Distal Cavity Fluctuations in Myoglobin: Protein Motion and Ligand Diffusion†

Mark L. Carlson,‡ Rebecca M. Regan, and Quentin H. Gibson*

Department of Biochemistry and Molecular and Cell Biology, Cornell University, Ithaca, New York 14853

Received July 31, 1995; Revised Manuscript Received November 2, 1995⊗

ABSTRACT: Experimentally, distal mutations in myoglobin substantially affect the contribution of fast and slow phases to picosecond geminate recombination of NO following flash photolysis. Earlier simulations of ligand diffusion among distal pocket mutants showed greatly differing rates of collisions between the ligands and the heme iron, suggesting that distal residues affect recombination by controlling ligand access to the iron [Gibson, Q. H., Regan, R., Elber, R., Olson, J. S., & Carver, T. (1992) *J. Biol. Chem. 267*, 22022−22034). In this work, molecular dynamics simulations of sperm whale myoglobin and mutations at positions 68 (E11) and 107 (G8) have been examined to investigate the structural mechanism that controls ligand diffusion and iron accessibility. Visualization of the distal ligand-accessible spaces shows a pattern of cavities (common to other hemoglobins and myoglobins) that fluctuate and interconnect due to protein motions. Access to the iron atom is highly sensitive to these fluctuations in the native structure, perhaps a reason for the strong conservation of distal residues. The positions of the helices surrounding the distal heme site were monitored to assess the involvement of more collective protein motions in ligand diffusion. Ligand migrations and collisions with the iron appear related to expansion of the distal protein matrix due to helix movements. The helices surrounding the distal site also make relative adjustments on the order of 0.5 Å to accommodate the presence of a mobile diatomic ligand, suggesting a mechanism for communication between the heme site and the exterior of the protein.

Structural fluctuations in myoglobin and hemoglobin are necessary for binding oxygen since crystal structures do not show an open passageway between the heme iron and the external solution (Perutz & Mathews, 1966). This is in contrast to some catalytic heme proteins whose crystal structures show channels leading to interior active sites (Murthy et al., 1981). The functionally relevant fluctuations of oxygen-binding heme proteins must at least transiently interconnect atom-sized cavities, which are sparse in crystal structures (Lee & Richards, 1971). Protein motion is also influential on time scales much shorter than for ligand escape and has been invoked to explain the complexity of ligand recombination following photolysis even on the picosecond time scale.

Molecular dynamics simulation is a tool well-suited for studying myoglobin dynamics. Sound starting structures for simulations are provided by high-resolution crystal structures which now exist for numerous mutant and wild-type hemoglobins and myoglobins in a variety of chemical states. Current computational power makes 50 ps simulations of systems containing several thousand atoms easily accessible. Consequently, numerous trajectories can be sampled for each of several related structures, allowing one to identify and compare representative behavior among them. [A single starting structure may produce a wide range of results due only to different initial atomic velocities (Gibson et al., 1992).] The overlap of the simulation time scale with experimental data permits molecular dynamics to suggest plausible molecular mechanisms for observed picosecond kinetics.

Experimentally, distal mutations have a profound effect upon the rebinding kinetics, substantially altering the contributions of fast and slow phases to geminate recombination (Carver et al., 1990; Gibson et al., 1992; Ikeda-Saito et al., 1993), and these differences are reflected in simulations. As the molecular dynamics algorithms used do not take the chemical reactions into account, this correlation suggests that residues in the distal pocket modify ligand access to the iron, and hence recombination, sterically. Significant distal effects do not, of course, exclude an influence of proximal residues on ligand binding, which should be regarded as well established; see e.g., Ahmed et al. (1991) and Duprat et al. (1995).

Given the success of diffusion simulations to represent picosecond geminate recombination, a detailed picture of the steric mechanism controlling ligand diffusion is needed. Experiments involving solvent viscosity (Austin et al., 1975) have implicated the movements of entire helices, as well as the gyrations of individual residue side chains, in facilitating ligand diffusion in Mb. Another important question is the extent to which a dissociated ligand influences the structure and dynamics of the protein. Following photolysis, the structure of Mb relaxes on the time scale of tens of picoseconds or longer based on various spectroscopic measurements (Richard et al., 1992; Dasgupta & Spiro, 1987). The driving force for structural reorganization in Mb after breaking the ligand bond was assumed to arise from the 0.5 Å displacement of the heme iron and related proximal effects (Takano, 1977; Kuriyan et al., 1986). As a result, little attention has been paid to the impact of a diffusing, diatomic ligand which could potentially act as a 1.5 Å wedge in the fluctuating distal protein matrix.

† This work was supported by U.S. Public Health Service Grant GM 14276.

* To whom correspondence should be addressed.

‡ Present address: Department of Physics, Illinois Benedictine College, 5700 College Road, Lisle, IL 60532.

⊗ Abstract published in *Advance ACS Abstracts,* January 15, 1996.

0006-2960/96/0435-1125$12.00/0

In this work, the ligand-accessible distal cavities are visualized and quantified as they fluctuate in time to better understand the steric constraints placed on ligand diffusion. This offers more information than do ligand positions alone since the cavities show directly when ligands are sterically prevented from encountering the iron and by which residues. It is also possible to compare actual and potential ligand movement. Previous efforts to describe distal barriers to diffusion have been limited to the conformations of key side chains or the separation distances between pairs of residues. In contrast, the fluctuations of the distal cavities reflect the motion of a greater portion of the protein since many residues from several helices contribute to these spaces.

METHODS

The molecular dynamics software package MOIL (version 6.0) (Elber et al., 1994) was employed, utilizing its implementation of the LES (locally enhanced sampling) algorithm. In these calculations, ten ligand copies experience the protein atoms (but not each other) while the protein sees their average force. The cutoff radius for nonbonded interactions was 10 Å and the 1−4 scaling factors were 2 for electrostatic and 8 for van der Waals interactions. All crystallographic waters, roughly 150−200, were included and modeled as TIP3 (Jorgensen et al., 1983). The starting coordinates came from crystal structures which had been energy minimized following insertion of polar hydrogens. The following PDB crystal structures (bound CO unless otherwise noted) were utilized: wild-type sperm whale Mb (2MKG) and its Val68Ala mutant (1MLF), horse met-Mb (1YMB), human Hb (1HCO), and homodimeric clam Hb (*Scapharca inaequivalvis*) (3SDH). As there are no crystal structures for the sperm whale Mb mutants Ile107Val and Val68Ala•Ile107Val, structures were created by energy minimizing wild-type and Val68Ala crystal structures after removing the $C_{\delta 1}$ atom of Ile107. Simulations of wild-type Mb without ligands used the bound CO structure with the ligand removed. Multiple (minimum 10) 50 ps simulation runs, differing only in the atoms' initial velocities, were examined for each protein. For human hemoglobin, only $\alpha\beta$ dimers were simulated. The temperature of the protein was raised to room temperature over 15 ps.

To quantify the access of the ligands to the binding site during the course of the simulations, the number of ligand atoms within 4 Å of the iron was counted at 1 ps intervals during each simulation.

The selection of atomic radii for use in identifying ligand-accessible cavities and estimating volumes within proteins is somewhat arbitrary. In MOIL, the potential function describing close encounters between atoms, although steep (r^{12}), is not infinite; atoms are not modeled as hard spheres. As a result, the distance of closest approach between colliding atoms depends on the kinetic energy. Values of atomic radius appropriate for kinetic energy = kT were used. (As in MOIL, pair distances for dissimilar atoms are the geometric mean of those for each atom.)

The surfaces that enclose the space accessible to the center (nucleus) of a carbon atom were computed to visualize the cavities available to the ligand. (In the figures, the color of the cavity surface is the same as that of the residue generating it.) For this purpose, the radii assigned to the protein atoms were the pair distances, which reflect the radii for both protein and ligand and are similar to a procedure of Richards (Lee & Richards, 1971). These distances range from 0.85 Å between an explicit hydrogen and a ligand carbon to 3.15 Å between a methyl carbon (united hydrogens) and a ligand carbon. Using pair distances excludes the volume of the ligand itself and leaves only the volume in which the ligand could move. One significant advantage is that discrete ligand sites can be seen by inspection (Figure 1). Superimposing the spherical shape of the atom on the space it can explore (Tilton et al., 1986) obscures the shapes of cavities and can spuriously make isolated locations appear connected.

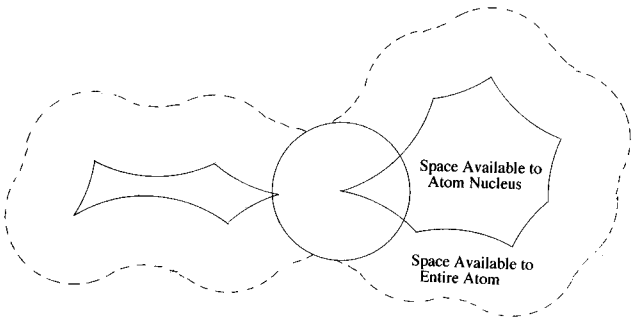

FIGURE 1: Volumes accessible to the nucleus vs the entire atom. The volume in which an atom's center can move is substantially smaller than that which would be swept out by the whole atom. Unlike the entire volume, the nuclear-accessible volume clearly shows when an atom cannot pass between cavities as the figure demonstrates.

Volumes within proteins were estimated by counting the points that remain on an embedded cubic lattice (10 Å in length) after removal of those points which are too close to protein atoms. Protein−ligand atom pair distances were again used to estimate the fractional volume available to the nucleus of a carbon atom while individual atomic radii were used to estimate the fractional void (interatomic space) volume, the fraction of the total volume not occupied by protein atoms.

To quantify the extent to which the B, E, and G helices and the CD interhelical region collectively impinge upon the distal site, the α-carbons of residues 29, 43, 68, and 107 (sperm whale numbering) were used to construct a pyramid (Figure 2). The change in the volume of this pyramid is taken to reflect the expansion or contraction of that portion of the protein which surrounds the distal site.

RESULTS

Although there is considerable interatomic (void) volume near the distal binding site in myoglobin, only a small fraction is available for ligands to move about during molecular dynamics simulations. Figure 3A shows the extent to which the region near the heme binding site is occupied by protein atoms at an arbitrary instant during a typical wild-type myoglobin simulation. The unoccupied (void) volume is roughly half the total (based on our estimates which are described in Methods) and is interwoven among the heme and all distal residue side chains. (Each group has some room in which to move before it rebounds off neighboring residues.) The fraction of void volume decreases to about 25% as one moves through the protein interior away from the heme and toward the N-terminus. Lower packing densities around protein active sites have previously been observed in crystal structures (Richards, 1977), but the value for the distal pocket of myoglobin is more extreme. Figure 3B,C shows the limited space in which a ligand is free to

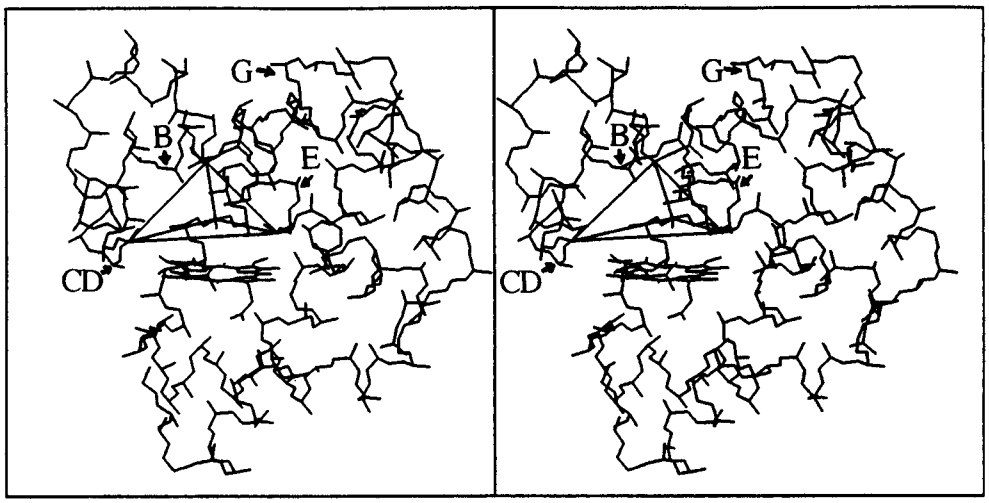

FIGURE 2: Distal pyramid measure of protein expansion (in stereo). The α-carbons of residues 29, 68, 107, and 43 in sperm whale Mb are connected to form a pyramid. The volume was used to measure the expansion or contraction of the distal region due to the movements of the B, E, and G helices and CD interhelical region, respectively.

move at an arbitrary time during a typical wild-type simulation. It is on the order of a few percent of the void volume. (The cavities shown represent the volume accessible to the center of a carbon atom.) Consequently, during simulations ligands are not free to diffuse among arbitrary groups of side chains but are instead confined to a few cavities which may be interconnected at times.

Estimates of volumes within proteins are inherently only relative. Since the separation distance at which colliding atoms rebound depends on the kinetic energy, the atoms do not have a definite radius. As a result, absolute occupied or unoccupied volumes do not exist. Although the typical repulsion potential for colliding atoms is a very steep function of separation distance, differential penetration of a few tenths of an angstrom occurs among atoms sampled from a room temperature Maxwell−Boltzmann distribution. The result is that a ligand from the high-energy end of the velocity distribution sees a significantly larger protein cavity in which to diffuse than one from the low-energy end. Figure 3B,C shows the spaces accessible to a carbon nucleus when the kinetic energy corresponds to the 15th and 85th percentiles in the energy distribution, respectively. In addition, the effective size of a cavity depends upon the chemical identity of the ligand atom in question. Cavities would appear slightly larger to oxygen than to carbon owing to the former's smaller radius. [The dependence of effective cavity volume upon test particle radius has been shown previously (Tilton et al., 1986).]

Figure 3D shows the positions of ligand clusters at the end of each of several runs for wild type and minor mutants (differing by one or two carbons) of residue 68 (Val in wild type). For this group of simulations, three regions are collectively defined. (The runs were selected to show the extent of the commonly visited regions and do not necessarily reflect the frequency of occupation.) One is just to the rear of the iron (rear being toward the vinyl bearing B and C pyrroles), and there are two more, each over a vinyl group. We have previously reported that, in our room temperature simulations, ligands reside in these general locations (Gibson et al., 1992; Carlson et al., 1994; Quillin et al., 1995). However, the constancy of these regions despite the steric and initial velocity differences of the various simulations is surprising. [It is possible to disrupt this pattern of three regions by mutation, but this typically requires the insertion or removal of large groups such as histidyl or phenyl rings (Carlson et al., 1994; Quillin et al., 1995).] Furthermore, visualization of the ligand-accessible spaces show that these few ligand locations dominate our observations because steric barriers more or less prevent further ligand exploration at room temperature. From this point of view, panels B and C of Figure 3 are very representative. Aside from connecting to each other, these three regions are nearly always isolated from the rest of the protein interior and the external solvent. Conspicuously absent are any cavities in the directions of either putative escape route (Case & Karplus, 1979), which might be expected to connect with the occupied ones given sufficient time. The most likely direction for ligand exploration at longer simulation times is deeper within the protein, and the prospective route is shown in Figure 3C. The cavity over the right rear vinyl can connect with another cavity to the interior of the protein, in the direction of the H helix, and while the actual connection is very infrequent, it is almost the only extension that is observed.

The key to describing ligand diffusion in the distal matrix of myoglobin is the significant time-dependent fluctuations of the distal, ligand-accessible cavities. The size, shape, and connectivity of these spaces change rapidly during room temperature simulations. Figure 4 follows typical behavior in the wild type over 5 ps consecutively. In this example trajectory, the ligand copies reside only in the central cavity. These few frames summarize observations for numerous wild-type simulations surprisingly well. Ligand copies spend most of their time in the central cavity, moving in and out of collision range (4 Å) of the iron. Examination of the shape of that central lobe shows that the portion extending over the iron (it dips down) varies from being the bulk of the cavity to forming a remote corner. When ligands do manage to diffuse away from the iron in wild type, they are almost always found over the right rear (B pyrrole) vinyl. The ligand-accessible spaces of Figure 3 show a persistent cavity there. However, the small size and short life of the connection to the central cavity explain the infrequent ligand passage between them that is observed in simulation. Additionally, ligands are rarely found over the left (C pyrrole) vinyl in wild-type simulations. This is quite understandable given the infrequent existence of the corresponding cavity. One can also see from the frames of Figure 4 that the visited cavities over the heme are close to making contact with

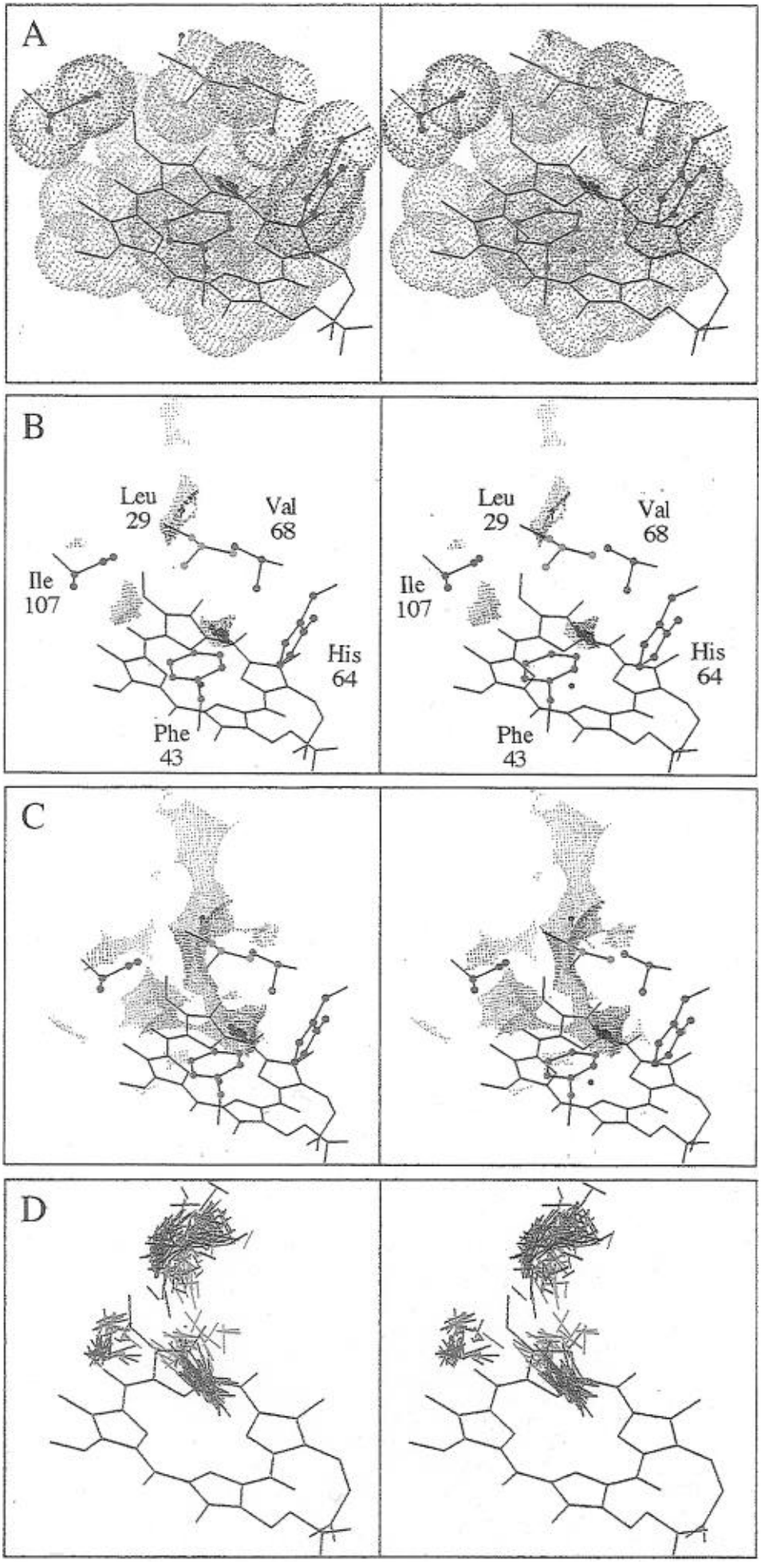
A
B
Leu
29
Val
68
Ile
107
His
64
Phe
43
Leu
29
Val
68
Ile
107
His
64
Phe
43
C
D

FIGURE 3: Ligand-accessible cavities vs void volume in Mb (in stereo). Five distal residues in sperm whale Mb are identified in panel B. Panel A shows the void (interatomic space) volume during a typical wild-type Mb simulation. Panel B shows the subregions available to a "cold" carbon atom's center (kinetic energy equal to the 15th percentile of the room temperature Maxwell−Boltzmann distribution). Panel C shows the same for a "hot" atom (85th percentile). Panel D shows the coincidence of typical ligand exploration for wild-type and mutant Mb: wild type (black), Val68Ala (blue), Val68Ile (red), and Val68Leu (green).

another cavity to the right, which, as previously mentioned, might serve to conduct ligands more toward the center of myoglobin. The tenuous, transient connections between ligand-accommodating cavities suggest the importance of gating residues. The fourth picosecond in Figure 4 demonstrates how residues 29, 68, and 107 control the passage between the central and right rear cavities.

To assess the ability to engineer changes in the passage between the central and right rear cavities in the distal matrix of myoglobin, simulations were done in which the side chains of residues 68, 107, or both were reduced in size. The expectation was that with shorter gating residues, the passage between the central and right rear cavities would be larger and more long-lived. This would in turn allow greater ligand movement out of the central distal cavity, decreasing the number of iron−ligand collisions. Tallies for the collisions accumulated during 50 ps simulations are shown in Figure 5 for 10 runs for each of the four cases examined: wild type, Val68Ala, Ile107Val, and Val68Ala·Ile107Val. As expected, the mutants have consistently fewer iron−ligand collisions. The only surprise was that Ile107Val had the fewest. Our preconception was that the growth in the passage would reflect the number of carbons removed from the wild-type Mb. Ile107Val differs by only 1, yet it showed the greatest impact upon the iron−ligand collisions. The answer again lies in visualizing the actual spaces available to the diffusion ligands. Figure 6 shows typical ligand-accessible spaces for the three mutants considered. For the mutants in which Val68 was trimmed to Ala, a stable passage exists as expected and is indeed larger when the Ile107 is also trimmed to Val. However, for Ile107Val, the passage (and nearly the central cavity) has ceased to exist once the ligands have entered the right rear cavity. In this mutant, the distal protein matrix has responded to our simulated mutation at 107, by moving Val68 toward the binding site. [We have reported similar encroachment on the heme site by trimmed residues or compensating waters before (Carlson et al., 1994).] In all three mutants, one might wonder why residue 107 does not impinge upon the heme site in an analogous manner when a smaller residue is substituted. The steric contours in Figure 3A reveal that the β-branched side chain of Ile107 is impeded by the edge of the heme plane and is unable move inward, unlike 68 which is free to slide along the heme's distal surface. In this fashion, computer simulation and visualization predict considerably different results for what were superficially similar mutations.

Simulations of wild-type myoglobin without ligands were conducted to assess the impact of the ligand itself upon the cavities in the distal protein matrix. As the distal residues are affected by the presence or absence of a single carbon or another side chain, it is reasonable to assume that they are also affected by the presence of a diatomic ligand molecule. Runs were carried out using the wild-type carbonmonoxy crystal structure from which the ligand was removed prior to simulation. (The bound structure without a ligand more fairly reflects the impact of the ligand than the deoxy structure since the latter has a water molecule in the distal matrix and structural differences which may persist on our time scale.) Figure 7 shows that, in the absence of a ligand molecule, the ligand-accommodating cavities have shrunk substantially, making connections rare and even further subdividing the cavities. However, the general placements and shapes of the cavities remain unchanged despite their partial collapse. Tables 1 and 2 list the volumes of the cavities pictured in Figures 4 and 7. As the figures imply, the numbers for the liganded case are larger. Notice that all the cavities are larger in the liganded case even though only the central one was occupied in these sampled frames. One might have expected unoccupied cavities to be as small as those in the unliganded simulation. Figure 8 shows the average collective distal cavity volume for the unliganded case as well as for various liganded structures. In spite of any compensatory protein movements, the overall distal volume in which ligands could move generally reflects the size of the distal residues. While reasonable, this is not necessarily true as most void volume is inaccessible to ligands.

Table 1: Distal Cavity Volumes in the Presence of Ligands[a]

	left	left center	center	right center	right	total
frame 1						20.48
frame 2		14.27			7.62	21.89
frame 3	0			14.46		14.46
frame 4						21.50
frame 5	0		5.18		10.05	15.23
frame 6						22.73

[a] Volumes (in Å^3) are listed for individual or united distal cavities as well as their combined totals for the wild-type simulation shown in Figure 4.

Table 2: Distal Cavity Volumes in the Absence of Ligands[a]

	left	left center	center	right center	right	total
frame 1	0		1.48		0.30	1.78
frame 2	1.43		0.27		3.54	5.24
frame 3	0		2.27		3.05	5.32
frame 4	0			4.75		4.75
frame 5	0		3.27		2.13	5.40
frame 6	0		0.92		2.40	3.32

[a] Volumes (in Å^3) are listed for individual or united distal cavities as well as their combined totals for the wild-type simulation shown in Figure 7.

In addition to static averages, there are dynamic differences between liganded and unliganded simulations. Figure 9A follows the fluctuation in the distal cavity volume for one liganded and one unliganded trajectory at 0.1 ps intervals. The fluctuation, on average, is greater when the diatomic ligand is present.

While the seemingly random gyration of the distal residue side chains modulates the shape of the distal cavities, the overall distal volume is related to more global motions of the protein, as entire residues move in and out following the movements of their supporting α-helices. The change in the volume of the pyramid (see Methods) is taken to reflect the expansion or contraction of that portion of the protein which engulfs the distal site. Panels B and C of Figure 9 show the distal cavity volumes and pyramid volumes,

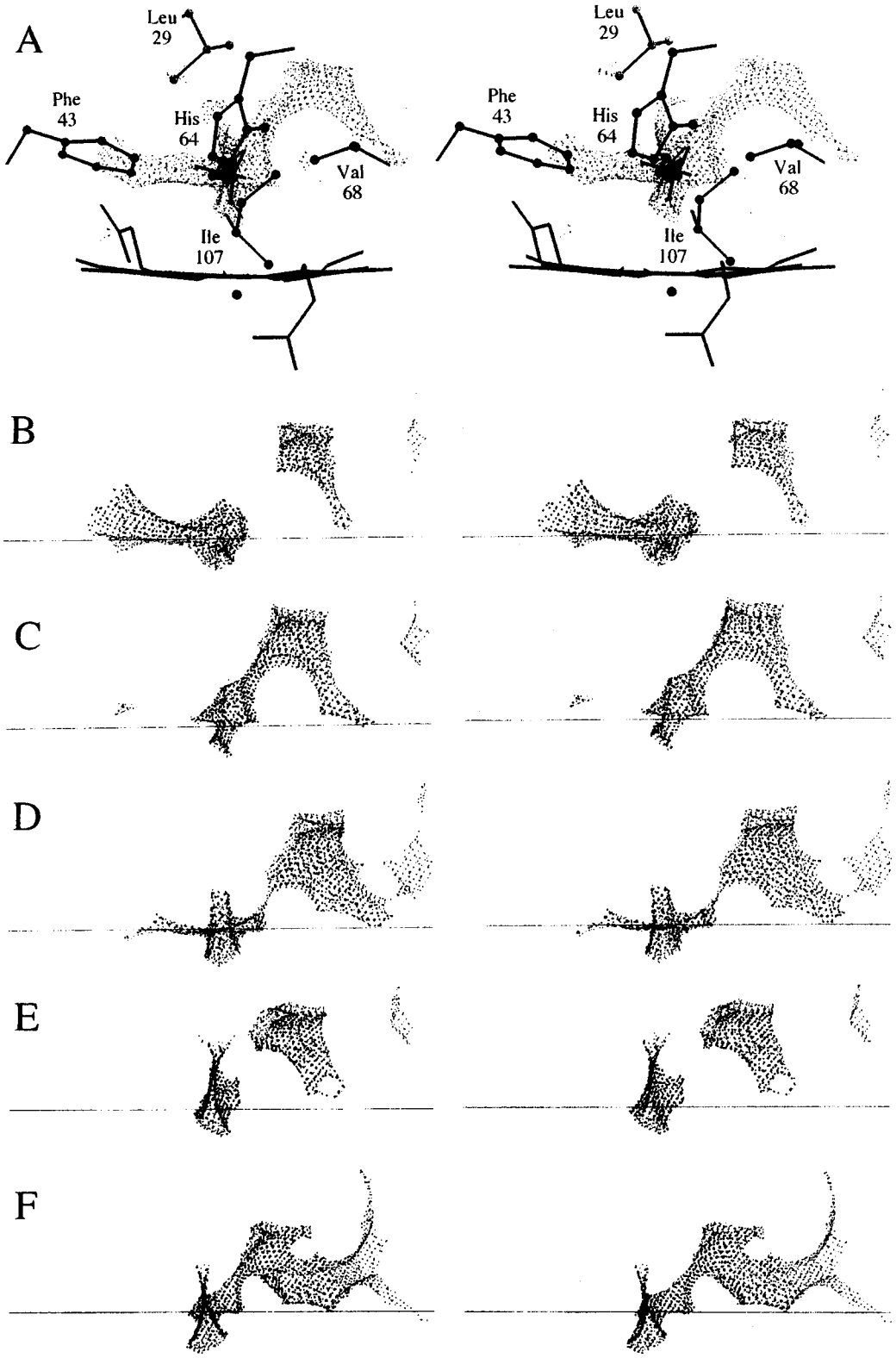

FIGURE 4: Distal cavity fluctuations in wild-type Mb (in stereo). Panel A shows the ligand-accessible spaces formed by the distal residues surrounding the heme site. In this orientation (front view), the heme is edge on with the propionates toward the viewer. Panels B−E show only the ligand-accessible cavities at five successive 1 ps intervals. All ligand copies (shown only in panel A) remain in the central cavity throughout this time interval. Horizontal lines approximate the top surface of heme and indicate the downward extension of the central cavity when over the iron.

respectively, for two extreme wild-type simulations with ligands. The similarity in the two sets of traces suggests that the changes in the distal cavity volumes are due more to expansion of the protein than to long-lived side-chain conformations, for example. Average differences in the lengths of some of the edges of the pyramid for liganded vs unliganded simulations are on the order of a half an angstrom (see Table 3). Surprisingly, the displacements of several of the helices in response to an unbound, diffusing diatomic molecule are of the same magnitude as the heme iron displacement following photodissociation. The two divergent wild-type simulations in Figure 9b,c also represent the functional extremes for ligand diffusion. In the smaller volume simulation, the ligands are held close to the heme

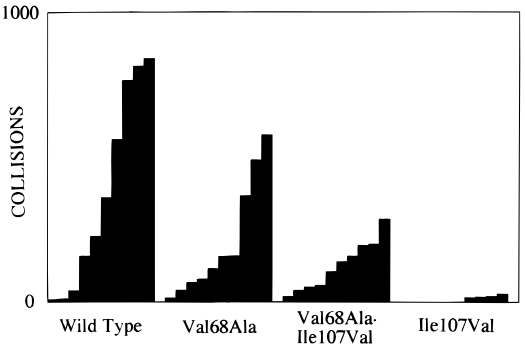

FIGURE 5: Iron accessibility for passage-gating mutants in Mb. The iron−ligand collisions accumulated during each of 10 simulations are shown for wild-type Mb and mutants in which one or both of residues 68 and 107, which gate a passage to the right rear cavity, are reduced in size. A collision is counted for each ligand atom within 4 Å of the iron at 1 ps intervals during these 50 ps simulations. (Note several 0 collision simulations for Ile107Val.)

Table 3: Ligand Impact on Distal Helix Arrangement[a]

	ligands	no ligands	difference
pyramid volume	131	118	+23
68−107 distance	11.1	10.7	+0.4
68−29 distance	9.8	9.1	+0.7
107−29 distance	8.1	7.8	+0.3

[a] Average interhelical distances (in Å) and pyramid volumes (in Å^3) are listed for the simulations of wild-type Mb with and without ligands as well as their differences. The distances between the B, E, and G helices measured at the α-carbons of residues 29, 69, and 107 are of particular interest since those residues' side chains gate ligand passage between the central and the right near cavities.

iron and collide frequently with it. In the simulation where the distal protein matrix expands, the ligands pass back and forth to the right rear cavity and have correspondingly fewer collisions with the iron throughout. These extremes illustrate the distal model for multiphasic rebinding. The fast phase results from conformations in which the ligands have no opportunity to leave the iron while slower phases arise from diffusing ligands returning from more remote locations. Figure 10 shows the number of collisions and migration events counted at 1ps intervals as a function of pyramid volume for 10 wild-type simulations. Overall, iron collisions fall and ligand migrations increase as the distal volumes increase. Regardless of protein and cavity expansion, the mechanism by which ligands can migrate from one site to another is intercavity connections. Helical movements that expand the protein greatly increase the fraction of time that cavities are connected. In Figure 9B,C, the central and right rear cavities are connected 60% of the time in the expanded structure compared to 1% in the other. As noted earlier, ligands typically do not utilize every opportunity to migrate.

To verify that the pattern of distal, ligand-accessible spaces observed in sperm whale myoglobin was not specific to it, other oxygen-binding proteins belonging to the "myoglobin-fold" family were surveyed. Panels A−C of Figure 11 show typical spaces from horse myoglobin, human α-chain hemoglobin, and one subunit of a homodimeric clam hemoglobin, respectively. Although the contour details reflect their differences, the general pattern of cavities persists. For horse myoglobin, this is not too surprising since the five immediate distal residues are the same as those in sperm whale myoglobin. Of the 19 substitutions throughout the protein, only one lines one of the distal cavities and that is Val28 (Ile28 in whale). Human hemoglobin has considerably more differences. The equivalent of residue 107 is Leu instead of Ile. Of the more remote residues that still contact the cavities in Figure 11B, all are different in human Hb compared to sperm whale Mb. On the other hand, clam Hb has two substitutions among the five immediate distal residues. The residues corresponding to 29 and 68 in Mb are Met and Leu instead of Leu and Val. Despite these differences, the cavity pattern is maintained. In Figure 11C, one can see that the cavity over the right (B pyrrole) vinyl is connected to the still more interior cavity. In contrast to Mb, this happens in both hemes in the dimer with such regularity that the ligands are frequently observed in this space (Figure 11B,C).

DISCUSSION

Although no hard and fast rule can be laid down, experiment has shown that characteristic patterns of ligand distribution usually develop during 50 ps, and an admittedly small number of simulations to 200 ps have not shown significant further change, presumably because a local energy minimum has been found. Repeated starts with randomly distributed initial velocities allow more efficient search of the possible energy minima. The multiple ligand copies available in the LES algorithm help to visualize the occupation of the cavities, at the expense of some loss of precision, which, however, is small as long as the ligands remain tightly bunched, as has usually been the case. Typically, the mean separation of the ligand centers of the copies is only between 0.2 and 1 Å, which is less than the distance between atomic centers in the ligand. Additionally, simulations with single ligands have given results within the envelope defined by analogous LES simulations.

The pattern of distal cavities in sperm whale myoglobin persists in most of the distal mutants thus far examined, and a similar pattern is found in other myoglobins and hemoglobins. The architecture of the protein matrix in the native protein appears to maximize the effect of typical structural fluctuations on ligand−iron access. Very minor steric modifications of the distal site can greatly alter the rate of ligand collisions with the heme iron, suggesting a reason for the strong conservation of distal residues. In simulation, collective motions of the protein have functional relevance as ligand migration and ligand−iron collisions correlate with expansion of the protein due to movements of the helices surrounding the distal site. These helices, in turn, are sensitive to the presence of a diffusing diatomic ligand and made relative adjustments on the order of 0.5 Å. Repositioning of the distal helices provides a mechanism for two-way communication between the heme site and the exterior of the protein.

Ligand diffusion supplements iron reactivity in explaining the nonexponential character of NO picosecond geminate recombination. The two models focus on different aspects of the rate equation as well as opposite sides of the heme. The iron model contends that the rate constant for ligand rebinding should decrease after dissociation as the heme iron assumes its deoxy position out of the heme plane and the protein relaxes (for example, Petrich et al. (1991)]. In contrast, the diffusion model suggests that ligand movement and distal residue fluctuations alter the effective concentration of ligands which are immediately available to rebind. Simulations show differences in ligand−iron collision frequencies on a time scale of tens of picoseconds, appropriate

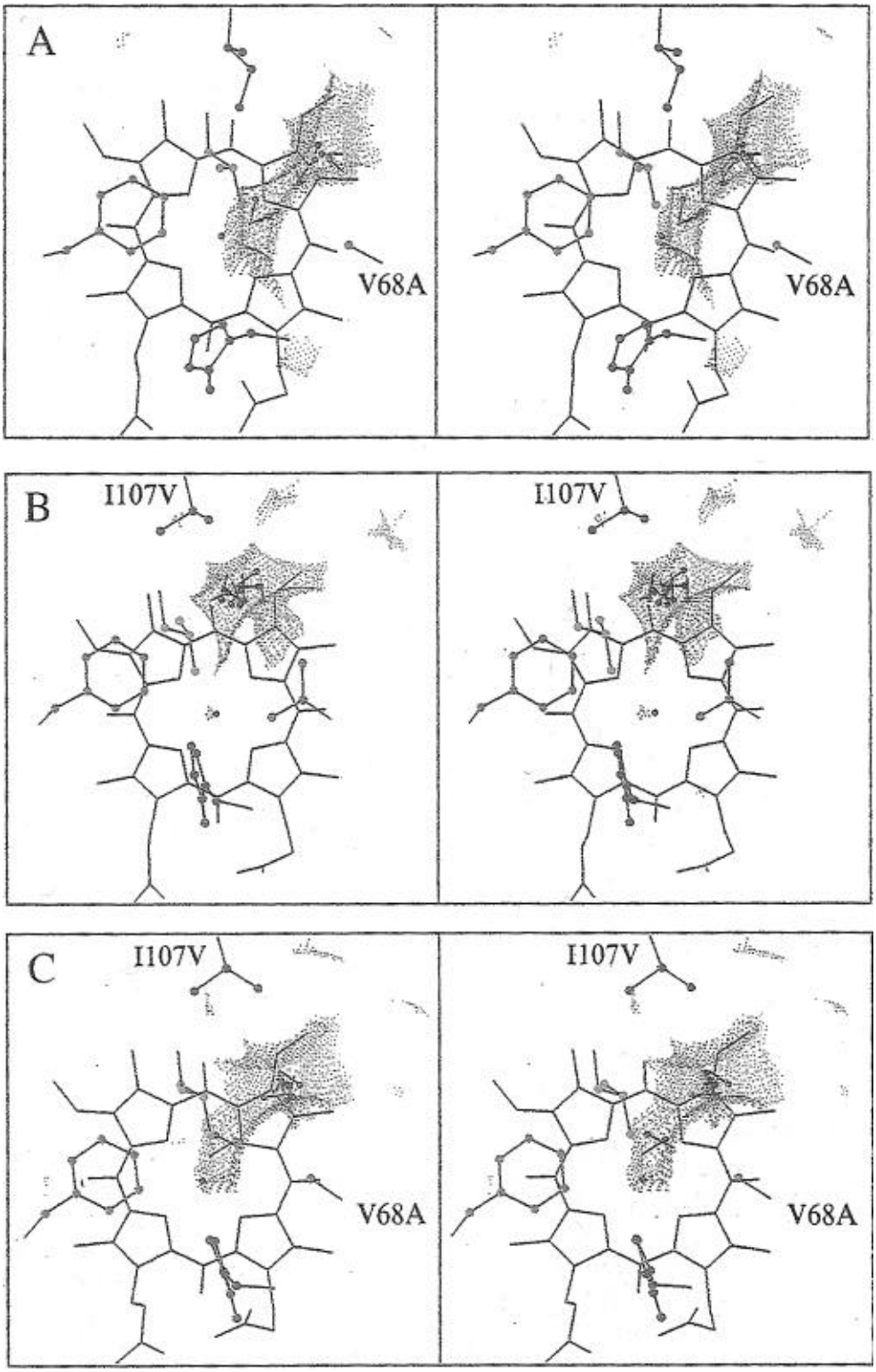

FIGURE 6: Cavity connectivity in Mb passage-gating mutants (in stereo). Typical cavity connectivity and iron accessibility are shown for Mb mutants in which one or both passage-gating residues, 68 and 107, have been trimmed. Panel A depicts Val68Ala, panel B depicts Ile107Val, and panel C depicts Val68Ala·Ile107Val. Note the Val68Ala mutants have stable passageways and iron access while Ile107Val does not.

in relation to experiment (Gibson et al., 1992). Although most movement of the iron out of the heme plane occurs in less than 1 ps (Henry et al., 1985), a 100 ps simulation has suggested further iron average displacement of two hundredths of an angstrom (Petrich et al., 1991), which is significantly less than its rms fluctuation. The distal diffusion model would appear to have several advantages in that it accounts directly for the marked impact of distal substitutions on recombination kinetics (Carver et al., 1990; Gibson et al., 1992) and that it applies equally to cobalt-substituted Mb's (Ikeda-Saito et al., 1993) in which the metal center is not expected to move significantly out of the porphyrin plane (Hoard & Scheidt, 1973).

The generality in the steric mechanism for ligand diffusion among sperm whale Mb mutants, horse Mb, and various Hb's is plausible, given the similarity of structure. The common folding pattern shared by these proteins brings very similar residues close together on the distal side of the heme. For

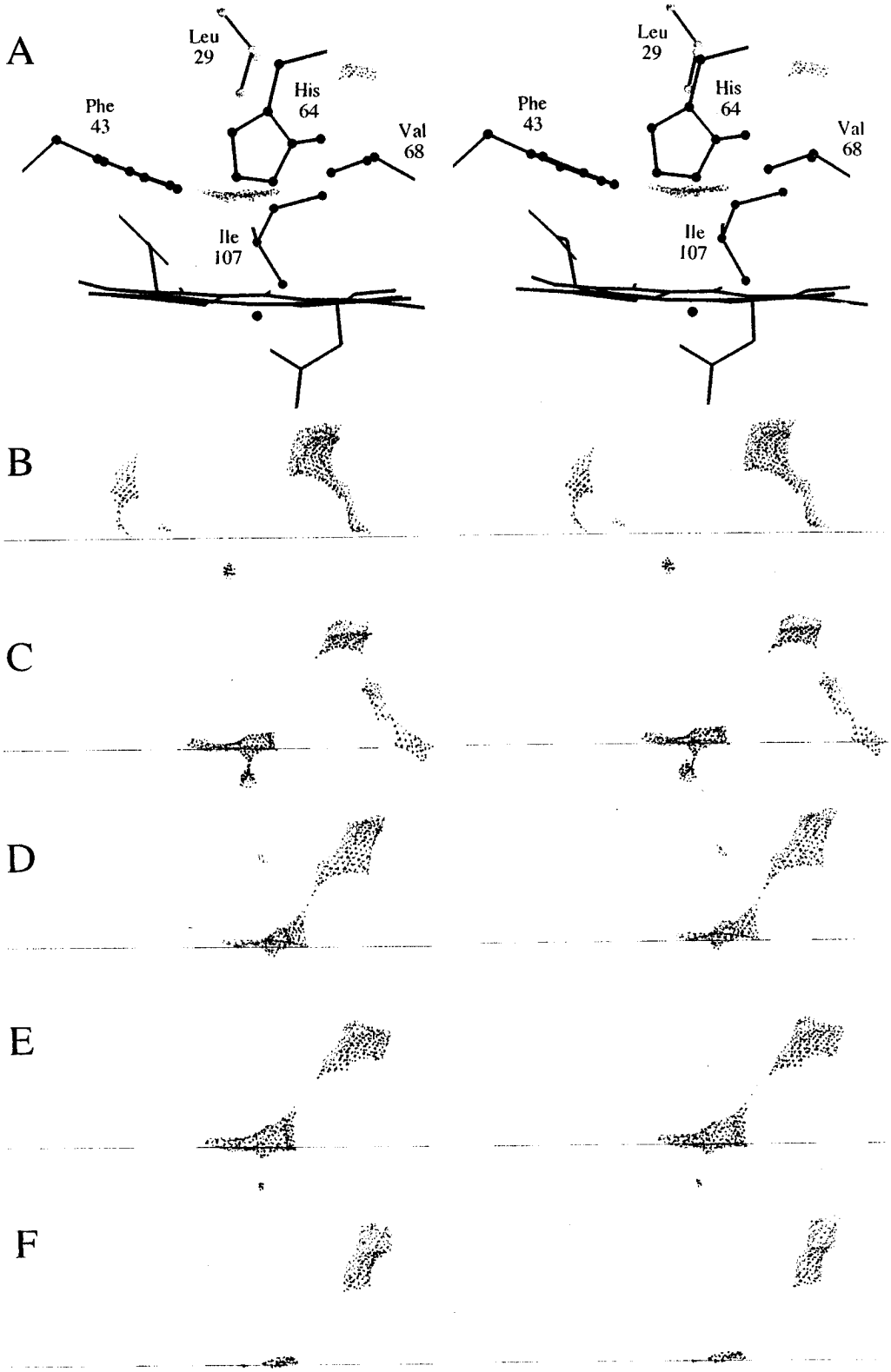

FIGURE 7: Distal cavities in the absence of ligands (in stereo). The size, shape, and connectivity of the distal cavities when no ligands are present show the impact of a diatomic ligand on Mb structure and dynamics. The panels of this figure are analogous to those in Figure 4.

a given protein, distal residues are conserved among species variants, especially the five residues which surround the binding site. Among related proteins, many of the distal residues are still conserved while those that are not are sterically quite similar, differing by a carbon or two or in their branching. For example, among the five immediate distal residues, vertebrate Hb's have only substituted Leu for the analogous Ile107 in Mb (Dickerson & Geis, 1983). In addition, distal substitutions among related proteins may serve to compensate for overall structural differences. The behavior of the distal cavities and ligands in the dimeric clam Hb is more like that of native Mb than that of the Mb mutant in which, like clam, Leu has been substituted for Val68 (unpublished observation). While the pattern of distal cavities is fairly resilient with respect to residue substitution, it can be disrupted if entire rings are inserted or deleted (Carlson et al., 1994; Quillin et al., 1995). These substitutions, however, are not typically found in native proteins.

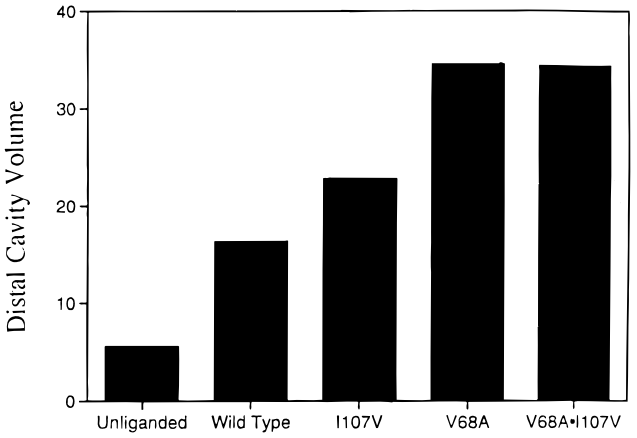

FIGURE 8: Distal cavity volume comparison. Insertion of a diatomic ligand or removal of distal side-chain carbons on average increases the combined volume of the distal atom-sized cavities. The volumes of all ligand-accessible distal cavities were combined and averaged over the length of the simulations for wild type and the passage-gating mutants of Mb (positions 68 and 107).

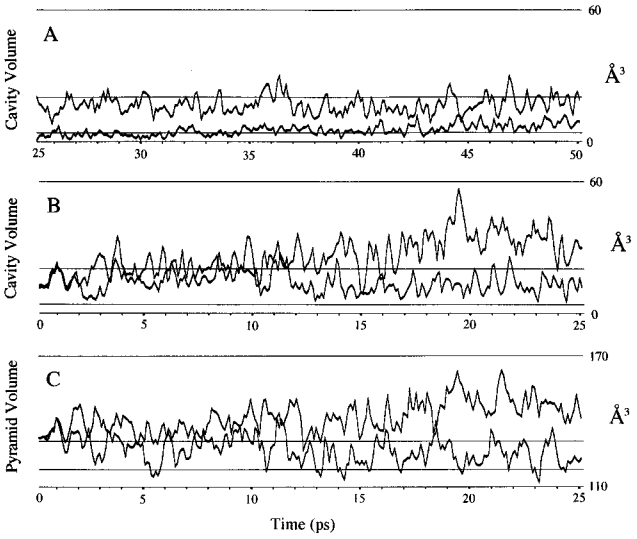

FIGURE 9: Distal volume fluctuations. Panel A shows the fluctuation of the distal cavity volume in Mb in the presence (upper trace) and absence (lower trace) of ligands at 0.1 ps intervals over 25 ps. The horizontal lines are the respecrtive average values from multiple simulations. Panel B shows the distal cavity volumes for two wild-type simulations which represent the extremes of observed ligand behavior. The lower trace corresponds to a simulation in which the ligands are held in a small central cavity directly over the iron, constantly colliding with it. The upper trace reflects a simulation in which the ligands frequently pass between the central and the right rear cavities and seldom encounter the iron. Panel C shows the interhelix pyramid volumes (refer to Figure 2) for the same simulations as in panel B. There is correspondence between expansion of the protein and growth in the distal cavities.

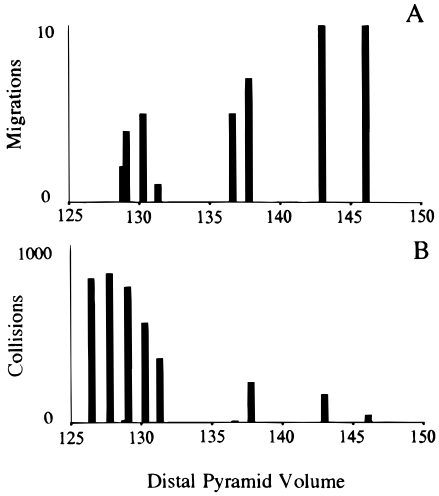

FIGURE 10: Ligand migration and collisions vs distal expansion. Ligand−iron collisions and ligand cavity migration events are plotted as a function of the distal interhelix pyramid volume (defined in Figure 2). In general, distal expansion in Mb correlates with greater ligand mobility between cavities and fewer ligand−iron collisions. Migration was quantified by the number of times one or more ligand copies moved to a different cavity when examined at 1 ps intervals during the 50 ps simulations. Collisions are defined as in Figure 5.

While the strong conservation of the distal histidine andresidue E11 (Val68 in Mb) is thought to be due to their roles in determining the bond strength of ligands, similar justification has not been advanced for the other distal residues. It may be that their role in sterically controlling ligand diffusion is responsible.

The construction of the distal matrix is such that the central cavity is just within reach of the iron and the other cavities are poised on the verge of connection with it. As a result, the normal thermal fluctuations of the protein are sufficient to move the ligands in the central cavity in and out of collision range with the iron and allow them to migrate periodically to and from the more remote cavities. As a result of this delicate balance, nearly all possible outcomes in terms of iron−ligand access are observed. In mutants, the structural differences seem to bias them significantly toward a smaller subset of outcomes. For example, the mutants Val68Ala and Ile107Val differ from wild type by only one or two carbons. For Val68Ala, the heme iron is rarely obstructed, and the passage to the right rear cavity is nearly always open. The reverse is true for Ile107Val (see Figure 6A,B).

In the crystal structures of Mb distal mutants, the E helix has been observed to impinge upon the binding site in response to substitutions with smaller side chains (Carlson et al., 1994). Our simulation results for Ile107Val show that helices can shift position in response to just one less side-chain carbon in a key location, and as the presence of the ligands affects the helices surrounding the distal site, there can be communication between the active site and the exterior of the protein. The relative displacements of the distal helices are comparable in magnitude to the displacement of the heme iron following photolysis and offer an alternate or complementary means for communication that is perhaps less localized.

In order to exchange the bound CO for the distal water molecule found in deoxy-Mb, intermediate structures must accommodate the presence of both or neither of the molecules and consequently should show greater distal perturbation than that between the bound CO and deoxy structures. Our simulations do indeed show distal reorganization when the ligand is removed, and it has been reported previously that a water molecule in the distal matrix significantly affects ligand diffusion and access to the iron (Carlson et al., 1994).

Simulation has also proved useful in cases where crystallography admits alternate structures of nearly equal probability. For example, simulation, combined with experimental results, offered a clear choice between alternate positions for C_δ of Val68Leu (Quillin et al., 1995). Although molecular dynamics is certainly capable of generating results at variance with experiment, it is probably much better than

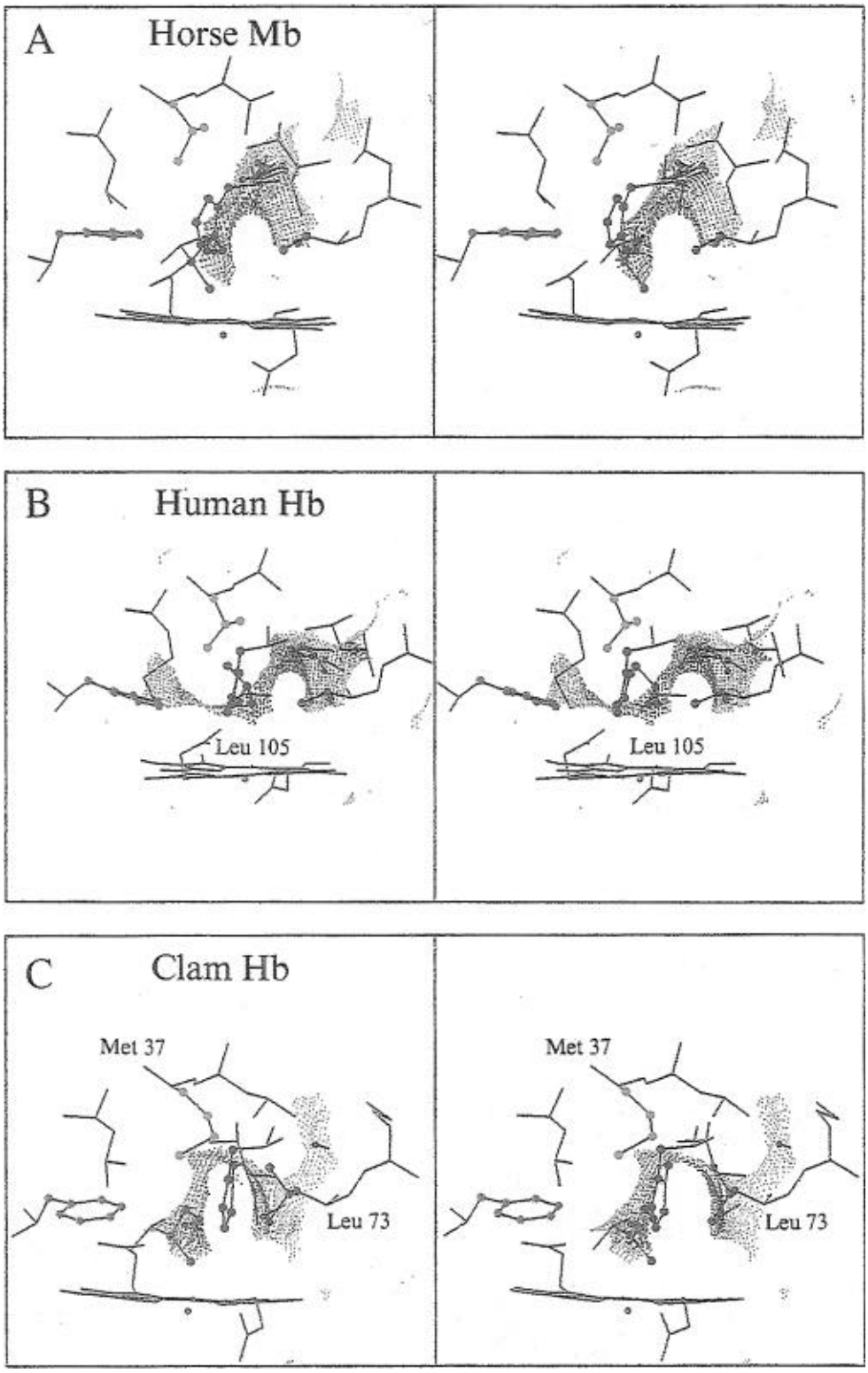

FIGURE 11: Distal cavities in other myoglobins and hemoglobins (in stereo). Cavity patterns are sampled from simulations of horse heart Mb (panel A), human Hb α-chain (panel B), and clam homodimeric Hb (panel C). The color coding of distal residues indicates equivalent positions in the various proteins despite residue substitutions and sequence numbering differences. Additional distal residues are included to show other residue substitutions where they exist. Note that, in the dimeric clam Hb (panel C), the cavity over the right rear pyrrole vinyl is connected to another to its right which is actually visited by a ligand.

intuition at predicting the effects of hypothetical mutations. Our attempt to enlarge the passage between the central and right rear cavities by two different mutations illustrates this point. Trimming either Val68 or Ile107 did indeed make the remote site more accessible but also led to the unanticipated result of substantially different simulated ligand−iron access in the two cases. On the basis of the isolation of the iron atom in Ile107Val, we would predict much more recombination in slower phases for it than for wild type or Val68Ala. Given the relatively small investment, it should be efficient to survey proposed mutations by simulation before synthesis is begun.

The fluctuating cavities of the distal protein matrix offer a bridge between protein structure and kinetic descriptions of ligand binding in the first picoseconds after photolysis. In general, we would expect the ligands in the central cavity

to rebind quickly and those which diffuse away to rebind more slowly although not at definite rates as the position of the central cavity and the frequency of intercavity connections change, sometimes slowly, on the picosecond time scale.

REFERENCES

Ahmed, A. M., Campbell, B. F., Caruso, D., Chance, M. R., Chavez, M. D., Courtney, S. H., Friedman, J. M., Iben, I. E. T., Ondrias, M. R., & Yang, M. (1991) *Chem. Phys. 185*, 329−351.

Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H., & Gunsalus, I. C. (1975) *Biochemistry 14*, 5355−5373.

Carlson, M. L., Regan, R., Elber, R., Li, H., Phillips, G. N., Jr., Olson, J. S., & Gibson, Q. H. (1994) *Biochemistry 33*, 10597−10606.

Carver, T. E., Rohlfs, R. J., Olson, J. S., Gibson, Q. H., Blackmore, R. S., Springer, B. A., & Sligar, S. G. (1990) *J. Biol. Chem. 265*, 20007−20020.

Case, D. A., & Karplus, M. (1979) *J. Mol. Biol. 132*, 343−368.

Dasgupta, S., & Spiro, T. G. (1987) *Biochemistry 26*, 5698−5692.

Dickerson, R. E., & Geis, I. (1983) *Hemoglobin: Structure, Function, Evolution, and Pathology*, Benjamin/Cummings Publishing Co., Menlo Park, CA.

Duprat, A. F., Traylor, T. G., Wu, G.-Z., Coletta, M., Sharma, V. S., Walda, K. N., & Magde, D. (1995) *Biochemistry 34*, 2634−2644.

Elber, R., Roitberg, A., Simmerling, C., Goldstein, R. F., Verkhivker, G., Li, H., & Ulitsky, A. (1994) MOIL: A molecular dynamics program with emphasis on conformational searches and reaction path calculations, *Statistical mechanics, protein structure and protein substrate interactions* (Doniach, S., Ed.) Plenum Press, New York.

Gibson, Q. H., Regan, R., Elber, R., Olson, J. S., & Carver, T. E. (1992) *J. Biol. Chem. 267*, 22022−22034.

Henry, E. R., Levitt, M., & Eaton, W. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2034−2038.

Hoard, J. L., & Scheidt, W. R. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3919−3922.

Ikeda-Saito, M., Dou, Y., Yonetani, T., Olson, J. S., Li, T., Regan, R., & Gibson, Q. H. (1993) *J. Biol. Chem. 268*, 6855−6857.

Jorgensen, W. L., Chandrasekhar, J., & Madura, J. D. (1983) *J. Chem. Phys. 79*, 7270−7286.

Kuriyan, J., Wilz, S., Karplus, M., & Petsko, G. A. (1986) *J. Mol. Biol. 192*, 133−154.

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.

Murthy, M. R. N., Reid, T. J., III, Sicignano, A., Tanaka, N., & Rossmann, M. G. (1981) *J. Mol. Biol. 152*, 465−499.

Perutz, M. F., & Mathews, F. S. (1966) *J. Mol. Biol. 21*, 199.

Petrich, J. W., Lambry, J. C., Kuczera, K., Karplus, M., Poyart, C., & Martin, J. L. (1991) *Biochemistry 30*, 3975−3987.

Quillin, M. L., Li, T., Olson, J. S., Phillips, G. N., Jr., Dou, Y., Ikeda-Saito, M., Regan, R., Carlson, M., Gibson, Q. H., Li, H., & Elber, R. (1995) *J. Mol. Biol. 245*, 416−436.

Richard, L., Genberg, L., Deak, J., Chiu, H. L., & Miller, J. D. (1992) *Biochemistry 31*, 10703−10715.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151−176.

Takano, T. (1977) *J. Mol. Biol. 110*, 569−584.

Tilton, R. F., Jr., Singh, U. C., Weiner, S. J., Connolly, M. L., Kuntz, I. D., Jr., & Kollman, P. A. (1986) *J. Mol. Biol. 192*, 443−456.

BI951767K