

# Computer simulation of water in cytochrome *c* oxidase

Xuehe Zheng, Dmitry M. Medvedev, Jessica Swanson, Alexei A. Stuchebrukhov\*

*Department of Chemistry, University of California, One Shields Avenue, Davis, CA 95616, USA*

Received 9 October 2002; received in revised form 11 December 2002; accepted 23 December 2002

## Abstract

Statistical mechanics and molecular dynamics simulations have been carried out to study the distribution and dynamics of internal water molecules in bovine heart cytochrome *c* oxidase (CcO). CcO is found to be capable of holding plenty of water, which in subunit I alone amounts to about 165 molecules. The dynamic characterization of these water molecules is carried out. The nascent water molecules produced in the redox reaction at the heme  $a_3$ –CuB binuclear site form an intriguing chain structure. The chain begins at the position of Glu242 at the end of the D channel, and has a fork structure, one branch of which leads to the binuclear center, and the other to the propionate d of heme  $a_3$ . The branch that leads to the binuclear center has dynamic access both to the site where the formation of water occurs, and to delta-nitrogen of His291. From the binuclear center, the chain continues to run into the K channel. The stability of this hydrogen bond network is examined dynamically. The catalytic site is located at the hydrophobic region, and the nascent water molecules are produced at the top of the energy hill. The energy gradient is utilized as the mechanism of water removal from the protein. The water exit channels are explored using high-temperature dynamics simulations. Two putative channels for water exit from the catalytic site have been identified. One is leading directly toward  $Mg^{2+}$  site. However, this channel is only open when His291 is dissociated from CuB. If His291 is bound to CuB, the only channel for water exit is the one that originates at E242 and leads toward the middle of the membrane. This is the same channel that is presumably used for oxygen supply.

© 2003 Elsevier Science B.V. All rights reserved.

**Keywords:** Molecular dynamics simulation; Water; Cytochrome *c* oxidase

## 1. Introduction

Despite the remarkable recent progress in structural and kinetic studies of cytochrome *c* oxidase (CcO), the molecular mechanism of proton pumping by this enzyme remains largely unknown [1–5]. Water plays a crucial role in the enzyme's function. First of all, water is produced in the enzyme's catalytic site, which is located deeply in the interior of the molecule, and there is an unresolved question about the mechanism and the pathways of water removal. Second, while the electrons required for oxygen reduction can be transferred over large distances between redox co-factors [6–10] without any “assistance”, by quantum mechanical tunneling in protein medium [6], the protons, also required for the oxygen reduction reaction, as well as those that are pumped by the enzyme, by themselves can move only short distances limited to one hydrogen bond, and therefore chains

of water molecules are required to assist proton translocation over long distances [11]. This occurs via the well-known Grotthuss mechanism [12,13]. Thus, water is a crucial functional component of the enzyme. Most of the internal water molecules are dynamic—otherwise, they would not be useful for the Grotthuss transport of protons, and therefore cannot necessarily be seen in the crystal structure. (The crystal structure does show some, but not all water molecules in the enzyme [14].) Thus, without the detailed data about water in the enzyme, the structural and dynamic characterization of the enzyme is incomplete. Computer simulations in this situation can be very helpful.

In this paper, we report the results of our computer simulations of water in bovine heart cytochrome oxidase. The questions that we address are: (1) How many water molecules are there in the enzyme [we consider here only subunits I and II—the central processing unit (CPU) of the enzyme], and what are their dynamic characteristics, that is, how do they move in the protein? (2) What is the stability and dynamic characteristics of water chains for proton transfer? (3) What is the fate of nascent water molecules produced in

\* Corresponding author. Tel.: +1-530-752-7778; fax: +1-530-752-8995.  
E-mail address: [stuchebr@chem.ucdavis.edu](mailto:stuchebr@chem.ucdavis.edu) (A.A. Stuchebrukhov).

the enzyme—that is, what are the mechanisms and channels of their removal? What is their role in proton translocation? The answers to these questions have considerable implications for understanding the CcO function.

Previously, several authors utilized computer simulations to investigate related issues in CcO [15–18]. In particular, Ref. [15] simulated solvation and dynamics of internal water in the bacterial enzyme, together with dynamics of oxygen in the vicinity of the catalytic site, and Refs. [16,17] investigated presence and dynamics of water in the D and K channels of the bovine CcO. None of the previous studies, however, addressed the question of what happens with water molecules produced by the enzyme.

## 2. Methods

The structure of bovine heart CcO due to Yoshikawa et al. [19] is taken from the protein data bank (PDB entry 2OCC). This structure corresponds to fully oxidized enzyme (resolution 2.3 Å). From this structure, we take subunits I and II. (These two subunits have essentially identical bacterial counterparts, and are believed to be sufficient for understanding both catalytic activity and proton pumping of the enzyme [1].) The hydrogen atoms are added using Amber program [20].

For correct simulation of interactions of water molecules with the enzyme, one need to know, among other things, its protonation state. If protonatable amino acids are taken in their standard state, the resulting total charge of the enzyme is  $-21$ . The electrostatic simulations [21,22] indicate, however, that most of the acidic groups will be protonated (resulting in a total charge about  $-5$ , in reduced form) and the distribution of charge qualitatively follows the principle of local electroneutrality. Following this principle, for simulations, we take the enzyme in a completely neutral state, which is obtained by protonating all acidic groups and deprotonating all basic groups. We should mention, however, that the most significant interactions for this study are the local Lennard–Jones interactions which determine the space available for water molecules within the enzyme. In this respect, the electrostatics should play only a minor role here.

In the published structure of bovine CcO, the water molecules are absent. Water molecules are placed in the hydrophilic pockets of the enzyme using Dowser program [23], which finds the cavities within the protein and evaluates the energy of water–protein and water–water interactions in the cavities. Previously, this method has been used in Ref. [15] for similar studies on a bacterial (*Paracoccus*) enzyme. The Dowser application is essentially the same as in Ref. [15] with minor variations in parameters described below. In the calculations, the effective chemical potential of bulk water was varied in the range  $-10$  to  $-5$  kcal/mol, to quantify the hydrophobicity of interior protein regions, in particular those around the catalytic site. After each dowser run, the newly placed water molecules are

fixed, and the dowsing is repeated until no new water is found.

The range of  $-10$  to  $-5$  kcal/mol for the effective chemical potential of bulk water is obtained as follows. The enthalpy of vaporization of water at 298 K is 10.5 kcal/mol, while the change of entropy from the liquid to gas phase at 298 K is  $14.3R$ , where  $R$  is the gas constant [24]. From the enthalpy of vaporization, one finds that the average potential energy of water molecules in liquid state is  $-9.9$  kcal/mol, or approximately  $-10$  kcal/mole (i.e., 5 kcal/mol/bond). If one assumes that the entropy of a water molecule in a protein site is the same as in the liquid state ( $8.4R$ ), the average potential energy of the molecule within the protein should be  $-10$  kcal/mol or less, for this site to be occupied. This is the lower boundary for the potential. On the other hand, it is clear that for hydrophobic sites, the entropy of the site can be much higher than that in the liquid state. The higher entropy results in the lower chemical potential of the site by  $T\Delta S$ . For an estimate, we assume that the average dimension of the site is about 3 Å, and allows for unhindered rotations in the site, as in the gas phase; then the entropy change from the liquid state to a state in such a protein site is approximately  $7.5R$ , or  $T\Delta S = 4.5$  kcal/mol. This change is largely due to the rotational entropy in the hydrophobic site,  $5.5R$ , which is practically absent in the liquid state; see relevant discussion in Ref. [12]. Hence, for the hydrophobic protein site to be occupied, its potential energy can be  $-5.5$  kcal/mol, or lower. Because the energy is evaluated approximately, we take the upper boundary for the potential to be  $-5$  kcal/mol. (Here we should point out that the actual accuracy of even sophisticated force fields, such as Amber, is much worse than 0.5 kcal/mol.) The Dowser program evaluates only the average potential energy of the site (no entropy), for which we set the cutoff in the range  $-10$  to  $-5$  kcal/mol to identify the occupied sites. This cutoff value can be understood as an effective chemical potential of bulk water relative to the protein interior. The values around  $-10$  kcal/mol or slightly lower for the bulk potential have often been used by other researchers in the past, for example, Refs. [15,25].

The dynamics of water molecules are simulated using TIP3P model [26] and Amber [20] forcefield for proteins. The metal center and heme parameters are taken from Charmm force field [27], with atomic charges obtained from quantum mechanical ZINDO [28] calculations [29]. The molecular dynamics (MD) simulations at constant temperature are carried out using Amber 5 program. The water molecules and other residue side-chain groups are placed for belly run under Shake constraint for 250 ps after equilibration. Coulomb forces are computed with a cutoff distance of 25 Å in pairwise evaluations. The integration time steps are 0.5 fs at room temperature and 0.1 fs at higher temperatures, with 0.5 ps relaxation time for  $T = \text{const}$  MD runs. Overall, the general computational technique is similar to that used in Ref. [15].

In the kinetic studies of water exit, the high-temperature simulations are used. Because the exit is an activated

process, the rate of exit  $k$  exponentially increases with increasing temperature. Thus, at sufficiently high temperature, the microsecond or even millisecond processes could be observed within our typical 250 ps dynamics runs. In the dynamics simulations, most of the protein structure is kept fixed; therefore, high temperature does not lead to any significant structural changes of the protein matrix. Thus, it was possible to perform simulations at temperatures as high as 5000 K. The increase in temperature is equivalent to lowering the activation barriers, because  $k \sim \exp(-V_a/T)$ ; therefore, a long-time event can be effectively simulated on a short time scale.

### 3. Results and discussions

#### 3.1. Water distribution and its general dynamic characterization

As predicted by the Dowser simulation, the presence of water is quite extensive in CcO. After four well-converged dowserings, a total of 165 water molecules are located in the subunits I and II (Fig. 1). The chemical potential of bulk was set to  $-5$  kcal/mol, to put initially as many as possible water molecules in the protein.

The bulk potential of water in the Dowser simulations was set to the upper boundary,  $-5$  kcal/mol, to reduce the possibility of missing some important water molecules, such as those in the catalytic site. The main goal here was to generate a typical water distribution in the protein for the subsequent dynamic simulations, and to see if any water molecules will be present in the catalytic site at a high bulk potential. We should stress that the Dowser program explicitly evaluates only the average potential energy of the site, and not the entropy (the entropy contribution in our simulation, however, is taken into account phenomenologically, see Methods section); moreover, the force fields used by the Dowser, and even by the more accurate Amber program, are not completely reliable; in particular, electrostatic interactions of atoms at close distances are problematic—such energies are high, but the partial charges and dielectric properties of the protein medium are not very well defined. On the other hand, the cavities within the protein can be predicted quite accurately. Given the difficulties with energy evaluations, the number of water molecules found by the Dowser program involves some degree of uncertainty. (Rigorously speaking, this number can be considered only as an upper limit.)

As expected, the greatest presence of water is observed in the hydrophilic cavity above the hemes, and near the Mg site, where we find three water molecules covalently bound to  $\text{Mg}^{2+}$  cation with bond lengths 2.24, 2.20 and 2.28 Å. This water coordination to  $\text{Mg}^{2+}$  compares well with locally refined X-ray structure (see Fig. 2). The distribution of water in both D and K proton channels are not continuous. Instead, it is concentrated in protein cavities in these two channels.

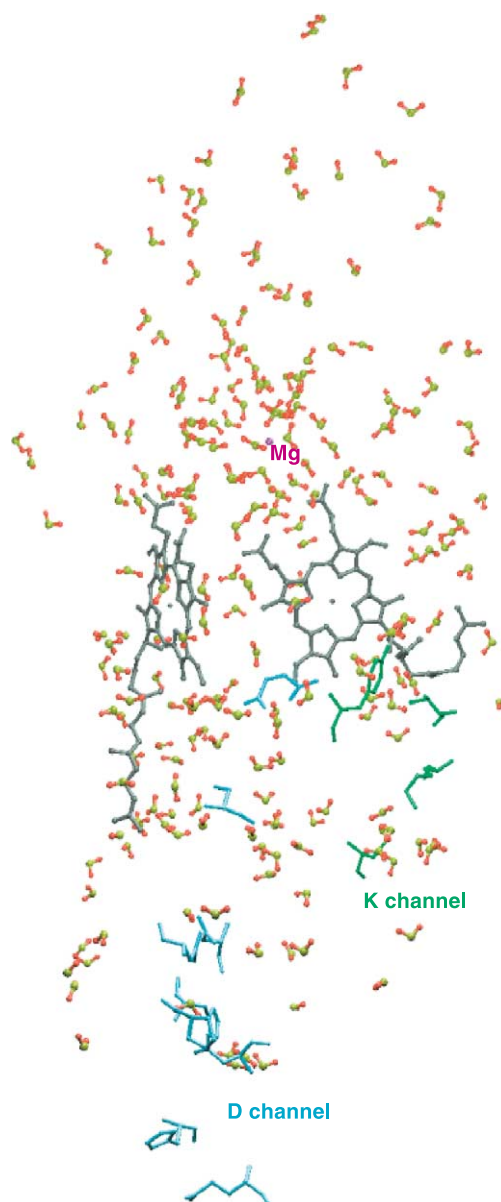


Fig. 1. Water distribution in subunits I and II of CcO.

Such water distribution suggests that protein residues play the role of intermediate retranslation sites in the proton translocation process. Continuous chains of water molecules are found in the H channel, however, and in some regions, along the inter-heme space.

In the Dowser simulation, the heme pockets, and the catalytic site hold no water even at  $-5$  kcal/mol potentials of the bulk. These sites are obviously highly hydrophobic. Given the structure of and near the catalytic site, its hydrophobicity is not surprising. The Dowser prediction of complete absence of water molecules in this important region of the protein, however, should be taken cautiously. We have already remarked about the reliability of the Dowser predictions and the discussion of this important issue will be continued in the next section. From the operation of the

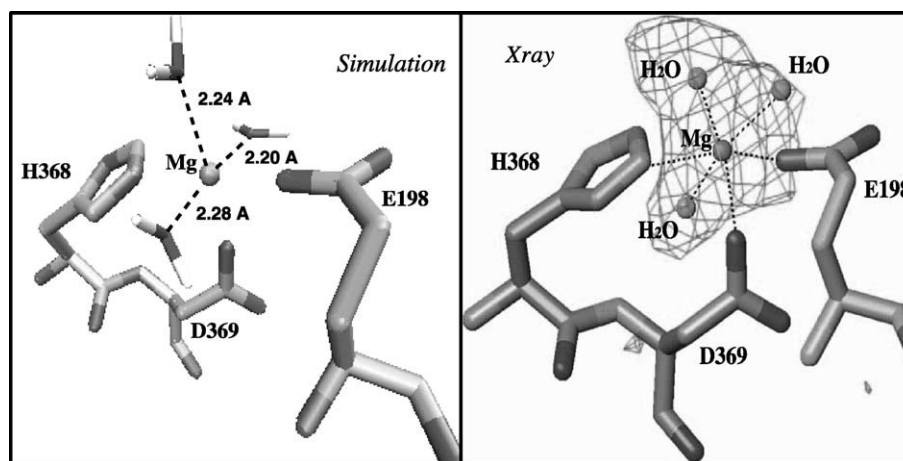


Fig. 2. Simulated water coordination in  $\text{Mg}^{2+}$  site, and comparison with X-ray data of Ref. [19].

enzyme, it is clear that there should be at least some thermodynamically stable water molecules present in this region. The earlier computational studies (with a different method) also indicate that this is likely the case [16]. Also, the hydrophobic pocket of heme  $a_3$  is the source of new water molecules, and this region must have some water molecules in a transient state. Both issues of stable and metastable water molecules in the catalytic region will be re-examined in the next section in a more accurate Amber simulation.

First, however, we would like to characterize dynamically those water molecules that are identified in the equilibrium Dowser simulations. The main question is how different/similar these water molecules are from those in the bulk. Here we focus on the hydrophilic water cluster region above the hemes, where plenty of water has been found.

The found equilibrium Dowser configuration is used as an input for subsequent energy minimization and dynamic calculations with Amber. In Fig. 3a, the average quadratic displacement of water in the enzyme and in the bulk is shown for comparison during 250 ps run at room temperature. The protein structure is kept fixed in these simulations. It is seen that the translational motion of water in the enzyme is much hindered compared with that in the bulk. It is clear that the time scale for the translational water movement in the enzyme is much greater than 250 ps, in agreement with the millisecond time scale observed in the related experimental studies [32].

In Fig. 3b, the rotational (or orientational) correlation function of a typical water molecule in the enzyme is compared with that in the bulk. We find that the time scale for reorientational water motion in the enzyme is shorter than that in the bulk. The reorientation dynamics in the bulk is determined by the dynamics of breaking of hydrogen bonds, where each water molecule has close to four hydrogen bonds [12]. The faster reorientational dynamics in the protein can be understood as the enzyme molecules have

less hydrogen bond constraints than in the bulk, which is reasonable, given much lower density of water in the enzyme and mainly hydrophobic environment of the protein interior. Also, the plateau in the enzyme correlation function indicates that water molecules in the enzyme have preferred orientation.

### 3.2. Water molecules in the region of catalytic site

We now discuss the behavior of water molecules in the region of the catalytic site. No such water molecules are predicted by the equilibrium Dowser calculations. The motivation here is 2-fold. First of all, we know that water is produced in the catalytic site, and the simulation in this section provides some insights into the dynamic nature of such water molecules. Second, the accuracy and the predictive power of the Dowser program used in the previous section is not very high. Here, we re-evaluate the energy of water molecules in the catalytic site by using a more sophisticated Amber program, and re-examine the equilibrium predictions.

We simulate the production of water in the enzyme by placing a water molecule in the catalytic site between Fea<sub>3</sub> and CuB atoms, equilibrating it, and running dynamic simulation at 298 K, as specified in the last section. At the end of a 250-ps run, a new water is placed in the heme  $a_3$  pocket. The energy evaluations are performed during the dynamic run. The initial equilibration redistributes water molecules, and the subsequent dynamics simulation arranges molecules into a quasi-stationary structure. We repeat this procedure until it visually becomes obvious that there is no much space left in the pocket, and that the nice structure of the formed hydrogen-bonded chains, described below, is destroyed when the pocket becomes overcrowded. In addition, the average potential energy is used to control the internal pressure (positive van der Waals energy), and to estimate the upper limit on the number of water molecules in the catalytic site more accurately.

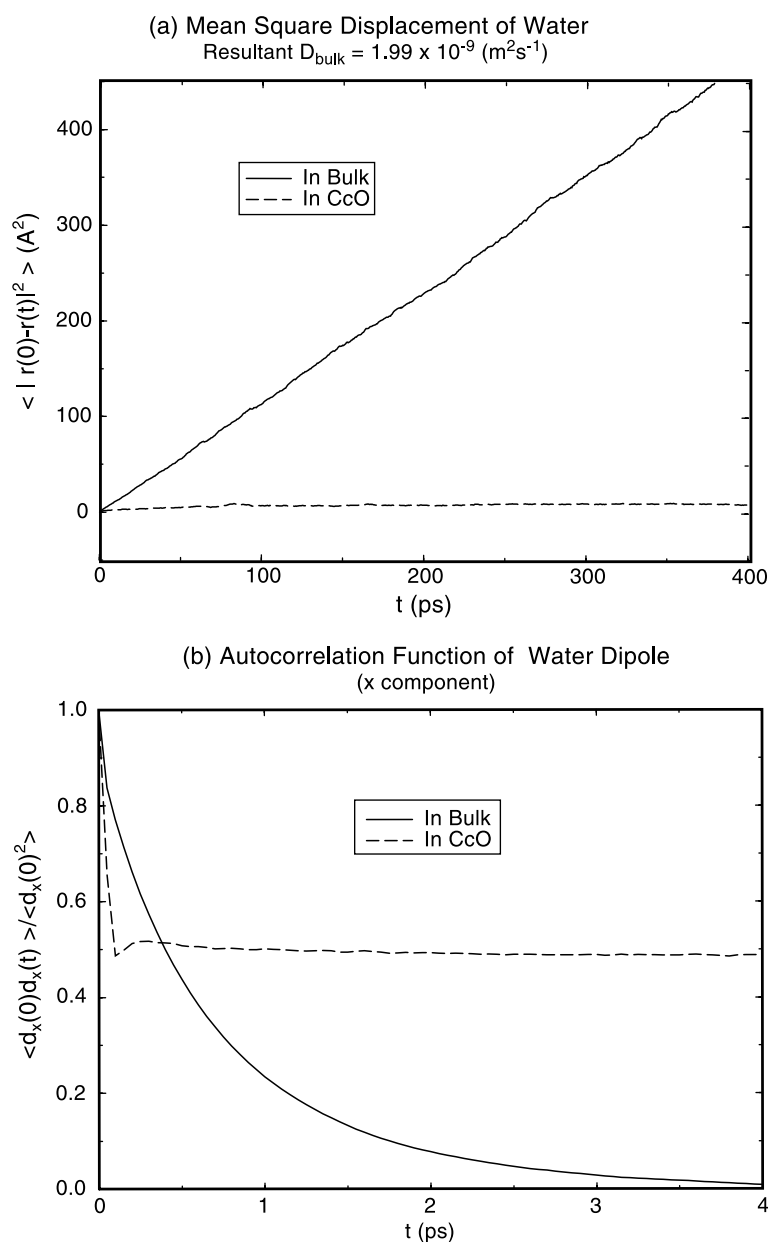


Fig. 3. Dynamic characterization of water in the hydrophobic cluster. (a) Quadratic displacement in bulk and in the enzyme. (b) Orientational correlation function in the bulk and in the enzyme.

Up to 9–10 water molecules could be placed in the described manner in the pocket that begins at Glu242 position, extends both toward propionate d of the hemes and toward the catalytic site, and from there continues toward the K channel. The formed structure of water in the catalytic site is shown in Fig. 4. The analysis of the MD trajectory reveals that the chains of water molecules shown in Fig. 4 form a hydrogen-bonded network. (Here, the hydrogen bond is defined as a bond with distance between oxygen atoms less than 4 Å, as in Ref. [20].)

Up to nine water molecules in the catalytic pocket, the average potential energy of most of the molecules varies in our calculations from  $-2$  to  $+2$  kcal/mol, relative to the

lower bound of the bulk potential of  $-10$  kcal/mol [30]. Given that most of such molecules are arranged in linear chains, with two loose hydrogen bonds of average energy 4 kcal/mol each, their entropy is not expected to be as high as in the gas phase; on the other hand, the molecular entropy should certainly be much higher than in the liquid state, where close to four tight hydrogen bonds restrict the motion of the molecule. It means that the effective bulk potential should be taken somewhere between  $-10$  and  $-5$  kcal/mol. The exact value is impossible to state within our present simulation technique, although the presence of hydrogen bonds would most likely put it somewhere close to  $-8$  kcal/mol. Relative to this level, then, all molecules will have



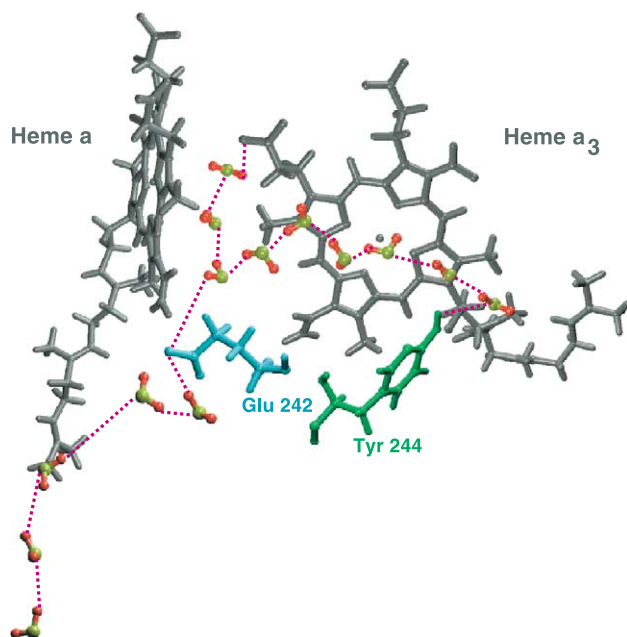


Fig. 4. Hydrogen-bonded chains of water molecules formed near the CcO catalytic site. The forked chain begins at the position of Glu242 (D channel), one branch connects propionate d of heme  $a_3$ , another leads to the catalytic site. Both the catalytic site itself, and the d-nitrogen of His291 can be accessed dynamically via this chain. The chain continues on the right and provides connection to the K channel.

negative potential and should therefore be considered as thermodynamically stable, in contrast to the Dowser predictions. We should admit, however, that even the accuracy of the present calculations is not sufficient to unequivocally state which of the molecules in the catalytic site are thermodynamically stable and which are produced by the enzyme. The fact that the Dowser simulation does not capture these molecules perhaps is not surprising, given that the error in energy evaluations at the very best is  $\pm 2$  kcal/mol (the energy in the Amber and Dowser calculations are evaluated differently).

We note in passing that the high entropy of a hydrophobic site (such as the catalytic pocket of COX) on the one hand results in a lower chemical potential of the site, and increased chances that water molecules will be present there; on the other hand, the same high entropy prevents water molecules in the site to be detected in the X-ray structures, because of their high mobility/entropy. As we pointed out earlier, in general, such entropic effects can give as much as 4.5 kcal/mol in the driving force to occupy the site. We should mention that none of the water molecules in the catalytic region discussed here have been observed in the structural studies. Given their mobility, this perhaps is not surprising.

As the number of water molecules increases above nine, their average potential energy increases (by 3–5 kcal/mol in our present calculations). This increase of energy can be utilized as a driving force for the removal of newly produced water molecules from the catalytic site. Because of the crude nature of our energy evaluations and entropy estimates, it is

difficult to calculate accurately the exact magnitude of this driving force. The unrelaxed structure of the protein utilized in our calculations also contributes to this uncertainty.

The energy landscape in the region of the catalytic site is not completely flat, but has minor variations. For example, the three upper water molecules between the heme edges in Fig. 4 have average potential energy of 2 kcal/mol higher than those in the vicinity of the metal sites. These energy variations are less than the energy of one hydrogen bond, and easily accessible for water generated by the enzyme, given the amount of energy released in the reduction reaction in the catalytic site. They are also within the range of dynamic variations of the energy of individual molecules along the dynamic trajectory.

It is obvious, that the chains of water molecules shown in Fig. 4 can be used for proton transport to catalytic site, both from the K channel and the D channel, via Glu242. Moreover, the chain of water molecules that begins at Glu242 position and has a fork structure, one branch of which runs toward the catalytic site, and the other toward the propionate d of heme  $a_3$ , can be utilized in the pumping. One can speculate that the pumped protons (all or some) coming through the D channel via Glu242 can use the branch that leads directly to propionate d, avoiding the catalytic site altogether.

The possible involvement of the delta propionate of heme  $a_3$  in proton pumping has been earlier suggested in the

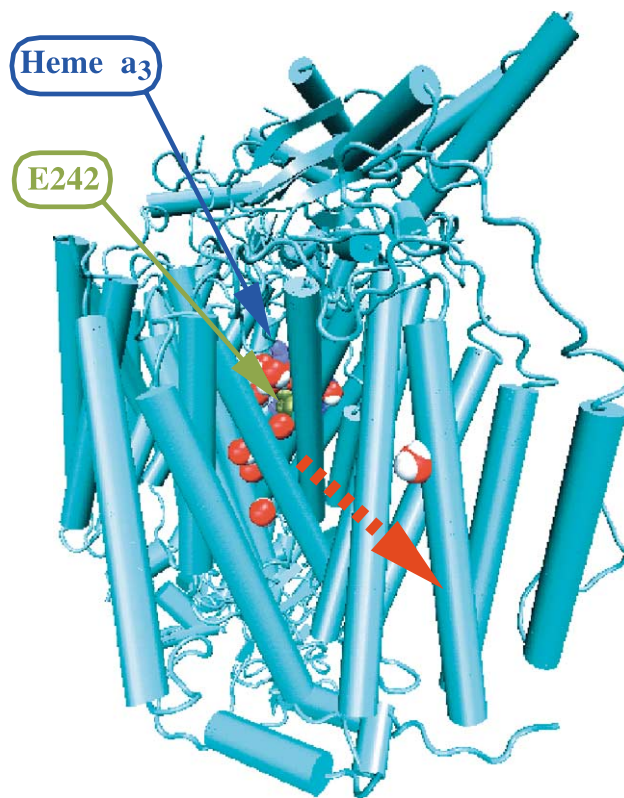


Fig. 5. Putative water exit channel 1. This is the same channel as that presumably used for oxygen supply to the enzyme. It begins at Glu242 position and leads to the middle of the membrane.

literature, for example, Ref. [3]; moreover, Puustinen and Wikström [31] provided some experimental evidence that this may actually be the case.

The branch that connects Glu242 and the catalytic site can also dynamically access the d-nitrogen of His291, and it is also conceivable that protons can be delivered to that position [16]. This coupling, however, is less stable and needs further examination.

Earlier, Ref. [16] described simulations of water in D and K channels. An important finding was the presence of water molecules “above” Glu242 that could provide a link to the catalytic site and to the d-nitrogen of His291. However, no connection of Glu242 and propionates of heme  $a_3$  has been observed in that simulation. As we mentioned earlier, no experimental evidence has been given so far for the presence of any water molecules in this hydrophobic cavity. The link between Glu242 and the d-propionates of the heme  $a_3$  obviously is the central result of the present paper.

In the MD simulations, we saw several events of conformational change of Glu242, similar to those reported earlier in Ref. [17]. As in this earlier work, the conformational change is such that the carboxylic group of the Glu is most stable when pointing toward the D channel (as shown in Fig. 4), while there is a metastable state (about 4 kcal/mol higher in energy [17]) in which this group points toward propionates of heme  $a_3$  (see Fig. 4).

We emphasize that at least some of the discussed water molecules here are produced by the enzyme. (A more accurate evaluation of thermodynamic properties of water molecules in the catalytic site, which would more accurately predict which of the molecules shown in Fig. 4 are thermodynamically stable, is in progress in this group.) This simulation supports earlier proposals that the newly formed water molecules may be used for proton translocation toward the catalytic site [1,4]. If so, the question arises as to how the enzyme begins to work, before few first molecules are produced. The present study points to a possible answer to this question. First of all, some of the mentioned water molecules leading to the catalytic site indeed are thermodynamically stable, and sufficient for the initial oxygen reduction chemistry. The high entropy of the site (large space) can be a driving force for the process. In fact, some of such molecules have been reported in Ref. [16]. The new water molecules resulting from oxygen reduction may add pathways for proton translocation.

A surprising aspect of the dynamic behavior of the newly formed water in the catalytic pocket is that they are not easily removed from it. It was possible, for example, to squeeze up to 15 molecules, with clearly high repulsion pressure inside the pocket, yet molecules would not be easily removed from the site. Obviously, there are barriers that keep a finite number of molecules – 8 to 10 in our estimate, in the pocket. The question then is how water is removed from that pocket, if the enzyme continues to reduce dioxygen and produce more water molecules.

### 3.3. Water exit channels

The simulation is carried out on a dynamically equilibrated structure of 10–12 water molecules in the catalytic site pocket shown in Fig. 4. Here, together with water, the following part of the protein is allowed to move: Glu242, the edges of both hemes, together with their propionate d, Phe377, and His291 (the reason for this selection is explained below). All bonds are kept at constant length and only angles (harmonic potentials) are allowed to change. The rest of the protein structure is fixed. Here the whole protein (all subunits) is used in the simulations. Neither the membrane nor the external bulk water is included in the simulation.

At room temperature, no exit events are observed during the first 250 ps of simulation. Obviously, the pocket has potential barriers that keep water in the region from immediate exit; therefore, the exit is an activation event. Because the time scale of this process can be as large as milliseconds (i.e., the overall turnover time), it is impossible to simulate such events in real time at room temperature.

To speed up the exit process, and to explore possible exit channels for water molecules, the simulation is carried out at the raised temperature of 5000 K. (This is a purely computational trick, no protein falls apart, and the high temperature is used only to increase the kinetic energy of water molecules, which is equivalent to lowering the activation barriers.)

In the high-temperature simulations, we mainly probe the space available for water molecules. This space is defined by the overall structure, which we keep fixed, and by the van der Waals radii of the protein atoms (the repulsive part of the potential), which for all practical purposes can be viewed as hard spheres.

At the raised temperature, within the 250-ps run, the exit of all water molecules is observed. It turns out, all water molecules exit via the same well-defined route. It begins roughly at the position of Glu242 (i.e., slightly beyond this residue if one moves from the catalytic site toward the D channel), and continues almost “horizontally” through the whole protein toward the middle of the membrane, as shown in Fig. 5. This is obviously the same channel that is presumably used for oxygen supply, as described in Refs. [15,33]. (In the simulation, there is no membrane; therefore, the exit from this channel is not blocked, and molecules evaporate from the protein when they reach the protein surface.)

The exit through this channel in our simulations is not surprising. For oxygen supply, in the protein, there has to be a channel which connects the hydrophobic interior of the membrane and the catalytic site. If there is such a connection, naturally, water is forced to this channel in the simulation. The question is will water indeed exit through this channel when the protein is surrounded by membrane?

It seems awkward that the produced water molecules would be expelled into the hydrophobic interior of the membrane. On the other hand, the site of water production itself as well as this putative water channel are hydrophobic;

then the more relevant question would be whether the space is available for water to be expelled into the membrane (or space between the membrane and the protein). The diffusion of water in the membrane is known to occur quite efficiently

[34]. Within the present simulation model, we cannot answer this question, and more detailed studies are obviously needed.

We do not observe any other exit events in the simulations under conditions described above (fixed bonds and fixed overall structure). Because the overall structure of the protein is fixed, we do not allow the breathing motion of the protein in the simulation. These motions can potentially provide other exits that we do not see here.

The recent experimental studies [32] indicate that the exit may occur via  $\text{Mg}^{2+}$  site, which is located just above the catalytic space between Fea<sub>3</sub> and CuB. There are two routes that can lead to this site. The first is to “slide” on the surface of the heme a<sub>3</sub>, between the propionates (a water molecule is predicted there by the Dowser simulation), directly to the Mg site. This route, however, is completely blocked by the bulky His291. The second route is to move toward the space between the heme edges (i.e., in the direction of Glu242), squeeze between the hemes, and access the Mg site on the other side of the heme a<sub>3</sub>. The passage between the hemes, however, is completely blocked by Phe377. The dynamic simulations of both Phe377 together with the hemes a–a<sub>3</sub> edges, and His291 (fixed bond lengths) does not change this picture.

If we assume, however, that the bond between His291 is broken, water molecules easily access the Mg site (Fig. 6). It is observed that the His291–CuB bond break is part of the assumptions list of the histidine cycle model by Wikström [2]. Although spontaneous metal–ligand bond break is chemically counterintuitive, and has never been observed experimentally (see further discussion of this issue in Ref. [2]), one can speculate about the effect of internal pressure within the catalytic site. It is also interesting to consider [35] the possibility that the expelled water can carry a proton in the form of the hydronium ion  $\text{H}_3\text{O}^+$  and is part of the pump mechanism.

#### 4. Conclusion

1. There are plenty of water molecules inside the protein. Most occupy thermodynamically stable states; others that are produced in the catalytic site are in quasi-stationary states with finite lifetimes. Simulations agree with available crystallographic data, in particular for water around Mg site and some water molecules in D and K channels. We gave dynamic characterization of these water molecules.

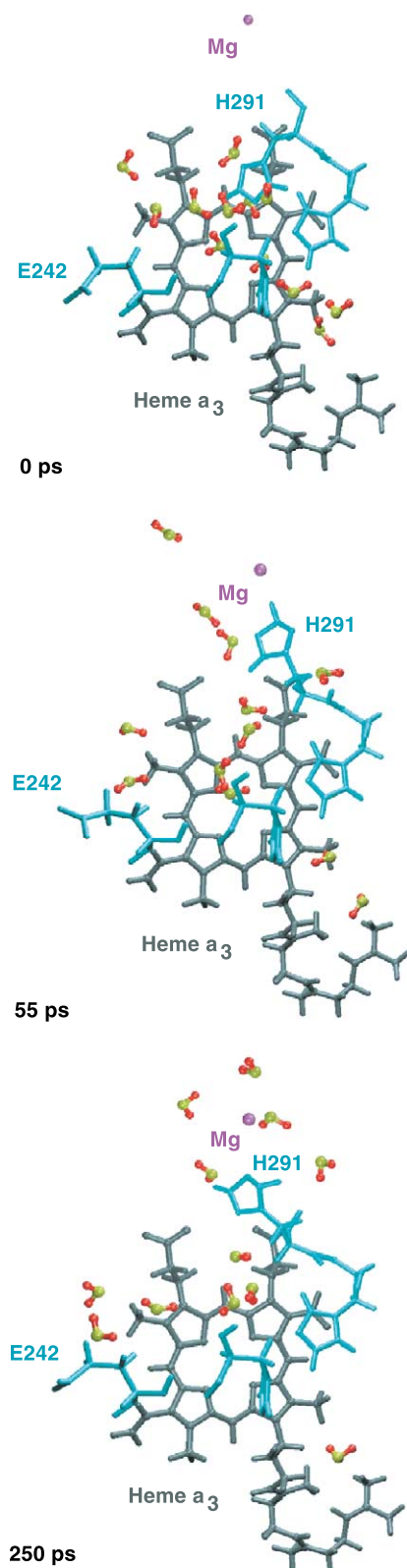


Fig. 6. Putative water exit channel 2. This channel leads directly toward  $\text{Mg}^{2+}$  site; however, it is open only when the bond between His291 and CuB is broken. In the figure, three snapshots of the MD simulation trajectory of water exit from the catalytic pocket are shown. In the upper image, His291 is in the locked position, and in the other two, the His is in the open position and water exits toward the Mg site. The simulation has been performed at an artificially high temperature (5000 K); therefore, the water exit is almost instantaneous when the gate His291 is open. The path beyond the Mg site has not been resolved.



2. There is a (quasi-) stable chain of water molecules, formed from water partially produced by the enzyme, connecting Glu242 (end of D channel), catalytic site, and the propionate d of heme  $a_3$  (Fig. 4) [37]. (A similar finding has been recently reported by Michel [36].) The chain continues through the catalytic site to connect Tyr244 and the K channel. Water molecules form a chain of hydrogen bonds along which protons can be transferred from D (Glu 242) and K (Tyr244) channels to catalytic site and to a position 'above' heme  $a$ –heme  $a_3$  redox pair level.
3. There are water molecules between the edges of heme  $a$  and heme  $a_3$ . These may facilitate electron transfer between the hemes (by providing stronger electronic coupling) and explain recent puzzling data on ET between the hemes [38]. Further study of this issue is underway.
4. Two putative channels for water exit from the catalytic site were identified. One is leading toward  $Mg^{2+}$ , directly from the  $Fea_3/CuB$  water formation site. However, this channel is only open when His291 is dissociated from CuB. If His291 is bound to CuB, the only channel for water exit is the one that originates at Glu242 and leads toward the middle of the membrane. This is the same channel which is believed to be used for  $O_2$  supply. We find that water cannot exit through the D channel.
5. Simulated kinetics in the  $H_2O$  exit channel provides information about the activation barriers for the exit.

## Acknowledgements

We thank Alexander Konstantinov, Joel Morgan, Denis Rousseau, Bob Gennis, and Marten Wikström for stimulating discussion of the results of this paper; Shelagh Ferguson-Miller and Bryan Schmidt for discussing with us their experimental work before publication. Thanks to Dr. Dat Nguyen for his help with MD simulations. This work was supported by NIH (GM54052).

## References

- [1] M. Wikström, *Curr. Opin. Struct. Biol.* 8 (1998) 480–488.
- [2] M. Wikström, *Biophys. Biochim. Acta* 1458 (2000) 188–198.
- [3] H. Michel, *Biochemistry* 38 (1999) 15129–15140.
- [4] R.B. Gennis, *Biophys. Biochim. Acta* 1365 (1998) 241–248.
- [5] D. Zaslavsky, R.B. Gennis, *Biochim. Biophys. Acta* 1458 (2000) 164–179.
- [6] A.A. Stuchebrukhov, *Adv. Chem. Phys.* 118 (2001) 1–44.
- [7] C.C. Page, C.C. Moser, X. Chen, P.L. Dutton, *Nature* 402 (1999) 47–52.
- [8] C.C. Moser, J.M. Keske, K. Warncke, R.S. Farid, P.L. Dutton, *Nature* 355 (1992) 796–802.
- [9] R. Langen, I. Chang, J.P. Germanas, J.H. Richards, J.R. Winkler, H.B. Gray, *Science* 268 (1995) 1733–1735.
- [10] H.B. Gray, J.R. Winkler, *Ann. Rev. Biochem.* 65 (1996) 537–561.
- [11] A.A. Stuchebrukhov, *J. Theor. Comp. Chem.* (2003, Jan.–Feb.) (in press).
- [12] N. Agmon, *Isr. J. Chem.* 39 (1999) 493; N. Agmon, *J. Chim. Phys. (Paris)* 93 (1996) 1714; N. Agmon, *Chem. Phys. Lett.* 244 (1995) 456–462.
- [13] (a) J.F. Nagle, H.J. Morovitz, *Proc. Nat. Acad. Sci. U. S. A.* 75 (1978) 298; (b) R. Pomes, *Isr. J. Chem.* 39 (1999) 195–387.
- [14] S. Yoshikawa, private communication.
- [15] I. Hofacker, K. Schulten, *Proteins* 30 (1998) 100–107.
- [16] S. Riistama, G. Hummer, A. Puustinen, R.B. Dyer, W.H. Woodruff, M. Wikström, *FEBS Lett.* 414 (1997) 275–280.
- [17] R. Pomes, G. Hummer, M. Wikström, *Biophys. Biochim. Acta* 1365 (1998) 255–260.
- [18] D.M. Medvedev, A.A. Stuchebrukhov, *J. Am. Chem. Soc.* 122 (2000) 6571–6582.
- [19] S. Yoshikawa, K. Shinzawa-Itoh, R. Nakashima, E.R. Yaono, N. Inoue, M. Yao, M.J. Fei, C.P. Libeu, T. Mizushima, H. Yamaguchi, T. Tomizaki, T. Tsukihara, *Science* 280 (1998) 1723–1729.
- [20] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, *J. Am. Chem. Soc.* 117 (1995) 5179–5197.
- [21] A. Kannt, C.R.D. Lancaster, H. Michel, *Biophys. J.* 74 (1998) 708–721.
- [22] D. Popovic, A.A. Stuchebrukhov, Electrostatic calculations of cytochrome oxidase, in preparation.
- [23] L. Zhang, J. Hermans, *Proteins: Struct., Funct. Genet.* 24 (1996) 433–438.
- [24] NBS tables for chemical thermodynamic properties, 1982.
- [25] (a) J. Hermans, private communication; (b) P.J. Goodford, *J. Med. Chem.* 28 (1985) 849–857; R.C. Wade, P.J. Goodford, *J. Med. Chem.* 36 (1993) 147–148.
- [26] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, *J. Chem. Phys.* 79 (1983) 926–935.
- [27] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, *J. Comp. Chem.* 4 (1983) 187–217.
- [28] M.C. Zerner, J.E. Ridely, A. Bacon, J. McKelvey, W. Edwards, J.D. Head, C. Culberson, M.G. Cory, X. Zheng, W. Parkinson, Y. Yu, A. Cameron, T. Tamm, G. Pearl, A. Broo, e.a.K. Albert, ZINDO, University of Florida, Gainesville, FL, 2002, Quantum Theory Project, distributed by Accelrys, 9685 Scranton Road, San Diego, CA 92121-3752, USA.
- [29] The charges, molecular coordinates, and some other details of the calculation are available in the supplementary material to this paper.
- [30] The energy evaluations within the protein are done with no dielectric screening of the Coulomb interactions of the partial charges, as no such screening is applied in the bulk state calculations. It is usually assumed that the partial charges are already rescaled to reflect the electronic polarizability of the medium. There is a well-known uncertainty of the Coulomb energy evaluation here, which in some cases may result in an additional factor of 1/1.4 to 1/2.
- [31] A. Puustinen, M. Wikström, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 35–37.
- [32] B. Schmidt, B.L. Florens, L. Qin, J. McCracken, S. Ferguson-Miller, *Biophys. J.* 82 (2002) 1391.
- [33] S. Riistama, A. Puustinen, A. Garcia-Horsman, S. Iwata, H. Michel, M. Wikström, *Biophys. Biochim. Acta* 1275 (1996) 1–8.
- [34] A. Finkelstein, *Water Movement Through Lipid Bilayers*, Wiley, 1987.
- [35] Joel Morgan, private communication.
- [36] H. Michel, EBEC2002 Lecture, Sept. 2002.
- [37] Results shown in Fig. 4 have been reported at *Biophys. Soc. Meeting*, San Francisco, Feb. 2002, and at *EBEC2002*, Sep. 2002, *BBA Short Reports* 12 (2002) 107.
- [38] M.I. Verkhovsky, A. Jasaitis, M. Wikström, *Biochim. Biophys. Acta* 1506 (2001) 143–146.