Theoretical Study of the Mechanisms of Substrate Recognition by Catalase

Susana G. Kalko,[§] Josep Lluis Gelpí,[§] Ignacio Fita,[†] and Modesto Orozco[§,*]

Contribution from the Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain, and Instituto de Biología Molecular de Barcelona, C.S.I.C., Jordi Girona 18-26, Barcelona 08034, Spain

Received February 26, 2001. Revised Manuscript Received June 25, 2001

Abstract: A variety of theoretical methods including classical molecular interaction potentials, classical molecular dynamics, and activated molecular dynamics have been used to analyze the substrate recognition mechanisms of peroxisomal catalase from *Saccharomyces cerevisiae*. Special attention is paid to the existence of channels connecting the heme group with the exterior of the protein. On the basis of these calculations a rationale is given for the unique catalytic properties of this enzyme, as well as for the change in enzyme efficiency related to key mutations. According to our calculations the water is expected to be a competitive inhibitor of the enzyme, blocking the access of hydrogen peroxide to the active site. The main channel is the preferred route for substrate access to the enzyme and shows a cooperative binding to hydrogen peroxide. However, the overall affinity of the main channel for H_2O_2 is only slightly larger than that for H_2O. Alternative channels connecting the heme group with the monomer interface and the NADP(H) binding site are detected. These secondary channels might be important for product release.

Introduction

The need for protection against reactive oxygen species such as hydrogen peroxide has led aerobic organisms to the development of different enzymatic systems for the degradation of these compounds. Among them, catalases are one of the most powerful mechanisms for the protection against hydrogen peroxide. This family of enzymes catalyzes the dismutation of hydrogen peroxide into water and oxygen. The importance of catalases is clearly established by the existence of a series of pathologies related to their malfunctioning which comprise, among others, increased susceptibility of thermal injury,[1] high rates of mutations,[2] inflammation,[3] and accelerated aging.[4] The importance of catalases explains its widespread presence in living organisms. Three types of catalases have been described in prokaryotes: (i) managanese-based,[5] (ii) catalase peroxidase,[6] and (iii) heme catalase.[7,8] Only heme catalases have been described for higher organisms.[7,8]

The mechanism of action of heme-catalases is complex, and not fully understood. Apparently, the preferred mechanism has two steps:[7,9−12] (i) the heme iron is oxidized to an oxoferryl (Fe(IV)═O) and the macrocycle is oxidized to a porphyryl radical cation (compound I) by a H_2O_2 molecule, which yields a H_2O molecule,[13,14] and (ii) compound I accepts two electrons from a second H_2O_2 molecule, generating H_2O and O_2, and recovering the resting form of the enzyme. The resting form can also be obtained using small aliphatic alcohols as substrates, yielding aldehydes. This secondary reaction has been related to the neurological effects of alcohol.[15−18] Other mechanisms for the reduction of compound I have been described for some catalases.[19,20] The nature and the precise location of the radical center in compound I are dependent on the species from which the catalase is isolated. Thus, using EPR and rapid-mix freeze-quench techniques, it has been shown that compound I of *Proteus mirabilis* catalase, formed by the reaction of the enzyme with peroxyacetic acid, contains a relatively stable phorphyrin π-cation radical.[21] On the contrary, the porphyrin π-cation radical of bovine liver catalase is short-lived and is replaced by a tyrosyl radical.[21] However, owing to the very fast turnover of

* Author to whom correspondence should be sent. E-mail: modesto@luz.bq.ub.es.
§ Universitat de Barcelona.
† Instituto de Biología Molecular de Barcelona, C.S.I.C.

(1) Leff, J. A. *Inflammation* **1993**, *17*, 199.
(2) Halliwell, B.; Aruoma, O. I. *FEBS Lett.* **1991**, *281*, 9.
(3) Halliwell, B.; Gutteridge, J. M. C. *Methods Enzymol.* **1990**, *186*, 1.
(4) Taub, J.; Lau, J. F.; Ma, C.; Hahn, J. H.; Hoque, R.; Rothblatt, J.; Chalfie, M. *Nature* **1999**, *399*, 162.
(5) Barynin, V. V.; Vagin, A. A.; Melik-Adamyan, V. R.; Grebenko, A. I.; Khangulov, S. V.; Popov, A. N.; Andrianova, M. E.; Vainshtein, B. K. *Sov. Phys. Crystallogr.* **1986**, *31*, 457.
(6) Fraaije, M. W.; Roubroeks, H. P.; Hagen, W. R.; Van Berkel, W. J. H. *Eur. J. Biochem.* **1996**, *235*, 192.
(7) Schonbaum, G. R.; Chance, B. In *The Enzymes*; Academic Press: New York, 1976; Vol. 13, p 363.
(8) Klotz, M. G.; Klassen, G. R.; Loewen, P. C. *Mol. Biol. Evol.* **1997**, *14*, 951.
(9) Chance, B.; Herbert, D. *Biochem. J.* **1950**, *46*, 402.
(10) Nicholls, P.; Schonbaum, G. R. In *The Enzymes*; Academic Press: New York, 1963; Vol. 8, p 147.
(11) Deisseroth, A.; Dounce, A. *Physiol. Rev.* **1970**, *50*, 319.
(12) Maté, M. J.; Murshudov, G.; Bravo, J.; Melik-Adamyan, W.; Loewen, P. C.; Fita, I. In *Handbook of Metalloproteins*; Messerschmidt, A., Huber, R., Eds.; John Wiley: New York, 2001.
(13) Dolphin, D.; Forman, A.; Borg, D. C.; Fajer, J.; Felton, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 614.
(14) Ivancich, A.; Jouve, H. M.; Sartor, B.; Gaillard, J. *Biochemistry* **1997**, *36*, 9356.
(15) Zimatkin, S. M.; Liopo, A. V.; Deitrice, R. A. *Alcohol Clin. Exp. Res.* **1998**, *22*, 1623.
(16) Hamby-Mason, R.; Schenker, S.; Chen, J. J.; Perez, A.; Henderson, G. I. *Alcohol Clin. Exp. Res.* **1997**, *21*, 1063.
(17) Hunt, W. A. *Alcohol* **1996**, *13*, 147.
(18) Aragon, C. M.; Amit, Z. *Neuropharmacology* **1992**, *31*, 709.
(19) Lardinois, O. M.; Rouxhet, P. G. *Free Radical Res.* **1994**, *20*, 29.
(20) Obinger, C.; Maj, M.; Nicholls, P.; Loewen, P. C. *Arch. Biochem. Biophys.* **1997**, *342*, 58.
(21) Ivanchich, A.; Jouve, H. M.; Sartor, B.; Gaillard, J. *Biochemistry* **1997**, *36*, 9356.

10.1021/ja010512t CCC: $20.00
Published on Web 09/11/2001

the enzyme on H_2O_2, the authors suggested that the tyrosine radical would not play a significant role in the catalatic function of catalase.[21]

Many structures of catalases have been described in their native forms. These include three prokaryote enzymes from *Micrococcus luteus*,[22] *Proteus mirabilis*,[23] and *Escherichia coli*,[24] and four eukaryote enzymes: *Penicillium vitale*,[25] bovine live catalase,[26] *Saccharomyces cerevisiae*,[27] and human catalase.[28] Recently, structural information of the intermediate compounds I and II (an alternative ferryl intermediate obtained by one-electron reduction of the porphyrin radical) was obtained using X-ray fast data-collection techniques at 2.7 Å resolution.[29] The data revealed only minor local rearrangements from the structure of the enzyme in its native form.[29]

Besides the large amount of experimental information on catalases, including high-resolution structures, there are many unsolved fundamental issues on these enigmatic enzymes. First: why are these enzymes so large, with domains of unknown function? Second: what is the reason for the NADP(H) binding site, when NADP(H) is not used? Third: what are the physical reasons for their unique enzymatic properties, with a very high apparent K_m and an extremely large k_{cat}?[10,11] Four: what are the reasons of the diversity of heme modifications found in different catalases?[30,31]

In this contribution we have used *state of the art* theoretical methods to complement our previous structural studies on heme catalases.[24,26b,27] A variety of techniques, derived mostly from classical dynamics, are used to gain insight into the mechanism of substrate recognition by catalases. Particularly, the importance of channels connecting the heme group and the exterior of proteins is analyzed using standard and activated molecular dynamics. A physical mechanism for substrate selection is derived, and a reasonable explanation for the unique enzymatic properties of heme catalases is found.

Methods

Standard Molecular Dynamics. Simulations were performed starting from the crystal structures of the native and Val111→Ala mutant of *Saccharomyces cerevisiae* enzyme solved previously at 2.8 Å resolution.[27] The active form is a homotetramer,[26b] but due to the enormous size of the protein, a reduced model is used here. Accordingly, the systems considered here consist of monomers A, plus a thin layer (those residues of the second monomers located at less than 25 Å from the main channel) of the contiguous monomer (D) facing the interface closest to the suggested main channel.[26b] Counterions were added using the classical molecular interaction potential (cMIP) program[29] to maintain neutrality and to avoid artifactual distortions in the trajectories. Interior cavities and channels in the two proteins were hydrated (when necessary) using the iterative cMIP program[32] to avoid collapses in the structure.[32] Finally, the system was immersed in a preequilibrated box of TIP3P water.[33] Waters located at less than 2.4 Å from any atom of the protein, and more than 35 Å from the center of the main channels, were removed. This procedure guarantees an excellent solvation of all the regions of interest. The final system contains the model protein, 22 Na^+, 18 Cl^-, and 3734 waters, that is, around 20000 atoms. These two systems are to our knowledge among the biggest studied by nanosecond-time scale molecular dynamics.

All of the residues of the protein located at more than 15 Å from the center of the main channel (around 13 Å from the heme group) were frozen at their crystal conformation. This allows us to mimic the effect of the tetrameric packing on the structural fluctuations of the active site and channels. All of the residues having at least one atom within 15 Å from the center were free to move. All of the waters and ions were free to move within the 35 Å sphere. A half harmonic potential retains the drop of water in the regions of interest.

The hydrated systems were optimized in two steps: (i) 5000 cycles for water with the protein and ions fixed and (ii) 5000 cycles of the protein and ions with the water fixed. The optimized system was then heated and equilibrated in four steps: (i) 200 ps of MD simulation of water (protein and ions fixed) to 100 K, (ii) 200 ps of MD simulation of proteins and ions (water fixed) to 100 K, (iii) mixing of velocities and heating of the entire mobile systems from 100 to 300 K for 200 ps, and finally iv) equilibration of the entire systems for 100 ps. Each (native and mutant) equilibrated structures were the starting point for 5 ns of MD simulations at constant temperature ($T = 300$ K). SHAKE[34] was used to keep all of the bonds at their equilibrium length, which allowed us the use of a 2 fs time step for integration of Newton equations. A residue-based cutoff of 12 Å was used in the evaluation of nonbonded molecular interactions.

The optimized structure in water was used to generate a starting model for the protein in the mixture of water and hydrogen peroxide (ca. 25−30% in a volume of hydrogen peroxide). For this purpose randomly selected water molecules were changed into hydrogen peroxide molecules. Water molecules with unfavorable steric interactions with the new inserted hydrogen peroxide molecules were removed (typically 2−3 molecules). The final system contains 560 hydrogen peroxide and 2096 water molecules, for an initial cap size of solvent similar to that of a pure water simulation. The molar ratio of hydrogen peroxide in the simulation was around 27%.

AMBER-95[35] and TIP3P[33] force-fields were used to describe the proteins and water. Bonded parameters for hydrogen peroxide were determined using the PAPQMD strategy,[36] using HF/6-31G(d) calculations as reference. Charges were determined using RESP[37] and HF/6-31G(d) wave functions, and van der Waals parameters were transferred from the water model.

Activated Molecular Dynamics. Standard molecular dynamics at the nanosecond time scale are not able to provide a complete representation of the diffusion of molecules along narrow channels. Thus, activated molecular dynamics (aMD) simulations were used to analyze the narrow channels connecting the heme group with the exterior of the protein. For comparison two types of activated runs were performed: (i) using the solvated system and (ii) using a gas-phase environment for the protein.

(22) Murshudov, G. N.; Melik-Adamyan, W. R.; Grebenko, A. I.; Barynin, V. V.; Vagin, A. A.; Vainshtein, B. K.; Dauter, Z.; Wilson, K. S. *FEBS Lett.* **1992**, *312*, 127.

(23) Gouet, P.; Jouve, H. M.; Dideberg, O. *J. Mol. Biol.* **1995**, *249*, 933.

(24) Bravo, J.; Verdaguer, N.; Tormo, J.; Betzel, C.; Switala, J.; Loewen, P. C.; Fita, I. *Structure* **1995**, *3*, 491.

(25) (a) Vainshtein, B. K.; Melik-Adamyan, W. R.; Barynin, V. V.; Vagin, A. A.; Grebenko, A. I. *Nature* **1981**, *293*, 411. (b) Melik-Adamyan, W. R.; Barynin, V. V.; Vagin, A. A.; Borisov, V. V.; Vainshtein, B. K.; Fita, I.; Murthy, M. R. N.; Rossmann, M. G. *J. Mol. Biol.* **1986**, *188*, 63.

(26) (a) Murthy, M. R. N.; Reid, T. J.; Sicignano, A.; Tanaka, T.; Rossmann, M. G. *J. Mol. Biol.* **1981**, *152*, 465. (b) Fita, I.; Silva, A. M.; Murthy, M. R. N.; Rossmann, M. G. *Acta Crystallogr., Sect. B* **1986**, *42*, 497.

(27) Mate, M. J.; Zamocky, M.; Nykyri, L. M.; Herzog, C.; Alzari, P. M.; Betzel, C.; Koller, F.; Fita, I. *J. Mol. Biol.* **1999**, *286*, 135.

(28) Putnam, C. D.; Arvai, A. S.; Bourne, Y.; Tainer, J. A. *J. Mol. Biol.* **2000**, *296*, 295.

(29) Gouet, P.; Jouve, H. M.; Williams, P. A.; Anderson, I.; Andreoletti, P.; Nussaume, L.; Hadju, J. *Nat. Struct. Biol.* **1996**, *3*, 951.

(30) Chiu, J. T.; Loewen, P. C.; Switala, J.; Gennis, R. B.; Timkovich, R. *J. Am. Chem. Soc.* **1989**, *111*, 7046.

(31) Murshudov, G. N.; Grebenko, A. I.; Barynin, V.; Dauter, Z.; Wilson, K. S.; Vainshtein, B. K.; Melik-Adamyan, W.; Bravo, J.; Ferrán, J. M.; Ferrer, J. C.; Switala, J.; Loewen, P. C.; Fita, I. *J. Biol. Chem.* **1996**, *271*, 8863.

(32) Gelpi, J. Ll.; Kalko, S. G.; Barril, X.; Cirera, J.; de la Cruz, X.; Luque, F. J.; Orozco, M. *Proteins* **2001**. In Press.

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

(34) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(35) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 11946.

(36) Alemán, C.; Canela, E. I.; Franco, R.; Orozco, M. *J. Comput.Chem.* **1991**, *12*, 664.

(37) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.

Activated molecular dynamics were performed by using an additional Berendsen's thermal bath[38] for the activated ligand molecule (a water molecule located near the heme group). This allowed us to have an individual molecule with much higher temperature ($T = 10000$ K) than that for the rest of the system ($T = 300$ K). The selection of a suitable relaxation time for the activated molecule allowed us to maintain its high temperature along the simulation. The extreme kinetic energy of the activated molecule allowed it to find pathways to leave the active site, which may not be easily accessible in a nanosecond time scale for a molecule at room temperature.

The aMD calculations were carried out using SHAKE,[34] and 1 fs integration step. The following simulation protocol was used: (i) the coordinates and velocities of a snapshot of a standard MD were used as starting configuration, (ii) a solvent molecule in the active site (near the heme group) was randomly chosen to be activated, (iii) the velocity of the activated molecule was scaled to 10000 K, (iv) the direction of the velocity vector of the activated molecule was randomly assigned, (v) 10 (aqueous environment) or 5 (gas phase) ps of constant temperature MD simulation with the two target temperatures (10000 K for the activated molecule and 300 K for the rest) and the two Berendsen's baths were run.

For both aqueous- and gas-phase aMD simulations steps (i) and (ii) were repeated four times, that is, four different structures of the protein, and four different positions of the solute in the active site were selected to start the aMD computations. For each of these four starting configurations steps (iii−v) were repeated (random velocities were assigned in each case) 100 times. Calculations were done using both the wild-type and mutant proteins in the gas phase, and using the trajectories obtained in aqueous solution and in the 30% mixture. The aMD simulations in the presence of solvent, which convinced us that the neglect of solvent in short aMD simulations does not lead to dramatic errors, were performed only for the mutant protein in aqueous solution.

Classical Molecular Interaction Potential. cMIP calculations were used to locate waters in hidden regions of the proteins (see above), as well as to determine the ability of the proteins to interact with water and hydrogen peroxide. For this purpose cMIP calculations[32,39] were carried out using H_2O and H_2O_2 as probes. For the flexible H_2O_2 up to 20 conformers were considered for each grid position The interaction energies for all of these conformers were corrected by their internal stability and were then Boltzman-averaged as noted elsewhere.[32] In all of the cases, the differential desolvation penalty between H_2O and H_2O_2 was introduced using the solvation free energies determined by the cMIP.[32] The grid was centered in the center of the main channel, and a spacing of 1 Å was used. A polar grid of 18 degrees was used to study the different orientations of a probe molecule in a given grid element.

Results and Discussion

Structure of the Protein in Aqueous Solution. MD calculations for the native and mutant proteins in aqueous solution provide stable trajectories for the 5 ns simulation time. The regions of the configurational space sampled along the trajectories are close to the crystal structure. This is shown in the very small values (all atoms-RMSd around 1.0 (wild-type) and 1.5 Å (mutant)) of the root-mean-square deviation (RMSd) between the flexible part of the protein in the two trajectories and the reference crystal structure (see Figure 1). These small RMSd values confirm the suitability of MD trajectories to analyze the structure of catalase in aqueous solution.

For our purposes it is useful to divide the mobile part of the protein into five regions (g1−g5 in Figure 2): (i) g1 corresponds the internal part of the main channel, and is defined by residues A66−A73, A109−A126, and A148−A156, (ii) g2 corresponds to the external part of the main channel and is defined by residues A157−A189, (iii) g3 defines the channel to access to the interface between monomers (residues D59−D63), (iv) g4 is the channel connecting with the NADP(H) binding site (residues A142−A147), and (v) g5 encloses the rest of mobile residues in the protein. The two proteins are quite well represented by the MD-averaged conformation, as noted in the very small RMSd in Table 1. Regions with particular importance in ligand binding and catalysis (g1−g4) are not more rigid than the rest of mobile regions of the protein (g5), as seen in the RMSd in Table 1, suggesting that flexibility can be important for the functionality of the protein. The external part of the main channel (g2) is the most flexible part for both the wild-type and mutant protein, and the channel connecting the heme group and the interface between monomers (g3) is also quite flexible. On the other side, the channel connecting the heme group and the NADP(H) binding site seems to be the less flexible part.

RMSd for regions g1−g5 with respect to the crystal conformation show the existence of small movements in some of these regions (see Table 1). The movements are in all cases small

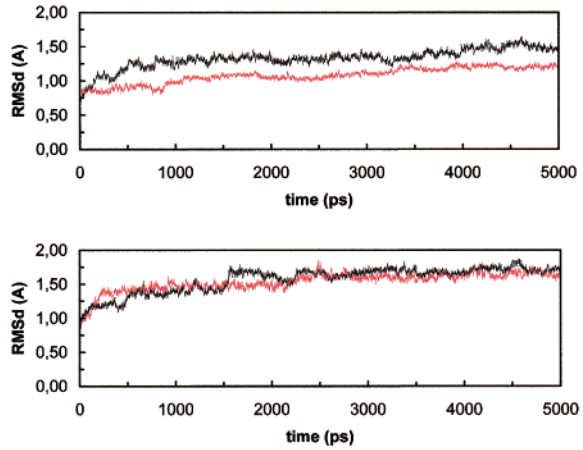

Figure 1. Root-mean-square deviation (RMSd in Å) plots between MD trajectories and the crystal structure for the wild-type (top) and mutant proteins (bottom). RMSd for the mixture of H_2O_2/H_2O (black) and for pure aqueous solution (red) is shown.

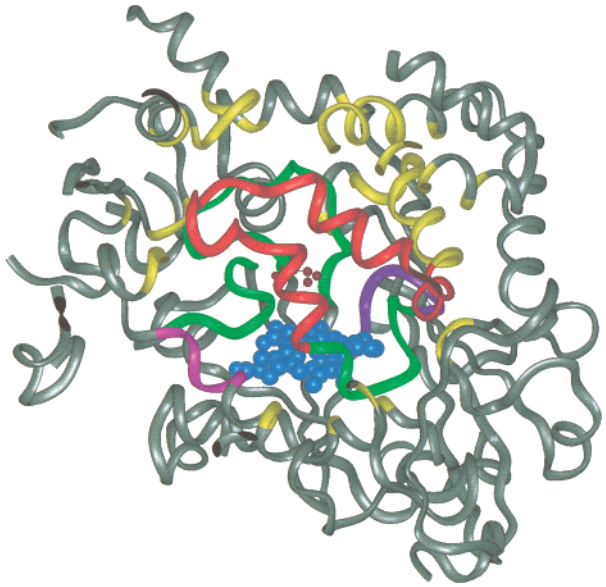

Figure 2. Ribbon representation of the protein system, indicating the five regions analyzed independently (see text): G1 (green), G2 (red), G3 (pink), G4 (violet), and G5 (yellow). Residue Val[111], replaced by Ala in the mutated protein, and the heme group are explicitly represented with balls and sticks.

(38) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(39) Shields, G.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7463.

Table 1. All Atoms Root Mean Square Deviation (RMSd in Å) between the MD-averaged, and Crystal Structures for the Native (plain) and Mutant (*Italic*) Trajectories in Water[a]

reference	G1	G2	G3	G4	G5
MD- av	0.7(0.1)	1.1(0.1)	1.2(0.3)	0.5(0.3)	0.9(0.1)
	0.9(0.2)	*1.3(0.1)*	*1.4(0.4)*	*0.5(0.1)*	*0.8(0.1)*
crystal	1.0(0.1)	1.0(0.2)	1.7(0.3)	1.1(0.3)	0.9(0.1)
	1.2(0.2)	*1.8(0.2)*	*1.4(0.3)*	*1.0(0.1)*	*1.1(0.1)*

[a] MD-averaged conformations (for native and mutant proteins) are obtained by averaging the entire trajectories followed by a restricted minimization. Standard deviations in parentheses.

Table 2. All Atoms Root-Mean-Square Deviation (RMSd in Å) between the MD-averaged, and Crystal Structures for the Native (Plain) and Mutant (Italic) Trajectories in the Mixture of Water and Hydrogen Peroxide[a]

reference	G1	G2	G3	G4	G5
MD- av	1.0(0.1)	1.7(0.3)	1.0(0.2)	0.5(0.1)	0.7(0.1)
	1.1(0.2)	*1.5(0.2)*	*1.0(0.2)*	*0.4(0.1)*	*0.8(0.1)*
crystal	1.0(0.1)	1.9(0.3)	0.8(0.2)	0.6(0.2)	0.9(0.1)
	1.5(0.2)	*2.2(0.3)*	*1.2(0.1)*	*0.7(0.1)*	*1.1(0.1)*

[a] MD-averaged conformations (for native and mutant proteins) are obtained by averaging the entire trajectories followed by a restricted minimization. Standard deviations in parentheses.

(typically less than 2 Å) and do not introduce major alterations in the global structure of the active regions. The channel that links the heme group to the interface (g3) deviates around 1.7 (native) and 1.4 Å (mutant) from the crystal structure. The mutant shows also nonnegligible (1.8 Å) displacements from the crystal conformation in the external part of the main channel (g2). However, these movements correspond to an external loop and do not introduce any important changes in the accessibility to the main channel.

In summary, catalase has a nonnegligible flexibility in aqueous solution, even in regions of catalytic importance. However, according to our 5 ns simulations in aqueous solution, protein movements in these regions do not introduce any remarkable change in the local arrangement of residues important for catalysis and ligand binding. Even though subtle differences are found in the wild-type and mutant trajectories, no major structural changes appear to be associated with the mutation Val111→Ala in the MD trajectories in aqueous solution.

Structure of the Protein in 30% Hydrogen Peroxide/Water Solution. MD simulations in mixtures of water and hydrogen peroxide for the native and mutant protein yield stable trajectories for the 5 ns simulation times (see Figure 1). All atoms-RMSd in Table 2 shows that the flexibility of the protein is similar in water and in the 30% mixture, as noted in the RMSd between the trajectories and their corresponding MD-averaged structures. As found in the simulations in water the external part of the main channel (g2) is the most flexible region for both the wild-type and mutant proteins in the mixture, while the channel connecting the heme group and the NADP(H) binding site is the less flexible region.

Comparison of all atoms-RMSd between the trajectories and the crystal structure (Table 2), and all atoms-RMSd between the trajectories in mixed solvent and the MD-averaged structure in water (Table 3) shows the existence of only small structural changes in both native and mutant proteins related to the change of solvent. Not surprisingly, the changes are mostly located in the g2 region. Analysis of protein−protein H-bonds in H_2O and H_2O/H_2O_2 solvents shows that (see Table 4) the number of unstable H-bonds (those with time-occupancy greater than 10% and lower than 90%) is increased in the presence of hydrogen peroxide, especially in the g2 region. On the contrary, the number of very stable H-bonds (time-occupancy greater than 90%) is decreased (see Table 4). These results suggest that the external loop in g2 is more flexible in the presence of H_2O_2 than in water and moves between different states, involving breaking and formation of new low-occupancy H-bonds.

Table 3. All Atoms Root-Mean-Square-Deviation (RMSd in Å) between the MD Trajectories H_2O/H_2O_2 Referred to the Water MD-averaged Conformations[a]

reference	G1	G2	G3	G4	G5
native	1.8(0.1)	2.5(0.1)	1.0(0.2)	0.6(0.2)	1.2(0.1)
mutant	1.3(0.2)	2.1(0.2)	1.1(0.2)	0.7(0.1)	0.9(0.1)

[a] Standard deviations in parentheses.

Table 4. Low (between 10 and 90%) and High (more than 90%) Time-Occupancy Intramolecular H-Bonds in the G2 Region for Wild-Type and Mutant Proteins in Water and the Mixture of Water and Hydrogen Peroxide

type of H-bond	wild-type H_2O	mutant H_2O	wild-type H_2O_2	mutant H_2O_2
low occupancy	49	54	71	65
high occupancy	14	13	8	8

Interestingly, RMSd between the structures sampled in the mixture and that obtained in the crystal structure or those sampled during the simulation in water are small for all of the internal parts of the protein (assuming no chemical reactions between the protein and the hydrogen peroxide). This suggests that even large quantities of H_2O_2 do not introduce dramatic (conformational) changes in the interior of the protein. There are, however, subtle differences in the interior of the main channel, which in the presence of hydrogen peroxide is slightly wider than in water, both for the native and mutant protein. The residues Val111 and Phe149 define the narrowest point of the main channel for the wild-type protein in water (shortest heavy-atom−heavy atom distance around 6 Å). As expected, the equivalent distance (Ala111−Phe149) is around 0.7 Å larger for the mutant protein. In the presence of large amounts of H_2O_2 the narrowest points of the main channel are increased in 0.8 Å (wild-type protein) and 0.4 Å (mutant protein). Owing to the large flexibility of the Phe149 side chain, the narrowest point of the channel in the mutant protein is defined by Pro124 and Phe159 in the presence of 30% H_2O_2.

In summary, catalase has a structure very similar to that of the crystal,[27] pure aqueous solution, and in the presence of large quantities of hydrogen peroxide. The structure of the enzyme resists very large concentrations of hydrogen peroxide (no potential chemical oxidations are considered in MD simulations). This ability to maintain the structure in the presence of large concentrations of H_2O_2 should be important from a biological point of view, since considering the high apparent K_m,[10] this enzyme might work under conditions of large H_2O_2 concentration, as those existing in situations of oxidative stress.

Analysis of the Water versus Hydrogen Peroxide Affinity of the Main Channel. CMIP calculations of the wild-type and mutant protein allowed us to obtain a picture of the ability of the proteins to interact with H_2O or H_2O_2 (note that due to the use of classical mechanics, this type of analysis cannot be accurately performed for radical species such as compound I). cMIP calculations were carried out using the corresponding MD-averaged structures in the H_2O/H_2O_2 mixture (similar results, data not shown, were obtained for the average structure obtained in pure water), and interaction values were corrected by the desolvation cost of transferring the water and hydrogen peroxide to a given position of the protein.

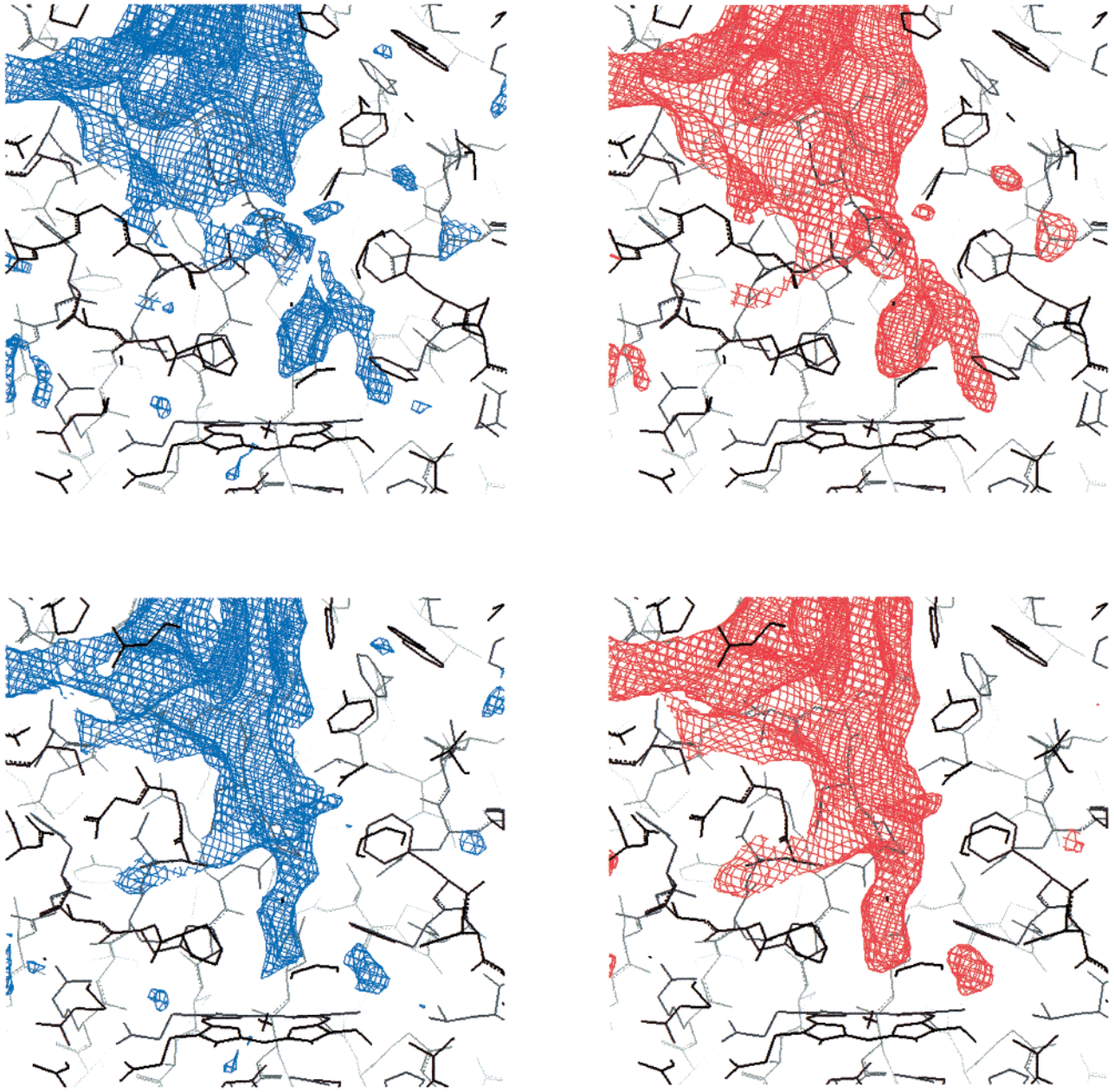

Figure 3. Classical molecular interaction plots (cMIP) representing the interaction of water (left) and hydrogen peroxide (right) in the region around the active site of catalase. Plots are given for the wild-type (top), and the mutant (bottom) proteins. The −1 kcal/mol contour is represented.

cMIP plots in Figure 3, which correspond to equivalent contour levels, clearly define the shape of the main channel, which shows a remarkable ability to interact with both water and hydrogen peroxide. A small extension of the main channel to the NADP(H) binding site might suggest the existence of an alternative channel, but this hypothetical channel, however, seems to be quite close when catalase is in its average "substrate-binding" conformation. Interestingly, the region corresponding to the best interaction for H_2O and H_2O_2 is placed in the vicinities of the heme group, suggesting that the substrate arrives very well oriented to the heme group, which would result in a large increase in catalytic efficiency.

Contours in Figure 3 show a small, but nonnegligible energetic preference of the enzyme for H_2O_2, which allows the enzyme to discriminate between water and hydrogen peroxide under conditions of large concentrations of H_2O_2. However, if the concentration of hydrogen peroxide is not very large, the enzyme may preferably bind H_2O instead of H_2O_2 due to the large amount of water in aqueous solutions. cMIP values suggest, then, an unexpected inhibitory role for water in the binding mechanism of catalase.

The Val111→Ala mutation leads to an enlargement of the channel, which according to cMIP calculations interacts better with small polar probes. Thus, a discontinuous spine of hydration is observed in the channel of the wild-type protein (Figure 3), but a continuous isodensity contour appears for the mutant protein for the same contour level. This means that more water can enter the channel as a consequence of the Val111→Ala mutation, and this water is better structured in a continuous network (see below). The mutation also leads to a slight increase in the intrinsic ability of the channel to bind H_2O_2 (see Figure 3), but no major alterations in the structure of H_2O_2 strings are found.

In summary, cMIP calculations are able to trace accurately the main channel, and suggest the existence of other alternative channels, as those previously suggested by inspection of X-ray structures.[40] Our cMIP calculations suggest that the main channel has only a small preference for H_2O_2 instead of H_2O. Finally,

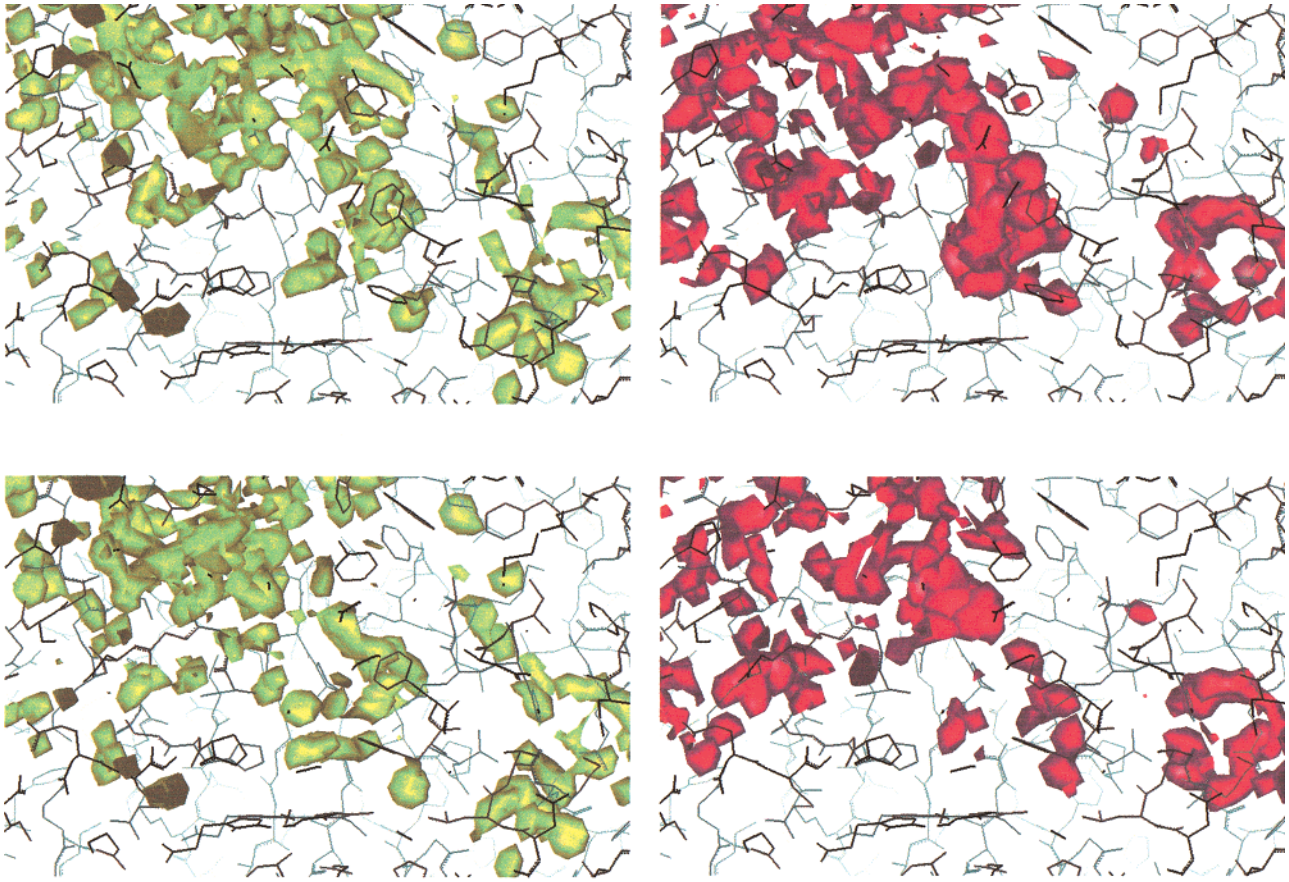

Figure 4. Regions of large apparent density of water (left) and hydrogen peroxide (right) obtained in the MD simulations of wild-type (top) and mutant (bottom) proteins. Trajectories represented here correspond to those obtained in the H_2O_2/H_2O mixture. Contours represented here correspond to a density of 2 times respect to the background density.

classical MIP calculations suggest that the mutation Val111→Ala does not lead to important changes in the intrinsic ability of the enzyme to bind hydrogen peroxide, while it improves the binding of water in the main channel.

The inspection of solvation maps (i.e., maps derived from the analysis of apparent solvent density along the trajectory) allows us to introduce in the analysis dynamic information that is not considered in the simpler cMIP calculations. Density (H_2O and H_2O_2) calculations were performed using the trajectory obtained in the MD trajectories of the wild-type and mutant proteins in the H_2O/H_2O_2 mixture (Figure 4). Similar contours (for H_2O) are found in the trajectory obtained in pure water (data not shown).

Solvation maps in Figure 4 show the large ability of the main channel to interact with both water and hydrogen peroxide. There are two other regions that concentrate H_2O and H_2O_2 above from the background level: one is located in the channel connecting the heme group and the NADP(H) binding site, and the other is located in a region connecting the heme group and the interface between the two monomers. It is worth noting that these two alternative channels are less populated than the main channel and are discontinuous, especially for H_2O_2. Therefore, although these regions can act as alternative channels (as previously suggested[40]), some conformational changes might be necessary to open completely these hypothetical channels, which are more likely to be alternative routes for product release rather than channels for substrate entry.[28,40]

Inspection of the density of H_2O and H_2O_2 in the wild-type protein in the 30% hydrogen peroxide solution shows the existence of more hydrogen peroxide than water in the main channel. The plotted contour (2 times density from background) is continuous for H_2O_2, while it is discontinuous for H_2O, supporting the suggestions derived from the cMIP analysis. Integration of water population in the channel region shows that around 4.2 H_2O_2 and 4.5 H_2O are found on average in the interior of the channel. Considering that the ratio H_2O/H_2O_2 is 3.7 in the system, our results suggest a preferential binding factor around 3.5, favoring the binding of H_2O_2 over H_2O. This moderate discriminating power of catalase between water and hydrogen peroxide found by density analysis agrees well with cMIP analyses. Both theoretical analyses suggest that the poor apparent binding of H_2O_2, found experimentally, is the result of the competition between water and hydrogen peroxide. We predict then, that the apparent K_m of the enzyme is, in fact, dependent on water concentration, and for diluted solutions of H_2O_2 in water, high apparent K_m (in the molar range) can be expected, in good agreement with experimental results.[9,10] Our calculations suggest that catalase is an enzyme designed to have a mediocre differential affinity toward its natural ligand. This implies that the enzyme is engineered by evolution to work only under high H_2O_2 concentration. Catalase allows, then, the existence of moderate concentrations of H_2O_2 in the cell, which might be necessary for many cellular processes.

Comparison of the density maps obtained in the 30% solution of water and hydrogen peroxide for the wild-type and mutant proteins illustrates the change in relative H_2O_2/H_2O binding affinity originated by the Val111→Ala mutation. It is clear that there is more water in the interior of the main channel in the mutant than in the wild-type protein (see Figure 4). Integration of water population reveals an increase of almost one water

(40) Sevinc, M. S.; Maté, M. J.; Switala, J.; Fita, I.; Loewen, P. C. *Protein Sci*. **1999**, *8*, 490.

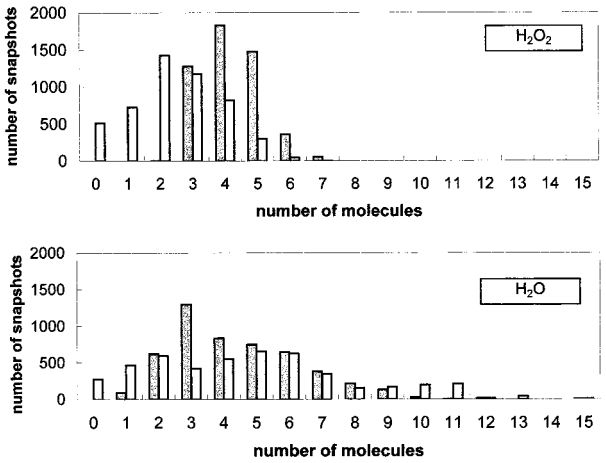

Figure 5. Population (number of molecules found simultaneously) plots for H_2O_2 (top) and H_2O (bottom) along the trajectory of the wild-type (grey boxes) and mutant (white boxes) proteins.

molecule (from 4.5 to 5.2 on average) in the main channel region due to the Val111→Ala mutation. The increase in water population is related to the enlargement of the main channel, which allows, in fact, the definition of a secondary strand of water molecules (Figure 4). There is, then, a good accordance with cMIP calculations, which suggests that the mutation leads to an increase in the binding of water in the main channel.

The Val111→Ala mutation decreases binding of hydrogen peroxide in the channel (Figure 4) around 1.8 H_2O_2 molecules. This means that the ratio of H_2O_2 versus H_2O molecules bound decreases (for the 30% solution) from 0.9 (wild-type) to 0.5 (mutant protein). This is the result of a similar affinity to hydrogen peroxide combined with a better affinity for water (see cMIP results). Since the affinity for water increases, more water molecules occupy the channel, removing hydrogen peroxide molecules from there. The poorer enzymatic efficiency of the mutant protein is then a consequence of the larger ability of the channel to bind water, which hinders the entry of hydrogen peroxide, as suggested by previous X-ray studies.[27−29] Once again, the inhibitory characteristics of water become evident. At this point we should note that very recently (while this contribution was under revision) Jones[41] has suggested an additional mechanism by means of which a water molecule placed near the active site can modulate the catalytic properties of the enzyme. It is clear then that water is not a simple spectator in the mechanism of action of catalase.

Analysis of the instantaneous population of H_2O_2 and H_2O molecules in the main channel allows us to determine possible cooperative effects in the binding of H_2O_2 in the main channel, that is, whether the regions of large apparent densities and favorable cMIP (Figures 3 and 4) are occupied simultaneously by H_2O_2. If cooperativity in the binding of H_2O_2 exists, this should be detected by a large number of snapshots where several molecules are detected simultaneously. The loss of one or a few of these molecules should lead then to a total disruption of the H_2O_2 atmosphere in the interior of the main channel. On the contrary, if the large apparent densities of H_2O_2 found in solvation maps are due to high, but independent populations of contiguous H_2O_2 binding sites (i.e., no cooperativity of binding exists), we should expect a more random distribution of population states in the channel.

Results in Figure 5 show very clearly that there is a cooperative effect in H_2O_2 binding for the wild-type protein.

(41) Jones, P. *J. Biol. Chem.* **2001**, *276*, 13791.

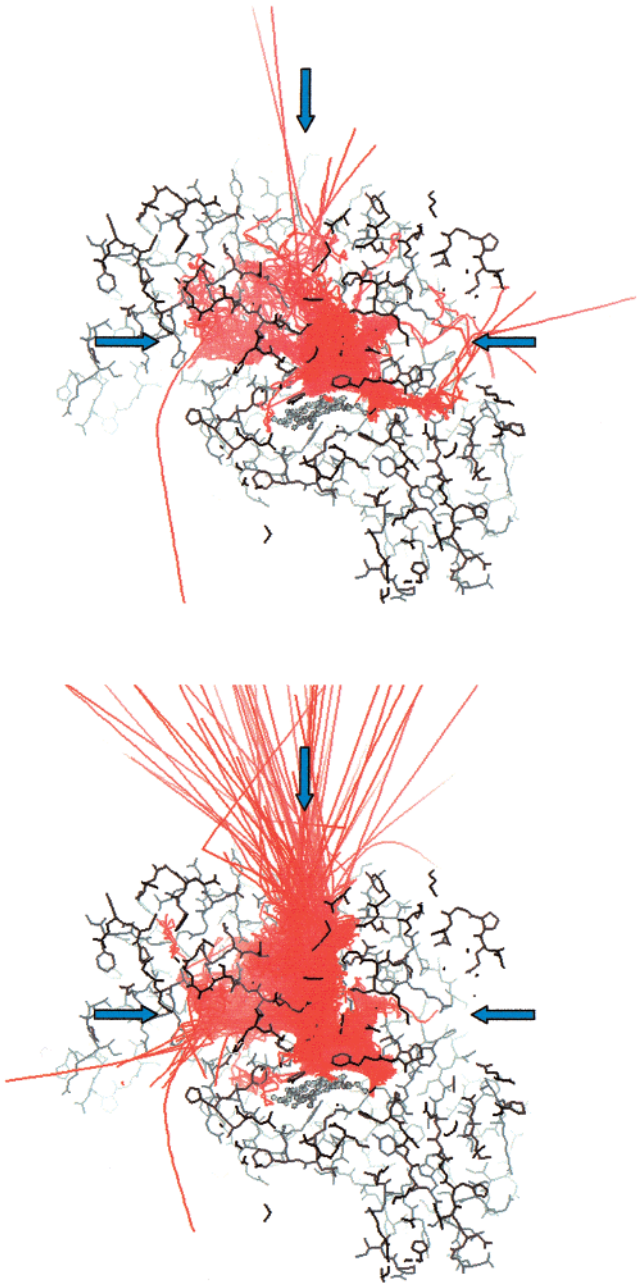

Figure 6. Trajectories obtained in the aMD simulations of wild-type (top) and mutant (bottom) proteins. For each structure the main channel, NADP(H) and interface channels are noted as arrows located at top (main), left (interface), and right (NADP(H)) channels.

The main channel binds 4 ± 1 H_2O_2 molecules simultaneously, and just very few snapshots are detected with higher or lower population. For the mutant protein the main channel binds only around 2−3 molecules simultaneously, and the flattening of the distribution plot indicates a clear loss of cooperativity. Similar plots for water (see Figure 5) show a more random distribution, suggesting the lack of clear cooperativity for the binding of water molecules along the main channel. MD simulations confirms, then, previous hypotheses on the cooperativity of H_2O_2 binding in the main channel.[27,29]

Preferred Routes for Substrate-Entry/Product-Release. Standard trajectories in water or H_2O/H_2O_2 mixtures, and cMIP calculations allowed us to define the main channel and to determine its ability to interact with H_2O and H_2O_2. However, the analysis of minor channels is difficult since the time length of opening of such channels can be larger than the length of

the simulation. Activated molecular dynamics can help to determine the existence of alternative channels for entry/release of substrate/products.

Up to 800 exit routes (100 trajectories started from 8 snapshots of the protein in pure aqueous and 30% mixture) were analyzed by 5 ps trajectories, which allows us to explore preferred routes for interchanging solutes/products between the protein and the solvent. We should note that the simulation time of each trajectory is limited by computational cost, as well as by the need to avoid artifacts arising from the hydrophobic collapse of channels in the absence of water. Fortunately, 5 ps of aMD seem enough to capture in many cases the exit of the substrate/product from the enzyme. Moreover, aMD trajectories seems to be largely independent of the random velocity assigned, and on the starting configuration selected, which allows us to integrate results for all the trajectories, increasing then the statistical quality of the results. For the mutant protein in water, where aMD were performed with and without explicit solvent, the agreement between both aMD runs was excellent (data not shown), even though the aMD simulations in the gas phase provide more "productive pathways" with less computational effort. The suitability of aMD protocols in the gas phase for the exploration of minor pathways for subtrate/product entry/release is then clear.

Activated trajectories for the wild-type and mutant protein are displayed in Figure 6, where as reference the crystal structure is shown in all cases. It is clear that the main channel (noted by the arrow in top of the structures in Figure 6) defines the main route for the interchange of solute/products, but other minor channels, similar to those suggested by solvation analysis also occur. One of these secondary channels (left arrow in Figure 6) connects the heme group with the interface between monomers, while another secondary channel (right arrow in Figure 6) connects the heme group with the NADP(H) binding site. As suggested from previous structural analysis[25] this latter channel appears to be bifurcated, with an upper path passing through Gln192 and Leu196 and a wider lower path through Asn143 and Pro146. We should note that the existence of the NADP(H) secondary channel agrees well with previous suggestions derived from X-ray analysis,[40] as well as with the high conservation of residues along this channel.[28]

Despite the general similarity between the set of aMD trajectories found for the wild-type and mutant proteins, there are nonnegligible differences between the two structures. Thus, the population of exit trajectories passing through the main channel is larger for the mutant than for the wild-type protein. This is not an unexpected result, since the main channel is wider for the mutant protein than for the wild-type (see above). There are not major differences between the two proteins in the interface channel, since the slightly larger number of exit trajectories found in the wild-type protein seems to be fortuitous. There are, however, clear differences between the wild-type and mutant proteins in the low-path of the NADP(H) channel, which is mostly blocked for the mutant protein, while it is open for the wild-type. Analysis of the structures shows that the opening of this path of the NADP(H) channel is related to the movement of the side chains of Asn143 and Thr145. It is unclear whether these movements are related to the mutation, or what is more likely is a fortuitous situation related to the dynamic of the protein. In any case, aMD simulations suggest the existence of this secondary channel in two conformations, open or closed, which can interchange easily as a consequence of slight changes in the side-chain orientation of a few residues.

The large catalytic efficiency of the enzyme under conditions of a large concentration of H_2O_2 implies that a huge amount of products (H_2O and O_2) are generated, and need to be released from the active site. This can be done partially through the main channel, but the channel might be too narrow to allow an efficient release of water molecules. The existence of secondary channels might be an alternative mechanism for a fast, noninhibitory release of products under conditions of high H_2O_2 stress. The fact that mutations that enlarge the putative NADP(H) channel lead to an increase of the catalytic efficiency of the enzyme, combined with an increase to the accessibility of the active site to inhibitors support[40] the hypothesis of a "back door" mechanism for catalase.

Acknowledgment. Thanks are due to Drs. J. C. Ferrer, F. J. Luque, P. C. Loewen, F. Koller, M. J. Mate, S. Wodak, and W. Melik-Adamyan for inspiring discussions. This work has been supported by the Centre de Supercomputació de Catalunya (CESCA; Mol. Recog. Project) and by the Spanish Ministerio de Ciencia y Tecnología PM99-0046 and BIO099-0865.

JA010512T