

Internal and Overall Peptide Group Motion in Proteins: Molecular Dynamics Simulations for Lysozyme Compared with Results from X-ray and NMR Spectroscopy

Matthias Buck and Martin Karplus*

Contribution from the Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02139, and Laboratoire de Chimie Biophysique, Institut le Bel, Université Louis Pasteur, 4, rue Blaise Pascal, 67000 Strasbourg, France

Received April 23, 1999. Revised Manuscript Received July 19, 1999

Abstract: The motions of the main-chain N–H vector and of the atoms which define the peptide plane in proteins have been analyzed by use of a 1.6 ns molecular dynamics simulation of hen lysozyme in explicit solvent. By use of both local and molecular superposition of the peptide plane, fluctuations of the peptide group atoms relative to one another are distinguished from motions of the group as a whole. Distortions of the peptide plane arise from changes in bond geometry at the nitrogen and carbonyl carbon center, as well as from twisting around the C–N bond. Probability distributions for the distortions obtained from the protein average structure were compared with the instantaneous distortions sampled in the simulation. For the peptide group ω angle, there is approximate agreement between these two ways of determining the potentials of mean force, providing support for the widespread use of the former to obtain information about the latter. Good agreement is obtained also with potentials of mean force derived from experimental data on peptide plane distortions in peptide crystal and high-resolution protein structures. By contrast, for the N–H vector probability distribution, the potentials of mean force do not agree with those obtained from the average distortion. The N–H vector, which is of primary interest for NMR, has an average position that is nearly in the peptide plane and antiparallel to the C=O bond vector. The N–H bond vector undergoes rapid in-plane and out-of-plane fluctuations with average amplitudes of $4.7^\circ \pm 0.1^\circ$ and $7.4^\circ \pm 0.4^\circ$, respectively. Since the N–H reorientational fluctuations occur on a subpicosecond time scale, their contribution to ^{15}N relaxation can be described by a local order parameter S_{loc}^2 whose average is 0.931 ± 0.005 for 126 peptide planes in hen lysozyme. For 16 of the N–H bond vectors, the calculated S_{loc}^2 values are within 0.02 units of the overall N–H order parameters S^2 . For smaller values of S^2 , as found for the majority of the N–H vectors, the dominant contribution comes from overall peptide group motion. These results suggest that a renormalization of experimental order parameters should be used to extract the peptide group motions. The N–H bond-stretching motions are on the order of ± 0.024 Å, and the average bond lengths are almost uniform along the protein sequence. Thus, given the correct average value, the N–H bond-stretching vibrations have a negligible effect on the calculation of order parameters from the simulation. Peptide planes which are involved in secondary structures show reduced fluctuations. They can also exhibit motions that are correlated with dihedral angle fluctuations involving the surrounding heavy atoms and adjacent peptide planes.

I. Introduction

An essential structural element of proteins is the peptide unit, the set of atoms consisting of the main-chain carbonyl carbon, oxygen, nitrogen, and amide hydrogen, which link successive amino acids. Theoretical studies^{1,2} and experimental observations^{3,4} have shown that these atoms are on average in a near (trans) planar arrangement in peptides and proteins; planar cis peptide bonds also occur, most commonly for proline residues.⁵ Both global motions (i.e., motion of the peptide plane relative

to a fixed reference frame in the molecule) and distortions from the planar average structure are of interest. An understanding of the departures from planarity is important in the refinement of X-ray and NMR structures^{6–8} and for the modeling of main-chain motion from NMR relaxation data.^{9–11}

Although global peptide group motions are most important for ^{15}N and ^1H relaxation in NMR, a knowledge of the contribution of the internal dynamics (i.e., motion of the N–H bond vector relative to the peptide plane) is essential for a full description. It is the purpose of the present paper to use molecular dynamics simulation results to examine the average

* Address correspondence to this author at Harvard University. Phone: 617-495-4018. Fax: 617-496-3204. E-mail: marci@tammy.harvard.edu.

(1) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L.; Pople, J. A. 1991 *J. Am. Chem. Soc.* **1991**, *113*, 5989–5997.

(2) Guo, H.; Karplus, M. *J. Phys. Chem.* **1992**, *96*, 7273–7287.

(3) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1996**, *264*, 1180–1195.

(4) Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 6360–6368.

(5) Weiss, M. S.; Jabs, A.; Hilgenfeld, R. *Nat. Struct. Biol.* **1998**, *5*, 676.

(6) Karplus, P. A. *Protein Sci.* **1996**, *5*, 1406–1420.

(7) Wilson, K. S.; et al. *J. Mol. Biol.* **1998**, *276*, 417–436.

(8) Reif, B.; Hennig, M.; Griesinger, C. *Science* **1997**, *276*, 1230–1233.

(9) Fischer, M. W. F.; Zeng, L.; Pang, Y.; Hu, W.; Majumdar, A.; Zuiderweg, E. R. P. *J. Am. Chem. Soc.* **1997**, *119*, 12629–12642.

(10) Bremi, T.; Brüschweiler, R. *J. Am. Chem. Soc.* **1997**, *119*, 6672–6673.

(11) Daragan, V. A.; Mayo, K. H. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *31*, 63–105.

structure, the dynamic internal distortions, and the overall motions of the peptide group in the protein lysozyme for which some experimental data are available. MacArthur and Thornton³ surveyed the Cambridge structural database of small molecules and the database of protein structures and showed that the dihedral angle, ω_1 , defined by atoms $C\alpha_{i-1}-C-N-C\alpha_i$, departs from trans planarity (180°) with a standard deviation of up to 6° . The distortions arise from twisting about the peptide bond and from pyramidalization at the carbonyl carbon and amino nitrogen.¹² Hu and Bax⁴ studied the deviation between planes defined by atoms $C_{i-1}-N_i-C\alpha_i$ and $H^{N_i}-N_i-C\alpha_i$ for 45 residues in ubiquitin by use of up to six different three-bond heteronuclear NMR coupling constants. An analysis showed that the average deviation was less than 8° , particularly in hydrogen-bonded secondary structure. The analysis of one bond $^1J_{NH^N}$ coupling constants indicated that the average position of the amide proton is no more than 5° above or below a plane formed by atoms $C_{i-1}-N_i-C\alpha_i$.¹³

In addition to the average structure of the peptide unit, the dynamics of the amide hydrogen with respect to the peptide plane are of interest. The data used to define atomic positions of the polypeptide chain from X-ray or neutron diffraction and NMR data are averaged over the measurement time scales,^{14,15} which are much longer than the picosecond to nanosecond time scales that dominate the motions of most atoms. However, NMR relaxation can be affected by the fluctuations of atoms involved in the peptide unit on the short time scale. Heteronuclear relaxation, in particular of ^{13}C and ^{15}N atoms, is widely used to study the dynamics of polypeptides and proteins (e.g., see review by Dayie et al.¹⁶). To interpret the relaxation data, the Lipari-Szabo model¹⁷ is frequently employed. This formalism, which is based upon reorientational correlation times, separates the dynamic contributions into fluctuations arising from overall (global) tumbling and from protein internal motions.^{11,18} In most analyses, the contribution of dynamic distortion of the peptide group is not considered explicitly.^{9,19} This contrasts with the interpretation of fluorescence depolarization measurements of tryptophan in peptides and proteins, where a correction is generally made for the ultrafast (ps) relaxation due to internal fluctuations.^{20,21}

Both the average structure and the dynamics of the peptide unit are ideally suited for exploration by molecular dynamics simulation. Recently, we have analyzed a 1.6 ns simulation of hen lysozyme (Buck and Karplus, manuscript in preparation) and found a good correspondence with the experimental order parameters²² for main-chain amides, which undergo low-amplitude fast motions. In this report we use the same simulation to analyze the internal picosecond dynamics of the atoms in

the peptide unit and the overall motion of the peptide unit with reference to the molecular frame.

Because of the importance of the peptide group, a large number of computational studies concerned with its motional properties have appeared.⁹⁻¹¹ Recent interest in the peptide group conformation is stimulated by the possibility of obtaining more detailed information from newly developed NMR experiments.^{4,9,19,23}

This report first examines the average and dynamical distortions of the peptide planes (for example, the dihedral angle ω_1 ; $C\alpha_{i-1}-C-N-C\alpha_i$) and the motion of the N-H bond vector with respect to the peptide plane in the protein, hen egg white lysozyme, by use of molecular dynamics simulations with the CHARMM all-atom potential function.²⁴ The results are employed to test a commonly used method for estimating potentials of mean force or effective free energies from the average distortions observed in the database of peptide crystal structures. This is based on a direct comparison of the potentials obtained in this way from the simulation with those calculated from the dynamics itself. Significant deviations are found in some cases. The relationship between different measures of the fluctuations involving the peptide group heavy atoms and the amide hydrogen is investigated. We examine to what extent these local motions are coupled to longer-range motions, involving the surrounding main chain. The effect of the intra-peptide fluctuations, including N-H bond-stretching motions, on the NMR order parameters from the trajectories is determined. It is suggested that renormalization of the experimental order parameters to account for the fast internal fluctuations, in analogy to the procedure used in the analysis of tryptophan depolarization measurements, would be useful in interpreting the experimental data.

II. Methods

(a) **Simulation.** The crystal structure of hen lysozyme, refined to 2.0 Å by Handoll,²⁵ served as the starting point for the simulation. The program CHARMM²⁶ with the all-atom Param22 potential function²⁴ was used for the simulations. Details of the simulation are reported elsewhere (Buck and Karplus, manuscript in preparation). The protein was solvated by placing it at the center of a 30 Å sphere of water molecules, using a modified TIP3P model for the latter.^{27,28} The solvent was subjected to a stochastic boundary potential.²⁹ The simulation temperature was maintained at 300 K by the NOSE-HOOVER thermostat,^{30,31} which has been implemented in CHARMM (Watanabe and Karplus, unpublished results). The atom positions were propagated by the multiple time step (MTS) ds-RESPA method using the velocity Verlet algorithm and updating bond and angle, improper and dihedral, and all other forces at time steps of 0.5, 2, and 4 fs intervals respectively.^{32,33} Short- versus long-range force selection was employed with a cutoff of 6.0 Å³⁴ by Watanabe and Karplus (unpublished results).

(12) Fischer, S.; Dunbrack, R. L.; Karplus, M. *J. Am. Chem. Soc.* **1994**, *116*, 11931-11937.

(13) Tjandra, N.; Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6264-6272.

(14) Petsko, G. A.; Ringe, D. *Annu. Rev. Biophys. Bioeng.* **1984**, *13*, 331-371.

(15) Dobson, C. M.; Karplus, M. *Methods Enzymol.* **1986**, *131*, 362-389.

(16) Dayie, K. T.; Wagner, G.; Lefevre, J.-F. *Annu. Rev. Phys. Chem.* **1996**, *47*, 243-282.

(17) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546-4559.

(18) Palmer, A. G.; Williams, J.; McDermott, A. *J. Phys. Chem.* **1996**, *100*, 13293-13310.

(19) Lienin, S. F.; Bremi, T.; Brutscher, B.; Brüschweiler, R.; Ernst, R. *J. Am. Chem. Soc.* **1998**, *120*, 9870-9879.

(20) Ichiye, T.; Karplus, M. *Biochemistry* **1983**, *22*, 2884-2893.

(21) Hansen, J. E.; Rosenthal, S. J.; Fleming, G. R. *J. Phys. Chem.* **1992**, *96*, 3034-3040.

(22) Buck, M.; Boyd, J.; Redfield, C.; MacKenzie, D. A.; Jeenes, D. J.; Archer, D. B.; Dobson, C. M. *Biochemistry* **1995**, *34*, 4041-4055.

(23) Dayie, K. T.; Wagner, G. *J. Am. Chem. Soc.* **1997**, *119*, 7797-7806.

(24) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.

(25) Handoll, H. H. G. *Crystallographic studies of hen lysozyme*. D. Philos. Thesis, University of Oxford, 1985.

(26) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187-217.

(27) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.

(28) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902-1921.

(29) Brooks, C. L.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312-6325.

(30) Nose, S. *J. Chem. Phys.* **1984**, *81*, 511-519.

(31) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695-1697.

(32) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990-2001.

(33) Watanabe, M.; Karplus, M. *J. Phys. Chem.* **1995**, *99*, 5680-5697.

(34) Humphreys, D. D.; Friesner, R. A.; Berne, B. J. *J. Phys. Chem.* **1994**, *98*, 6885-6892.

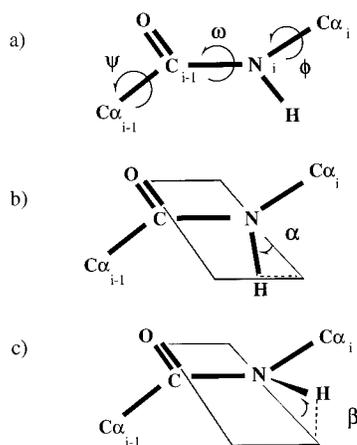


Figure 1. Schematic representation of the peptide group and adjacent C_{α} atoms: (a) dihedral angles; (b) in-plane projection angle α of the N-H vector; (c) out-of-plane projection angle β . The geometric construction of these planes and projections is described in the Methods section.

A force switched cutoff between 9.0 and 11.5 Å³⁵ and a constant dielectric of unity consistent with Param22 were used. The SHAKE algorithm³⁶ was applied to the TIP3P water hydrogens but not to the protein hydrogens, since it was deemed important that the protons, which play an essential role in ¹⁵N and ¹³C NMR relaxation, be treated in detail. This was achieved without additional computer time by use of the multiple time step algorithm. The simulation was run for 1.6 ns at 300 K after 14 ps of equilibration. Coordinates were saved every 0.2 ps for analysis. To examine motional events that occur on shorter time scales, we continued the trajectory for a further 12 ps with coordinate frames saved every 4 fs.

(b) Analysis. The dihedral angle ω , associated with the peptide bond $C_{i-1}-N_i$ (Figure 1a), can be defined in a number of ways. It was calculated here in four different ways: that is, as the angle between the vectors $C_{\alpha_{i-1}}-C_{i-1}$ and $N_i-C_{\alpha_i}$, $O_{i-1}-C_{i-1}$ and N_i-H_i , $C_{\alpha_{i-1}}-C_{i-1}$ and N_i-H_i , and $O_{i-1}-C_{i-1}$ and $N_i-C_{\alpha_i}$. These four angles are denoted as ω_1 to ω_4 , respectively; the most common definition corresponds to ω_1 . We note that the peptide planes are numbered by their nitrogen atoms (N_i) to allow straightforward comparison with the N-H order parameter (N_i-H_i). A geometrical analysis of N-H bond fluctuations in the peptide plane and perpendicular to it was carried out in terms of two angles (α , β) which are defined by the cross product and dot product of normalized bond vectors (Figure 1b,c): The angle α in the peptide plane is defined by

$$[(C=O) \times (C-N)] \times (C=O) \cdot (N-H) = \cos(90 - \alpha) \quad (1a)$$

and the angle β , which measures displacement out of the peptide plane, is defined by

$$[(C=O) \times (C-N)] \cdot (N-H) = \cos(90 - \beta) \quad (1b)$$

The C, N, and O atoms were chosen to define the plane to avoid involving C_{α} atoms that would include distortions of non-peptide unit geometries, such as variations in the angle $C-N-C_{\alpha}$ (see comparison of α , β with dihedral angles below). The $O=C$ bond vector was chosen as the reference for the in-plane angle, α , since the N-H vector is nearly parallel to it.

The motional behavior of 5 peptide planes that are representative of the different structural and dynamic environments of the protein were examined in detail (Figure 2). Cys30 is at the center of the B-helix, is deeply buried in the protein, and is hydrogen bonded throughout the simulation; Ser36 is at the C-terminal end of the B-helix and its amide hydrogen is bonded to Ala32; Tyr53 is at the center of the second

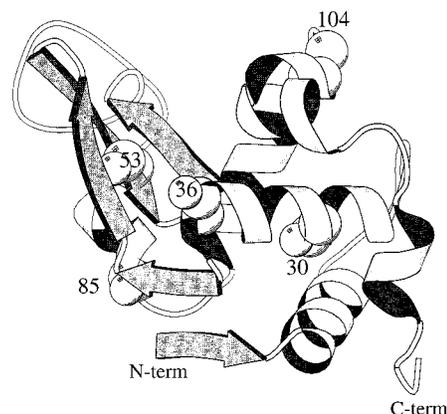


Figure 2. Schematic representation of the fold of hen lysozyme using Molscrip.⁷¹ The locations of 5 peptide planes that are representative of the different structural and dynamic environments in the protein are indicated (see Method and Table 1).

β -strand of the antiparallel β -sheet and is hydrogen bonded throughout the simulation by both its amide and carbonyl oxygen to Ile58 in the third β -strand; Ser85 is at the end of a 3_{10} helical segment and is rarely hydrogen bonded to this main-chain segment, although it is not consistently solvated by hydrogen bonding to water; Gly104 is located in a short, highly mobile loop and its amide interacts intermittently with main-chain acceptor atoms of Ile98 and Val99, as well as with solvent molecules. In addition a representative set of 20 peptide planes was used in some of the analysis; it consists of residues 4, 18, 21, 23, 30, 33, 36, 49, 53, 55, 59, 71, 78, 85, 97, 99, 100, 102, 104, and 112.

The distributions which are sampled in the simulation were used to calculate a potential of mean force or effective free energy.³⁷ Assuming an equilibrium distribution we have for the free energy change, $\Delta G(\Delta\omega)$,

$$\Delta G(\Delta\omega) = -RT \ln[p_i(\Delta\omega)] \quad (2)$$

where $p_i(\Delta\omega)$ is the relative probability of the distortion $\Delta\omega$ of an angle ω from its average value $\langle\omega\rangle$; R is the gas constant and T is the absolute temperature; and ΔG and RT are in kcal/mol. A temperature of 300 K was used in the simulation. The value of $p(0)$ is normalized to 1.³⁸ The quantity $p_i(\Delta\omega)$ was obtained from the molecular dynamics simulation by sampling $\Delta\omega$ every picosecond over the 1.6 ns trajectory and by plotting the distribution as a histogram with a bin size of 2°. If the fluctuations are assumed to be harmonic, an effective force constant (K_{ω}), at the temperature T , can be defined by

$$\Delta G(\Delta\omega) = \frac{1}{2}K_{\omega}\Delta\omega^2 \quad (3)$$

with

$$\Delta\omega = \sqrt{(RT/K_{\omega})} \quad (4)$$

The ¹⁵N-¹H order parameters, S^2 , were calculated from the final 1.6 ns of the trajectory as the plateau value of the autocorrelation function¹⁷

$$S^2 = \lim_{t \rightarrow \infty} C_2(t) \quad (5a)$$

where

$$C_2(t) = A \langle P_2[\mu(\tau)\mu(\tau+t)]/r^3(\tau)r^3(\tau+t) \rangle \quad (5b)$$

Here A is a constant such that $C_2(0) = 1$; P_2 is the second Legendre polynomial, $P_2[x] = \frac{1}{2}(3x^2 - 1)$; and $r(\tau)$ is the N-H bond length at time τ . The angle brackets $\langle \rangle$ represent the average over the trajectory. The unit vectors $\mu(\tau)$ and $\mu(\tau+t)$ describe the orientation of the N-H

(35) Loncharich, R. J.; Brooks, B. R. *Proteins: Struct. Funct., Genet.* **1989**, *6*, 32-45.

(36) Rychaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1979**, *23*, 327-341.

(37) McQuarrie, D. A. *Statistical mechanics*; Harper Collins Pub.: New York, 1976.

(38) McCammon, J. A.; Wolynes, P. G.; Karplus, M. *Biochemistry* **1979**, *18*, 927-941.

vector at time τ and $\tau + t$ in relation to a fixed reference frame. To construct this frame, we removed global translational and rotational motions from the trajectory by a root-mean-square optimized superposition of all the main-chain atoms (N, C, and C α) on the first coordinate frame.³⁹ Thus, S^2 reflects the internal motion of the protein. In addition, a local order parameter, S^2_{loc} was defined analogously for the N–H vector relative to the plane defined by atoms C $_{i-1}$, O $_{i-1}$, and N $_i$ of the peptide unit. An alternative reference frame for the N–H vector is based on atoms C α_{i-1} , C $_{i-1}$, and N $_i$ defining S^2_{loc} .^{28,40} For the calculation of S^2_{loc} and S^2_{loc} , the structures were superimposed on the corresponding three atoms of each peptide group.

The linear correlation coefficient (R) between two datasets x_i and y_i is given by

$$R = \frac{\sum (x_i - \langle x \rangle)(y_i - \langle y \rangle)}{(\sum (x_i - \langle x \rangle)^2)^{1/2} (\sum (y_i - \langle y \rangle)^2)^{1/2}}$$

where $\langle x \rangle$ and $\langle y \rangle$ are the average values of the datasets.

III. Results and Discussion

(a) Peptide Group Distortion and the Intrapeptide Group N–H Motion. Prior to analyzing the dynamics of the peptide unit, we examine its mean (equilibrium) distortion in crystal structures and in the simulation average structure. Figure 3a [i] shows the torsion about the C–N bond as defined by the dihedral angle, ω_1 (C α_{i-1} –C $_{i-1}$ –N $_i$ –C α_i), for each residue in the tetragonal crystal structure of hen lysozyme used at the start of the calculations (see Methods). The distortions from planarity are very small for the majority of residues (standard deviation of $\pm 1.9^\circ$), reflecting the use of torsion angle restraints in the refinement protocol.²⁵ Recent tetragonal and triclinic structures of hen lysozyme (pdb accession codes 2lym and 3lzt, respectively) have been refined without such restraints, and the deviations in ω_1 are $\pm 5.3^\circ$ and $\pm 6.6^\circ$ respectively.^{7,41} Although the distortions in the Handoll structure are very small, comparison with the pattern of the distortion in the 2lym and 3lzt crystal structures yields correlation coefficients of 0.80 and 0.62, respectively. The mean ω_1 value of all residues in the simulation average structure is 179.2° for the entire protein. The average distortion of peptide planes follows approximately a normal distribution with a standard deviation of $\pm 5.6^\circ$. The distortions in the average simulation structure are very close (within an rms difference of 0.24°) to the mean values of ω_1 obtained by determining them in picosecond intervals and averaging. This emphasizes that there is a symmetric distribution of ω_1 about the average structure value (see below). The correlations of ω_1 between the distortions in the simulation average structure and those in the 3lzt and 2lym structures are 0.59 and 0.63, respectively. This is to be compared with the correlation coefficient of 0.57 between ω_1 values in the 2lym and 3lzt structures.

The spread of ω_1 values in the simulation is also similar to that found in the database of 287 small linear peptide crystal structures of the Cambridge Structure Database for which MacArthur and Thornton³ reported a mean of 178.8° and a standard deviation of $\pm 5.6^\circ$. A recent analysis⁷ of eight protein structures with resolutions better than 1.2 \AA gives a value of $179.0^\circ \pm 5.6^\circ$. Thus, there is good agreement in the average deviations of ω_1 between the experimental structures and the structure after it has been subject to simulation with the CHARMM potential function. In both the triclinic crystal structure of hen lysozyme (3lzt) and the simulation some of

the largest distortions occur at or near the ends of regular secondary structure (e.g., Asn37, Trp63, and Arg114) or in coil regions (e.g. Arg21 and Ile78) shown in Figure 3a [ii]. However, there are some large distortions elsewhere in the structure, for example, for Asp52, which is located in a β -strand.

Figure 3b [ii] shows the rms fluctuation in ω_1 of each residue during the simulation. Summed over the protein, the average rms fluctuation is 7.0° with a standard deviation of $\pm 1.1^\circ$. There is no correlation between the average peptide plane distortion and the magnitude of the fluctuations, though the magnitude of the overall rms fluctuation and the standard deviation from the mean distortion are similar (see Section b). The magnitude of the fluctuations are reduced for the majority of residues that are located in regular secondary structure. The correlation coefficient between the average time the main-chain amide is hydrogen bonded to protein acceptors or water, and the average fluctuation in ω_1 is -0.41 (using a geometric cutoff criterion for hydrogen bonding of 2.5 \AA for the hydrogen donor and acceptor distance and 90° for the hydrogen bond angle). Average distortions from planarity are less in central parts of the α -helical segments (Figure 3b [i]). These trends are in agreement with the experimental observations made from X-ray crystallographic data³ and from measurements of NMR coupling constants.⁴ Hydrogen-bonded secondary structure has been shown to increase the double bond character of the C–N bond,² which reduces the expected deviation of the peptide group from planarity.^{3,42} Peptide units which undergo the largest fluctuations in ω_1 are frequently located at or near Gly residues (Gly4, 16, 49, 102, 104), though large fluctuations also occur for peptide planes of Pro70 and Asp119 (71 and 117 are Glycines).

The torsional distortion and fluctuations involving the amide hydrogen are of particular interest, primarily because of their role in NMR relaxation (see Section d below). They also contribute to the vibrational spectra (amide bands II and III) of peptides and proteins.⁴³ Since angles ω_2 and ω_3 depend also on the amide hydrogen position, unlike ω_1 , they cannot be obtained from X-ray structures of proteins, in which the positions of the hydrogens are not determined; in principle, high-resolution neutron structures can be used for the proton positions but relatively few of these are available. The hydrogens atoms were built into the lysozyme structure using standard methods⁴⁴ and were then employed in the dynamics calculations. The results of the simulation are used to determine the different measures of peptide group planarity and to determine the relationships among them. Correlation between changes in ω_1 and the alternative measures ω_2 , ω_3 , and ω_4 of peptide plane distortion are low in all cases (the magnitude of the correlation coefficient R is less than 0.15), demonstrating that, over the 1.6 ns time-scale, the amide hydrogen H N_i and carbonyl oxygen O $_{i-1}$ fluctuate largely independently of the main-chain C α atoms when the intrapeptide group motions are considered. There is a modest correlation between ω_3 and ω_4 ($|R| = 0.1-0.4$), suggesting a slight relationship between the fluctuations of the H N_i and O $_{i-1}$ atoms as a result of peptide unit distortions that are mediated through bond angles around N $_i$ (pyramidalization, see below). Measures which include the H N_i atom (ω_2 and ω_3) are highly correlated ($R = 0.7-0.9$) since the fluctuations of the hydrogen atom dominate the motions.

To examine more specifically the fluctuation of the N–H bond vector that is involved in NMR relaxation of the ^{15}N atom, we introduce the out-of-plane (β) and in-plane (α) displacement

(39) Van Gunsteren, W. F.; Karplus, M. *Biochemistry* **1982**, *21*, 2259–2274.

(40) Fushman, D.; Ohlenschläger, O.; Rüterjans, H. *J. Biomol. Struct. Dyn.* **1994**, *11*, 1377–1402.

(41) Kundrot, C. E.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 157–170.

(42) Milner-White, E. J. *Protein Sci.* **1997**, *6*, 2477–2482.

(43) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181–364.

(44) Brünger, A. T.; Karplus, M. *Proteins: Struct. Funct., Genet.* **1988**, *4*, 148–156.

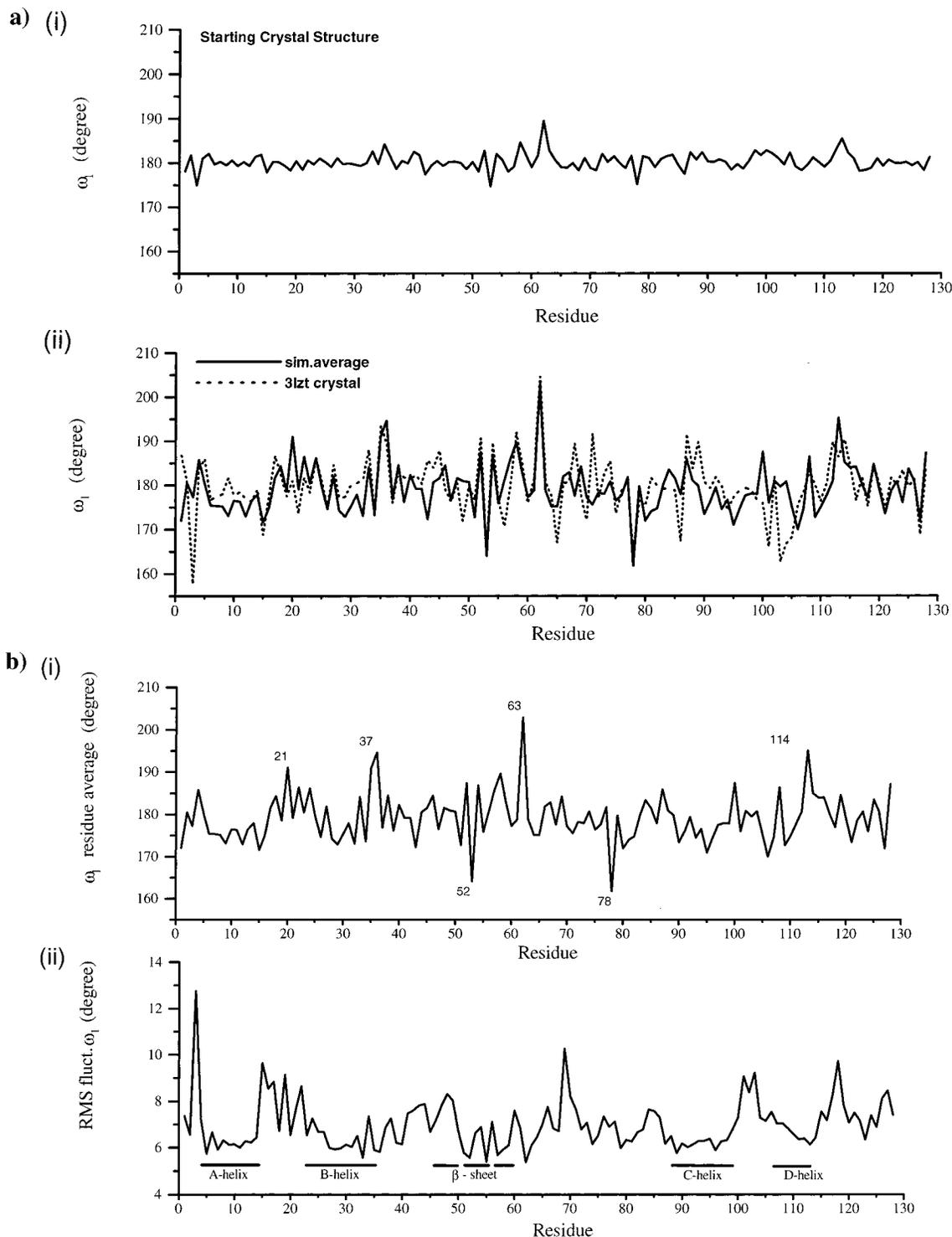


Figure 3. Dihedral angle ω_1 ($C\alpha_{i-1}-C_{i-1}-N_i-C\alpha_i$) (a) (i) the starting crystal structure and (ii) the simulation average structure (thick line) and a triclinic lysozyme structure (3lzt)⁷ (dotted line); (b) (i) mean angle ω_1 from the simulation average structure and (ii) the standard deviation from that average showing the rms fluctuation in ω_1 over the course of the simulation as a function of protein sequence. The location of the four α -helices and the strands of the main β -sheet of hen lysozyme is indicated.

angles (see Figure 1b,c and Methods). Figure 4a shows the mean (equilibrium) values of α and β for the 126 peptide N-H bonds. The mean value for α is $1.23^\circ \pm 1.10^\circ$, and that for β is $0.47^\circ \pm 2.34^\circ$. Thus the average position of the N-H vector is in a plane defined by atoms N_i , C_{i-1} , and O_{i-1} , and it is nearly antiparallel to the $C_{i-1}=O_{i-1}$ vector. The ranges in these average distortions are much less than that of ω_1 , which is $179.2^\circ \pm 5.7^\circ$. The mean distortions of α and β are not correlated for the individual peptide planes.

The mean deviation of the N-H vector from an in-plane and C=O antiparallel orientation is small, as seen above. However, there are larger fluctuations in the position of the amide hydrogen (angles α and β) over the course of the trajectory for each of the peptide units, and the distributions of both angles are nearly Gaussian (see section b). The mean value of the rms fluctuations of 126 N-H vectors are $4.7^\circ \pm 0.1^\circ$ in α and $7.4^\circ \pm 0.4^\circ$ in β . A relationship exists between the extent of in- and out-of-plane motions (rms fluctuations with a correlation

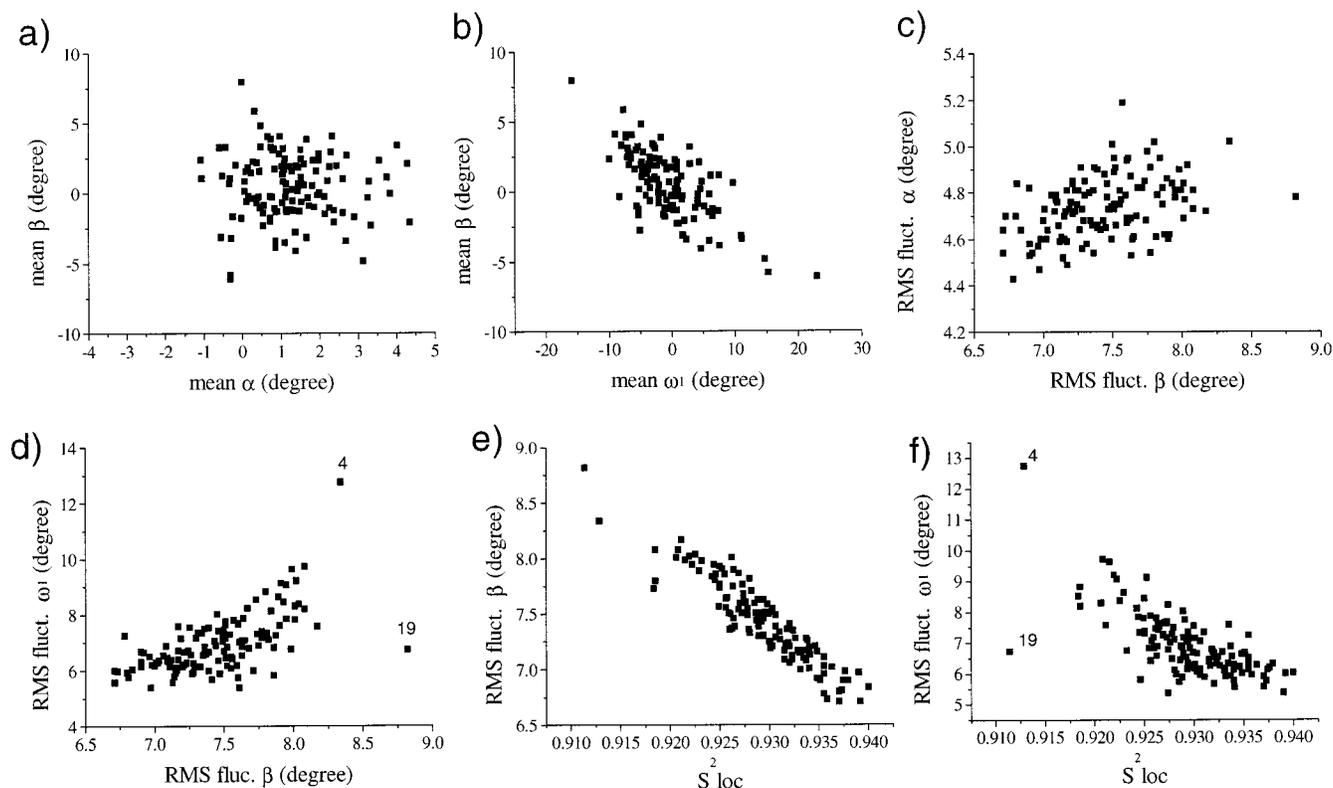


Figure 4. Correlations of peptide plane angles: (a,b) correlation of equilibrium peptide plane distortions for all 126 peptide planes; (c,d) correlation of rms fluctuations; (e,f) correlations of the order parameter for motion internal to the peptide plane, S_{loc}^2 , and the rms fluctuations in β and ω_1 , respectively. Gly4 and Asn19, which deviate from the general behavior, are labeled in d and f.

Table 1. Order Parameters and Geometrical Measures ($\mu \pm \sigma$, Mean \pm rms Fluct. in Degree) for Representative Peptide Planes^a

residue	Cys30	Ser36	Tyr53	Ser85	Gly104
S^2	0.92	0.87	0.91	0.67	0.42
S_{loc}^2	0.94	0.93	0.94	0.93	0.92
angles					
in-plane, α	-1.1 ± 4.7	0.3 ± 4.9	-1.5 ± 4.6	-0.7 ± 5.0	-1.3 ± 4.8
out-of-plane, β	2.5 ± 6.7	-3.2 ± 7.3	1.2 ± 6.7	-2.0 ± 7.5	-1.2 ± 8.0
ω_1	172.7 ± 5.7	190.9 ± 7.6	187.3 ± 7.3	183.5 ± 7.6	179.6 ± 9.2
ω_2	177.1 ± 7.8	176.3 ± 8.5	189.3 ± 8.0	182.3 ± 8.7	181.4 ± 9.3
ω_3	-5.9 ± 7.7	5.2 ± 8.3	-11.3 ± 7.9	2.1 ± 8.5	1.2 ± 9.2
ω_4	-4.4 ± 6.6	9.2 ± 7.2	-14.2 ± 7.0	3.3 ± 8.2	-0.6 ± 9.2
force constants					
in-plane, α	89.3 ± 1.9	72.2 ± 3.6	90.0 ± 3.3	78.8 ± 2.3	84.0 ± 1.9
out-of-plane, β	42.0 ± 1.6	36.1 ± 1.2	40.7 ± 2.3	34.1 ± 1.7	29.5 ± 1.0
dihedral, ω_1	59.8 ± 1.3	48.6 ± 1.3	60.4 ± 2.0	31.5 ± 1.0	21.7 ± 0.7

^a Force constants are in units of kcal mol⁻¹ rad⁻².

coefficient of 0.39; Figure 4c) but, similarly to ω_1 , not between the mean displacement (values of the equilibrium position) and rms fluctuations. Although the structural context of the N-H vectors and the dynamics of the surrounding main-chain vary considerably (for example, in terms of hydrogen-bonding or packing interactions), the variations in the magnitude of the local rms fluctuations ($\pm 0.1^\circ$ and $\pm 0.4^\circ$ for α and β , respectively) involving the amide hydrogen are smaller than those of ω_1 ($\pm 1.1^\circ$).

It is of interest to determine whether the out-of-plane angle β is related to peptide plane distortions as measured by the angles ω_1 – ω_4 . Instantaneous values for angle ω_1 are generally uncorrelated with β over the course of the trajectory, as are the values of α for 5 peptide planes, which were examined in detail (see Table 1 and Figure 2). By contrast, the dihedral angle ω_2 shows near perfect correlation with the out-of-plane displacement of the hydrogen. This is not surprising since ω_2 is the angle between vectors connecting atoms $C_{i-1}=O$ and $H-N_i$.

The angle β also correlates well with ω_3 , the dihedral angle between $C_{i-1}-C\alpha_{i-1}$ and $H-N_i$; there is a weak correlation with ω_4 , the $C_{i-1}=O$ and $N_i-C\alpha_i$ dihedral, since $C_{i-1}=O$ is part of the peptide plane definition.

Instantaneous displacements of the N-H vector as measured by angles α and β of the individual peptide planes are not correlated with one another. A correlation of the motion could be achieved by a common distortion in the reference frame or if a certain range of α , β angular distortions were related to a particular structural feature, such as hydrogen bonding, for example. We find that the former distortion is small (e.g., $C_{i-1}-N_i-C\alpha_i$ bond angles fluctuate with an rms of 3.2°) and that there are no examples of the latter for the five peptide planes examined. Fourier transforms of the correlation functions for the dynamics of the α and β angles show the dominant in- and out-of-plane bending frequencies for the N-H group in the range 1205–1240 and 748–798 cm⁻¹, respectively, close to those obtained from a vibrational analysis of *N*-methylacetamide

with the CHARMM potential function;²⁴ the values are 1267 and 797 cm^{-1} respectively. Relaxation times of 0.5–2 ps were calculated for the displacements in α and β for the five residues examined in detail. The values of the relaxation times are not correlated with the amplitude of the motion.

Correlations of any of the ω angles with the in-plane angle α are small, as expected (given that all dihedral angles require some distortion perpendicular to the C–N axis). However, there is a good correlation between the mean distortion in β and ω_1 (correlation coefficient -0.71 , as shown in Figure 4b). The magnitude of the correlation of β with dihedral angle values involving the C α atoms is consistent with the comparisons of the correlation between the different ω angles over the course of the simulation. This makes clear that out-of-plane displacements, as defined by some ω , are not simply due to a rotation about the axis formed by the C–N bond with the elements of the peptide group (i.e., C α_{i-1} , O $_{i-1}$, C $_{i-1}$, N $_i$, and C α_i) moving as a rigid entity. The dominant deviation from rigidity is the motion of H N_i out of the plane defined by atoms O $_{i-1}$, C $_{i-1}$, and N $_i$. However, other motions also contribute. For example, the C α can move relative to the peptide plane by a change in the C $_{i-1}$ –N $_i$ –C α_i bond angle. The mean angle is 123.7° with a $\pm 1.1^\circ$ standard deviation for the 126 peptide planes. Observations of MacArthur and Thornton³ show that ω_1 and the deviation from ideal bond angles at atom centers (for example, the nitrogen pyramidalization angle θ) are closely correlated in the Cambridge Database of small molecule crystal structures (Figure 11 of their paper). Earlier Dunitz and Winkler⁴⁵ showed that both bond angle variation and plane twisting contribute to peptide plane distortion. Although the particular motions of the atoms may not be correlated for individual peptide groups (e.g., instantaneous values in ω_1 and β), the mean distortion and the rms fluctuation of ω_1 correlate well with the rms fluctuation of the H N_i atom out of the peptide plane (Figure 4, parts b and d, with correlation coefficients of -0.71 and 0.63 , respectively). Thus, the magnitudes of the fluctuations appear to be related through context-dependent effects, such as average geometries of neighboring atoms or hydrogen bonding.

Figure 5 shows a projection of the N–H vector for three representative peptide planes (Cys30, Ser85, and Gly104; see also Figure 2) onto a coordinate frame defined as parallel and perpendicular to a plane constructed from the average orientation of the N–H and C–N bonds over the course of the trajectory. The projections are made both relative to the average orientation of the vectors when both internal and global motions of the peptide plane are included and relative to the average vectors that result when only distortions of the individual peptide plane are considered. For residues that are located in regular secondary structure (e.g., Cys30), the fluctuations appear to originate largely from the displacement of the N–H bond vector relative to the peptide plane, with little additional contribution from other motions; that is, the two projections give essentially the same result. For the more flexible regions of the protein, this is not the case, as illustrated by Ser85 and Gly104; these are two of the most mobile peptide planes. Because the extent of fluctuation of the N–H vector relative to the peptide plane is similar to that of residues whose peptide plane is located in rigid regions of the protein (comparing with the right-hand side of Figure 5), the additional motion of the N–H bond vectors is due to the collective reorientation of the atoms defining the peptide plane (C $_{i-1}$, O $_{i-1}$, N $_i$). This is consistent with the observation that changes in the dihedral angles ϕ and ψ of the adjacent main-

chain groups are found not to be correlated with the in- and out-of-plane motion of the N–H vector over the course of the trajectory (see Section a). The N–H vector fluctuations relative to the peptide plane are therefore relatively independent of their dynamical and structural context.

Figure 5 also shows the fluctuations as a function of time over two separate segments of 50 ps of the simulation (linked points at successive 1 ps intervals). The results demonstrate that the fluctuations are diffusion-like (erratic both in direction and in amplitude). They sample the region corresponding to the N–H vector motion relative to the peptide plane on the 50 ps time scale (Figure 5c [iv]). The relaxation times for these local motions are on the picosecond time scale (see section d). However the reorientation of the peptide plane, involved in sampling the two regions shown in Figure 5c [iii] is on a much slower time scale and does not really correspond to diffusion in a cone. The importance of this aspect of the N–H vector motion for NMR parameters is presented below.

Several other molecular dynamics studies have been concerned with the description of the motion of the N–H vector. One such report⁴⁶ used global superposition on the trajectory average structure. Unless the fluctuations were substantial, involving transitions between main-chain dihedral angles, the displacements could be interpreted in terms of a model in which N–H vectors undergo motions which are confined to a cone of elliptical shape. A number of different projections of the N–H vector were considered in an analysis of a molecular dynamics trajectory of Ribonuclease T1 calculated with GROMOS.⁴⁰ These included a local superposition on individual peptide units involving atoms C α_i , N $_i$, and C $_{i-1}$. The amplitudes of in-plane and out-of-plane displacements, as well as the magnitude of an order parameter, that were reported for Ribonuclease T1 are in good agreement with those derived for hen lysozyme in the present study. This is of interest since different molecular mechanics potentials were used in the two studies.

(b) Energetics of Intrapeptide Group NH Motion and Peptide Plane Distortion. As shown above, displacements of the peptide unit N–H bond vector from its equilibrium position originate from several types of distortions involving dihedral angles as well as bond angles. The peptide group distortions are controlled in the CHARMM22 energy function by a number of different bond angle, dihedral angle, and improper dihedral angle parameters; the force constants involved are given in MacKerell et al.²⁴ The bonding terms for the peptide group are identical for all residues so that differences in average angles and dynamics are due to the nonbonded van der Waals and electrostatic nonbonded interactions. Since there is considerable variation in the peptide group average structure and dynamics, the nonbonded interactions with the rest of the protein make a significant contribution to the effective peptide group potential energy. To determine the overall contribution of the various terms to the dynamics, we evaluated the potential-of-mean-force for distortions as measured by angles α , β , and ω_1 for a representative set of 20 peptide groups as well as for the average of all peptide plane fluctuations (see Methods).

The effective potentials for α , β , and ω_1 of residue 85 are shown in Figure 6b,c. The results are well fit to a quadratic function, demonstrating that the potential of mean force is essentially harmonic over the accessible range. Corresponding harmonic fits are applicable to the other residues. The force constants obtained from the quadratic relationship (eq 3 in Methods) are listed for 5 residues in Table 1. The values for

(45) Dunitz, J. D.; Winkler, F. K. *Acta Crystallogr., Sect. B* **1975**, *31*, 251–257.

(46) Yamasaki, K.; Saito, M.; Oobatake, M.; Kanaya, S. *Biochemistry* **1995**, *34*, 6587–6601.

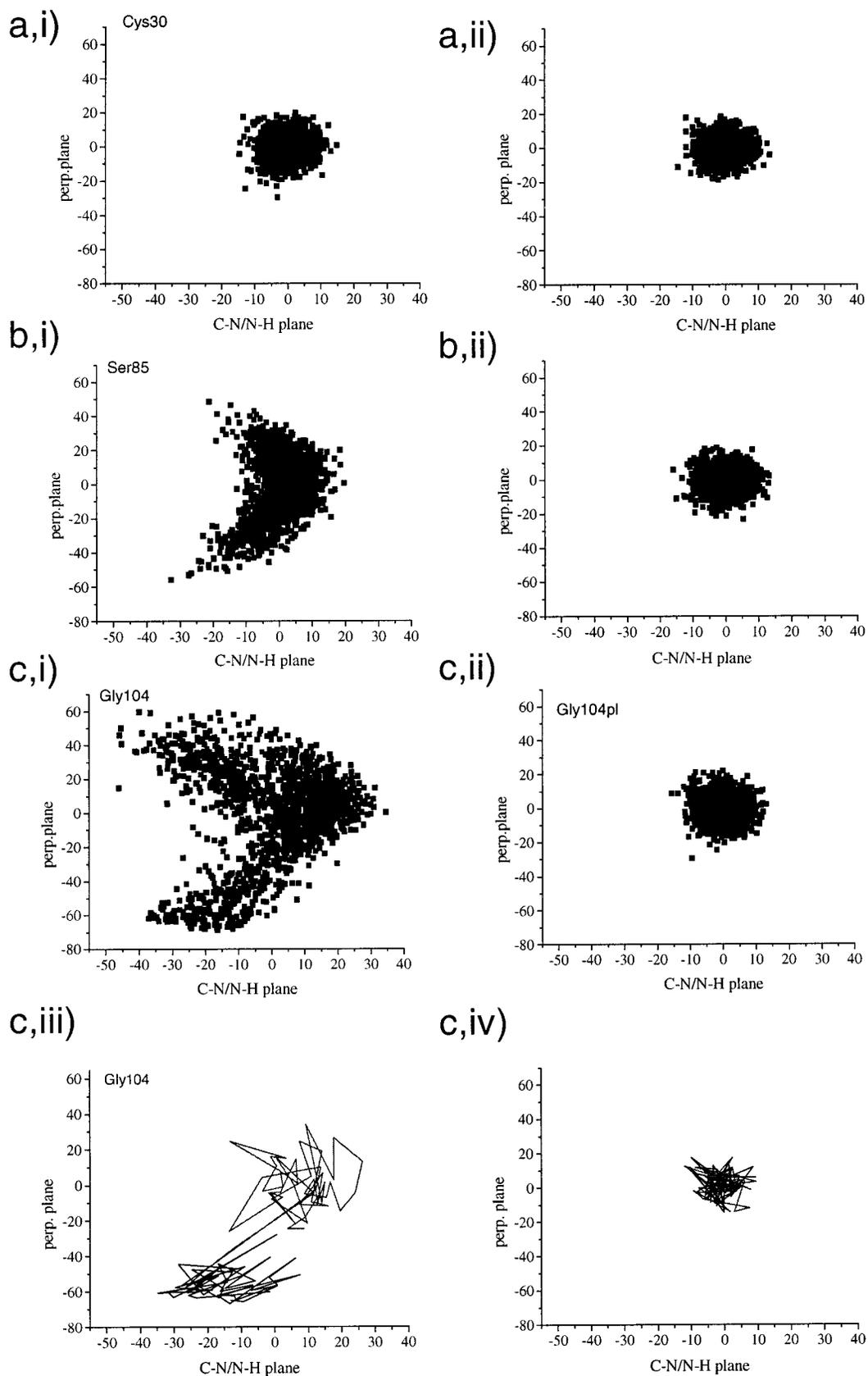


Figure 5. N–H vector projections (see text). (a–c) Projection of the N–H vector for Cys30, Ser85, and Gly104 onto a coordinate frame defined by the average orientation of the N–H vector (as z-axis, into the paper) and a vector that is perpendicular to it and lies in the plane of atoms N, H, and C in the direction of the average N–C bond vector (x-axis to the left). The y-axis is defined as perpendicular to these two average vectors. The simulation average and instantaneous orientation of the N–C vector is used to construct projections of the motion (i) relative to the entire main chain of the protein and (ii) relative to the individual peptide plane, respectively. The axes are angles in units of degree. c,iii and c,iv show motions of the Gly104 N–H vector relative to the same coordinate frames, taken from two 50 ps segments of the simulation (100–150 and 800–850 ps in black and gray, respectively).

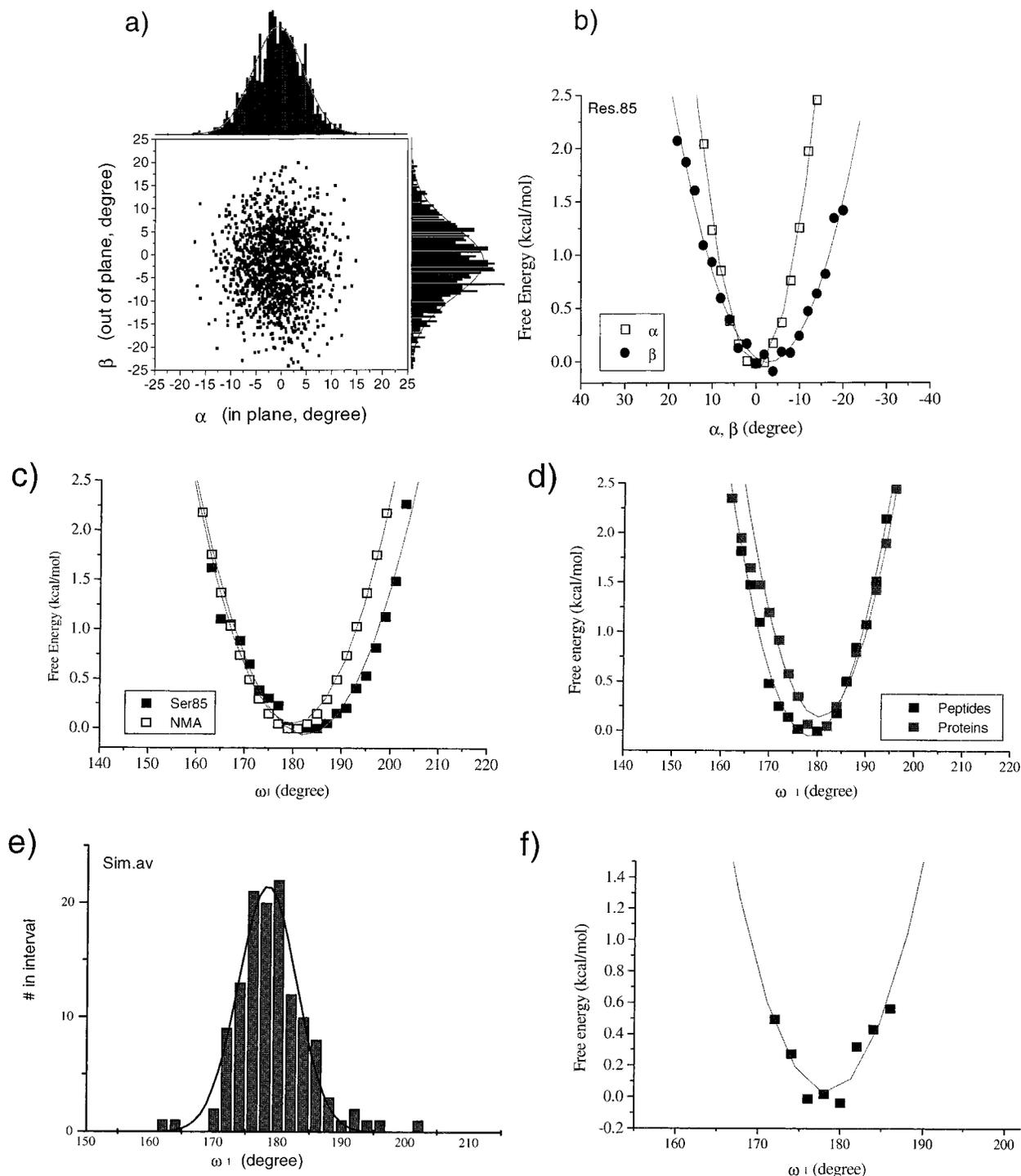


Figure 6. Probability distributions and potentials of mean force. (a) In-plane and out-of-plane displacement angles (α and β , respectively) of the N–H vector sampled every 0.1 ps for residue Ser85. The equilibrium position is close to the origin ($0^\circ, 0^\circ$), a vector in the plane and antiparallel to the C=O bond (see Figure 1). Histograms of the probability distribution of α and β are also shown. (b) Effective potential of mean force derived from the statistical distribution of the in-plane and out-of-plane displacement angles, α and β , for residue 85 fit with a harmonic function. (c) Torsion angle ω_1 of Ser85 together with the energy function of ω_1 in NMA, derived by Kuchnir and Karplus (unpublished), fit with a harmonic function. Force constants corresponding to these functions are given in Table 1. (d) Data taken from the paper by the EU 3-D structure validation network⁷ and fits of the data to a harmonic function as described in the text. (e) Distribution of peptide plane average distortions (data from Figure 3b [i]) in the simulation, fit to a normal distribution. (f) Potential of mean force derived from the data in panel e, fit with a harmonic function.

the 20 representative residues range from 72 to 96 kcal mol⁻¹ rad⁻² for α , from 27 to 42 kcal mol⁻¹ rad⁻² for β , and from 15 to 60 kcal mol⁻¹ rad⁻² for ω_1 . The average fluctuations of all peptide planes yield force constants equal to 89, 38, and 42 kcal mol⁻¹ rad⁻² for α , β , and ω_1 , respectively. In correspondence with Figure 4d, the force constants for ω_1 and β are similar for many of the residues (e.g., Ser85), but they are

greater for displacements in ω_1 than in β for the peptide planes, which are hydrogen bonded in secondary structure throughout the simulation (e.g., Cys30 and Tyr53). The force constants for ω_1 are smaller, as compared with β , for many of the residues that undergo large rms fluctuations (e.g., Gly104 in Table 1). This is consistent with the fact that the range of rms fluctuations is greater for ω_1 than β (Figure 4d).

The results found here for lysozyme can be compared with the potential energy for deviations of the peptide unit from planarity for an isolated *N*-methylacetamide (NMA) molecule. The results for the ω_1 angle using adiabatic mapping with the CHARMM potential function Param22 are shown in Figure 6c (Kuchnir and Karplus, unpublished results) together with those for the ω_1 angle of residue Ser85. The functional form for NMA is harmonic for small angular displacements and corresponds to an effective force constant of 40 kcal mol⁻¹ rad⁻². This is close to the mean force constant for ω_1 for all residues in the lysozyme simulation (42 ± 11 kcal mol⁻¹ rad⁻²). The large range of the protein force constants confirms that effects from the surrounding atoms have an important influence on the dynamics of peptide plane distortion, as mentioned above.

The simulation results for the potential of mean force for the ω_1 fluctuations of the peptide group can be compared with the distribution of mean distortions of all peptide groups in hen lysozyme from the simulation; the latter are shown in Figure 3b [i] and as a histogram in Figure 6e. Although the data are somewhat sparse, the potential of mean force is approximately harmonic (Figure 6f) and can be fit with a force constant of 71 ± 14 kcal mol⁻¹ rad⁻². This predicts an average distortion of 5.3° at 300 K, close to the mean peptide group distortion of ±5.6° that is obtained over the course of the simulation. Since the experimental values for the two lysozyme structures (2lym and 3lzt) are similar (±5.3° and ±6.6°, respectively), the simulation results for the average distortion are in good agreement with the experiment, indicating that the simulations are meaningful. However, the important point is that the simulation results permit a self-consistent comparison between the actual potentials of mean force, as obtained from the dynamics, and those estimated from the average structure from the same simulation. The former value is ±7.0° for the mean rms fluctuation, compared with the latter value of ±5.6°. Thus the simulation and "database"-derived force constants, although similar, are not the same for ω_1 . There is a priori no reason they should be. In fact, a more extreme discrepancy exists between the rms fluctuations of α and β (4.7° and 7.4°, respectively) and the distribution of mean N-H vector displacements from the simulation. The equilibrium distortions for α and β are only 1.1° and 2.3°, respectively.

The above comparison is of great interest because of the widely used methodology pioneered by Dunitz and co-workers for estimating potentials of mean force from distributions observed in crystal structures.^{45,47} For the peptide group distortion angle ω_1 , such analyses have been made by Thornton and co-workers.³ Figure 6d shows the quadratic fit to the observed ω_1 angle for 492 peptide bonds in the Cambridge Structure Database. A force constant of 60 ± 2 kcal mol⁻¹ rad⁻² is obtained. As the majority of the peptides are hydrogen-bonded, this value can be compared with the force constants obtained for the peptide planes in the most regular secondary structures of lysozyme; for example, in Table 1, the peptide planes of Cys30 and Tyr53 have a force constants for ω_1 with values close to 60 kcal mol⁻¹ rad⁻², as do peptide planes of Asn59, Lys97, and Arg112, which are also located in regular secondary structure.

The potential of mean force derived from the distribution of 8 high-resolution protein structures,⁷ shown in Figure 6d, does not fit well to a quadratic function implying that the potential of mean force is not harmonic. This suggests that additional

data are needed, as well as careful consideration of the refinement protocols and their effects on the "observed" distortions; that is, even if no restraints on torsion angles are included, as in the 8 high-resolution refinements using SHELXL,⁴⁸ restraints on bond angles may lead to a distortion of the normal distribution, as is proposed by several papers.^{3,7,49}

The present self-consistent analysis of the determination of a potential of mean force is also relevant to the use of structural databases to derive, for example, probabilities of distances between residues or atoms (e.g., Miyazawa and Jernigan,⁵⁰ and Sippl⁵¹). The suggestion that a special folding temperature should be introduced⁵² is not applicable to our statistical analyses, where a single simulation is used, though it may be for the more general statistical analyses of the type mentioned here. However, it should be noted that, even though statistical potentials of mean force for residues (or atoms) may be meaningful within the given context (i.e., probabilities arising in the presence of the polypeptide backbone), adding them together on the assumption that they are independent is certainly not, since considerable overcounting would be involved. Some of these points have been discussed recently.⁵³⁻⁵⁶

(c) Correlations of Peptide Group Distortion with the Motion of Surrounding Atoms. By contrast to the lack of correlations between the N-H vector orientation and neighboring ϕ , ψ dihedral angle motions, correlations between peptide plane angle ω_1 , and main-chain dihedral angle fluctuations are significant. In fact, the correlation between $\omega_{1,i}$ and ϕ_i can approach in its extent that of the well-known and frequently found anticorrelation between ψ_{i-1} and ϕ_i across the peptide plane.⁵⁷ Data for two segments of secondary structure are shown in Figure 7a,b and Table 2. The angles 56 ψ and 57 ϕ are anticorrelated (correlation coefficient of -0.83), as are 56 ω_1 and 56 ϕ (correlation coefficient of -0.27). The fluctuations of ω_i and ϕ_i are also correlated significantly in helical (e.g., res. 94-99) and in turn regions (res. 55-56) but appear not to be correlated in extended regions of β -strands.

Correlation and anticorrelation of the motions of neighboring peptide planes occur in some cases. Correlation coefficients of 0.2 to -0.3 are found for 94 ω_1 and 95 ω_1 , 96 ω_1 and 97 ω_1 , as well as 57 ω_1 and 58 ω_1 , for example (not shown). No significant correlations are detected for peptide planes with separations greater than one residue, including cases in which persistent hydrogen bonding exists between them (for example residue 29 and 33 in the B-helix). Long-range correlations of motions in ϕ and ψ are low for the residues that are located in the most ordered region of the protein, in part because the fluctuations in these dihedrals is small. However, for more flexible regions of the protein, long-range correlations exist (Buck and Karplus in preparation). They have been characterized in several other protein studies.⁵⁸ These findings are in general agreement with

(48) Sheldrick, G. M.; Schneider, T. R. *Methods Enzymol.* **1997**, 277, 319-343.

(49) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, 286, 291-304.

(50) Miyazawa, S.; Jernigan, R. L. *Macromolecules* **1985**, 8, 534-552

(51) Sippl, M. J. *Comput.-Aided Des.* **1993**, 7, 473-501.

(52) Finkelstein, A. V.; Badretinov, A. Y.; Gutin, A. M. *Proteins: Struct. Funct., Genet.* **1995**, 23, 142-150.

(53) Thomas, D.; Dill, K. A. *J. Mol. Biol.* **1996**, 257, 457-465.

(54) Ben-Naim, A. *J. Chem. Phys.* **1997**, 107, 3698-3706.

(55) Zhang; Skolnick, J. *Protein Sci.* **1998**, 7, 112-122.

(56) Miyazawa, S.; Jernigan, R. L. *Proteins: Struct. Funct., Genet.* **1999**, 34, 49-68.

(57) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, 267, 585-590.

(58) Swaminathan, S.; Ichiye, T.; van Gunsteren, W.; Karplus, M. *Biochemistry* **1982**, 21, 5230-5241.

(47) Dunitz, J. D. In *X-ray analysis and the structure of organic molecules*; Cornell Univ. Press: New York, 1979; pp 328-337.

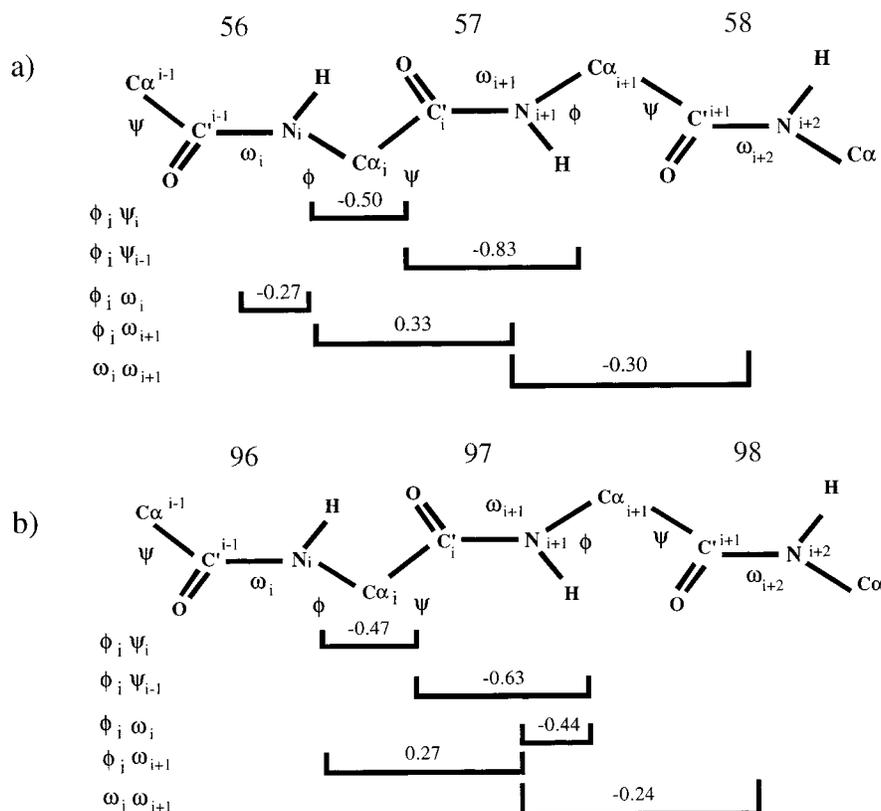


Figure 7. Correlations between dihedral angle measures involving main-chain heavy atoms for two representative segments of the protein: (a) residues 56–58 located in a turn of the β -sheet and (b) residues 96–98 located in the center of the C-helix. Significant equal time correlations of the fluctuations over the 1.6 ns simulation are indicated with the angles involved.

Table 2. Correlation between Dihedral Angles in Selected Regions of the Protein (Figure 7)

residue and dihedral angle	average and rms fluct. of angle	ψ_{i-1} vs ϕ_i [ϕ_i vs ψ_i]	ϕ_i vs $\omega_{1,i}$ [ψ_i vs $\omega_{1,i+1}$]	ϕ_i vs $\omega_{1,i+1}$ [ψ_i vs $\omega_{1,i}$]
(a) Turn between β -Sheet Strands 2 and 3				
56 ω_1	175.7 \pm 5.2	-0.49		
56 ϕ	-100.1 \pm 9.4	[-0.50]	-0.27	0.33
56 ψ	1.7 \pm 13.3		[-0.31]	[-0.08]
57 ω_1	180.9 \pm 7.1	-0.83		
57 ϕ	58.0 \pm 12.5	[0.06]	-0.08	-0.07
57 ψ	75.1 \pm 10.9		[0.08]	[0.41]
58 ω_1	185.7 \pm 5.9	-0.81		
58 ϕ	-100.0 \pm 11.5	[-0.04]	-0.33	-0.27
(b) C-Helix: Refer to Figure 7				
96 ω_1	174.2 \pm 5.7	-0.63		
97 ϕ	-72.1 \pm 9.2	[-0.35]	-0.44	0.27
97 ψ	-41.8 \pm 9.2		[-0.12]	[0.21]
97 ω_1	177.3 \pm 6.2	-0.68		
98 ϕ	-60.5 \pm 9.3	[-0.51]	-0.45	0.29
98 ψ	-45.9 \pm 8.4		[-0.34]	[0.35]

the results of normal-mode analysis of helical peptides⁵⁹ and of proteins.⁶⁰

(d) Implications for NMR Order Parameters. Internal and collective motions of the peptide plane unit are of considerable importance for the interpretation of NMR order parameters and the development of models for the motions. Many models used in the interpretation of relaxation or coupling constant data make the assumption of a rigid peptide plane.^{9,11,61} The Gaussian Axial Fluctuation (GAF) model by Bremi and Brüschweiler¹⁰ does

not consider the motion of the N–H bond vector relative to the peptide group, as discussed in Lienin et al.¹⁹ From the analysis given in the earlier section, the motions of the N–H vector relative to peptide plane are small but not negligible and in some cases dominate the motion of that vector. To investigate the effect of this on NMR relaxation data, we compare order parameters derived with respect to local reference frames, such as groups of atoms belonging to the peptide unit and those relative to a molecular reference frame. Autocorrelation functions P_2 of the N–H vector, which are involved in the calculation of NMR order parameters, are shown for three residues in Figure 8. The entire protein or individual peptide planes formed by the three atoms (O_{i-1} , C_{i-1} , N_i ; see Methods) were taken as the reference frame for the motion of the N_i –H vectors, shown as $C_2(t)$ and $C_2(t)_{loc}$, respectively. Differences in $C_2(t)$ for different residues are apparent both in the time scale of convergence and in the plateau value, which is associated with S^2 . For example, the simulation-derived order parameters with respect to a molecular reference frame are 0.92, 0.67, and 0.42 for residues 30, 85, and 104, respectively. Not all of the correlation functions (e.g., Gly104) have converged at 200 ps, suggesting that rare transitions occur (Buck and Karplus, manuscript in preparation). The correlation functions generally have much longer time developments when evaluated relative to the entire main chain than relative to the peptide plane. We therefore calculated $C_2(t)_{loc}$ and define separately a quantity S_{loc}^2 , an order parameter due to local motion. The $C_2(t)_{loc}$ functions decay on a subpicosecond time scale to plateau values, S_{loc}^2 , near 0.93. The S_{loc}^2 values are essentially the same for all residues, in accord with the earlier analysis of the N–H vector fluctuations with respect to the peptide plane (see Section a). These local motions give rise to functions that show an oscillatory behavior on a subpicosecond time scale and have

(59) Levy, R. M.; Karplus, M. *Biopolymers* **1979**, *18*, 2465–2495.

(60) Brüschweiler, R. *J. Chem. Phys.* **1995**, *102*, 3396–3403.

(61) Reif, B.; Steinhagen, H.; Junker, B.; Reggelin, M.; Griesinger, C. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 1903–1906.

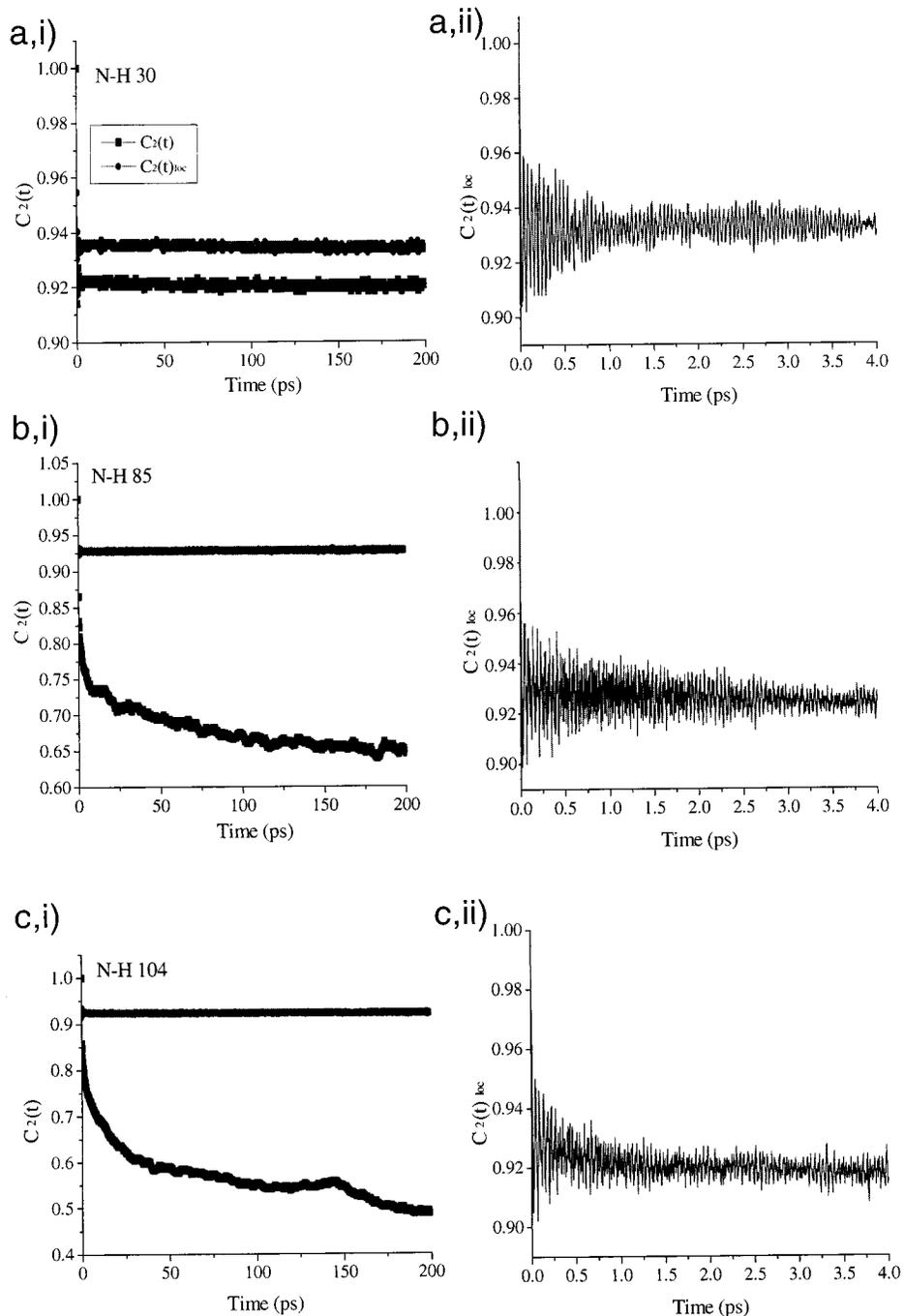


Figure 8. (a–c,i) N–H vector autocorrelation functions for Cys30, Ser85, and Gly104, respectively, calculated as $C_2(t)$ (squares, black) using the molecular frame as reference, and as $C_2(t)_{loc}$ by local superposition on atoms O_{i-1} , C_{i-1} , and N_i (circles, gray) for each of the peptide planes (see text). The initial decay of $C_2(t)_{loc}$ for the three vectors is shown in a–c,ii. N–H vectors of Ser36 and Thr53 behave similarly to Cys30.

decays which are slightly different for different N–H bond vectors. Williams and McDermott⁶² found increased T_1 relaxation times for amides flanked by alkyl groups in the α_R helical conformation. These differences were attributed to variations in the global motion of the peptide plane. No relationship between the decay of $C_2(t)_{loc}$ and the environment of the peptide amide group could be discerned in the present study (data not shown).

Figure 9 shows main-chain sites of hen lysozyme which have simulation-derived order parameters S^2 ranging from 0.87 to 0.93 for the motion of the N–H vectors relative to a molecular

reference frame. Also plotted are values of S_{loc}^2 and S_{loc}^{2*} for motions relative to the peptide plane defined by atoms O_{i-1} , C_{i-1} , and N_i and atoms $C\alpha_{i-1}$, C_{i-1} , and N_i respectively; the latter includes fluctuations involving $C\alpha_{i-1}$ as explained above. Close agreement between S_{loc}^2 and S_{loc}^{2*} implies that distortions of the different set of atoms which form the reference frames for the N–H fluctuation are similar. The trend in S_{loc}^2 (and S_{loc}^{2*}) as a function of sequence is similar to that seen for the extent of fluctuations in ω_1 (Figure 3b [ii]). These parameters are well correlated (correlation coefficient of -0.71 for 126 peptide planes; Figure 4e) as are S_{loc}^2 and the standard deviation of in-plane and particularly out-of-plane displacement (Figure 4f; correlation coefficients of -0.50 and -0.92 respectively).

(62) Williams, J. C.; McDermott, A. E. *J. Phys. Chem. B* **1998**, *102*, 6248–6259.

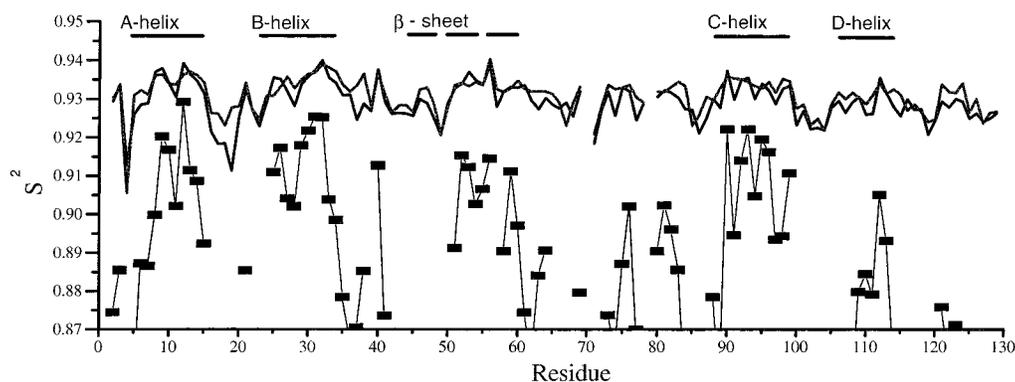


Figure 9. Local order parameters S_{loc}^2 (black line) and S^2 (gray line) as a function of protein sequence. S^2 values are shown if greater than 0.87 (black rectangles). Refer to Methods for definitions and derivation.

S_{loc}^2 is nearly equal to S^2 (within 0.02) for 16 residues with high S^2 . Values of S^2 that are smaller than 0.93 arise from motions which lead to displacements of the entire peptide group (compare Figure 5). An example of such motions are fluctuations of the adjacent dihedrals ϕ and ψ , but other factors are also involved. The fact that the difference between S^2 and S_{loc}^2 can be small shows that N–H motions relative to the peptide group may be the dominant contribution to NMR order parameters, particularly in regions of secondary structure. This finding contradicts the assumption that the peptide unit can be modeled as a rigid entity and that S^2 for the N–H vector may be estimated by consideration only of the fluctuations of ϕ and ψ dihedral angles adjacent to the peptide plane.^{11,60} However, the magnitude of the variation of S_{loc}^2 appears to be within the uncertainty with which order parameters can be derived from experimental relaxation measurement in many cases.⁶³ Consequently, the Gaussian Axial Fluctuation model of Brüschweiler⁶⁴ can yield accurate results in accord with the accuracy of the data. Nevertheless, it is appropriate to rescale S^2 by a factor $S_{loc}^2 \sim 0.93$ due to the neglect of local N–H motion. Similar corrections for local motions have been proposed for the interpretation of dynamics measurements made by time-resolved fluorescence depolarization,²⁰ although the correction is on average much greater than proposed here for the motion of N–H vectors relative to the peptide plane.

Bond Length Effects. The near independence from the environment of the motion of the N–H vector relative to the peptide group suggests that the N–H bond lengths (i.e., the mean bond lengths and rms fluctuations as a function of residue number) may behave similarly. Although the fluctuations of N–H bond length are considerable over the time course of the simulation (with a rms fluctuation of 0.022 to 0.030 Å), the variation in the mean N–H bond distance is very small (1.0027 ± 0.0025 Å) between peptide groups. Bond length fluctuations could be studied in the present simulations because they were carried out without the use of restraints, such as SHAKE,³⁶ typically applied to hydrogens bound to protein heavy atoms. To determine the effective order parameters, we factored S^2 in eq 5a,b into components $S^2 = S_{angular}^2 S_{vibrational}^2$. This is a good approximation if the order parameter is large.⁶⁴ The value of $S_{vibrational}^2$ is obtained by comparing S^2 values that are calculated using the instantaneous bond length in the simulation (i.e., allowing for N–H bond length variations) with S^2 values derived with an average fixed N–H bond length (i.e., $S_{vibrational}^2 = 1.0$).

Alternatively, Brüschweiler et al.⁶⁵ have defined $S_{vibrational}^2 = \langle 1/r^3 \rangle^2 / \langle 1/r^6 \rangle$. By using both methods, we found $S_{vibrational}^2$ to be in the range 0.992–0.995 for N–H bonds, in agreement with calculations of Philippopoulos and Lim,⁶⁶ who carried out simulations of Ribonuclease H1 also without the use of SHAKE; they used a time step of 1 fs, while in the present case multistep dynamics are used and X–H bonds are updated every 0.5 fs (see Methods). The bond length fluctuations occur on a time scale that is sufficiently fast not to affect the NMR relaxation parameters¹⁵ so that an average bond length can be used. When N–H order parameters are derived from experimental relaxation data (T_1 , T_2 , NOE), there is a $(1/r^6)$ dependence on N–H bond length variations.¹⁷ Thus, a vibrational lengthening of the bond by 1% would cause an increase of approximately 6% in the value of S^2 to obtain the same relaxation parameter values.⁶⁷ By using the simulation-derived average bond lengths in the calculation of order parameters from the experimental relaxation data, we expect an average variation of at most $\pm 1.5\%$ for S^2 of the N–H groups. The data confirm the assumption that a common value for the N–H bond length is appropriate in the calculation of heteronuclear order parameters (see also Ottinger and Bax⁶⁸). This has been implicit in previous theoretical and experimental studies.^{69–70}

Conclusion

The distortions of the peptide plane and the motions of the main-chain N–H bond vector in molecular dynamics simulations of lysozyme have been analyzed in terms of local and molecular frame superpositions and by use of the peptide plane dihedral angles. The distortions originate from changes in bond geometry at the nitrogen and carbonyl centers, as well as from twisting around the C–N bond. The distortions are of the same magnitude as those found in databases of high-resolution peptide structures. A self-consistent analysis shows that average distortions of the peptide plane measured by the dihedral angle ω_1 and the instantaneous fluctuations involve similar but not identical potential of mean force surfaces. For the displacement of the N–H vector, the deviation between the two types of surfaces is much larger. This finding suggests that caution is

(65) Brüschweiler, R.; Roux, B.; Blackledge, M.; Griesinger, C.; Karplus, M.; Ernst, R. R. *J. Am. Chem. Soc.* **1992**, *114*, 2289–2302.

(66) Philippopoulos, M.; Lim, C. *J. Mol. Biol.* **1995**, *254*, 771–792.

(67) Kay, L. E.; Torchia, D. A.; Bax, A. *Biochemistry* **1989**, *28*, 8972–8979.

(68) Ottinger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 12334–12341.

(69) Henry, E. R.; Szabo, A. *J. Chem. Phys.* **1985**, *82*, 4753–4761.

(70) Tjandra, N.; Szabo, A.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6986–6991.

(71) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946–950.

(63) Philippopoulos, M.; Mandel, A. M.; Palmer, A. G.; Lim, C. *Proteins: Struct. Funct., Genet.* **1997**, *28*, 481–493.

(64) Brüschweiler, R.; Wright, P. E. *J. Am. Chem. Soc.* **1994**, *116*, 8426–8427.

required in the now widespread use of statistically derived potentials of mean force.

The extent of peptide plane distortions involving heavy atoms, particularly the fluctuation amplitudes, can be correlated with the motions of the surrounding groups. By contrast, the N–H vector orientation relative to a local reference frame reveals a more uniform behavior. The motion of the N–H vector is very rapid and less correlated with the protein environment, such as hydrogen bonding or secondary structure. The nature of the fluctuations is such that an overall renormalization of the N–H order parameter is suggested for the analysis of NMR data. For

N–H bond-stretching motion, which are also decoupled from the rest of the system, an average value is shown to be appropriate.

Acknowledgment. This work was supported in part by a grant from the National Institute of Health. M.B. was a fellow of the Human Frontiers Science Program Organization and funded subsequently by a Marie Curie Fellowship in BioTechnology of the European Commission. Some of the calculations were done at the Pittsburgh Supercomputer Center.

JA991309P