PROTEINS: Structure, Function, and Genetics 5:337–354 (1989)

A Molecular Dynamics Analysis of Protein Structural Elements

Carol Beth Post,[1] Christopher M. Dobson,[2] and Martin Karplus[3]
[1]Chemistry Department, Purdue University, West Lafayette, Indiana 47907; [2]Inorganic Chemistry Laboratory, Oxford University, Oxford OX1 3QR, England; [3]Chemistry Department, Harvard University, Cambridge, Massachusetts 02138

ABSTRACT **The relation between protein secondary structure and internal motions was examined by using molecular dynamics to calculate positional fluctuations of individual helix, β-sheet, and loop structural elements in free and substrate-bound hen egg-white lysozyme. The time development of the fluctuations revealed a general correspondence between structure and dynamics; the fluctuations of the helices and β-sheets converged within the 101 psec period of the simulation and were lower than average in magnitude, while the fluctuations of the loop regions were not converged and were mostly larger than average in magnitude. Notable exceptions to this pattern occurred in the substrate-bound simulation. A loop region (residues 101–107) of the active site cleft had significantly reduced motion due to interactions with the substrate. Moreover, part of a loop and a 3_{10} helix (residues of 67–88) not in contact with the substrate showed a marked increase in fluctuations. That these differences in dynamics of free and substrate-bound lysozyme did not result simply from sampling errors was established by an analysis of the variations in the fluctuations of the two halves of the 101 psec simulation of free lysozyme. Concerted transitions of four to five mainchain ϕ and ψ angles between dihedral wells were shown to be responsible for large coordinate shifts in the loops. These transitions displaced six or fewer residues and took place either abruptly, in 1 psec or less, or with a diffusive character over 5–10 psec. Displacements of rigid secondary structures involved longer timescale motions in bound lysozyme; a 0.5 Å rms change in the position of a helix occurred over the 55 psec simulation period. This helix reorientation within the protein appears to be a response to substrate binding. There was little correlation between the solvent accessible surface area and the dynamics of the different structural elements.**

Key words: computer simulation, fluctuations in proteins, secondary structural dynamics, lysozyme, protein-substrate complex

INTRODUCTION

Mainchain hydrogen bonding in proteins defines the elements of secondary structure—helix, β-sheet, coiled loop. Since protein internal dynamics are determined in part by the local nonbonding interactions, mobility should be correlated with the type of structural element. Spatial constraints imposed by a regular pattern of hydrogen bonding and electrostatic interactions are expected to lead to smaller amplitude, higher frequency motions, while less regular interactions would allow larger, lower frequency displacements. Most simply, this suggests that helical and β-sheet structures would be relatively rigid and coiled loops more flexible. Crystallographic thermal factors,[2,12] and fluctuations from molecular dynamics (MD) simulations,[20,27] are indeed generally smaller for helices and β-sheets than for coiled loops. A relationship between fluctuations and the number of mainchain hydrogen bonds has been suggested, based on crystallographic results.[13]

A detailed analysis of the magnitude and time dependence of the atomic fluctuations of individual structural elements of hen egg-white lysozyme determined in molecular dynamics simulations is reported in this paper. The results provide a characterization of the relationship between secondary structure and dynamics. Lysozyme comprises all secondary structure types and, as such, is well suited for such an analysis. There are four α-helices, of which one is mainly buried in the protein interior, a three-stranded β-sheet, and a long stretch of coiled loop. Moreover, extended molecular dynamics simulations have been made for free lysozyme and for lysozyme bound with the substrate hexa-*N*-acetylglucosamine, $(GlcNAc)_6$.[23] Thus, a comparison of the dynamic properties of the two systems can be made.

While spectroscopy and crystallography provide information on certain internal motions in proteins, an advantage of molecular dynamics simulations is the possibility for investigating the individual motion of all atoms. It is of interest, for example, to be able to study the experimentally elusive enzyme–

Received July 5, 1988; revision accepted December 21, 1988.
Address reprint requests to Dr. Carol Beth Post, Chemistry Department, Purdue University, West Lafayette, Indiana 47907.

substrate complex to determine how substrate binding perturbs the dynamic behavior. Moreover, proteins are known to exist in multiple conformations in the neighborhood of the native structure; conformational heterogeneity has been observed in several crystallographic studies,[3,4,6,8,14,22,25,27] and a large number of conformational substates have been suggested to account for spectroscopic results.[1,3,19] The experimental data indicating that the multiple conformations exist are supplemented by molecular dynamics simulations that provide a detailed picture of their characteristics.[10] Conformational transitions in the densely packed protein environment are complex and the simulations provide information concerning the multidimensional dynamic process necessary for concerted motions. A mainchain transition of an eight-residue loop (residues 100–107) occurred in the free lysozyme simulation, and the motional mechanism is described in this report. The thermal factors of this loop are large[2,6] and substantial deviations occur in the coordinates of this loop obtained from different crystallographic structures.[15]

The Methods section briefly describes the simulations and the method of calculation of the fluctuations and the structural changes; a general description of the simulations, which included strongly bound water molecules (so-called "structural" waters), has been given.[23] The Results section begins with a description of the lysozyme structure, including the secondary structure, solvent accessible surface, and substrate binding. The time development of the fluctuations is then analyzed. We consider the whole molecule and the individual structural elements. Comparison is made with crystallographic thermal factors and the relationship of mobility with solvent exposure is examined. The extent to which fluctuations vary due to statistical error is estimated from the 101 psec simulation of free lysozyme. To determine the effects of the substrate on the protein dynamics, specific motions and structural changes in the free and bound simulations are analyzed. The conclusions are then presented.

METHODS

Two molecular dynamics simulations were used for the present analysis. A description of the simulations has been published[23]; it reports the details of the trajectory calculations, a comparison of the dynamic structures with the crystal structure, and an overall view of the protein fluctuations and of the dynamic behavior of the waters. The 101 psec simulation of free lysozyme included lysozyme (1001 heavy atoms, 265 polar hydrogen atoms) and 53 waters (265 atoms for the five-center ST2 water model). The 55 psec bound simulation included lysozyme, the substrate $(GlcNAc)_6$ (85 heavy atoms, 20 polar hydrogen atoms), and 35 waters (175 atoms). Polar hydrogens were explicitly included to allow a more accurate hydrogen-bonding model, while the aliphatic hydrogens were accounted for with the extended-atom model.[9] Several of the 53 structural waters included in the simulation of free lysozyme are in the active cleft and, as such, are displaced by $(GlcNAc)_6$ in the bound simulation. Inclusion of these active-site waters in the free lysozyme simulation provides a more reliable comparison of the protein dynamics between the free and the bound state.

Initial atomic positions for the free and bound simulations were obtained from the crystallographic coordinates of tetragonal lysozyme refined to 1.6 Å resolution and of tetragonal lysozyme bound with $(GlcNAc)_3$ (Handoll, Artymiuk, Cheetham, and Phillips, unpublished results), respectively. As previously reported,[23] the coordinates for the first three sugars in sites A, B, and C are from the crystallographic results,[7] while the remaining three sugars in sites D, E, and F were model built into the structure. The crystal waters included in the simulations were twice hydrogen bonded to the protein either directly or through an intermediate water molecule.

The trajectories were calculated on a CRAY-1S computer using a version of the program CHARMM.[9] The empirical potential energy is a function of the atomic positions and includes quadratic dependences on bond lengths and angles, two- and three-periodic torsional potentials, coulombic and van der Waals interactions, a hydrogen bond potential with both radial and angular dependences, and a weak quartic constraint on the waters. Newton's equations of motions were solved using the Verlet algorithm with a timestep of 0.98×10^{-15} seconds. The average temperature of the free trajectory was 295 K, while that of the bound trajectory was 304 K. The coordinate sets used for this analysis were taken from the trajectories after completion of the equilibration period (17 psec in the free trajectory; 28 psec in the bound trajectory) at approximately 0.1 psec intervals.

To determine the time development of the atomic fluctuations, the length of the time interval over which the fluctuations are calculated is varied;[28] e.g., for a 10 psec average fluctuation value, the root mean square (rms) displacement of an atom from its average position is calculated using a 10 psec window from the trajectory, and the results for all 10 psec windows are then averaged. Averaging times used included 1, 5, 10, 20, 25, 50, and 100 psec windows.

To determine the possibility that conformational transitions are the cause of some of the variations, dynamically averaged structures from points along the trajectory were compared to either the crystallographic structure or other simulation average structures. Translation and rotation of the coordinates to give the best least-squares fit of the structures were performed before calculating the rms de-

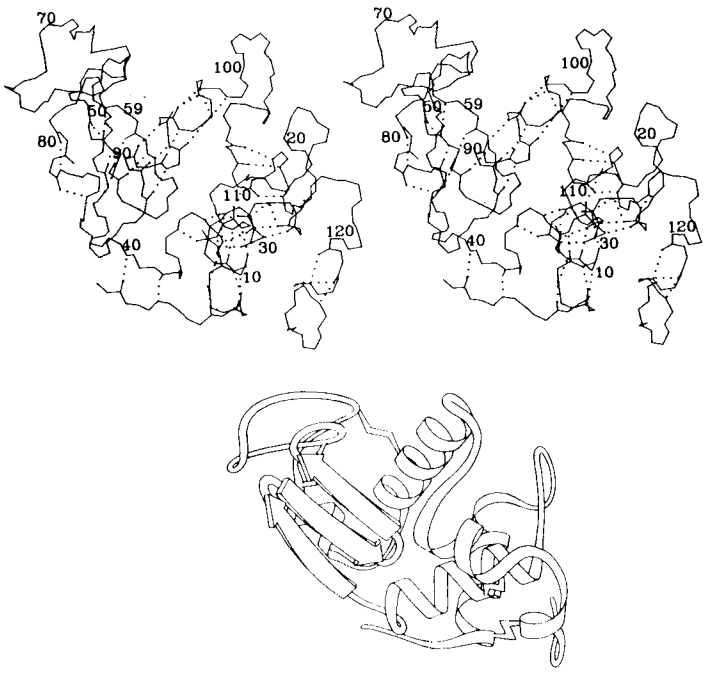

Fig. 1. Mainchain atoms (N, Cα, C) of lysozyme. **a**: Stereodrawing; secondary structural elements are delineated by hydrogen bonds (dotted line) and include the mainchain atom O: αA (4–15), αB (24–36), αC (88–99), αD (108–115), β-sheet (1–3 and 41–60), $3_{10}-1$ (79–84), and $3_{10}-2$ (120–124). **b**: Schematic secondary structure (drawing courtesy of Jane Richardson).

viations. This least-squares fit included all protein heavy atoms or only the heavy atoms of specific structural elements. In this way a conformational change internal to a structural element could be distinguished from a displacement of the element relative to the rest of the protein.

RESULTS

Lysozyme Structure and Exposed Surface Area

Lysozyme has 129 amino acids and consists of two domains separated by a cleft in which hexasaccharides are bound (Fig. 1) Overall, lysozyme consists of 15% β-sheet, 34% α-helix, 9% 3_{10} helix, and 42% loop. The elements considered in the present analysis are listed in Table I. Not all lysozyme residues are included in an element of secondary structure; that is, sequences less than five residues in length and the five residues at the carboxyl terminus were not considered. In addition, it should be noted that the elements do not have an equal number of residues; this has some effect on the statistics. One domain comprises four α-helices and one 3_{10} helix connected by short loops. The helices are αA (residues 4–15), αB (24–36), αC (88–99), αD (108–115), and $3_{10}-2$ (120–124). Helices αA, αC, and αD are relatively exposed while αB is more in the interior with only the ends partly exposed to solvent. Short loops connect αA to αB (16–23) and αC to αD (100–107). The second domain comprises a three-stranded, antiparallel β-sheet (41–60), a long coiled-loop region (61–78), and $3_{10}-1$ (79–84). One of the three β-sheet strands (42–45) runs along the protein surface near the reducing end of the active site cleft. The other two strands (50–53 and 58–60) are in the protein interior. Residues 46–49 form an exposed β-turn while the β-turn 54–57 is fully buried, an unusual environment for a β-turn.[24] However, both 55 NH and 56 NH hydrogen bond to interior waters. The long loop region begins in the protein interior at the carboxy terminus of the β-sheet (residue 61), winds out to the surface, forming a coil along the surface, and ends at the short $3_{10}-1$ helix on the side of the enzyme away from the cleft. The two domains are joined between $3_{10}-1$ and αC by residues 85–87 and by a small antiparallel β-sheet structure, residues 1–3 and 38–40.

The solvent accessible surface area of each structural element was calculated from the crystallographic coordinates for a 1.4 Å probe[18,26] and is also given in Table I. Calculations were carried out for an element in the environment of the rest of the protein and for the isolated element by itself. Both main chain and side chain atoms are included. The ratio of exposed surface of the element within the protein to that of the isolated element is the fraction shown in Table I. This comparison cancels effects of the change in exposed surface area due to the formation of the individual element from an extended chain. The surface areas of the elements for free lysozyme are listed in Table Ia and for lysozyme with bound $(GlcNAc)_6$ in Table Ib. All of the structural elements have some fraction of exposed area.

Although nearly one-half the surface of the whole β-sheet is exposed, parts of the β-sheet have the smallest exposure; one strand (58–60) and the internal turn (54–57) are nearly inaccessible to solvent, and the fraction of accessible area of one other strand (50–53) is only 0.25 (Table Ia) without substrate and 0.13 with substrate (Table Ib). The helices have fractions of exposed area which range from 0.13 to 0.42, while the loops have larger fractions ranging from 0.4 to 0.64.

The bound hexasaccharide lies perpendicularly across the interior end of the β-sheet of one domain, and adjacent to the loop connecting αC to αD and the carboxyl end of αB of the other domain. Lysozyme contacts the substrate (atomic separations of less than 4.0 Å) at αB (residues 34, 35), αD (108–110, 112), the loop connecting αC and αD (residues 101–103, 107), the β-sheet (residues 43, 44, 46, 52, 56–59), and loop residues 62, 63, and 73. Relative to the free X-ray structure, the bound-state structure of lysozyme changes very little in solvent accessibility when $(GlcNAc)_6$ is excluded in the surface calculation. The values in Table Ia and Ib do

TABLE Ia. Surface Area of the Secondary Structural Elements of Free Lysozyme Accessible to a 1.4 Å Radius Spherical Probe[18,*]

Structural element	Residues	Calculated surface (Å^2)		Fraction exposed
		Lysozyme	Element	
α Helix A	4–15	618	1461	0.423
α Helix B	24–36	198	1475	0.134
α Helix C	88–99	286	1377	0.208
α Helix D	108–115	452	1207	0.374
3_{10}–1 Helix	79–84	225	801	0.281
3_{10}–2 Helix	120–124	247	856	0.289
β-sheet	41–60	850	1924	0.442
Strand	41–45	431	617[†]	0.699
β-Turn	46–49	333	397[†]	0.839
Strand	50–53	42	169[†]	0.249
β-Turn	54–57	12	420[†]	0.029
Strand	58–60	33	320[†]	0.103
Loop αA to αB	16–23	489	1106	0.442
αC to αD	100–107	419	1030	0.407
Long	61–78	1185	1865	0.635
Total		5820	15025	0.387

*The surface area is calculated from the crystallographic coordinates for the elements in the protein or for the isolated elements.
†Surface area of the isolated β-sheet contributed by this component.

TABLE Ib. Surface Area of the Secondary Structural Elements in Lysozyme Bound With $(GlcNAc)_6$ Accessible to a 1.4 Å Radius Spherical Probe (as in Table Ia)

Structural element	Residues	Calculated Surface (Å^2)		Fraction exposed*
		Lysozyme*	Element	
α Helix A	4–15	580	1455	0.399
α Helix B	24–36	143 (206)	1479	0.097 (0.139)
α Helix C	88–99	320	1372	0.233
α Helix D	108–115	340 (457)	1202	0.283 (0.380)
3_{10}–1 Helix	79–84	232	799	0.290
3_{10}–2 Helix	120–124	261	869	0.300
β-Sheet	41–60	619 (830)	1920	0.322 (0.432)
Strand	41–45	433	626	0.692
β-Turn	46–49	249 (310)	383	0.650 (0.809)
Strand	50–53	21 (42)	167	0.126 (0.251)
β-Turn	54–57	0 (15)	427	0.0 (0.035)
Strand	58–60	1 (30)	316	0.003 (0.095)
Loop αA to αB	16–23	508	1162	0.437
αC to αD	100–107	234 (405)	1040	0.225 (0.389)
Long	61–78	989 (1126)	1824	0.542 (0.617)
Total		4930 (5755)	15041	0.328 (0.383)

*Values in parenthesis were calculated without including $(GlcNAc)_6$.

not deviate by more than 5%, except for αC which has a 10% larger surface area when the protein is in the bound state. Upon binding, the elements covered by $(GlcNAc)_6$ are αB and αD, the β-sheet, the loop αC to αD, and the long loop, as shown by the values in parentheses in Table Ib.

For comparison the average MD structures of the free and bound simulations were also used to calculate the fractional accessibility of the folded-element surface area in the protein (the rms differences between the dynamics average structure and the X-ray structure for the main chain N, Cα, C coordinates were 1.5 and 1.4 Å for the free and bound simulations, respectively.[23]) Although most of the fractions from the average dynamic structure differed by less than 10% of the crystal structure values, three elements did show larger changes. There was a decrease in the accessibility of 15% for the β-sheet in the free simulation and αC in the bound simulation, and of 30% for 3_{10}–2 in the bound sim-

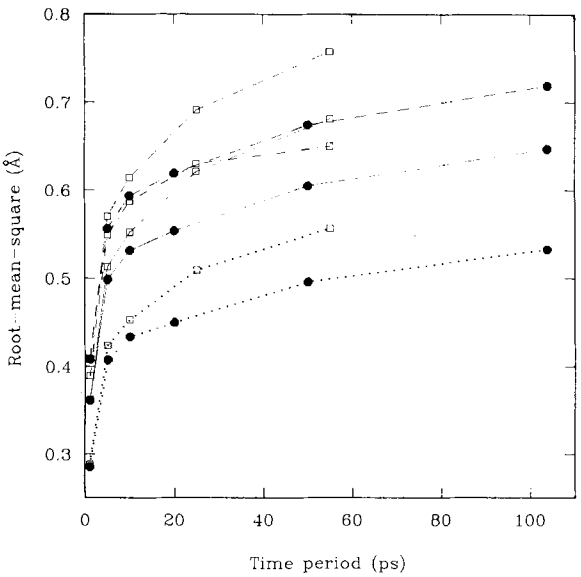

Fig. 2. Time development of the positional fluctuations of lysozyme and the substrate $(GlcNAc)_6$. Fluctuations were calculated for the indicated time period of averaging, and ensemble averaged over the full simulation. Time development values were then averaged over (solid) all 1001 lysozyme heavy atoms, (dotted) main chain atoms N, Cα, C, (dashed) side chain heavy atoms O, Cβ, and beyond, and (dot–dash) all 85 $(GlcNAc)_6$ heavy atoms. The simulations of free lysozyme (●) and substrate-bound lysozyme (□) are compared.

ulation. These changes result from small side chain displacements toward other protein atoms, and are likely to be a consequence of the lack of solvent and crystal contacts.

Overall Fluctuations

The time development of the fluctuations for the free and bound simulations are shown in Figure 2 and Tables II and III. Average values of the rms positional fluctuations are given for all lysozyme heavy atoms, main chain heavy atoms (N, Cα, C), side chain heavy atoms (O, Cβ, and beyond), and $(GlcNAc)_6$ heavy atoms. The increase in fluctuations is rapid up to 5 psec, and then begins to slow down; the fluctuations have reached approximately 80% of their total value in the 10 psec averages. The fluctuations averaged over all lysozyme atoms do not reach an asymptotic plateau value, although they are only increasing very slowly during the last 50 psec in the free simulations. The convergence behavior for the main chain and side chain atoms is essentially identical in spite of the fact that the fluctuations of the latter are significantly larger than those of the former. For as short an averaging interval as 5 psec, the fluctuations for the side chain atoms are about 0.15 Å greater than fluctuations in the main chain atoms. The $(GlcNAc)_6$ fluctuation magnitudes are intermediate between the protein main chain and side chain values, and very close to the lysosome average value.

Although 1 psec fluctuations averaged over the entire molecule in the bound simulation are equal to those in the free simulation, the fluctuation magnitudes diverge when longer time scale motions are considered. The difference in the fluctuation magnitudes cannot be accounted for by the slightly higher temperature of the bound simulation. In the harmonic approximation, the rms fluctuations are proportional to the square root of the temperature, and the ratio of the average square root of the temperatures is only ~3%, significantly less than the 10% difference in the fluctuation magnitudes of the two trajectories. The statistical error of the simulations can be a source of the difference since the discrepancy between the two simulations is of a magnitude similar to that between the two halves of the free simulation. Fluctuations calculated from the first and second 50 psec of the free simulation differ by 0.06 Å, and from the bound simulation are approximately 0.07 Å larger than from the free simulation. However, the fluctuations of certain individual structural elements show significantly larger differences between the free and bound simulations and these substrate-altered motions are discussed below.

Fluctuations by Structural Elements

The relationship between dynamics and structure is considered by averaging the rms positional fluctuations over the residues composing each structural element. Relative mobilities so obtained characterize helices, sheet, and loops in terms of flexibility. In contrast to a comparison between two independent trajectories, relative fluctuations from a single trajectory are more reliable; the mobilities in separate regions of the molecule are less affected by systematic variations, such as initial conditions and temperature, which can obscure small differences.

For atoms sampling a region of configuration space corresponding to a single minimum, the time development of the fluctuations is expected to show a smooth increase to a limiting value as the averaging time is increased. The rms fluctuation magnitude plotted as a function of the time of the subinterval would be a curve continuously rising to a plateau value corresponding to the equilibrium fluctuation. Motions of atoms whose configuration space is restricted to a small region exhibit high-frequency fluctuations and the plateau is reached rapidly. If the local potential surface is relatively broad, the space accessible to the atom is larger and a longer time is required before a plateau value is reached. The occurrence of a structural transition, involving the displacement of one or more atoms between two minima, results in a discontinuous increase in the slope of the time development curve. An atom is most often in either of the two wells and spends only

TABLE II. Positional Fluctuations for the Free Lysozyme Trajectory, Calculated Over Varying Time Periods and Averaged Over the Indicated Residues for Each Structural Element*

Structural element	Residues	Residues					
		1 psec	5 psec	10 psec	20 psec	50 psec	100 psec	X-Ray
All atom		0.361	0.499	0.531	0.554	0.605	0.647	0.602
Main chain								
Molecule	1–129	0.286	0.407	0.434	0.450	0.496	0.533	0.531
Helix								
αA	4–15	0.294	0.418	0.443	0.467	0.511	0.541	0.540
αB	24–36	0.259	0.353	0.376	0.391	0.417	0.425	0.422
αC	88–99	0.275	0.403	0.424	0.438	0.479	0.504	0.514
αD	108–115	0.268	0.387	0.407	0.419	0.438	0.444	0.470
β-Sheet	41–60	0.264	0.364	0.380	0.390	0.399	0.406	0.460
$3_{10}-1$	79–84	0.273	0.362	0.379	0.387	0.406	0.419	0.500
$3_{10}-2$	120–124	0.283	0.443	0.476	0.499	0.524	0.540	0.603
Loop								
αA–αB	16–23	0.336	0.492	0.524	0.544	0.615	0.707	0.539
Long	61–78	0.283	0.401	0.423	0.434	0.494	0.557	0.623
αC–αD	100–107	0.340	0.498	0.553	0.580	0.741	0.893	0.588
Side chain								
Molecule	1–129	0.408	0.556	0.593	0.619	0.674	0.719	0.647
Helix								
αA	4–15	0.400	0.540	0.572	0.600	0.645	0.681	0.651
αB	24–36	0.368	0.485	0.514	0.539	0.570	0.591	0.491
αC	88–99	0.387	0.534	0.565	0.587	0.644	0.685	0.607
αD	108–115	0.385	0.526	0.556	0.576	0.604	0.626	0.557
β-Sheet	41–60	0.386	0.503	0.529	0.547	0.566	0.578	0.586
$3_{10}-1$	79–84	0.386	0.495	0.528	0.551	0.593	0.628	0.568
$3_{10}-2$	120–124	0.421	0.601	0.639	0.662	0.698	0.716	0.680
Loop								
αA–αB	16–23	0.452	0.628	0.666	0.701	0.795	0.916	0.668
Long	61–78	9.420	0.564	0.594	0.610	0.679	0.747	0.712
αC–αD	100–107	0.464	0.650	0.711	0.748	0.923	1.068	0.706

*The main chain atoms are N, Cα, C and the side chain atoms are O, Cβ, and beyond.
†Values calculated from crystallographic thermal factors according to Eq. (1) and then reduced by 0.283 Å according to a minimum-function method.[22]

a small fraction of the time in the barrier region separating the minima, providing the barrier is sufficiently high. Such a transition is reflected by a sharp increase in the slope of the time development plot at an averaging interval large enough to sample the atom in both conformations.

The time development of the fluctuations for the main chain heavy atoms of the various structural elements was calculated from the free and bound trajectories (Fig. 3). The values averaged over all 129 residues are also given in the figure (dotted-line curve) to allow comparison of the fluctuation values of an element with the overall average. The time-development values for the free and bound simulations are listed in Tables II and III, respectively.

Examination of Figure 3 shows that the main chain atoms of the β-sheet and α-helical structures have fluctuations smaller than or close to the average fluctuation value, whereas most of the coiled-loop regions have slightly greater than average fluctuations. In the 101 psec free trajectory, fluctuations for main chain atoms of the helical and β-sheet structures (solid line curves in Fig. 3a) have reached or nearly reached asymptotic values. By contrast, the loop regions clearly have not converged. In the shorter (55 psec) bound trajectory (Fig. 3b), the fluctuations of helices and sheet have not converged; however, the time dependence is similar to that in the free trajectory which suggests that they would converge by 100 psec. The one case where a helical structure has large fluctuations is the short $3_{10}-1$ helix in the bound trajectory; a coiled-loop stretch, residues 100–107, in contact with the substrate in the bound trajectory has very small fluctuations. These two aspects of the bound trajectory are discussed below.

In general, the fluctuation characteristics of the main chain atoms are also evident for the side chain atoms. However, the increases with time of averaging are larger for the side chain atoms and their magnitudes are not quite converged even for the α-helices and β-sheet at 100 psec.

Some of the curves in Figure 3 cross, so that there is not a strict correlation between the magnitude and the rate of the time development of the displacements. This is most clearly illustrated by the $3_{10}-1$ helix in the bound trajectory (Figs. 3b). The fluctuations of this structural element have a slow time development; they are of average magnitude up to the 10 psec averaging window, but increase to val-

TABLE III. Positional Fluctuations for the Bound Lysozyme Trajectory, as in Table II

Element	Residues	Structural				
		1 psec	5 psec	10 psec	25 psec	50 psec
All atoms		0.362	0.513	0.552	0.621	0.681
Main chain						
Molecule	1–129	0.286	0.424	0.453	0.509	0.557
Helix						
αA	4–15	0.295	0.429	0.442	0.495	0.506
αB	24–36	0.260	0.368	0.391	0.436	0.475
αC	88–99	0.268	0.369	0.400	0.463	0.519
αD	108–115	0.280	0.416	0.457	0.491	0.498
β-Sheet	41–60	0.281	0.424	0.445	0.463	0.477
3_{10}–1	79–84	0.300	0.423	0.442	0.521	0.683
3_{10}–2	120–124	0.276	0.441	0.466	0.524	0.541
Loop						
αA–αB	16–23	0.297	0.439	0.462	0.555	0.586
Long	61–78	0.288	0.479	0.531	0.608	0.670
αC–αD	100–107	0.267	0.382	0.408	0.466	0.490
Side chain						
Molecule	1–129	0.409	0.570	0.614	0.691	0.755
Helix						
αA	4–15	0.423	0.599	0.629	0.704	0.734
αB	24–36	0.358	0.486	0.516	0.597	0.669
αC	88–99	0.374	0.484	0.520	0.584	0.660
αD	108–115	0.385	0.549	0.602	0.651	0.686
β-Sheet	41–60	0.402	0.564	0.598	0.630	0.655
3_{10}–1	79–84	0.424	0.567	0.603	0.714	0.857
3_{10}–2	120–124	0.407	0.593	0.619	0.678	0.709
Loop						
αA–αB	16–23	0.421	0.559	0.585	0.688	0.737
Long	61–78	0.437	0.645	0.725	0.830	0.944
αC–αD	100–107	0.390	0.516	0.542	0.611	0.639
$(GlcNAc)_6$						
Molecule	1–6	0.389	0.549	0.587	0.630	0.650

ues much greater than the average when calculated over the full simulation. Thus, 3_{10}–1 has a substantial component of low-frequency, large-magnitude motion while the component of faster motion (<20 psec) is near the average. The dynamics of other structural elements with large fluctuations over the full simulation, such as the loops 16–23, 100–107 (free), and 61–78 (bound), show large-magnitude, short-time motions, as well.

Although most of the time development curves increase smoothly, there are exceptions. The most striking examples of a change in the slope come from the fluctuations of the αC to αD loop and of the long loop in the free simulation. There is an increase in the slope at 20 psec for both of these structural elements (Fig. 3a). (When averaging over the entire molecule, this discontinuity is not evident due to the local nature of the displacements involved.)

The motional behavior of the β-turns is different from that of the β-sheet strands. Although fluctuations averaged over the full β-sheet are smaller than for other structural elements, the short-time fluctuations at the exposed β-turn 46–49 are of much greater than average amplitude, as shown in Figure 4. The buried turn 54–57 and the strands have smaller magnitude fluctuations, but all components are clearly coupled from the similar time development curves. Also shown for comparison in Figure 4 are the fluctuations of a loop outside the β-sheet, residues 16–23 connecting αA and αB. The exposed β-turn fluctuations are as large as those of the loop for times up to 20 psec, yet the dynamics of the β-turn plateau at 20 psec, while those of the loop increase up to 100 psec.

The motions of the small neighboring β-sheet (1–3, 38–40) are coupled to the large β-sheet as shown by their identical time development curves in Figure 4. This small β-sheet structure has hydrogen bonds between residues 1 and 40 and between 3 and 38. Residues 38–40 are adjacent to the 20-residue β-sheet and appear to transmit dynamic information to the more distant residues 1 and 3.

The time development of the four α-helices differs somewhat. In both the free and bound trajectories, the fluctuations of the internal αB helix are smaller than those of αA, αC, or αD (Fig. 3). In addition, the slopes of αB and αC in the bound trajectory are less steep than those of αA and αD (Fig. 3b), similar to the slowly developing fluctuations of 3_{10}–1 helix mentioned above. The motions giving rise to the behavior of αB and αC fluctuations are discussed below.

In addition to the overall time development of the different helices, the variations in the fluctuations

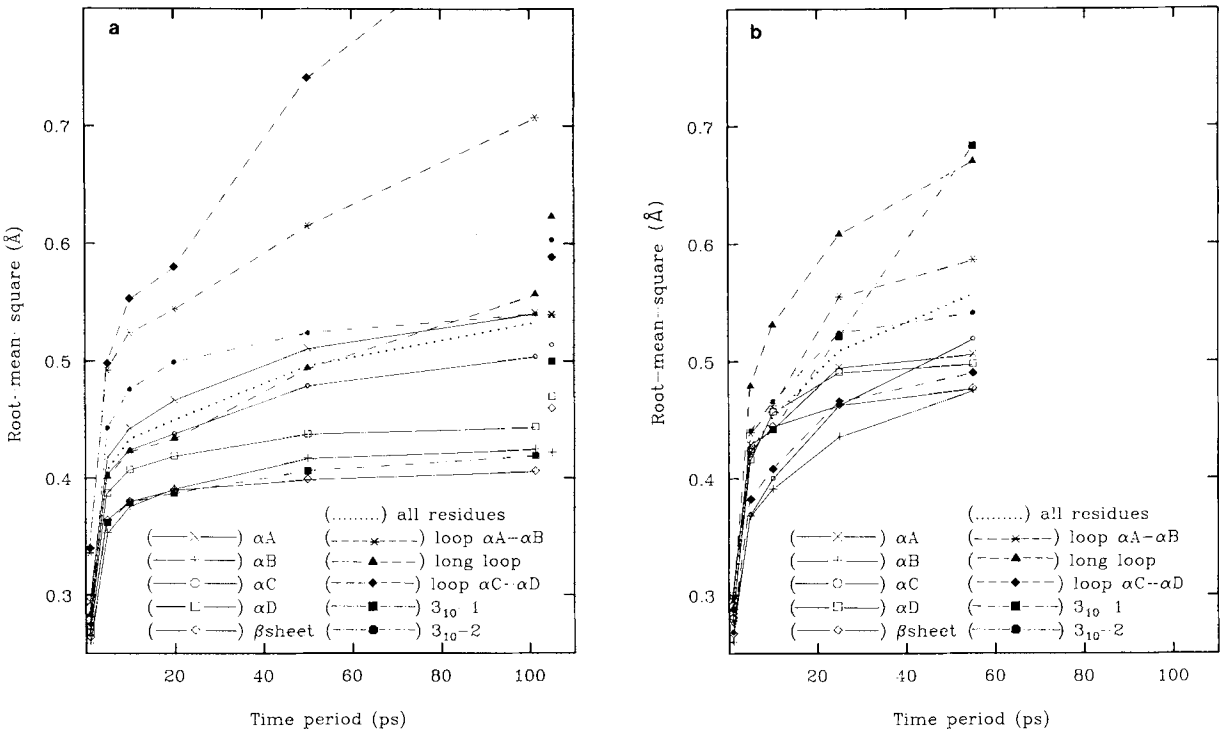

Fig. 3. Time development of the main chain positional fluctuations for each structural element of lysozyme. Fluctuations were calculated over the indicated time period and ensemble-averaged over **a:** the 101 psec free simulation and **b:** the 55 psec bound simulation of lysozyme. The time-averaged fluctuations were group-averaged over the main chainatoms (N, Cα, C) of the indicated elements and are compared to the average over all residues. The individual values plotted at 104 psec in **a** are fluctuations calculated from crystallographic thermal factors averaged over the same atoms.

along the helix were studied by examining individual main chain atom fluctuations. For both trajectories, the full-simulation fluctuations of the helical main chain atoms are plotted in Figure 5. Most often helices exhibit end-effects where the atoms at the ends of the helix have larger fluctuations than those in the middle. Helices αA and αC have the most notable end-effects, while the fluctuations of the terminal atoms of αD do not increase. One factor that is likely to reduce end-effects in αD is that the carboxyl terminal residue is a cystein that forms a disulfide bridge to residue 30 in αB.

Comparison With Crystallographic Thermal Factors

Atomic fluctuations calculated from the individual thermal factors of lysozyme refined to 1.6 Å resolution (Handoll, Artymiuk, and Phillips, unpublished results) are compared with the fluctuations in the simulation. Fluctuations were calculated from the thermal factor, B, according to the equation

$$<\Delta R^2> = 3B/8\pi^2 \quad (1)$$

and averaged over the main chain and side chain atoms of the structural elements. The $<\Delta R^2>$ values from the thermal factors are generally greater than those from the simulation. To compare the two sets of $<\Delta R^2>$, the crystallographic values were adjusted according to a minimum-function method[22] so that the smallest values are on an approximately identical scale as those in the simulation; i.e., the correction was made by requiring that the average of $<\Delta R^2>$ over the 50 smallest values obtained from the thermal factors be equal to the corresponding average from the MD fluctuations. This procedure introduces an overall reduction in the crystallographic values of 0.283 Å. Since only the relative values of the structural element fluctuations is relevant to the discussion here, the cause for the overall larger rms values from the thermal factors is not examined; from other studies it is likely that crystal disorder makes an important contribution.[22]

The scaled crystallographic values are listed in Table II and are also plotted in Figures 3a and 4a as isolated symbols at 104 psec. The range of rms thermal-factor fluctuations for the structural elements (main chain 0.422–0.623 Å and side chain 0.491–0.712 Å) is smaller than the range in the simulation (0.406–0.893 Å and 0.578–1.068 Å, respectively). The general order of the structural element fluctuations determined from temperature factors is the same as that from the simulation; the α-helices and β-sheet have smaller fluctuations than the coiled loops. Some differences, however, are seen in the

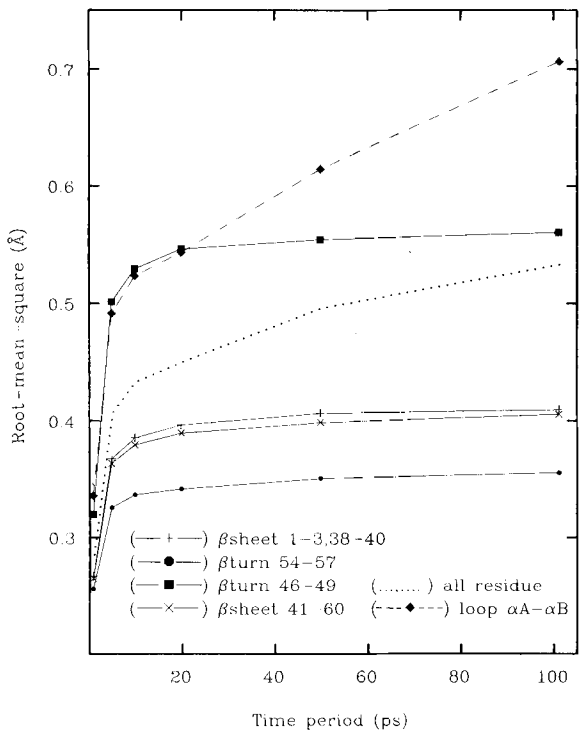

Fig. 4. Time development of the main chain atom fluctuations for the full β-sheet (41–60), the exposed β-turn (46–49), the interior β-turn (54–57), and the small adjacent β-sheet (1–3, 38–40). The time development for loop αA to αB is shown for comparison, as well as that for the whole molecule.

order among the loops. From the thermal factors, the long loop and $3_{10}-2$ show the largest fluctuations, while in the simulation the largest fluctuations come from the loops αA to αB and αC to αD. The fluctuations of the αC to αD loop are due to main chain dihedral transitions (described below). These transitions produce a displacement that is large relative to the picosecond fluctuations, though they are rare events on the time scale of the simulation. A rare event is not adequately sampled over the time period of the simulation, and any large displacement contributed by it to the fluctuation average is likely to be weighted too heavily. This sampling error can produce an overestimate in the calculated fluctuation value for this loop. Another possible cause for the observed relatively smaller scaled thermal-factor is that there is crystal contact involving this loop.[2]

Relationship of Fluctuation Magnitude With Solvent Exposure

The spatial restriction on an element imposed by the protein matrix is expected to be reflected in the fluctuations. One measure of the degree of confinement is the exposed surface area. Buried elements often have low mobility (e.g., αB) and surface-exposed elements often have high mobility (e.g., β-turn 46–49). (See Table I for accessible surface areas and Tables II and III or Fig. 3 for fluctuation values.) However, only a rough correlation exists between the fluctuation magnitudes and the degree of exposure to solvent. Some of the more exposed parts of lysozyme, including the exterior β-strand 41–45 and long loop in the free trajectory, have small or averaged-size fluctuations. An example of an element with a relatively small fraction of solvent exposed surface area but large fluctuations is $3_{10}-1$ in the bound trajectory.

No correspondence between relative mobility and exposure exists for the backbone atoms within a given α-helical structure. The panels of Figure 5 compare the full-simulation fluctuations and the crystallographic thermal fluctuations with the fraction of exposed surface for the individual main chain atoms of the helices. While the solvent exposure shows the helical periodicity of the surface helices A, C, and D, such periodicity does not exist in the fluctuation magnitudes or thermal factors. There is some indication of periodicity for αD in the bound simulation. This feature contrasts with the helical periodicity found in the thermal factors of myohemerythrin, an α-helical protein,[25] and cytochrome c'.[11]

Fluctuation Sampling and Convergence

Before comparing the fluctuations of the free lysozyme simulation with those of the $(GlcNAc)_6$-bound simulation, the statistical variation in the fluctuation values that results from the finite time of the trajectory needs to be considered. The magnitude of the statistical error can be estimated by comparing the fluctuation time development from the two halves of the 101 psec simulation. A pair of curves is shown in Figure 6 for the rms displacement of the main chain atoms in each structural element; the dotted curves are calculated from the atomic displacements of the first 50 psec and the solid line curves from those of the second 50 psec. The dash–dot and dash curves are rms displacements for all residues from the first and second 50 psec, respectively. For some elements, such as the β-sheet, the form of the time development is similar in the two halves, but the actual magnitudes of the fluctuations are different; this results in curves that are parallel. For other elements, such as the αC to αD loop, the time development *and* magnitude of the fluctuations are different and the pair of curves is not parallel. For the nonparallel curves, the difference between the fluctuations in the first and second half of the trajectory increases for the longer time windows of averaging. A third pattern involves curves that cross; this is the case for the αA to αB loop and for αD. Elements with larger fluctuation values do not show greater variations between the two halves of the trajectory.

The differences in the two 50 psec values for a

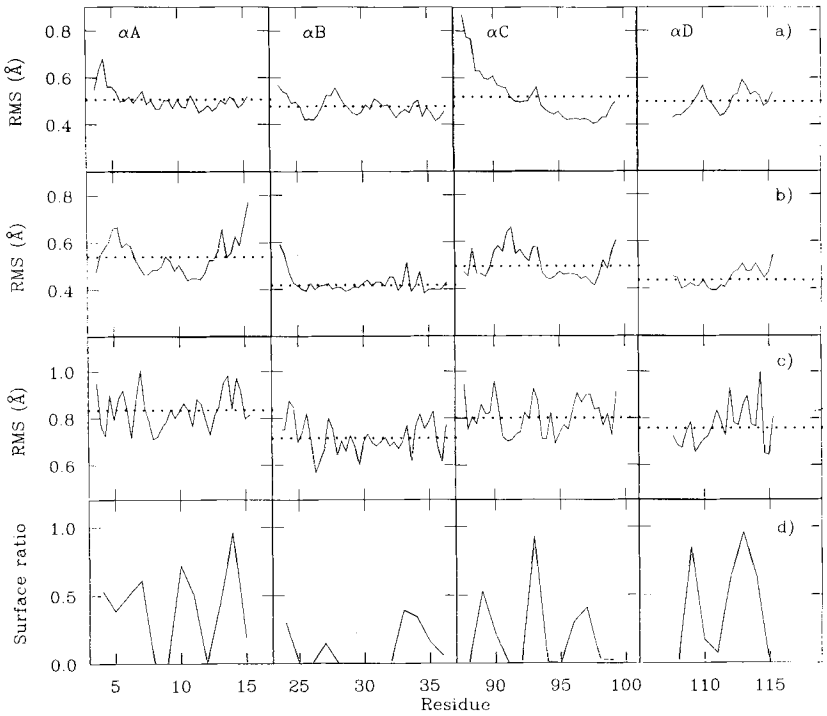

Fig. 5. Positional fluctuations for individual main chain atoms of the four α-helices in lysozyme. For the residue *i* indicated along the abscissa, the values corresponding to N, Cα, and C are plotted at $i-0.3$, i, $i+0.3$, respectively. The dotted line is the average value for the entire helix. **a**: The 55 psec fluctuation values for the bound simulation; **b**: the 101 psec values for the native simulation; **c**: the rms fluctuation calculated from the crystallographic temperature factor; **d**: the ratio of solvent accessible surface of a residue in the protein to that in the isolated helix.

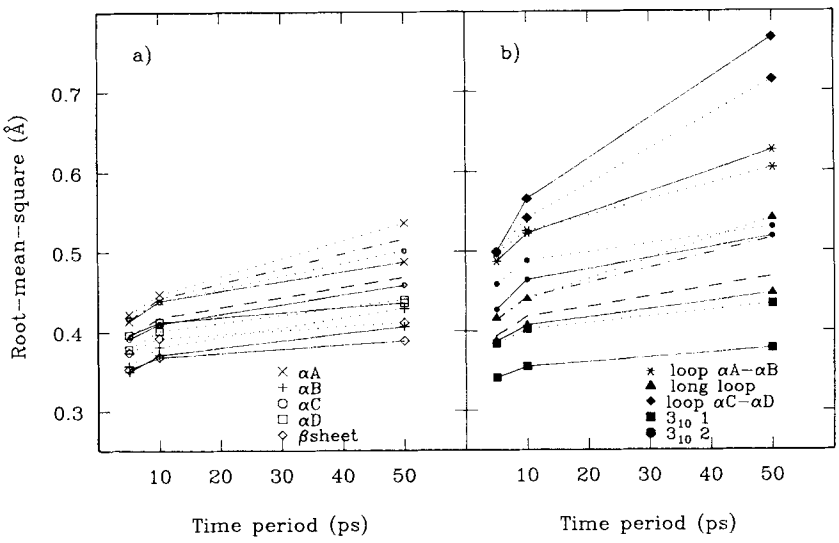

Fig. 6. Comparison of the time development of the main chain atom fluctuations from the two 50 psec halves of the free simulation. **a**: The α-helices and the β-sheet and **b**: the 3_{10} helices and loop structures. Fluctuations were ensemble-averaged from 0 to 50 psec (dotted) or from 51 to 100 psec (solid) for each element. The two curves without symbols correspond to values averaged over main chain atoms of all residues for 0–50 psec (dot–dash) and for 51–100 psec (dashed).

TABLE IV. Absolute Differences Between the 50 psec Fluctuations in the Two Halves of the Simulation of Free Lysozyme and Between Those of the Free and Bound Simulation

Structural element	Residues	50 psec fluctuation difference (Å)			
		Two halves free		Free vs bound	
		Main chain	Side chain	Main chain	Side chain
Molecule	1–129	0.048	0.059	0.061	0.081
α Helix					
A	4–15	0.049	0.056	0.005	0.099
B	24–36	0.023	0.006	0.058	0.099
C	88–99	0.043	0.055	0.040	0.016
D	108–115	0.004	0.010	0.060	0.082
β-Sheet	41–60	0.023	0.032	0.078	0.089
$3_{10}-1$	79–84	0.056	0.025	0.277	0.264
$3_{10}-2$	120–124	0.012	0.001	0.017	0.011
Loop αA–αB	16–23	0.022	0.042	0.029	0.058
Loop long	61–78	0.094	0.116	0.176	0.265
Loop αC–αD	100–107	0.053	0.032	0.251	0.284

given structural element are listed in Table IV and range from less than 0.01 to 0.1 Å for averages over either main chain and side chain atoms. It can be concluded that for differences to be significant they must be greater than 0.1 Å, the maximum deviation in sampling determined from the two halves of the 100 psec trajectory.

Effect of $(GlcNAc)_6$

The magnitudes of the fluctuations of certain structural elements (the αC to αD loop, the long loop, and $3_{10}-1$) are significantly altered in the substrate-bound state compared with the free state. The main chain and side chain fluctuations of these elements change by as much as 0.2–0.3 Å between the free and bound simulation. They are the only elements for which the fluctuations differ by an amount greater than the apparent sampling error of 0.1 Å. There also appear to be substrate effects on the dynamics of αB and αC. Although the difference in the magnitude of the 50 psec fluctuations is smaller than the 0.1 Å sampling error, a clear difference in the slope of the time development curves was found. The slope is changed for $3_{10}-1$ as well. A description follows of the structural aspects of the motions in the elements whose dynamics are affected by $(GlcNAc)_6$.

Loop αC to αD

The loop connecting αC and αD, residues 100–107, includes three reverse turns and interacts extensively with the substrate, as shown in Figure 7. An alteration of the dynamics of the αC to αD loop in the presence of substrate is of particular interest since this element could play both an enthalpic and entropic role in binding.[23] Bannerjee and Rupley[5] measured the entropy and enthalpy of binding at pH values spanning the pK_a of Asp-101 whose side chain hydrogen bonds with the sugars in sites A and B.[16] It was found that while the enthalpy of binding increased by 6 kcal/mol on protonation of Asp-101, consistent with the loss of hydrogen bonding interactions, the entropy of binding was more favorable at the lower pH. Because of these compensating effects, the free energy of binding changes by only 1 kcal/mol.

Large positional changes in the main chain from 100 to 107 occur twice during the free trajectory. These give rise, as already noted, to the low frequency time dependence of the fluctuations and the discontinuous slope in the time development curve (Fig. 3a). The time at which the transitions take place is shown by a time series of the rms coordinate deviations between the 5 psec average structures and the 101 psec average structure for the mainchain atoms N, Cα, C of residues 100–107 (Fig. 8, top panel). The jumps in the rms deviation between 15 and 25 psec and between 65 and 70 psec result primarily from changes in the mainchain dihedral angles, as seen from the time series of φ and ψ values for residues 101–104 in the lower panels of Figure 8. Though ϕ_{101} and ψ_{102} are strongly anticorrelated, there is still a large displacement of the loop. These conformational changes are elucidated by comparing structures before and after each transition. The positional change in the main chain atoms N, H, Cα, C, O between the 5 psec average structure from 10 to 15 ps and that from 20 to 25 ps (Fig. 9a) shows that the first transition involves displacement of main chain atoms from 101 Cα to 103 carbonyl. A similar plot comparing the 60–65 and 65–70 psec structures for the second transition (Fig. 9b) shows positional changes from 103 NH to 105 NH. Figure 9 illustrates the localized nature of the transitions in that the combined rotations of several φ and ψ angles produce large displacements of only a few residues.

A pictorial view of the two transitions is given in Figure 10. The main chain of several structures from the free trajectory is pictured in snapshots be-

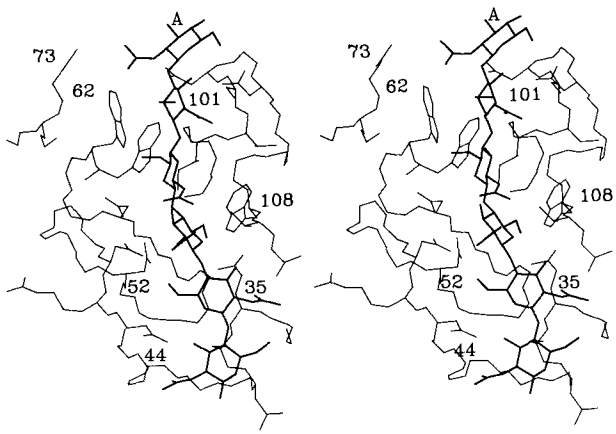

Fig. 7. The active site region of lysozyme. Side chains of the residues in contact with $(GlcNAc)_6$ are shown. These residues are E35, S36, N37, N44, R45, N46, D52, W62, W63, R73, I98, V99, S100, D101, G102, N103, W108, V109, and A110. Binding site A is labeled.

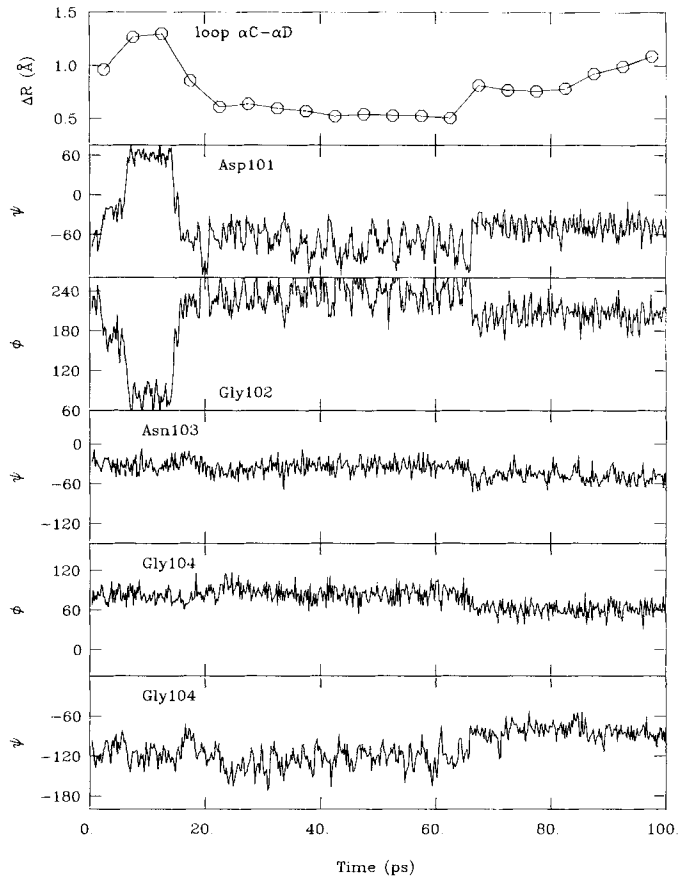

Fig. 8. Analysis of the motion in the free simulation of the loop 100–107, connecting αC and αD. Top panel: time series of the rms coordinate deviation of the main chain atoms of residues 100–107 between the 5 psec average structures and the full simulation average structure. Structural transitions occur near 18 and 65 psec. Lower panels: time series of dihedral angle values of several φ and ψ angles show that the mechanism of the motion for the two transitions is primarily the main chain dihedral rotations of Asp-101, Gly-102, and to a lesser extent the rotations of Asn-103 and Gly-104.

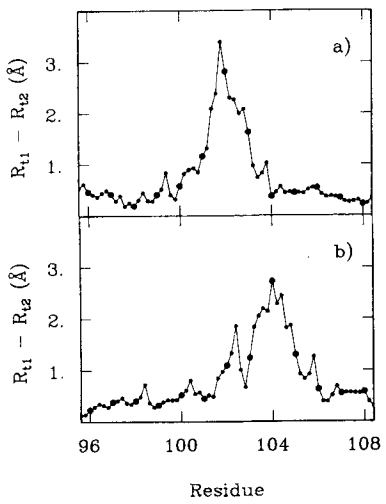

Fig. 9. Displacement of the main chain resulting from the structural transitions at 18 and 65 psec (see Fig. 8). The coordinate difference of individual main chain atoms between two 5 psec average structures is shown, with the abscissa values for N, H, Cα, C, O of residue *i* being *i*−0.4, *i*−0.2, *i*, *i*+0.2, *i*+0.4, respectively. The larger symbols represent the Cα values. **a**: The 18 psec transition, comparing the structure averaged over the period 10–15 psec with that over 20–25 psec. **b**: The 65 psec transition, comparing the structure averaged over 60–65 psec with that over 65–70 psec.

fore, during, and after the transitions. The rotation about a few main chain dihedral angles produces a very clear two-stage displacement of this surface loop.

In the presence of bound substrate, no dihedral transitions in the 100–107 loop are observed and the fluctuations are greatly diminished. It appears that the hydrogen bonds between the carboxylate group of Asp-101 and the saccharides in sites A and B, as well as other enzyme–substrate interactions, anchor the peptide chain so as to eliminate the mobility found in the free enzyme. This loss of motional freedom on binding suggests an interpretation for the experimental observation that the entropy of binding increases for pH values below the pK_a of Asp-101, as described in the beginning of this section; protonation of Asp-101 not only eliminates the hydrogen bonding interaction but also reduces the order in the system by allowing the loop to retain some of the flexibility of the free state. A simulation to test this suggestion with a protonated Asp-101 would be of interest.

Long Loop and 3_{10}−1

Altered fluctuations in the bound simulation were also found for the long loop and for the following one-turn helix, 3_{10}−1. In contrast to the αC to αD loop, significantly larger fluctuations are observed in the *bound* simulation (compare Fig. 3a and 3b). The specific residues involved are 67–87, with the N-terminus of αC showing a slight increase in fluctuations (see Fig. 5a). Residues 85–87 connect 3_{10}−1 to αC and have 50 psec fluctuation values of 0.424 and 1.053 Å for the free and bound simulation, respectively. It should be noted that the time development of 3_{10}−1 is very slow compared with other structural elements in the simulation (see Fig. 3b). These differences between bound and free lysozyme cannot be due to direct interaction with $(GlcNAc)_6$ since the substrate makes almost no contact with this part of the enzyme (Fig. 7); a single hydrogen bond from this region of the protein is formed between the Arg-73 guanidinium group and acetamide group in site A. Some evidence that this region undergoes a conformational change when substrate binds comes from the crystallographic structure for a trisaccharide bound in sites B, C, D, in which the conformation of residues 70–75 differs from that of the free structure.[17]

Time series of the rms coordinate deviations for the long loop and 3_{10}−1 in the bound trajectory between the 5 psec average structures and the energy-minimized crystal structure are shown in Figure 11. To analyze the nature of the motions involved two types of comparisons are made in Figure 11; first coordinate deviations are calculated after a fit of the entire protein (solid line), and second, after a fit of only the main chain atoms of either the long loop or 3_{10}−1 (dashed line).

For 3_{10}−1, it is clear that the large increase in the rms deviations at 30 psec is a movement of the whole helix relative to the rest of the protein, and is not a disruption of the internal helical structure. A fit of only the main chain atoms of 3_{10}−1 leads to a time progression of the rms coordinate deviations that is smaller and essentially constant throughout the trajectory. In contrast to the rigid body type motion of 3_{10}−1, a change of the internal structure of the long loop is indicated by the time series of rms coordinate deviations shown in Figure 11. Both types of fit yield an increase in rms deviation near 18 psec. The change at 35 psec, however, is not seen when fitting the loop structure alone, so that it corresponds to a rigid body motion of the loop relative to the protein, similar to that of 3_{10}−1.

The dynamics of main chain dihedral angles of the long loop were analyzed to determine the mechanism of the internal structural change near 18 psec. They involve the two peptide groups flanking Gly-71. Unlike the main chain dihedral transitions in loop 100–107 that occur over approximately 1 psec, the peptide Gly-71:Ser-72 transition takes place over a 5 psec period. In the free simulation, these dihedral angles oscillate in the neighborhood of their average values, showing no transitions between potential minima in accord with the small positional fluctuations in that trajectory.

The main chain displacements in the long loop internal transition near 18 psec involves Gly67, Gly-71, and Ser-72 while the rigid-body displace-

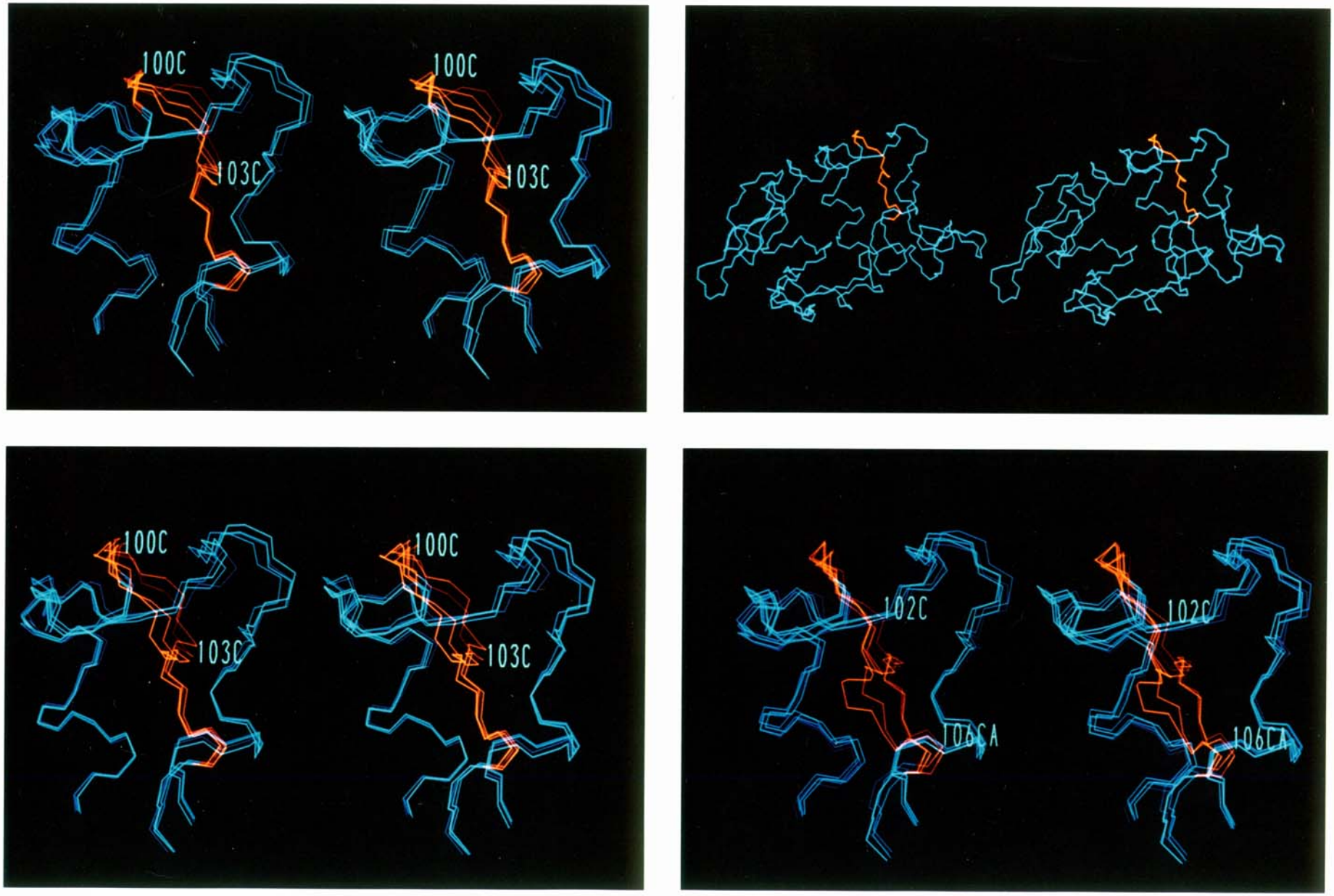

Fig. 10. A pictorial view of the movement of the loop 100–107 (orange) showing mainchain atoms, N, Cα, C. Top right: single structure, whole molecule. The expanded loop region with sequential 1 psec average structures is drawn for the transitions shown in Figure 8. Top left: 14–17 psec. Bottom left: 18–21 psec. Bottom right: 63–67 psec.

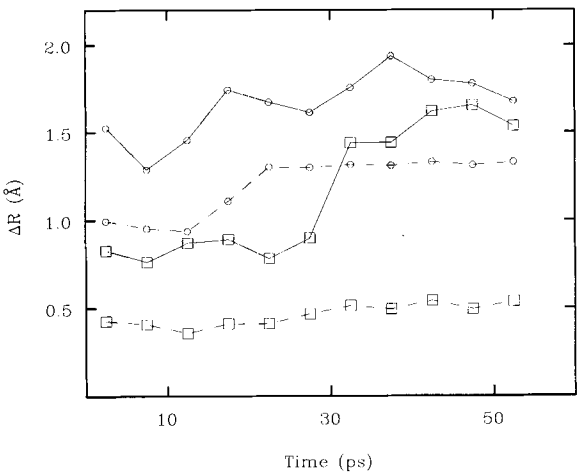

Fig. 11. A time progression of the main chain displacement for N, Cα, C in the long loop (61–78) (○) and $3_{10}-1$ (79–84) (□) from a comparison of the 5 psec average structures from the substrate-bound simulation and the optimized crystal structure. Differences were calculated after a least-squares fit of the entire 1001 lysozyme heavy atoms (solid) and after a fit of the structural element (dashed).

ment of $3_{10}-1$ near 30 psec involves the entire 3_{10} helix plus adjacent residues (81–86). The rigid-body displacements are smaller than those of residues 67, 71, and 72 and of the αC to αD loop of the free enzyme in Figure 9. All three involve main chain motion that is localized in six or fewer residues. The mechanism of motion for the rigid displacement of $3_{10}-1$ relative to the protein does not involve abrupt transitions, as was found for the long loop and loop αC to αD. Dihedral angle time series for ϕ and ψ or residues 75–88 were examined and no transitions (a change in dihedral angle greater than the fluctuations of the angle) occurred during the period 25–35 psec. Instead, the structural change at 30 psec in Figure 11 was found to be the result of small, cooperative rotations of ψ87 and ψ88. This feature of the dynamics is seen in the differences between ϕ and ψ in the 5 psec average structures (Fig. 12a). There are significant differences in ψ87 (16°) and ψ88 (10°), resulting in the main chain displacement from 81 to 86 (Fig. 12b). Over a 5 psec time period, these dihedral angles oscillate around slightly different values, yet remain within a single dihedral-energy minimum. These subtle changes in the angles can be seen in the dihedral time series in Figure 12c and d.

Other differences in the behavior of main chain dihedral angles in the long loop and $3_{10}-1$ between free and bound lysozyme were found. First, the peptide Arg-73:Asn-74 and ϕ86 underwent short-lived excursions of 60° or more in the bound simulation. These transitions were of a transient jump-and-return nature, unlike the behavior in Figure 8 where the new dihedral values remained and resulted in a change in the average structure. Second, there are alternative configurations in the two simulations for Cys-76 ϕ (−125° free; −75° bound) and Asn-77 ψ (50° free; −50° bound). The larger oscillations in a single dihedral potential well (in contrast to the transitions between wells described above) and the possibility of less curvature in the potential

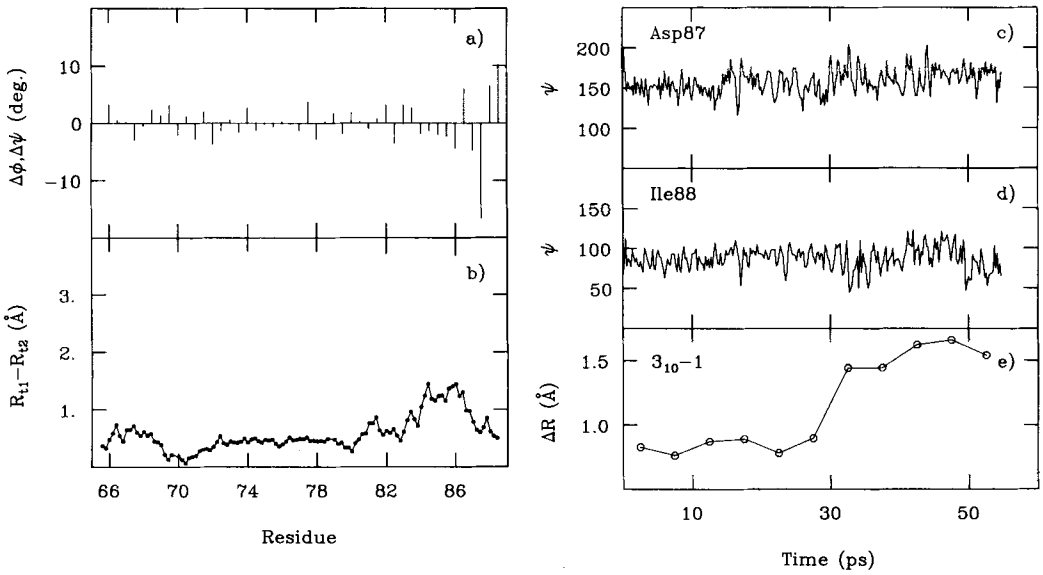

Fig. 12. Analysis of the rigid-body transition of $3_{10}-1$ at 30 psec in the bound simulation. Differences are examined in **a** main chain dihedral angles and **b** main chain atomic displacements (see Fig. 8 caption) between the 5 psec average structures from 25 to 30 psec and from 30 to 35 psec. The rigid-body transition is produced by small, concerted rotations of ψ87 and ψ88, the largest deviations in (**a**). These subtle differences in the average structure dihedral angles can be seen in the dihedral time series (**c**) and (**d**), although the change is not as obvious as the transitions in Figure 8. The time series of coordinate deviations is shown (**e**) to indicate the position of the transition along the trajectory.

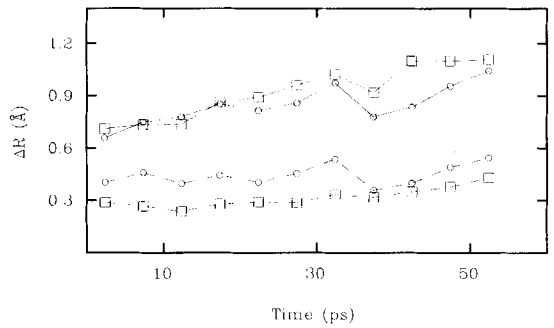

Fig. 13. The time progression of the coordinate deviations for the main chain atoms (N, Cα, C) of αB (○) and αC (□) in the bound simulation from comparing the 5 psec average structures with the optimized crystal structure. The positional differences were calculated after fitting either all lysozyme heavy atoms (solid) or fitting only the main chain atoms of the helix (dashed). A slow displacement of the two helices within the molecule frame occurs but the internal structure of the helix is not altered.

surface associated with an alternative conformation also relate to the larger fluctuations in the bound simulations.

Helices αB and αC

The fluctuations in the bound simulation (Fig. 3b) of αB and αC increase slowly and continuously up to 55 psec, unlike those of αA, αD, or the β-sheet, e.g., for αC the magnitude is less than that of the β-sheet up to 20 psec and then becomes larger. In the free simulation, the time development curves of these secondary structural elements all have similar slopes (Fig. 3a). The slow time development of αB and αC helices is similar to that for $3_{10}-1$. These long time scale fluctuations result from a change in helical position within the enzyme that takes place rather uniformly over the time of the simulation. A time series of rms deviations comparing 5 psec average structures with the energy-minimized crystal structure after a least-squares fit of the protein is a slowly increasing function (Fig. 13). Nevertheless, as was found for $3_{10}-1$, the internal helical structure of αB and αC is maintained; after fitting only the corresponding helix main chain atoms, the rms deviation value is constant. Consequently, Figures 3b and 13 reflect a gradual, rigid-body departure of the helices from their original positions.

The movement of αB and αC within the framework of the protein is pictured in Figures 14 and 15, which compare the initial and final 5 psec average dynamic structures. The fitting of the two structures is with respect to the main chain atoms of the β-sheet since the rms deviations in the β-sheet between the crystal structure and all of the 5 psec average structures remained constant. As illustrated in Figure 14, the N-terminus of αB moves in the direction of the loop αC to αD. The movement of the C helix during the simulation is seen in Figure 15 to be in the direction of its N-terminus, away from the active site and toward $3_{10}-1$. Positions of $3_{10}-1$ and residues 85–87, connecting $3_{10}-1$ and αC, are also shown in Figure 15. The larger fluctuations of residues 85–87 in the bound simulation are a result of it being the link between the mobile structural elements $3_{10}-1$ and αC.

It is possible that the continuous, relative displacements of αB and αC are adjustments to the bound $(GlcNAc)_6$. The carboxyl termini of both helices contact the substrate. The initial structure for the bound simulation was determined from crystals soaked with $(GlcNAc)_3$; thus, crystal contacts would have precluded significant conformational changes. Moreover, the addition of the three saccharide units in sites D, E, and F by building onto the crystal structure could also induce the changes observed in the bound simulation.

Fig. 14. Change in the relative position of αB during the bound simulation. The Cα positions of the first 5 psec average structure (0–5 psec) plus the main chain N and C atoms of helix B are connected by thin lines. Superposed, after a rotation by a least-squares fit of the β-sheet, are the N, Cα, C atoms of αB from the last 5 psec average structures (50–55 psec) (thick line).

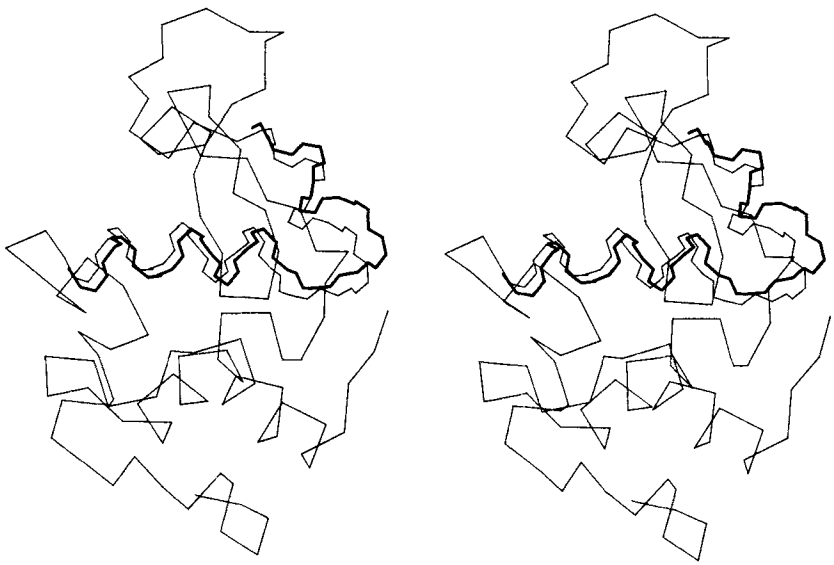

Fig. 15. Change in the relative position of αC and $3_{10}-1$ during the bound simulation. See Figure 14. The last 5 psec average structure main chain (thick line) is shown for residues 79–99.

CONCLUSIONS

Structural transitions observed in the present simulations are of two types. For loops mainly internal dihedral angle changes are involved, while for helices the displacements involve the entire secondary structural element, in accord with the analysis of the multiple minima of a myoglobin simulation.[10] For example, concerted rotation of main chain dihedral angles resulted in a rather localized displacement of the αC to αD loop. The dynamics of this loop are consistent with the conformational heterogeneity in residues 100–104 among the crystal structures of tetragonal and triclinic hen lysozyme and of human lysozyme[15]. In the simulation, the structure of this loop is closer to that in the tetragonal crystal structure (the initial coordinates for the simulation) during the period from 30 to 65 psec (i.e., between the two transitions) than near the beginning and end of the simulation. Thus, the simulation results have allowed visualization of an interesting dynamic process, although the low probability of such an event does not make it possible to determine the equilibrium between the different conformations without a much longer simulation.

Substrate binding altered the dynamics of the coiled loops of lysozyme more than the motions of the helical and sheet structures. It is reasonable that the regular hydrogen-bonding pattern in helices and sheets preserves the low-magnitude fluctuations both in the presence and absence of substrate. By contrast the more irregular nonbonding interactions in the coiled-loop regions allow a larger number of configurations, which are readily altered by the substrate. A modulation of the populations could be exploited to optimize binding of substrate and/or release of product by altering the entropy of formation of the complex. The two significant changes in the fluctuations of $(GlcNAc)_6$-bound lysozyme, relative to free lysozyme, were a decrease in fluctuations at the active site and an *increase* in the fluctuations of a region away from the active site. In the absence of substrate, the active site residues experienced main chain dihedral transitions, consistent with the observed crystallographic heterogeneity in this region among different lysozyme structures. The loss of flexibility in the presence of substrate is readily understood in terms of the specific enzyme–substrate interactions and offers an explanation of experimentally measured binding entropies. We do not have a complete explanation for the substrate-enhanced fluctuations away from the active site, although we have described here some of the motions involved.

The absence of solvent, other than the strongly bound waters, in the lysozyme simulations is unlikely to alter the relationship between dynamics and structural elements described here. While results from simulations in water environments indicate that the dynamics of surface residues are influenced by solvent, the solvent effect on fluctuations of interior atoms is small[28,29]; differences in main chain fluctuations found for the various secondary structural elements in vacuum simulations were confirmed by simulations in the presence of van der Waals or aqueous solvent. Further, the degree of solvent exposure in lysozyme was found not to correlate with the mobility of a structural element. Thus, the present analysis of the relations between dynamics and structure is unlikely to be an artifact

of the vacuum simulation and suggests that the much greater expense in computer time required for a fully solvated simulation is not justified for the present problem. With regard to the mobility of the solvent-exposed loop (αC to αD), solvent would not eliminate the freedom observed when the loop is not constrained by substrate contacts. However, the relaxation time of the loop motion obtained from the simulation is likely to be too short.[27]

Further analyses of the dynamic and structural features related to $(GlcNAc)_6$–lysozyme interactions is in progress. It is hoped that experimental studies, including NMR (which has already been employed in the study of lysozyme)[30,31] and site-directed mutagenesis, will be of use in examining the origin of the changes in mobility predicted by the simulation to result from saccharide binding to lysozyme.

ACKNOWLEDGMENTS

We gratefully acknowledge assistance with the calculations and helpful discussions with B.R. Brooks, P.J. Artymiuk, J.C. Cheetham, and D.C. Phillips. The simulation on which the present results are based was performed at Daresbury with support from the Science and Engineering Research Council. Work at Harvard was supported by the National Institutes of Health. C.M.D. is a member of the Oxford Center for Molecular Sciences.

REFERENCES

1. Ansari, A., Berendsen, J., Bowne, S. F., Frauenfelder, H., Iben, I. E. T., Sauke, T. B., Shyamsunder, E., Young, R. D. Protein states and proteinquakes. (1985) Proc. Natl. Acad. Sci. U.S.A. 82:5000–5004, 1985.
2. Artymiuk, P. J., Blake, C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C., Sternberg, M. J. E. Crystallographic studies of the dynamic properties of lysozyme. Nature (London) 280:563–568, 1979.
3. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H., Gunsalus, I. C. Dynamics of ligand binding to myoglobin. Biochemistry 14:5355–5373, 1975.
4. Baker, E. N. Structure of actinidinim after refinement at 1.7 resolution. J. Mol. Biol. 141:441–484, 1980.
5. Bannerjee, S. K., Rupley, J. A. Temperature and pH dependence of the binding of oligosaccharides to lysozyme. J. Biol. Chem. 248:2117–2124, 1973.
6. Berthou, J., Lifchitz, A., Artymiuk, P., Jolles, P. An X-ray study of the physiological-temperature form of hen egg-white lysozyme at 2Å resolution. Proc. R. Soc. Lond. B 217: 471–489, 1983.
7. Blake, C. C. F., Johnson, L. N., Mair, G. S., North, A. C. T., Phillips, D. C., Sarma, V. R. Crystallographic studies of the activity of hen egg-white lysozyme. Proc. Roy. Soc. Lond. B 167:378–388, 1967.
8. Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C., Kraut, J. Crystal structures of Escherichia coli and Lactobacillus casei dihydrofolate reductase refined at 1.7 resolution. J. Biol. Chem. 257:13650–13662, 1983.
9. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. 4:187–217, 1983.
10. Elber, R., Karplus, M. Multiple conformational states of proteins: A molecular dynamics analysis of myoglobin. Science 235:318–321, 1987.
11. Finzel, B. C., Weber, P. C., Hardman, K. D., Salemme, F. R. Structure of ferricytochrome c′ from Rhodospirillum molischianum at 1.67 resolution. J. Mol. Biol. 186:627–643, 1985.
12. Frauenfelder, H., Petsko, G. A. Fluctuations in macromolecular structure. Biophys. J. 32:465–483, 1980.
13. Gavish, B. Protein main-chain atomic displacements and density of stabilizing interaction. Biophys. Struct. Mech. 10:31–45, 1983.
14. Hendrickson, W. A., Teeter, M. M. Structure of the hydrophobic protein crambin determined directly from anomalous scattering of sulphur. Nature (London) 290:107–113, 1981.
15. Ichiye, T., Olafson, B. D., Swaminathan, S., Karplus, M. Structure and internal mobility of proteins: A molecular dynamics study of hen egg white lysozyme. Biopolymers 25:1909–1939, 1986.
16. Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., Rupley, J. A. Vertebrate lysozymes. The Enzymes 7:665–868, 1972.
17. Kelly, J. A., Sielecki, A. R., Sykes, B. D., James, M. N. G. X-Ray crystallography of the binding of the bacterial cell wall trisaccharide NAM-NAG-NAM to lysozyme. Nature (London) 282:875–878, 1979.
18. Lee, B., Richards, F. M. The interpretation of protein structure: Estimation of static accessibility. J. Mol. Biol. 55:379–400, 1971.
19. Nadler, W., Schulten, K. Theory of Mossbauer spectra of proteins fluctuating between conformational substrates. Proc. Natl. Acad. Sci. U.S.A. 81:5719–5723, 1984.
20. Northrup, S. H., Pear, M. R., McCammon, J. A., Karplus, M., Takano, T. Internal mobility of ferrocytochrome c. Nature (London) 287:659–660, 1980.
21. Petsko, G. A., Kuriyan, J., Gilbert, W. A., Ringe, D., Karplus, M. "Structural Biological Applications," Barturnik, H., Chance, B. (eds.). New York: Academic Press, 1984:99–109.
22. Petsko, G. A., Ringe, D. Fluctuations in protein structure from X-ray diffraction. Annu. Rev. Biophys. Bioeng. 13:331–371, 1984.
23. Post, C. B., Brooks, B.R., Dobson, C.M., Artymiuk, P.J., Cheetham, J.C., Phillips, D.C., Karplus, M. Molecular dynamics stimulations of native and substrate-bound lysozyme. J. Mol. Biol. 190:455–479, 1986.
24. Rose, G. D., Young, W. B., Gierasch, L. M. Interior turns in globular proteins. Nature (London) 304:654–657, 1983.
25. Sheriff, S., Hendrickson, W. A., Stenkamp, R. E., Sieker, L. C., Jensen, L. H. Influence of solvent accessibility and intermolecular contacts on atomic mobilities in hemerythrins. Proc. Natl. Acad. Sci. U.S.A. 82:1104–1107, 1985.
26. Shrake, A., Rupley, J.A. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. J. Mol. Biol. 79:351–371, 1973.
27. Smith, J.L., Hendrickson, W.A., Honzatho, R.B., Sheriff, S. Structural Heterogeneity in Protein Crystals. Biochemistry 25:5018–5027, 1986.
28. Swaminathan, S., Ichiye, T., van Gunsteren, W.F., Karplus, M. Time dependence of atomic fluctuations in proteins: Analysis of local and collective motions in bovine pancreatic trypsin inhibitor. Biochemistry 21:5230–5241, 1982.
29. van Gunsteren, W. F., Berendsen, H. J. C. Computer simulation as a tool for tracing the conformational differences between proteins in solution and in the crystalline state. J. Mol. Biol. 176:559–564, 1984.
30. van Gunsteren, W. F., Karplus, M. Protein dynamics in solution and in a crystalline environment: A molecular dynamics study. Biochemistry 21:2259–2274, 1982.
31. Redfield, C., Dobson, C.M. Sequential ^1H NMR Assignments and Secondary Structure of Hen Egg White Lysozyme in Solution. Biochemistry 27:122–136, 1988.
32. Karplus, M., Dobson, C.M. Internal Motions of Proteins: Nuclear Magnetic Resonance Measurements and Dynamics Simulations. Methods. Enzymology 131:362–389, 1986.