 3D coordinates

whose distances are a best fit to the rms distances, and hence a canonical average structure. The alignment used for the comparisons was then obtained by computing the rms deviation (rmsd) of the well-defined residues to this canonical average structure, where the rmsd is defined as in Rao and Rossmann (1973):

$$\text{rmsd} = \left[\frac{1}{n} \min_{i=1 \text{ to } n} \sum |\mathbf{r}_i - R\mathbf{r}_i'|^2\right]^{1/2}.$$

Here, $\mathbf{r}_i$ and $\mathbf{r}_i'$ are the vectors from the common centroid to the ith atom in the two structures. The translation and rotation transformation is selected that makes the sum a minimum. The sum runs over all backbone atoms of the residues 8–38 and 50–68 selected for alignment.

For comparing structures locally, we use the same alignment as above and sum over a section of atoms; we denote this by rmsd(selection). In particular, we use rmsd(residue i) when we sum over the atoms of the ith residue:

$$\text{rmsd(selection)} = \left[\frac{1}{k} \sum_{\text{selection}} |\mathbf{r}_i - R\mathbf{r}_i'|^2\right]^{1/2}.$$

Here, k is the number of atoms in the selected substructure; when $k = 1$, we mean a single atom. In addition, we use the average rmsd, $\langle\text{rmsd}\rangle$, and $\langle\text{rmsd(selection)}\rangle$ when we average over m pairs of structures.

$$\langle\text{rmsd}\rangle = \frac{1}{m} \sum_{\text{pairs}} \text{rmsd}_{\text{pair}}.$$

Figure 2A–C shows stereo views of the 66 structures that were accepted after the DG-II calculations; Kinemage 1 shows a subset of these structures. The structures were aligned using residues 8–38 and 50–68, which excludes the poorly defined regions of the protein, the N-terminal heptapeptide, and the proteinase-binding loop.

Distribution of violations and differences in the DG-II structures

Figure 3A shows an overview of the accumulated violations of NOE distance constraints vs. residue number. Figure 3B gives the absolute number of NOE distance constraints per residue, and Figure 3C shows the average violation, i.e., the sum of violations divided by the number of NOE constraints. This figure shows that the violations are concentrated mainly in the proteinase-binding loop. This observation indicates that the protons are in different environments due to structure fluctuations and sample NOEs from different conformations that cannot be satisfied by a single structural model. The larger spread in the binding loop of the ensemble of structures displayed in Figure 2 and Kinemage 2 provides an impression of the

mobility of this region of the protein. However, one has to be aware that each individual structure of the ensemble has some violations of the constraints and hence is slightly "wrong." Thus, the ensemble of structures is not an entirely correct description of the dynamic ensemble of structures, though it is probably indicative.

Figure 4 is a plot of the $\langle\text{rmsd(residue)}\rangle$ for the 66 accepted structures to the canonical structure, calculated with respect to the global alignment of the structures obtained using the well-defined segments 8–38 and 50–68 as described above. The solid line gives the $\langle\text{rmsd(residue)}\rangle$ selecting only the backbone atoms N, C^α, and C'; the broken line is for all heavy atoms in the residue, including the entire side chains. We have calculated both the $\langle\text{rmsd}\rangle$ values between the structures, and to the canonical structure, which give values of 0.61 Å and 0.43 Å, respectively, for the backbone of the well-defined segments 8–38 and 50–68. Also, all $\langle\text{rmsd}\rangle$ values reported are for a global alignment of the structures. This differs from practices used earlier, where stretches of three residues were aligned to show the differences of local structures (Wagner et al., 1987). Instead, we analyze the precision of the local structures with plots of an angular order parameter S, which are entirely independent of the alignment procedure (see below).

Figure 4 shows that the N-terminal heptapeptide, the binding loop, and the hairpin turn of the antiparallel β-sheet at around residue 60 are much less well defined than the rest of the molecule. The $\langle\text{rmsd(residue)}\rangle$ to the mean coordinates for the backbone of the N-terminal heptapeptide ranges from ca. 12 Å at residue 1 to 1 Å at residue 7. The core of the protein has values of the $\langle\text{rmsd(residue)}\rangle$ smaller than 0.8 Å, and for most of the residues it is below 0.4 Å (Fig. 4). It is significantly higher at the turn of the β-sheet (57–62) with values of 0.4–0.8 Å, and around the 3_{10}-helix at the residues 12–17 with values of 0.4–0.6 Å. In the proteinase-binding loop including the residues 37–51 the $\langle\text{rmsd(residue)}\rangle$ ranges from 0.4 to 3.4 Å, with peak values at residues 40, 44, and 46.

Order parameters of the dihedral angles

We define an order parameter S for each dihedral angle to describe how well the angles are defined. For example, the order parameter $S(\alpha_i)$ for the angle α_i of residue i (where $\alpha = \phi, \psi, \omega, \chi^1$, or χ^2, etc.) is defined as:

$$S(\alpha_i) = 1/N \left\|\left(\sum_{j=1}^{N} \alpha_i^j\right)\right\|.$$

Here, N is the total number of structures and α_i^j ($j = 1, \ldots, N$) is a 2D unit vector with phase equal to the dihedral angle α_i. Here, i represents the residue number, and j stands for the number of the calculated structure.
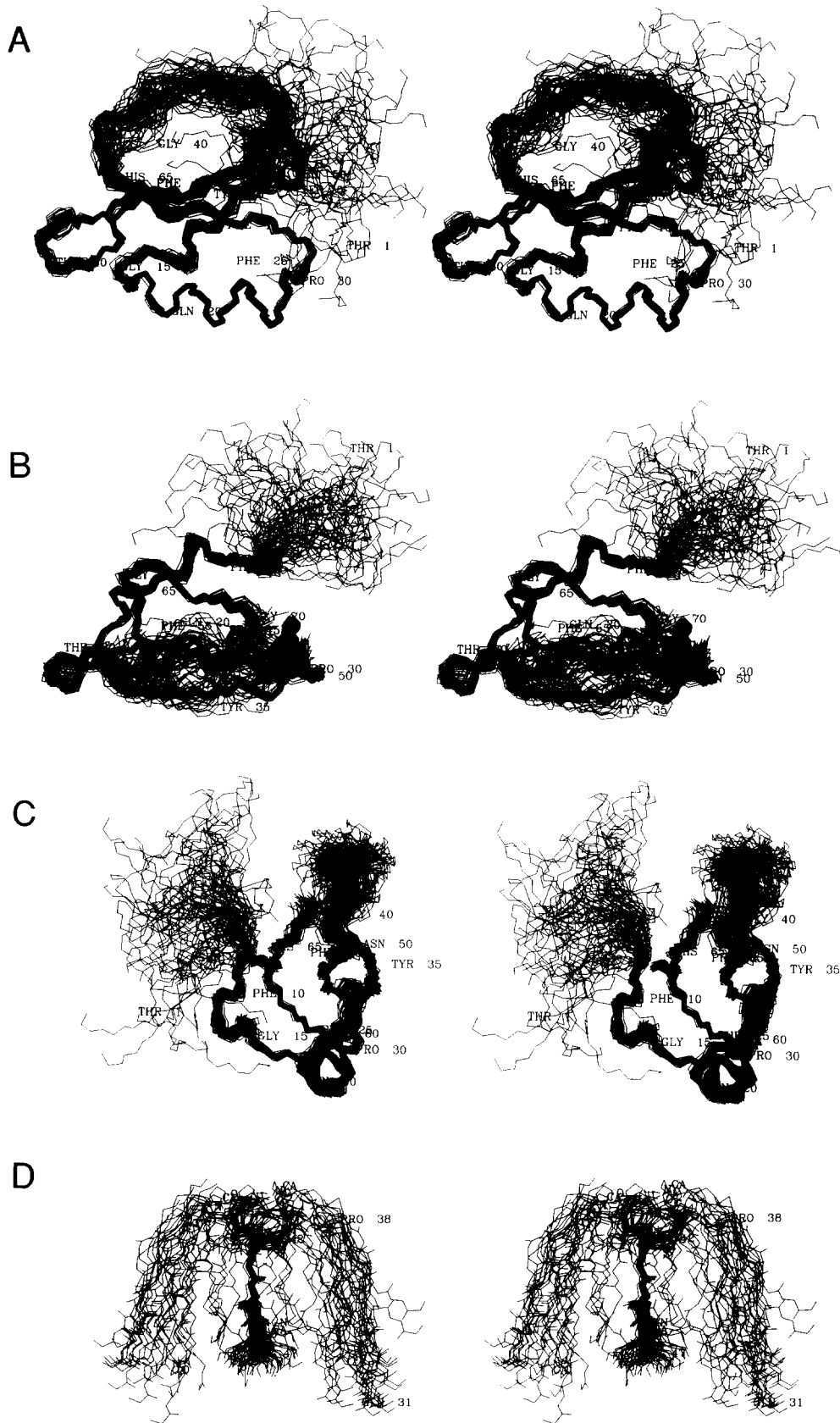
Fig. 2. Stereo representation of the 66 accepted structures of eglin c. Three different orientations are shown (A–C). D: Alignment of the binding loop residues 42–46. As reference for the orientation of the core of the molecule relative to the binding loop, the strand 31–38 is shown.
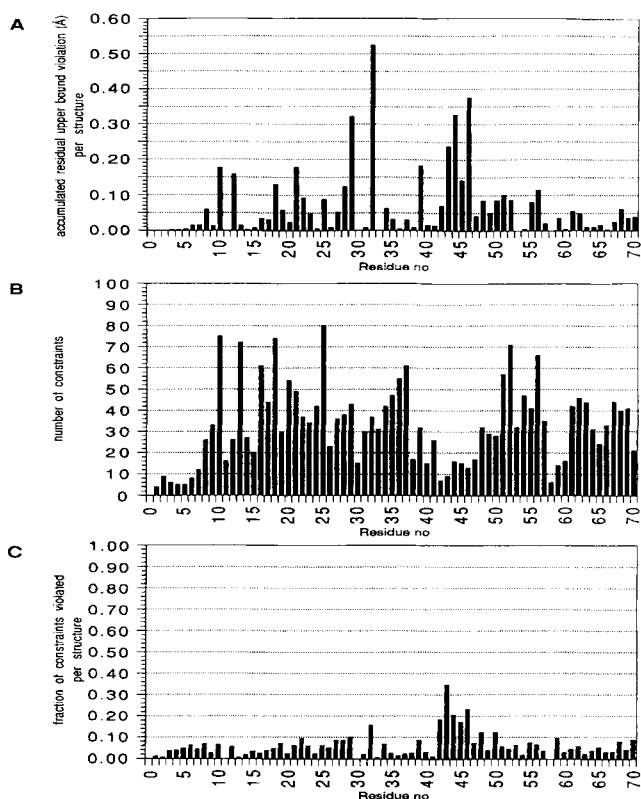
Fig. 3. Plot of (A) the average accumulated violations of NOE-derived distance constraints per residue, (B) number of constraints per residue, and (C) average violation of NOE constraints per residue.
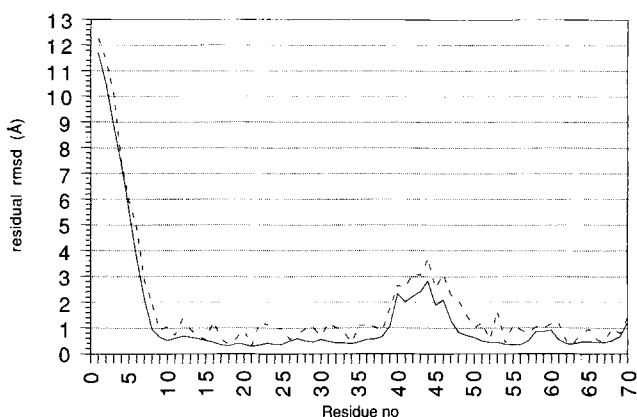

Fig. 5. Relation between S^{angle} and the standard deviation of the dihedral angle (see text).

the logarithm of S is plotted vs. the standard deviation, σ, of the dihedral angles (small circles). This relation was obtained by determining each S (each small circle) from a synthetically generated set of 100,000 unit vectors in the plane with a known Gaussian phase distribution. The solid line represents an empirically found approximate relation given by:

$$1 + 0.5 \lg(S) = \cos(\sigma/2)$$

$$\sigma = 2 \arccos[1 + \lg(S)/2].$$

For example, a value of $S = 0.9$ corresponds to a standard deviation σ of $\pm24°$, whereas a value of $S = 0.95$ corresponds to a σ of $\pm17°$, and a value of $S = 0.99$ corresponds to a σ of $\pm7.5°$.

The order parameter S, defined in this way is related to the standard deviation of the dihedral angles (Fig. 5). In our opinion, the angular order parameter is a better quantity to describe the precision of dihedral angles than a standard deviation. It clearly defines the two limits of an exactly defined angle ($S = 1$) and a completely random distribution of the angle ($S = 0$). For a completely random angle distribution the standard deviation of the angles is not defined. Furthermore, S is easier calculated than a standard deviation.

Figure 6 shows the values of the dihedral angles and the angular order parameters versus residue number for the dihedral angles ϕ, ψ, χ^1, and χ^2. The backbone dihedral angles ϕ and ψ are very well defined except in the N-terminal heptapeptide, at the beginning (around residue 40) and the end (around residue 47) of the binding loop, and at the turn of the antiparallel β-sheet at around residue 58. Most of the χ^1 angles are well defined except for Thr 1, Ser 9, Glu 12, Val 13, Thr 17, Glu 23, Tyr 32, Leu 37, Glu 39, Ser 41, and Asn 61. Remarkably, Val 13

Fig. 4. Plot of ⟨rmsd(residue)⟩ for the 66 accepted structures versus sequence. The solid line gives the values for the backbone atoms N, C$^\alpha$, and C'; the broken line gives the values for all heavy atoms.

If the angle is the same in all structures, then S has the value of 1, whereas a small value indicates a disordered structure. The relation of the parameter S to the standard deviation of the dihedral angle is given in Figure 5, where
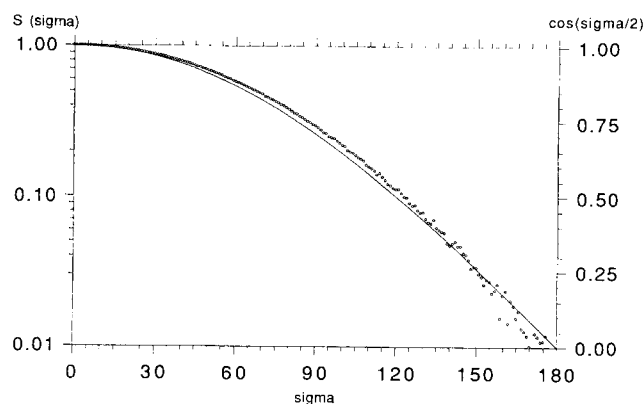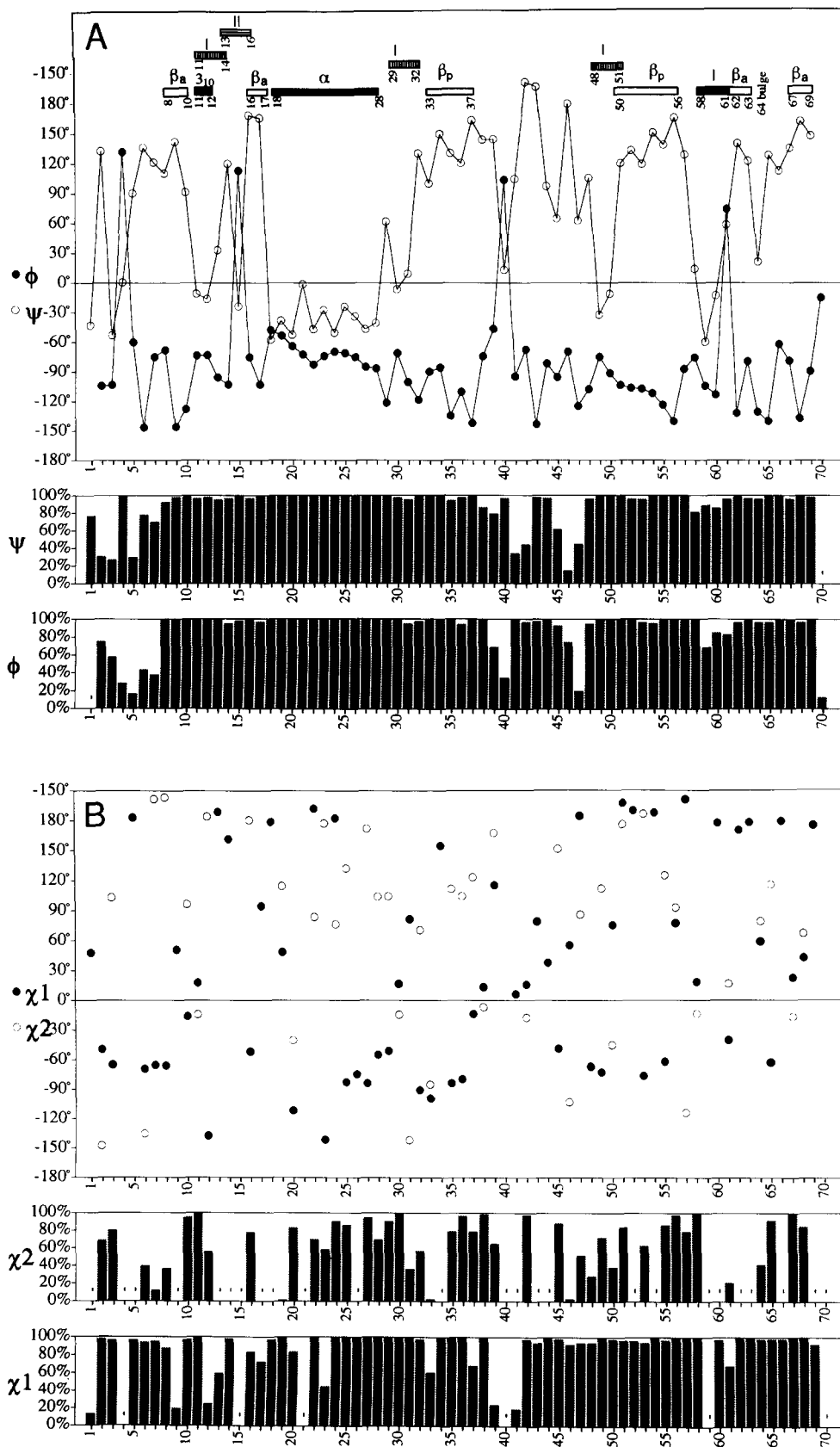
Fig. 6. Plot vs. sequence of the average dihedral angles ϕ and ψ (**A**), and χ^1 and χ^2 (**B**). The angular order parameters S^{angle} are shown at the bottom of the figures. At the top of Figure 6A, the secondary structures are indicated as derived by comparison with the patterns of Figure 7.
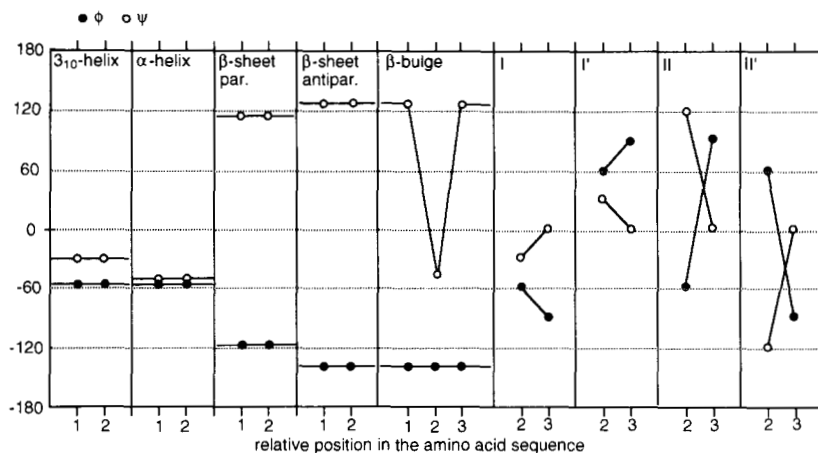
Fig. 7. Patterns of dihedral angles ϕ and ψ for the α-helix, the 3_{10}-helix, the parallel and antiparallel β-sheet, and the turns of type I, I', II, and II', and the β-bulge, using values for the angles as given by Richardson (1981). The horizontal axis indicates the relative position in the amino acid sequence. In particular for the turns, the angles are given for the positions 2 and 3. This plot was used as an optical aid to identify secondary structures in particular turns and the exact beginnings and ends of helices.

has its side chain in the interior of the protein (additional evidence for its mobility was given in the Results). Also, most of the χ^2 angles are relatively well defined.

The dihedral angle order parameter S as well as the average rmsd was used as a means of checking the completeness of the sampling of the allowed conformational space. It should be emphasized that it is desirable to *maximize* the average rmsd and to minimize S for a given set of constraints without increasing violations of the constraints, as we want to have a complete and unbiased sample of the conformations compatible with the constraints. To achieve this goal, the number of structure calculations was increased as long as the order parameters S continued to decrease, or the average rmsds continued

to increase. After ca. 40–50 converged structures, the order parameter did not decrease anymore for any part of the structure, and the average rmsd no longer increased. Therefore, we conclude that a rather complete sampling of the conformational space consistent with the constraints was achieved.

In order to check whether the dihedral angles ϕ and ψ fall within the allowed regions of the Ramachandran diagram, we have plotted the ϕ and ψ angles for all 66 converged structures. They are shown in Figure 8A. If we plot only ϕ/ψ angles with values of the order parameter S larger than 0.9 (Fig. 8B) we have a distribution of angles clearly consistent with the allowed ranges in the Ramachandran diagram.
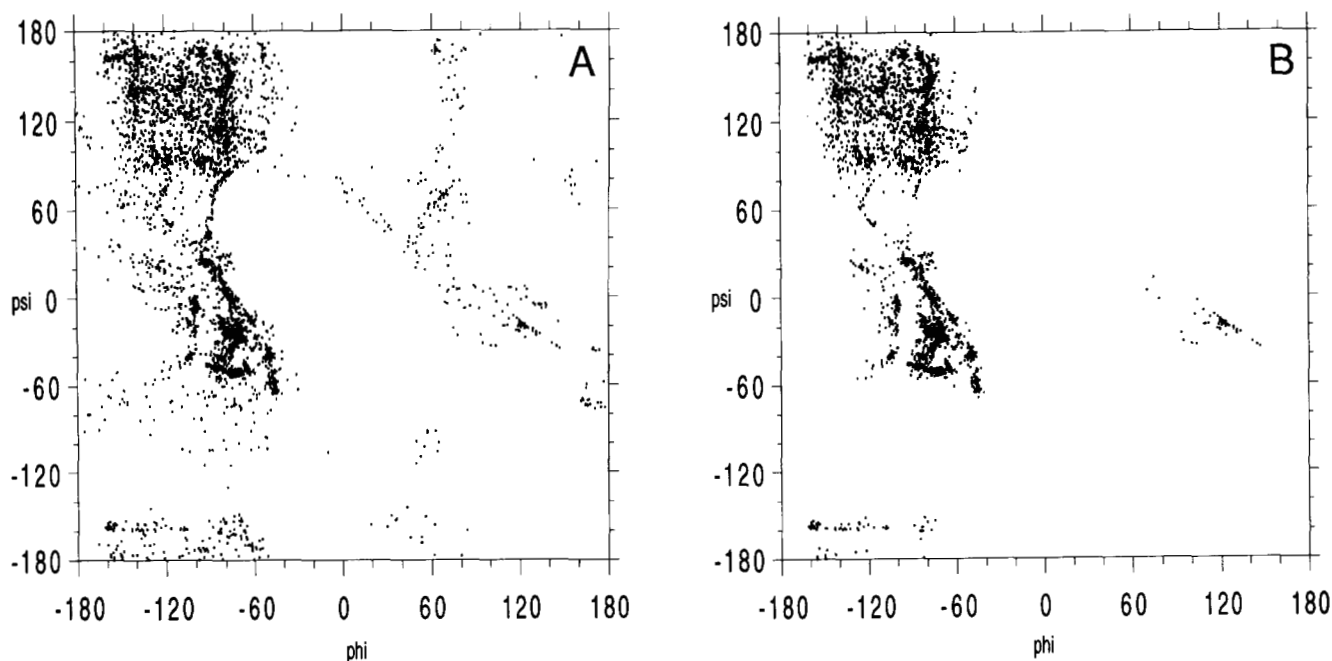

Fig. 8. Ramachandran plot for (**A**) all ϕ/ψ angles for the 66 structures and (**B**) only residues with the angular order parameter $S > 0.9$.

Description of the structure of eglin c

The secondary structural elements of eglin c were assigned by comparison with the known patterns of dihedral angles ϕ and ψ in the various types of secondary structures. This is outlined as follows. In Figure 7 we have plotted these patterns of dihedral angles ϕ and ψ using the ideal angles (Richardson, 1981) for the α-helix, the 3_{10}-helix, the parallel and antiparallel β-sheet, and the turns of type I, I', II, and II', and the β-bulge. This visual inspection of the dihedral angle patterns was very valuable, in particular for the location of turns, as well as the exact identification of the beginnings and ends of the helices. Some slight differences to the crystal structure were detected in this process (see below). The upper part of Figure 6A shows the secondary structure elements identified in this way. A representative structure is shown in Kinemage 3.

The N-terminal heptapeptide is essentially undefined; there are no long-range constraints in this region. Independently, relaxation time measurements of ^{15}N nuclei indicate a high mobility of this region (Peng et al., 1991; Peng & Wagner, unpubl.). Residues Lys 8–Phe 10 have an extended backbone and form an antiparallel β-sheet with the C-terminal residues Pro 67–Val 69. The ϕ and ψ angles are around $-100° \pm 40°$ and $+120° \pm 30°$, respectively (Fig. 6A). Residues Pro 11–Glu 12 form a 3_{10}-helix with ϕ and ψ angles around $-90° \pm 20°$ and $+0° \pm 20°$, respectively (Fig. 6A). This means that the NH of Val 13 is involved in a hydrogen bond with the carbonyl of Phe 10. The ψ angles around $0°$ (compared to $-60°$ to $-30°$ for α-helices) make this secondary structure clearly distinct from the α-helix (see below). The characterization of this piece of secondary structure was not possible from the primary data (NOEs and coupling constants) but it was unambiguous from the results of the distance geometry runs and the inspection of the resulting dihedral angles (Fig. 6A).

In addition, the residues 11–16 show the characteristics of two interlocked tight turns. These turn and helix characteristics are actually quite common for 3_{10}-helices (Richardson, 1981). The residues Pro 11–Glu 12–Val 13–Val 14 form a type I turn with dihedral angles $\phi_2 = -70°$, $\psi_2 = -10°$, $\phi_3 = -100°$, $\psi_3 = +40°$ (see Fig. 6A). The residues Val 13–Val 14–Gly 15–Lys 16 form a type II turn with dihedral angles $\phi_2 = -100°$, $\psi_2 = +120°$, $\phi_3 = +110°$, $\psi_2 = -20°$ (see Fig. 6A). The H^N of Lys 16 is slowly exchanging (Hyberts & Wagner, 1990), being involved in a hydrogen bond with the carbonyl of Val 13. Residues Lys 16–Val 18 form an antiparallel β-sheet with residues Asn 61–Val 63 with amide protons of residues 16, 18, and 62 being involved in three interstrand hydrogen bonds. There is an α-helix ranging from Val 18–His 28 with the typical ϕ and ψ angles (Fig. 6A). The H^N of Tyr 29 is involved in the last H bond of the helix and is slowly exchanging (Hyberts & Wagner, 1990). The ϕ and ψ angles are all (except for Ala 21) around $-60°$, which makes this

secondary structure distinct from the 3_{10}-helix observed for residues 11–12. Tyr 29–Pro 30–Gln 31–Tyr 32 form a type I turn (see Fig. 6A). The H^N of Tyr 32 is slowly exchanging (Hyberts & Wagner, 1990), being involved in a hydrogen bond with the carbonyl of Tyr 29. Residues Asp 33–Leu 37 form a parallel β-sheet with residues Asn 50–Tyr 56. The first H bond is Asp 33(H^N)–Asn 50(CO), and the last H bond is Leu 37(CO)–Tyr 56(H^N).

The conformation of the proteinase-binding loop is much less well defined than the rest of the protein. The dihedral angles ψ(Glu 39), ψ(Gly 40), ϕ(Ser 41), and ϕ(Pro 42) have low angular order parameters, S, indicating that the orientation of the backbone conformation is not well defined. Generally the binding loop between residues Ser 41 and Arg 48 is in an extended conformation with ψ angles between $+60°$ and $+180°$, and ϕ angles between $-60°$ and $-150°$. However, the ϕ and ψ angles are poorly defined at around residues 39–42, as well as around 46 and 47. This indicates that these residues may function as kind of a hinge. The backbone dihedral angles are relatively well defined for the residues Val 43–Thr 44–Leu 45, where Leu 45 is the P1 residue preceding the scissile bond. Thus, it appears that the backbone conformation of the P3–P2–P1 residues is well defined but moving relative to the rest of the molecule via the flexible hinges around residues 40 and 46. The low values of the angular order parameters for some regions of the binding loop do not directly prove that there is increased internal mobility, they just show that the structures are not well defined. This might also be due to the lower number of constraints that could be identified. However, measurements of ^{15}N relaxation parameters clearly indicate a significantly higher mobility for this part of the molecule (Peng et al., 1991; Peng & Wagner, unpubl.).

At the end of the proteinase-binding loop, residues Arg 48–Tyr 49–Asn 50–Arg 51 form a type I turn with dihedral angles $\phi_2 = -80°$, $\psi_2 = -29°$, $\phi_3 = -101°$, $\psi_2 = +0°$ (see Fig. 6A). The amide proton of Arg 51 shows slow isotope exchange rates in accordance with a hydrogen bond with the carbonyl of Arg 48. The residues Asn 50–Tyr 56 form a parallel β-sheet pairing with the residues Asp 33–Leu 37, as mentioned above. In addition, they form two pieces of an antiparallel β-sheet, where Gly 70(HN) and Val 69(CO) form hydrogen bonds with Arg 51(CO) and Val 52(H^N), as evidenced by slow exchange rates and characteristic NOEs. Residue Asn 57 forms two hydrogen bonds to residue Val 62 (Hyberts & Wagner, 1990).

The residues Pro 58–Gly 59–Thr 60–Asn 61 form a type I turn. Residue 61 has a slowly exchanging amide proton that seems to be involved in a hydrogen bond with the CO of Pro 58. The angular S-factors for this hairpin are relatively low (Fig. 6A), which is also manifested in the higher spread of structures in this region (Fig. 2). Residues Val 62–Gly 70 form the central strand of the triple-

stranded antiparallel β-sheet. All residues have ϕ-angles around $-100°$ and ψ-angles around $+120°$, except for Asn 64, which is in the center of a β-bulge.

Side-chain orientations

Although generally the χ^1 angles are well defined, some side chains appear to be rotating around the $C^\alpha-C^\beta$ bond. These are Thr 1, Ser 9, Val 13, Gln 20, Glu 23, Asp 33, Leu 37, Glu 39, Ser 41, and to some degree Asn 61. Figure 9 shows a stereo diagram of the 10 structures that had the lowest error function after the DG-II calculations. All heavy atoms including the side chains are shown except for the N-terminal heptapeptide. It can be seen that almost all side chains are well defined and very similar to each other as far out as to the C^γ atom. All aromatic side chains in eglin c flip rapidly on the NMR time scale, and no individual assignments of 2 vs. 6 protons and 3 vs. 5 protons could be made. Thus, the fact that most of the aromatic side chains have well-defined χ^2 angles is a consequence of the protein packing, not of direct experimental data.

Discussion

Performance of the structure calculations

The structure calculations described above have a high rate of convergence and fulfill almost perfectly all experimentally derived constraints. It is likely that the large number of calculations performed yielded a nearly complete sampling of the conformational space compatible with the experimental data. The few structures that were not accepted show local or global mirror images of the chain fold in the accepted structures. In one nonaccepted structure, the strand of residues 8-20 was trapped be-

tween the two strands of the antiparallel β-sheet. The high precision of the structure determined in solution enabled us also to characterize secondary structures in detail, such as the type of tight turns, and thus to identify subtle differences between the X-ray structure and the NMR structure. We have calculated the values of the $\langle rmsd \rangle$ for all pairs of the NMR structures, and between the 66 NMR structures and each X-ray structure. The values are listed in Table 1. Figure 11 shows a comparison of the backbone structures of the four X-ray structures (solid lines) with the 66 NMR structures (dotted lines).

Backbone conformation

We paid attention to characterizing the surface and compared the dihedral angles with those of the four crystal structures accessible to us. Figure 11 shows that the binding loop is well defined in the crystal structures of the complexes with proteinases but has high variability in the solution structure of the free inhibitor. The N-terminal

Table 1. *Values of the $\langle rmsd \rangle$ in Å between the different NMR and X-ray structures of eglin c*[a]

		X-ray			
	NMR	B	H	G	M
NMR	0.61	0.85	0.87	0.85	0.84
B			0.48	0.56	0.23
H				0.51	0.50
G					0.45

[a] Only the well-defined regions 8-38 and 50-68 are considered. The letters B, H, G, and M stand for the structures determined by Bode et al. (1987), McPhalen and James (1988), Gros et al. (1989), and Heinz et al. (1991).
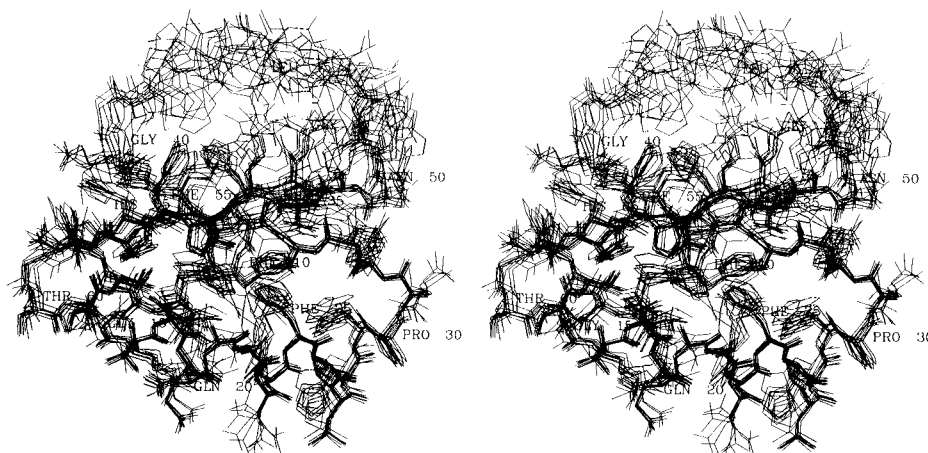

Fig. 9. Stereo diagram of the 10 structures that had the lowest error function after the DG-II calculations. The whole protein is shown, including side chains but excluding the N-terminal heptapeptide.
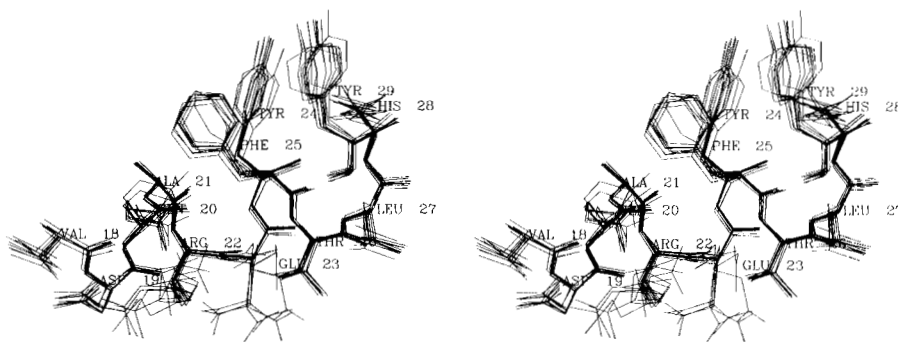
Fig. 10. Stereo diagram of the α-helix including side chains from the same 10 structures shown in Figure 9.
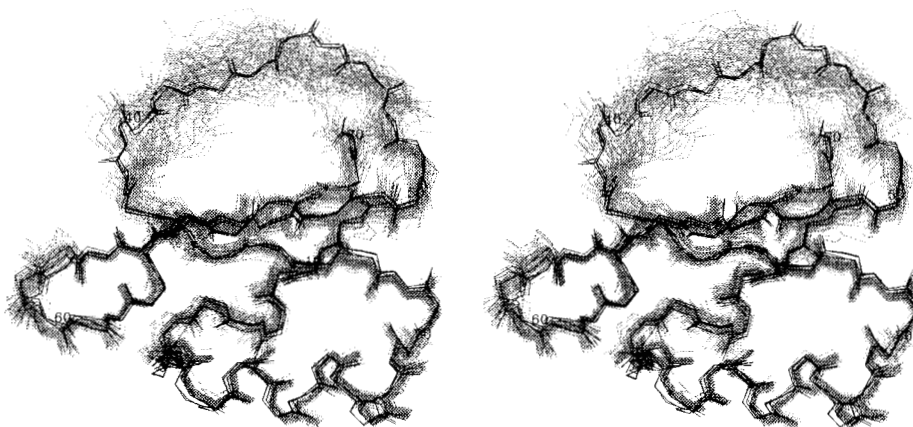

Fig. 11. Comparison of one of the NMR structures with the four X-ray structures for which we had access to the coordinates. The structures were aligned the same way as described for the alignment of the NMR structures. This is relative to the canonical NMR structure, and considering only residues 8–38 and 50–68.

residues 1–7 are not seen in the crystal structures and cannot be compared. There are also some differences at the turn of the antiparallel β-sheet that have been described above. For the rest of the molecule, the backbone conformations are virtually identical. In the region of the α-helix the X-ray structures seem to be slightly displaced from the NMR structures. Comparing the separation of the α-helix from the strand 32–38 (Fig. 11), the NMR structures appear to be slightly more compact than the X-ray structures. This appears to be the main reason why the ⟨rmsds⟩ between the NMR structures and the X-ray structures are larger than those between the X-ray structures. The local conformations are very similar. The local structure is most clearly manifested in dihedral angles. Therefore, Figure 12 shows a comparison of the ranges of dihedral angles in the NMR structures and the X-ray structures. To avoid a heavy zigzag, deviations from the mean angles in the NMR structures are plotted. We are aware that, in some cases, the mean angles may represent

impossible conformations but they provide a convenient reference for comparison. The solid lines are the deviations of the angles of the four X-ray structures of Bode et al. (1986), McPhalen and James (1988), Gros et al. (1989), and Heinz et al. (1991) from the mean angles in the NMR structures.

Generally, the ϕ and ψ angles of the X-ray structures are within the ensemble of the NMR structures with a few exceptions. The ranges of the ϕ and ψ angles for the 66 NMR structures are significantly larger than for the four X-ray structures. However, in a number of cases, the maxima or minima in the distribution are determined by angles in one or two NMR structures. Therefore, this representation has to be compared with the plot of the S-values (Fig. 6). Interesting differences between the X-ray structures and the NMR structures exist for the residues Pro 58 and Gly 59. The residues Pro 58–Gly 59–Thr 60–Asn 61 form a type I turn in the majority of the NMR structures. Residue 61 has a slowly exchanging amide
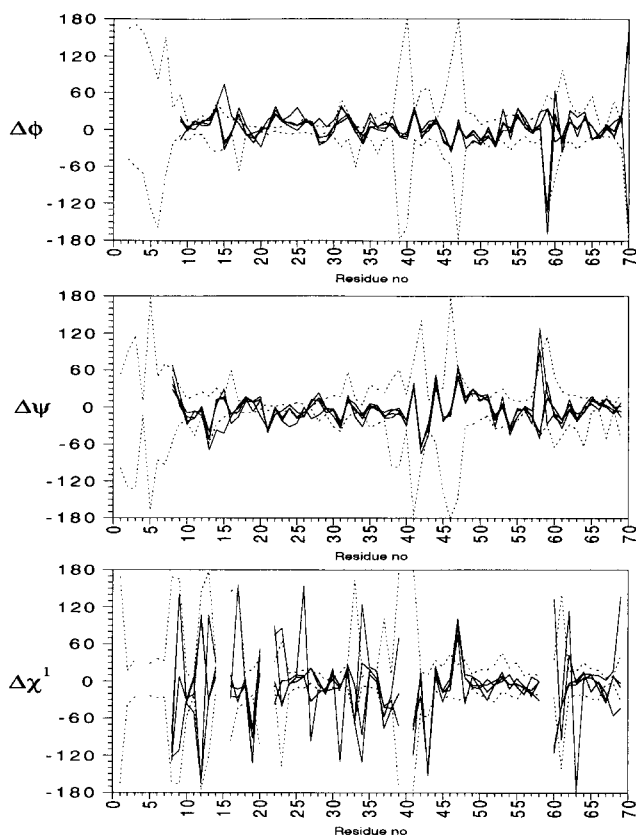
Fig. 12. Comparison of the ranges of dihedral angles ϕ, ψ, and χ^1 observed in the NMR structures with the angles in the four crystal structures. The ranges of deviations from the average dihedral angles of the 66 NMR structures (Fig. 6) are drawn with dashed lines. The deviations of the dihedral angles of the four crystal structures from the average angles of the NMR structures are drawn with solid lines. The difference of the angle in the NMR structure minus the value in the X-ray structure is shown.

Table 2. *Comparison of dihedral angles χ^1 between the four X-ray structures and the NMR structures*[a]

A. χ^1 angle differs between different X-ray structures and is undefined in the DG-II structures:

   Ser 9, Glu 12, Val 13, Thr 17, Glu 23, Glu 39

B. χ^1 angle is well defined in X-ray and NMR structures but is different:

   Lys 8, Asp 19, Val 43, Leu 47

C. χ^1 angle inconsistent in X-ray but well defined in NMR:

   Leu 27, Gln 31, Val 34, Thr 60, Val 62, Val 63, Val 69

D. χ^1 angle consistent in X-ray but large spread in NMR:

   Asp 33, Leu 37, Ser 41, Asn 61

[a] For underlined residues, only one of the four crystal structures differs from the NMR structures.

proton, which seems to be involved in a hydrogen bond with the CO of Pro 58. In two of the crystal structures, there are dihedral angles resembling a type II turn, shifted by one residue number, involving Asn 57–Pro 58–Gly 59–Thr 60, whereas the other two crystal structures have similar dihedral angles as the NMR structure. The angular S-factors for this hairpin are relatively low (Fig. 6A), which is also manifested in the higher spread of structures in this region (Fig. 2). However, although Asn 61 has a slowly exchanging amide, Thr 60 does not.

Side-chain conformations

The comparison of the ranges of the χ^1 angles is interesting. On average, it seems to be larger in the X-ray structures than in the NMR structures. Most of the side chains with a well-defined χ^1 angle have the same orientation as in the crystal structure. There are six residues where the χ^1 angle is not defined (low S) in the ensemble of NMR structures, and the individual X-ray struc-

tures exhibit definite values that are, however, different in the four crystal structures (Table 2A). This indicates, that these angles may be undefined in the crystal structures. This group includes the residues Ser 9, Glu 12, Val 13, Thr 17, Glu 23, and Glu 39. However, there are some residues where all four crystal structures agree on χ^1 but are different from the angle in solution, which is also well defined (Table 2B): Lys 8 has a χ^1 angle of 180° in all crystal structures whereas in the solution structure it is $-60°$; this residue is the first one that is structured following the flexible N-terminal heptapeptide. No dihedral angle constraints were imposed on χ^1 of Lys 8, and the resultant angle is mainly a consequence of the NOE constraints. Because there is more mobility in the N-terminal end of the protein, there is a possibility that the average NOEs are stronger than those corresponding to the average distance. Because we did not use a different calibration for the NOEs in the N-terminal residues, there is the danger that the distance constraints are too tight in this protein area, Asp 19 has a χ^1 angle of around +60° in all 66 NMR structures, whereas the four crystal structures exhibit values of $-84°$, $-78°$, $-43°$, and $-28°$, respectively. In the X-ray structures, one side-chain oxygen forms a hydrogen bond with the H^N of the same residue. In the NMR structure, the χ^1 angle is around +60° so that both side-chain oxygens can form hydrogen bonds with the two H^N's of Asp 19 and Gln 20. This is energetically favorable, compensating the positive side of the helix dipole. Figure 10 shows a stereo diagram of the α-helix from the 10 structures that had the lowest error function after the DG-II calculations. All heavy atoms including the side chains are shown. The orientation of the side chain of Asp 19 can readily be seen. Val 43 and Leu 47 belong to the proteinase-binding loop, and the different side-chain orientations in the crystal

structure may be a consequence of the contact with the proteinase. This, however, also means that these side chains reorient upon proteinase binding. As has been described above, the side chain of Val 43 shows indications of mobility around the C^α-C^β bond, with a preferred χ^1 angle around $+60°$.

There are seven residues where the χ^1 angles differ between the crystal structures, while the NMR structures have a well-defined value (Table 2C). For five out of the seven (underlined in Table 2C), three crystal structures agree with the NMR structures, and only one crystal structure (Gros et al., 1989) disagrees. For Val 34 the crystal structures sample all three rotamer states and one nonrotamer state, and for Thr 60 three crystal structures give values of $+60°$, one quotes $-60°$. In contrast, the NMR structures show that this angle is $180°$ in solution. This is also based on direct measurements of coupling constants. Thr 60 points out in the solution so that the different orientations could be explained by crystal contacts, while Val 34 is oriented toward the protein interior (Fig. 9).

There are four residues with a narrow distribution of χ^1 angles in the crystal structures but a wide distribution in the NMR structures (Table 2D). In all these cases, the average angle in the NMR structures is close to that in the X-ray structures. Out of these, however, Ser 41 has a very low S value of ca. 0.2.

In the X-ray structures, the N-terminal residues 1–6 are not seen, and only in the structure of McPhalen and James (1988) is residue 7 visible. In the NMR analysis, all resonances of this N-terminal strand were assigned, sequential NOEs were observed, but no long-range NOEs are present. This shows that this protein moiety is disordered. Independently, ^{15}N relaxation measurements show that this peptide is highly mobile (Peng et al., 1991). Details of these investigations will be shown elsewhere (Peng & Wagner, unpubl.). The proteinase-binding loop is very well defined in the X-ray structure of the complex. The strand Gly 40 to Leu 45 forms the central strand of a triple-stranded antiparallel β-sheet where the two outer strands are from the proteinase. Because this scaffold is absent for the free inhibitor, this strand is mobile as derived from the larger ⟨rmsd(residue)⟩ (Fig. 4), residual violations of distance constraints (Fig. 3), and studies of relaxation phenomena (Peng et al., 1991).

When we compare the values of the angular order parameter S for the backbone dihedral angles ϕ and ψ with those of χ^1 (Fig. 6), we make a noteworthy observation. In regions of high backbone mobility (or disorder, respectively) as detected by the order parameter S, in particular at the N-terminus and in the binding loop, the dihedral angles χ^1 are significantly better defined than ϕ and ψ, indicating that the side-chain orientation is better defined than the backbone conformation. Obviously, there are more coupling constant data to constrain χ^1 (see above) than ϕ, and the latter angle is mainly defined by NOEs. It remains to be seen if this is of functional relevance, or if we were simply unable to obtain a sufficient number of constraints in these regions of the protein.

Hinge bending motions of the binding loop

The angular order parameters of the backbone dihedral angle ϕ and ψ show minima at residues 40 and 47. In between, the angular order parameters are high for ϕ and for some ψ angles (Fig. 6A,B). Alignment of the residues 42–46 shows that this part of the molecule has internal rigidity. It appears that this strand moves as a rigid entity relative to the core of the protein. One could argue that the apparent hinges of this motion might be an artifact due to missing constraints in the hinge regions. However, studies of relaxation times and heteronuclear ^1H-^{15}N NOEs clearly show that the binding loop is significantly more mobile than the core of the molecule (Peng et al., 1991; Peng & Wagner, unpubl.). The internal rigidity of the binding loop seems to be necessary for the molecule to be an inhibitor. The segmental mobility may facilitate the fit of the inhibitor into the active site of the proteinase and enhance the adaptability of the inhibitor to the binding pocket.

Materials and methods

Sample preparation

Eglin c was produced by Ciba-Geigy, Basel, Switzerland, from a synthetic gene cloned and expressed in high yield in *Escherichia coli*. Structural analysis showed that the expressed eglin c is modified to N^α-acetyl eglin c in a posttranslational event (Märki et al., 1985). For the NMR measurements 0.006 M solutions in H_2O or ^2H_2O were used. For the measurements in ^2H_2O all labile protons were exchanged with deuterons prior to the measurements. All experiments were carried out at pH 3.0. ^{15}N-enriched protein ($>95\%$) was produced by growing *E. coli* on minimum media with ^{15}N ammonium chloride as the sole nitrogen source.

NMR experiments

The NMR spectra were measured on a Bruker AMX-500 and a General Electric GN-500 spectrometer. For the data used in this study, two NOESY spectra (Jeener et al., 1979; Kumar et al., 1980) were recorded with mixing times of 200 ms and 50 ms, respectively. The spectrum with the short mixing time was recorded as outlined in Macura et al. (1981) and Rance et al. (1985) to minimize zero-quantum contributions in NOESY cross peaks of coupled protons. This was achieved by shifting a 180° pulse systematically from the center of the mixing time by the amount τ. Thirteen spectra were recorded with $\tau = 0$, $\Delta\tau, 2\Delta\tau, \ldots, n\Delta\tau$, where $\Delta\tau$ was adjusted to $1/(\Delta\omega_{max} +$

$\Delta\omega_{min}$), and $\Delta\omega_{max}$ and $\Delta\omega_{min}$ are the maximum and minimum range of zero-quantum frequencies to be suppressed. Here a value of $\Delta\tau = 0.25$ ms was used. The spectrum with $\tau = 0$ was recorded separately with half the number of scans. All other spectra were coadded during acquisition.

^1H-^{15}N coupling constants

Vicinal ^1H-^{15}N coupling constants were measured in 2D total correlation spectra (TOCSY) (Müller & Ernst, 1979; Braunschweiler & Ernst, 1983; Bax & Davis, 1985) of ^{15}N-enriched eglin c in H_2O. The isotropic mixing was performed with a DIPSI-2 sequence (Shaka et al., 1988). The coupling constants were measured from the separation of the two ^{15}N doublets of the H^N-H^β cross peaks as described in Montelione et al. (1989). To resolve overlap of cross peaks in the 2D spectra a 3D ^{15}N-dispersed NOE spectrum was recorded without ^{15}N decoupling during detection as described by Wider et al. (1989). Homonuclear H^α-H^β coupling constants were estimated as small or large, respectively, from simulation of double-quantum-filtered COSY spectra and from multiplet patterns of the H^N-H^β cross peaks in the TOCSY spectra.

Structure calculations

Distance geometry calculations were performed using the program DG-II (Havel, 1991), running on a Sun SPARCstation-2. Each single structure calculation required approximately 4 h. The display of structures was carried out on Silicon Graphics IRIS 4D/35 and 4D/240 computers using the INSIGHT/DISCOVER software package from BIOSYM Technologies, Inc.

References

Bax, A. & Davis, D.G. (1985). MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magnet. Reson.* 65, 355–360.

Billeter, M., Braun, W., & Wüthrich, K. (1982). Sequential resonance assignments in protein 1H nuclear magnetic resonance spectra. Computation of sterically allowed proton–proton distances and statistical analysis of proton–proton distances in single crystal protein conformations. *J. Mol. Biol.* 155, 321–346.

Bode, W., Papamokos, E., & Musil, D. (1987). The high resolution X-ray structure of the complex formed between subtilisin Carlsberg and eglin c, an elastase inhibitor from the leech *Hirudo medicinalis*. *Eur. J. Biochem.* 166, 673–692.

Bode, W., Papamokos, E., Musil, D., Seemüller, U., & Fritz, H. (1986). Refined 1.2 Å crystal structure of the complex formed between subtilisin Carlsberg and the inhibitor eglin c. Molecular structure of eglin c and its detailed interaction with subtilisin. *EMBO J.* 5, 813–818.

Braunschweiler, L. & Ernst, R.R. (1983). Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J. Magnet. Reson.* 53, 521–528.

Bystrov, V.F. (1976). Spin coupling and the conformational states of peptide systems. *Prog. NMR Spectrosc.* 10, 41–81.

Clore, G.M., Gronenborn, A.M., James, M.N.G., Kjaer, M., McPhalen, C.A., & Poulsen, F.M. (1987a). Comparison of the solution and X-ray structures of barley serine proteinase inhibitor 2. *Protein Eng.* 1, 313–318.

Clore, G.M., Gronenborn, A.M., Kjaer, M., & Poulsen, F.M. (1987b). The determination of the three-dimensional structure of barley serine proteinase inhibitor 2 by nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *Protein Eng.* 1, 305–311.

Crippen, G.M. & Havel, T.F. (1978). Stable calculation of coordinates from distance information. *Acta Crystallogr.* A34, 282–284.

Dauter, Z., Betzel, C., Höhne, W.E., Ingelman, M., & Wilson, K.S. (1988). Crystal structure of a complex between thermistase from *Thermoactinomyces vulgaris* and the leech inhibitor eglin c. *FEBS Lett.* 236, 171–178.

Gros, P., Betzel, C., Dauter, Z., Wilson, K.S., & Hol, W.G.J. (1989). Molecular dynamics refinement of a thermistase–eglin c complex at 1.98 Å resolution and comparison of two crystal forms that differ in calcium content. *J. Mol. Biol.* 210, 347–367.

Havel, T.F. (1990). The sampling properties of some distance geometry algorithms applied to unconstrained polypeptide chains: A study of 1830 independently computed conformations. *Biopolymers* 29, 1565–1585.

Havel, T.F. (1991). An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. *Prog. Biophys. Mol. Biol.* 56, 43–78.

Heinz, D.W., Priestle, J.P., Rahuel, J., Wilson, K.S., & Grütter, M.G. (1991). Refined crystal structures of subtilisin novo in complex with wild-type and two mutant eglins. Comparison with other serine proteinase inhibitor complexes. *J. Mol. Biol.* 217, 353–371.

Hyberts, S.G., Märki, W., & Wagner, G. (1987). Stereospecific assignments of side chain protons and characterization of torsion angles. *Eur. J. Biochem.* 164, 625–635.

Hyberts, S.G. & Wagner, G. (1990). Sequence-specific ^1H NMR assignments and secondary structure of eglin c. *Biochemistry* 29, 1465–1474.

Jeener, J., Meier, B.H., Bachmann, P., & Ernst, R.R. (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* 71, 4546–4553.

Kline, A.D., Braun, W., & Wüthrich, K. (1988). Determination of the complete three-dimensional structure of the α-amylase inhibitor Tendamistat in aqueous solution by nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* 204, 675–724.

Kumar, A., Ernst, R.R., & Wüthrich, K. (1980). A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton–proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* 95, 1–6.

Macura, S., Huang, Y., Suter, D., & Ernst, R.R. (1981). Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J. Magnet. Reson.* 43, 259–281.

Märki, W., Rink, H., Schnebli, H.P., Liersch, M., Raschdorf, F., & Richter, W. (1985). Isolation and characterization of "native" and rec. eglin c from *E. coli*, selective proteinase inhibitors from human leucocyte elastase, cathepsin G and chymotrypsin. In *Peptides: Structure and Function, Proceedings of the 9th American Peptide Symposium* (Deber, C.M., Hruby, V.J., & Kopple, K.D., Eds.), pp. 385–388. Pierce Chemical Company, Rockford, Illinois.

McPhalen, C.A. & James, M.N.G. (1987). Crystal and molecular structure of the serine proteinase inhibitor CI-2 from barley seeds. *Biochemistry* 26, 261–269.

McPhalen, C.A. & James, M.N.G. (1988). Structural comparison of two serine proteinase inhibitor complexes: Eglin-c–subtilisin Carlsberg and CI-2–subtilisin Novo. *Biochemistry* 27, 6582–6598.

McPhalen, C.A., Schnebli, H.P., & James, M.N.G. (1985). Crystal and molecular structure of the inhibitor eglin from leeches in complex with subtilisin Carlsberg. *FEBS Lett. 188*, 55-58.

Montelione, G.T., Winkler, M.E., Rauenbuehler, P., & Wagner, G. (1989). Accurate measurements of long-range heteronuclear coupling constants from homonuclear 2D NMR spectra of isotope-enriched proteins. *J. Magnet. Reson. 82*, 198-204.

Müller, L. & Ernst, R.R. (1979). Coherence transfer in the rotating frame. Application to heteronuclear cross-correlation spectroscopy. *Mol. Phys. 38*, 963-992.

Pardi, A., Wagner, G., & Wüthrich, K. (1983). Protein conformation and proton nuclear magnetic resonance chemical shifts. *Eur. J. Biochem. 137*, 445-454.

Peng, J.W., Thanabal, V., & Wagner, G. (1991). 2D heteronuclear NMR measurements of spin-lattice relaxation times in the rotating frame of X nuclei in heteronuclear HX spin systems *J. Magnet. Reson. 94*, 82-100.

Rance, M., Bodenhausen, G., Wagner, G., Wüthrich, K., & Ernst, R.R. (1985). A systematic approach to the suppression of J cross-peaks in 2D exchange spectroscopy. *J. Magnet. Reson. 62*, 497-510.

Rao, S.T. & Rossmann, M.G. (1973). Comparison of super-secondary structures of proteins. *J. Mol. Biol. 76*, 241-256.

Richardson, J. (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem. 34*, 167-339.

Seemüller, U., Eulitz, M., Fritz, H., & Strobl, A. (1980). Structure of the elastase-cathepsin G inhibitor of the leech *Hirudo medicinalis*. *Hoppe-Seyler's Z. Physiol. Chem. 361*, 1841-1846.

Shaka, A.J., Lee, C.J., & Pines, A. (1988). Iterative schemes for bilinear operators; application to spin decoupling. *J. Magnet. Reson. 77*, 274-293.

Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Go, N., & Wüthrich, K. (1987). Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol. 196*, 611-639.

Wagner, G., Pardi, A., & Wüthrich, K. (1983). Hydrogen bond length and ¹H NMR chemical shifts in proteins. *J. Am. Chem. Soc. 105*, 5948-5949.

Wider, G., Neri, D., Otting, G., & Wüthrich, K. (1989). A heteronuclear three-dimensional NMR experiment for measurements of small heteronuclear coupling constants in biological macromolecules. *J. Magnet. Reson. 85*, 426-431.

Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*. J. Wiley and Sons, New York.