Protonation of the Chromophore in the Photoactive Yellow Protein

Elske J. M. Leenders,[§] Leonardo Guidoni,[†] Ursula Röthlisberger,[‡] Jocelyne Vreede,[§] Peter G. Bolhuis,[§] and Evert Jan Meijer*[,§]

Van 't Hoff Institute for Molecular Sciences, Universiteit van Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands, Physics Department, Università degli Studi di Roma "La Sapienza", Roma, Italy, and Laboratory of Computational Chemistry and Biochemistry, Institute of Chemical Sciences and Engineering, Swiss Federal Institute of Technology, EPF Lausanne, Switzerland

Received: October 31, 2006; In Final Form: January 23, 2007

The photoactive yellow protein (PYP) acts as a light sensor to its bacterial host: it responds to light by changing shape. After excitation by blue light, PYP undergoes several transformations, to partially unfold into its signaling state. One of the crucial steps in this photocycle is the protonation of *p*-coumaric acid after excitation and isomerization of this chromophore. Experimentalists still debate on the nature of the proton donor and on whether it donates the hydrogen directly or indirectly. To obtain better knowledge of the mechanism, we studied this proton transfer using Car−Parrinello molecular dynamics, classical molecular dynamics, and computer simulations combining these two methods (quantum mechanics/molecular mechanics, QMMM). The simulations reproduce the chromophore structure and hydrogen-bond network of the protein measured by X-ray crystallography and NMR. When the chromophore is protonated, it leaves the assumed proton donor, glutamic acid 46, with a negative charge in a hydrophobic environment. We show that the stabilization of this charge is a very important factor in the mechanism of protonation. Protonation frequently occurs in simplified ab initio simulations of the chromophore binding pocket in vacuum, where amino acids can easily hydrogen bond to Glu46. When the complete protein environment is incorporated in a QMMM simulation on the complete protein, no proton transfer is observed within 14 ps. The hydrogen-bond rearrangements in this time span are not sufficient to stabilize the new protonation state. Force field molecular dynamics simulations on a much longer time scale have shown which internal rearrangements of the protein are needed. Combining these simulations with more QMMM calculations enabled us to check the stability of protonation states and clarify the initial requirements for the proton transfer in PYP.

1. Introduction

Light governs some of the most important processes in life and does this, as with most biological processes, with proteins. The photoactive yellow protein (PYP) is such a light-sensitive protein, acting as a blue light sensor in the bacterium *Halorhodospira halophila*.[1] Excitation triggers a series of events that eventually lead to a negative phototactic response. As it is small and readily available, PYP has been studied extensively, using numerous experimental and computational techniques.[1−16] In the 125-amino acid protein, a deprotonated and negatively charged *p*-coumaric acid (pCA) is attached to Cys69 with a covalent thioester bond. Upon excitation, this chromophore trans−cis isomerizes (see Figure 1), leading to a chain of reactions. One of the steps in this process is a proton transfer, most likely from Glu46, to the deprotonated chromophore.[17,18] Within microseconds after the isomerization, the protein unfolds into its signaling state.[3] The photocycle of PYP is similar to that of rhodopsin, the protein in our eyes responsible for vision.[19] A chromophore covalently bonded to the protein is excited and rotates over a double bond. After this isomerization, a proton-transfer reaction can occur and the protein changes its structure.

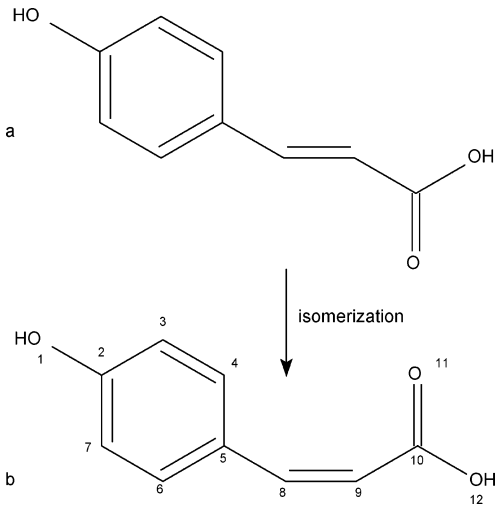

Figure 1. (a) The chromophore, hydroxycinnamic or *p*-coumaric acid (pCA), in its ground-state conformation. The double bond in the tail has a trans configuration. When pCA is bonded to the protein, the OH of the acidic group is replaced by the sulfur of Cys69, forming a thioester bond. (b) The cis state of pCA, after excitation and isomerization. The atom numbers are used in Figures 3 and 4.

With use of UV−vis spectroscopy, several intermediates are identified along the photoreaction pathway. This starts with the ground state, pG. After isomerization the protein visits a red-shifted state, pR, followed by the blue-shifted signaling state,

* Address correspondence to this author. Phone: +31-20-525-6448. E-mail: ejmeijer@science.uva.nl. Fax: +31-20-525-5604.
[§] Universiteit van Amsterdam.
[†] Università degli Studi di Roma "La Sapienza".
[‡] EPFL.

10.1021/jp067158b CCC: $37.00
Published on Web 03/22/2007

pB, after protonation and partial unfolding. Experiments show that pCA is deprotonated and Glu46 protonated in the pG and the pR state, whereas the protonation states for these groups are reversed in the pB state: in the signaling state, pCA is protonated, and Glu46 has lost its proton and is now carrying the negative charge.[4,17,18] Little is known on which intermediates are involved in the transition of pR to pB. The transition from pG to pR occurs on a nanosecond time scale and only affects the chromophore binding pocket, while the process of protonation and unfolding to pB takes microseconds and involves the entire protein. The exact mechanism remains unclear, as the different time and length scales make experimental and computational investigations difficult.

Unlike rhodopsin, which is membrane-bound, PYP is a small water-soluble protein, which results in a wealth of high-resolution structural information. In addition, PYP is a relatively compact protein, which makes it well-suited for computational studies. The isomerization of the chromophore has been described in a number of computational studies; interesting is the work by Groenhof et al.[7−9] and by Sergi et al.[10] Sergi performed quantum chemical geometry optimizations of the chromophore in different charge states. His very accurate structural results are used in this work for validation purposes. He studied the chromophore during the isomerization in vacuum and inside a model protein environment and calculated the energy barrier for this process. Sergi also looked into the excitation of the chromophore. Groenhof defined a force field for the chromophore in all photocycle states and looked into both the isomerization of the chromophore and the subsequent unfolding. We use his force field in this paper. In his later work,[9] Groenhof forced the protonation of the chromophore with thermodynamic integration, using semiempirical PM3 for the chromophore and its direct environment and classic molecular dynamics for the rest of the solvated protein and the water. Other studies of the proton transfer treat the chromophore environment in a very simplified environment.[20,21] Vreede et al.[3] have published interesting computational work on the partial unfolding of PYP after proton transfer, predicting the signaling state of the protein. They used parallel tempering with a classical force field. Antes et al.[22] studied the protein in different states of the photocycle using classical molecular dynamics. Both presented results on the large-scale fluctuations of the protein (backbone root-mean-squared deviations) and quantitated the breaking of hydrogen bonds.

Experimentalists suggest different mechanisms and intermediates for the proton transfer. Genick et al. state that after partial unfolding of the protein, the deprotonated chromophore is solvated in water and takes a proton from one of the solvent molecules.[23] In contrast, Hendriks et al. conclude from their measurements that it is most likely that pCA is protonated by the proton from Glu46. At low pH values, an additional group might be involved.[4] In this paper we address the question if we can find computational evidence to confirm one of these mechanisms.

We use different computational methods to bridge the various time and length scales involved in the photocycle. The larger protein movements are modeled by using force field molecular dynamics (FFMD), whereas we use Car−Parrinello molecular dynamics (CPMD)[24] for the actual reaction site. This density-functional theory (DFT) method allows us to study the electronic structure of the chromophore and its surroundings, enabling us to describe reactions involving chemical bond breaking ab initio. Due to computational limits, CPMD is affordable only for restricted system size and simulation time. Recently developed techniques such as the quantum mechanics/molecular mechanics (QMMM) method[25,26] allow us to combine FFMD and CPMD, enabling the first-principle investigation of reactions in larger systems. In this work, we treat the reaction center of the protein with DFT and use FFMD to describe the rest of the protein and the solvating water. With these methods we study the phenomena involved in the protonation reaction in PYP and we discuss the requirements for proton transfer to occur.

2. Computational Details

To study the initial events in the transformation from the pR to the pB state in PYP, the main region of interest is the chromophore binding pocket (consisting of pCA and the amino acids Tyr42, Glu46, Thr50, Arg52, and Cys69). We simulated two different systems: the relevant amino acids in vacuum and the complete protein solvated in water. For the former, we used CPMD; for the latter we used FFMD and QMMM. The Protein Data Bank (PDB) provided us with the protein structures for FFMD and QMMM. We used 1NWZ[5] as the pG structure and 1OT9[6] as the pR structure; both structures have a high resolution (0.82 and 1.00 Å, respectively).

For the molecular dynamics (MD) simulations, we employed Gromacs[27] and Gromos[28] software with the Gromos96 force field,[29] adjusted with the partial charges and dihedrals proposed by Groenhof et al.[7] for the chromophore. The time step was 2 fs and a Nosé−Hoover thermostat[30] with $\tau = 0.1$ ps controlled the temperature. The protein was solvated in approximately 9000 SPC water molecules in a cubic periodic box with a box length of 66.3 and 66.6 Å for the pG and pR state, respectively. Six Na^+ ions counterbalanced the negative charge of the protein. In the FFMD simulations, the LINCS[31] and SETTLE[32] algorithms constrained the bond lengths and angles. In the QMMM simulations, SHAKE[33] performed this task.

The CPMD simulations were done with the CPMD software package.[34] It employs the Kohn−Sham formulation of DFT, expanding the Kohn−Sham orbitals in plane waves up to a cutoff of 70 Ry. The time evolution of the electronic degrees of freedom is calculated with an MD-like algorithm, using a fictitious electron mass of 500 au. The simulations in vacuum were done with a time step of 0.1 fs; in the protein the time step was 0.12 fs. We employed the gradient-corrected BLYP functional[35,36] to account for the exchange correlation energy. Semilocal norm-conserving Troullier−Martins pseudopotentials[37] were used with cutoff radii of 0.26, 0.59, 0.65, 0.59, and 0.71 Å for H, N, C, O, and S, respectively. A Nosé−Hoover thermostat at 1500 cm^{-1} (∼22 fs) controlled the temperature. At the start of every simulation, we added 8 Å to the largest atomic distance in each direction to determine the size of the periodic box. To avoid rotation of the system during the CPMD simulations, the position of one pCA atom in the phenolic ring and two coordinates of the opposite atom in the ring were constrained, fixing the system in the box with a minimal loss of degrees of freedom (only 5). In these simulations, the hydrogen atom mass was set to 2 amu.

For the QMMM simulations,[38] we used a cutoff of 12.0 Å for the nonbonded interactions. Up to 30.0 Å around the quantum box, the electrostatic interactions were taken into account by using ESP charges.[26] The hydrogen atoms had a mass of 1 amu here. When the QM/MM border was in between two chemically bonded atoms, a link hydrogen atom[25] capped the QM atom. As QMMM simulations require an isolated box, we applied the Tuckerman[39] Poisson solver.

Figure 2 displays the system used for the CPMD (vacuum) calculations, and is referred to as the chromophore binding

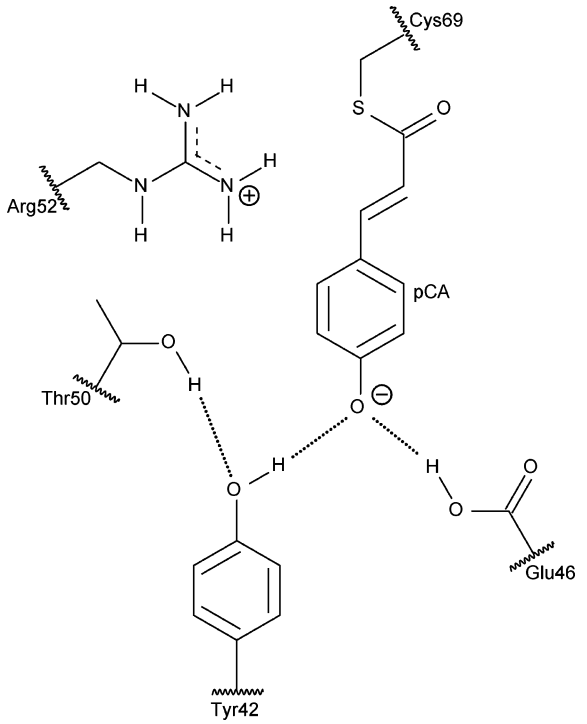

Figure 2. Schematic picture of the chromophore binding pocket in the pG state. The dotted lines are hydrogen bonds. The wiggling lines show where the quantum systems ends; the system was truncated at C−C bonds that are not part of the conjugated electron system and capped with hydrogens. Please note that in reality the positive charge is not localized, but spread over the entire guanidinium group of Arg52. This group is mostly oriented parallel to the chromophore ring at a distance of less than 4 Å.

pocket. It contains the amino acids Tyr42, Glu46, Thr50, and Cys69, which has a thioester bond with pCA. In the pG state, the phenolic oxygen of pCA accepts a hydrogen bond from Tyr42 and Glu46; Thr50 donates a hydrogen bond to Tyr42. Arg52 neutralized the system charge; its guanidinium group is in-plane with the phenolic ring of pCA. We included all these residues in the quantum system, truncating the amino acids at the first C−C bond that was not part of a conjugated π-system. We saturated the dangling bonds by capping with hydrogens. In particular, Tyr42, Thr50, and Cys69 were cut between C_α and C_β and Glu46 and Arg52 between C_γ and C_δ. The quantum system in the QMMM simulations was very similar, but here we did not include Arg52 in the quantum part; it was described by using the force field. Next to these dynamical simulations, we also performed geometry optimizations on the isolated chromophore (protonated and deprotonated, cis and trans) with the default settings in the CPMD software.

All simulations in this paper were done in the canonical ensemble at 300 K (unless stated otherwise). Table 1 gives an overview of the systems we studied and the different simulation methods used for each system; we will refer to the different systems in the rest of this paper according to the designations in this table. Molecular figures in this paper are made with VMD.[40]

3. Results and Discussion

3.1. Structure. For validation purposes, we performed extensive calculations verifying the structural properties of our system. Here we present only the main results; for more detailed and quantitative information we refer to the Appendix.

Figures 3 and 4 compare calculated pCA bond lengths to experimental data for pG and pR crystal structures, respectively. Our calculated bond length values are all in, or very close to, the area that is defined by the experiments. Our DFT-BLYP geometry optimizations of the isolated deprotonated trans and cis chromophore and the neutral cis chromophore yield bond lengths that are within 0.02 Å of the values obtained in DFT-BP,[10] DFT-B3LYP, and CASSCF[41] calculations.

During simulations of the pG state of PYP, we studied the hydrogen-bond network around the chromophore with and without the complete protein environment using FFMD, CPMD, and QMMM. Table 2 shows the calculated time-averaged hydrogen-bond lengths. Our observations are qualitatively consistent with the NMR measurements[11] that show the pCA-Glu46 and Tyr42-Thr50 hydrogen bonds existing in most but not all frames, and the pCA-Tyr42 hydrogen bond being present in all cases. The X-ray measurements[6] suggest that in the protein crystal, the hydrogen-bond network around the phenolic oxygen is rather static, with the pCA-Glu46, pCA-Tyr42, and Tyr42-Thr50 hydrogen bonds all stable. On average, the measured O−O distances are somewhat shorter than the ones resulting from our dynamical FFMD simulations of the solvated protein. Here we should note that the hydrogen bonds to pCA in the crystal structure are extremely short.[6] In the employed force field, these distances yield high energies and are therefore very unlikely to be accessed in the FFMD simulation. Although the difference with experiment is smaller, the quantum chemical simulations show larger hydrogen-bond lengths as well. Still, they do give a good estimate of the length ratio of the hydrogen bonds: almost all simulations show correctly that the hydrogen bond from Tyr42 to pCA is the shorter one. Only when the QM/MM border is placed through the hydrogen bonds are the lengths similar.

We determined the dihedral angles of the chromophore from our QMMM simulations of the protein-embedded chromophore binding pocket (systems E) and from our CPMD simulation of the binding pocket in vacuum (system F). We compared these angles to our PDB starting structure, 1OT9, that has a flat chromophore. Figure 5 shows a typical chromophore configuration from each of the simulations. In our simulations, the chromophore relaxes by bending itself within a couple of picoseconds. In most crystal structures of the pR state (1S4R, 1UWP, 1TS7, 1S1Y), the dihedral angles deviate from zero as well, but with very different magnitudes and directions. In our simulations, the direction of curvature is determined by the possibility to form stabilizing hydrogen bonds, most often with Arg52.

3.2. Protonation of pR. Our particular interest is in the proton transfer to pCA. We will first discuss it using a QM model system of the cis chromophore in the binding pocket in vacuum. We will compare the simplified results to those from QMMM simulations of the protein-embedded pocket. FFMD simulations provide us with further insight.

3.2.1. Binding Pocket in Vacuum. In the QM simulation of the isolated binding pocket (system F), the initial configuration, taken from the pR state, changes within 1 ps. Arg52 rotates in the plane parallel to the chromophore ring and moves toward Glu46 forming a stable hydrogen bond (Figure 6). The other hydrogen bonds in the binding pocket remain intact. Figure 7 shows the subsequent time evolution of the distances between the proton and the Glu46 and pCA oxygen, respectively. It is representative for all QM simulations of the chromophore binding pocket. The figure shows that there is frequent proton transfer between the pCA and the Glu46 oxygen. Within 8 ps, the proton transfers more than 10 times. The proton is always chemically bonded to one oxygen and hydrogen bonded to the

TABLE 1: The Different Systems and Simulation Methods in This Paper

	system	starting point	simulation method
A	pG, protein in water	1NWZ[5]	FFMD
B	pG, protein in water	A, equilibrated	QMMM
C	pG, binding pocket in vacuum[a]	A, equilibrated	CPMD
D	pR, protein in water	1OT9[6]	FFMD
E	pR, protein in water	D, equilibrated	QMMM
F	pR, binding pocket in vacuum	D, equilibrated	CPMD
G	constrained binding pocket in vacuum[b]	D, equilibrated	CPMD
H	*trans*-pCA in vacuum	modified 1NWZ	geometry optimization[c]
I	*cis*-pCA in vacuum	modified 1OT9	geometry optimization[c]
J	pR, as G, without Arg52	D, equilibrated	CPMD
K	as E, proton constrained on pCA[d]	E, equilibrated	QMMM

[a] See Figure 2. [b] The amino acids were truncated between C_α and C_β , fixing the position of the capping hydrogen atom. [c] The default geometry-optimization algorithm implemented in the CPMD software[34] was used. We started with the coordinates from the PDB structures and replaced the sulfur bridge by an OH group. [d] The proton from Glu46 is put on pCA. It stays there because the distance from the pCA oxygen to the proton is constrained at 1 Å.

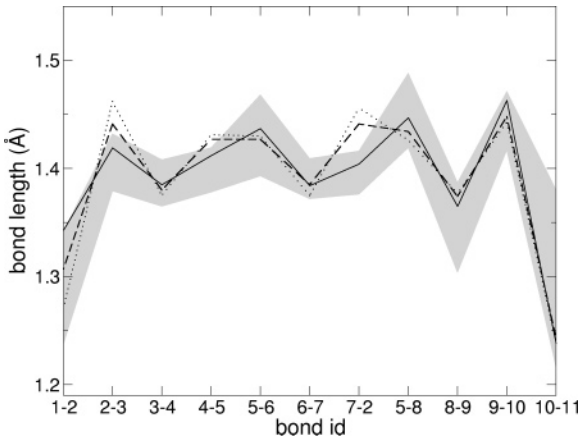

Figure 3. Bond lengths of the negatively charged trans chromophore in different simulations (at 300 K) and in PDB structures. All lengths are in Å. The different lines represent the following systems: CPMD of the binding pocket in vacuum (system C, straight line); QMMM of the binding pocket (system B, broken line); and geometry optimization of the chromophore in vacuum with CPMD (system H, dotted line). The gray area indicates the range of X-ray diffraction measurements[6,13−16] that have a resolution of 0.85 to 2.00 Å and are measured at temperatures from 85 to 295 K. In the case of dynamical simulations (systems B and C), the values are averages over all frames. Figure 1 defines the numbering of the atoms.

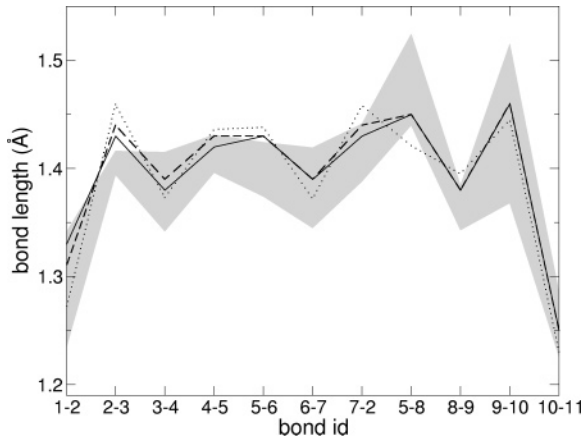

Figure 4. Bond lengths of the negatively charged cis chromophore, defined as in Figure 1, in different simulations (at 300 K) and in PDB structures. All lengths are in Å. The different lines represent the following systems: CPMD of the binding pocket in vacuum (system F, straight line); QMMM of the binding pocket (system E, broken line); and geometry optimization of the chromophore in vacuum with CPMD (system I, dotted line). The gray area indicates the range of X-ray diffraction measurements[6,12,14,15] that have a resolution of 1.00 to 1.90 Å and are measured at temperatures from 85 to 287 K. In the case of dynamical simulations (systems E and F), the values are averages over all frames. Figure 1 defines the numbering of the atoms.

other. The frequent proton transfer allows for estimation of the associated free-energy profile from the probability distribution of a proton-transfer reaction coordinate.[42] Figure 8 shows the free-energy profile as a function of the asymmetric stretch (the difference of the two OH distances), computed from three independent QM simulations of the chromophore binding pocket of 19.7 ps in total. The barrier has a height of around $2.5k_BT$ and the relative stability between the two protonation states is $0.5k_BT$: the proton is on pCA roughly one-third of the time.

Arg52 also hydrogen bonds to the pCA carbonylic oxygen. To accommodate these two Arg52 hydrogen bonds (to Glu46 and to pCA), the chromophore has to bend itself, as can be clearly seen in Figure 5. Figure 6 shows a snapshot where Glu46 and pCA accept one hydrogen bond each from the same NH_2 group. During the simulation, the two hydrogen bonds could be donated by different NH_2 groups as well. Furthermore, Glu46 often received two hydrogen bonds on the same oxygen, from different NH_2 groups (next to the one from pCA on the other oxygen). In the isolated binding pocket, proton transfer from Glu46 to pCA is relatively easy, as the negative charge on Glu46 can be stabilized by Arg52. Within the confinement of the protein, this stabilization is most likely not possible. Although this provides a free-energy barrier and proton-transfer frequency that is not representative for the protein-embedded binding pocket, it does provide information on what conditions are important for the protonation reaction.

To model the confinement of the binding pocket by the protein environment, we performed a 7-ps QM simulation of a "confined'' isolated chromophore binding pocket with the capping hydrogens on fixed positions (system G). The system is larger than the one presented in Figure 2, since all amino acids are now truncated between C_α and C_β, including the "tail" of Glu46 and Arg52 in the system. In contrast to the simulation of the "free'' isolated binding pocket, no proton transfer occurs, indicating a significantly higher barrier for protonation. Also, a "free'' QM simulation of a slightly different system, without Arg52 (system J), does not show the protonation reaction. In both cases, the hydrogen bond between Arg52 and Glu46 cannot be formed. As in the latter simulation the hydrogen bond of Arg52 to the carbonylic oxygen of pCA is not present either, the chromophore tail is much flatter in this case. It seems that stabilization of Glu46 is a prerequisite for proton transfer in this model system in vacuum.

3.2.2. Protein-Embedded Binding Pocket. A 14-ps QMMM simulation of the protein-embedded cis chromophore (pR state,

TABLE 2: Oxygen−Oxygen Distances (in Å) between Important Hydrogen-Bond Donors and Acceptors in the Chromophore Binding Pocket in the pG State[a]

system	method	pCA-Glu46	pCA-Tyr42	Tyr42-Thr50	time	no. of samples
A, 100 K	FFMD	2.73 ± 0.16	2.62 ± 0.13	2.72 ± 0.15	1.00 ns	1000
A, 300 K	FFMD	3.12 ± 1.49	2.70 ± 0.28	3.44 ± 1.02	0.923 ns	923
C, 100 K	CPMD	2.65 ± 0.18	2.63 ± 0.12	2.85 ± 0.31	9.3 ps	958
B, ffG43a1[b]	QMMM	2.75 ± 0.29	2.68 ± 0.19	3.12 ± 0.43	5.72 ps	9449
B, AMBER[c]	QMMM	2.73 ± 0.16	2.65 ± 0.12	3.31 ± 0.32	7.55 ps	12478
B, QM = pCA[d]	QMMM	2.71 ± 0.11	2.72 ± 0.10	3.22 ± 0.23	7.81 ps	16127
X-ray, 110 K[e]	exptl	2.59 ± 0.01	2.50 ± 0.01	2.85 ± 0.01		1
X-ray, 295 K[f]	exptl	2.58 ± 0.01	2.51 ± 0.01	2.89 ± 0.01		1
NMR[g]	exptl	3.44 ± 1.76	2.40 ± 0.37	4.50 ± 2.00		26

[a] Hydrogen bonds are measured as oxygen−oxygen distances in six different simulations and three different experiments from literature. We averaged over all frames available; frames in which the hydrogen bond was nonexistent are also included. The simulation time and number of samples used to calculate the average are listed. Please note that for the X-ray literature results, we give the error. For the averaged calculated and NMR results, we show the variance (double standard deviation). [b] QMMM with the chromophore binding pocket without Arg52 as the QM part (300 K). [c] As before, but using the AMBER force field at 100 K. [d] As before, but with only pCA as the QM part. [e] X-ray crystallography at 110 K.[6] [f] X-ray crystallography at 295 K.[6] [g] NMR measurements at 311 K in solution, averaged over 26 configurations.[11] In a majority of the frames, most O−O and O−H distances were within the hydrogen-bond range. Especially the O−O distance from pCA to Tyr42 was very short, down to 2.21 Å. But officially, hydrogen bonds were not present because of very small O−H−O angles (down to 33°). We assume that this is an artifact of the method used to predict the protein structure from the NMR data. The PDB file shows some very unusual conformations that are in our view unphysical.

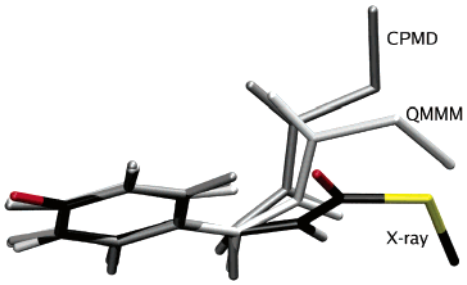

Figure 5. The chromophore from PDB structure 1OT9 (bottom), from the QMMM simulation (middle, white), and from the CPMD simulation of the isolated binding pocket (top, gray). The phenolic rings are fitted on top of each other. The simulated structures are snapshots from the trajectories; this figure shows the snapshot with the dihedral angles closest to the average values. The X-ray structure is almost flat. In the dynamical simulations, the carbonylic oxygen accepts a hydrogen bond from Arg52. To do this, it must bend upward. As the C_β next to sulfur in the QM simulation is not constrained, it may point in any direction.

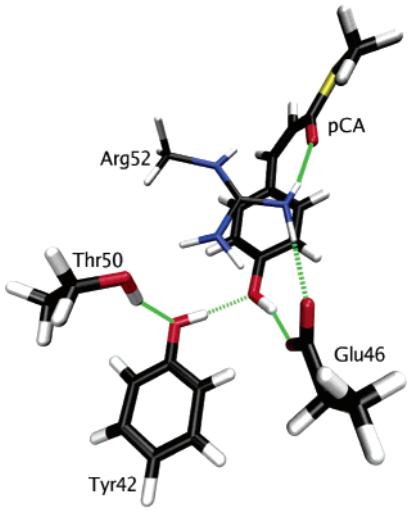

Figure 6. Hydrogen-bond network in a CPMD vacuum simulation with a protonated chromophore. The main difference with the pG state is that Arg52 hydrogen bonds to the carbonylic oxygen of pCA and to Glu46.

system E) shows no occurrence of spontaneous proton transfer. There is also no onset of proton transfer: the proton-to-pCA distance is never smaller than 1.32 Å. To further investigate possible pathways of proton transfer, we manually transferred

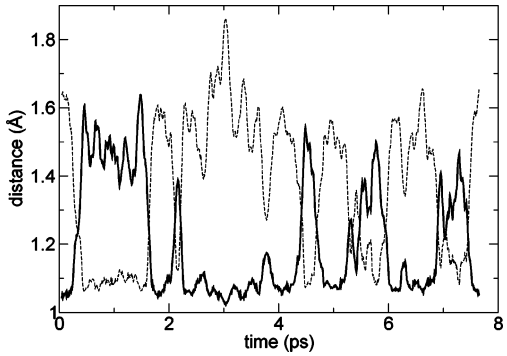

Figure 7. The distance between the proton and the Glu46 oxygen (solid line) and between the proton and the pCA oxygen (broken line) as a function of simulation time. The two lines show the running average (over ten points) for better visibility.

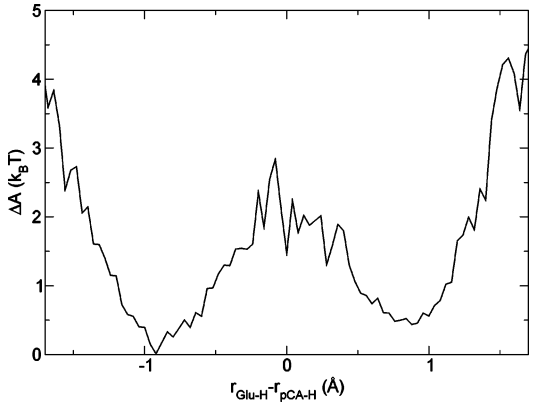

Figure 8. The free energy of protonation in the CPMD vacuum simulations as a function of the asymmetric stretch. The asymmetric stretch is 0 when the proton is exactly in between the two oxygen atoms (one from Glu46 and the other the phenolic oxygen of pCA), around −1 when the proton is on Glu46, and around 1 when the proton is on pCA.

the proton from Glu46 to pCA and constrained it to remain at a distance of 1.00 Å from the phenolic oxygen (system K). In the initial 3 ps of the simulation, the Thr50−Tyr42 hydrogen bond breaks and Thr50 forms a stable hydrogen bond with the deprotonated Glu46 (see Figure 10). After formation of this new hydrogen bond, it does not break during the entire simulation. Its length is 2.80 ± 0.32 Å. The Arg52 residue moves in the

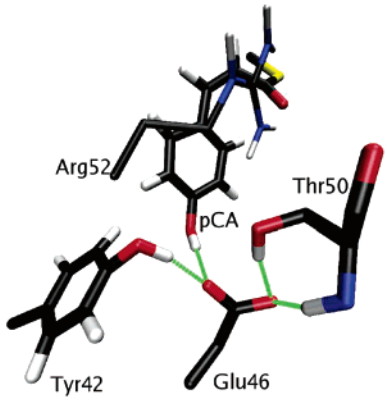

Figure 9. Hydrogen-bond network in the FFMD simulation after manual protonation of the chromophore in the pR state. This is a snapshot at 1.2 ps after the protonation. Glu46 is stabilized by four hydrogen bonds. Please note that this figure also shows the backbone atoms of Thr50, as the backbone nitrogen is now donating a hydrogen bond to Glu46. To do this, Thr50 has moved to the right.

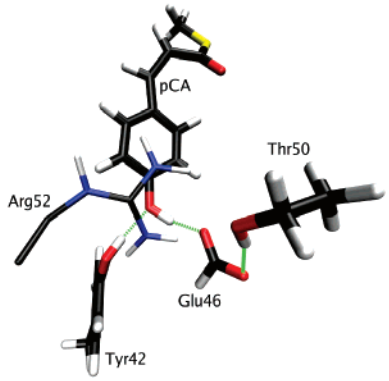

Figure 10. Hydrogen-bond network in the QMMM simulation after manual protonation of the chromophore in the pR state. The OH bond is constrained. This is a snapshot taken 14.5 ps after the protonation. Glu46 is only stabilized by two hydrogen bonds: the one with pCA that was already present before protonation and a new one formed with Thr50 that was hydrogen bonded to Tyr42 before.

direction of Glu46, but it does not form any stable, lasting, hydrogen bonds. It switches between Tyr42, Glu46, and Thr50, which all seem to compete for this hydrogen-bond donor. Over the entire time span of this QMMM simulation (11 ps), the constraint force shows that the proton was continuously pushed back to Glu46. Every 1.2 ps a snapshot was taken from this QMMM trajectory to start a new unconstrained simulation. Indeed, even after 11 ps of relaxation, the proton moves back to Glu46 within a picosecond after releasing the constraint. Apparently, proton transfer adjustment to the new protonation state is much more difficult in the confining protein environment than in the "free'' isolated binding pocket.

Comparison of the simulations of the "free'' binding pocket, the "confined'' binding pocket, and the full protein-embedded system shows that the most relevant difference is the motion of Arg52. Fixation of Arg52 to the protein backbone restricts the motion of Arg52 and prohibits it from reaching Glu46 and forming a stable, strong hydrogen bond. We suggest that this hydrogen bond provides for stabilization of the negative charge on Glu46 and the occurrence of relatively frequent proton transfer observed in the simulation of the "free'' binding pocket. The stabilization by only Thr50, which we see in the protein-embedded system, is apparently not sufficient. Although the large movements of Arg52 observed in our gas-phase dynamics might be prevented by steric constraints in the protein environment, we have shown that a displacement of this residue can significantly improve the proton transfer to the chromophore.

3.2.3. Stabilization by Hydrogen-Bond Rearrangement. To obtain further insight into the possible mechanism of proton transfer and charge stabilization, we studied the protein-embedded chromophore for much longer time scales with FFMD simulations. Initial configurations for these simulations were nine different snapshots of the FFMD simulation of the pR state (system D); in each configuration we manually moved the proton from Glu46 to pCA. In all these simulations, the hydrogen-bond network changes within a few picoseconds to a configuration with Glu46 accepting four hydrogen bonds (i.e., the maximum amount, two on each oxygen) from pCA, Tyr42, Thr50, the backbone nitrogen of Thr50, and, in a later stage, Arg52. During 100 ps of sampling, hydrogen bonds break and (re)form, but most of the time Glu46 accepts four. Figure 9 shows a typical snapshot. To verify the stability of these configurations within the QMMM setup, several of the FFMD configurations were taken as starting points for QMMM simulations. In these simulations the proton stays on the chromophore and no proton transfer occurs. When the proton is put back on Glu46, it returns to the chromophore within a picosecond. This shows that the state with the negative charge on Glu46 is now stable in the QMMM setup. We also took eight intermediate states from the FFMD simulations to check their stability in a QMMM simulation. They show that three hydrogen bonds (from pCA and two from Thr50 in general) are sufficient: the proton remains on pCA. When Glu46 receives on average two hydrogen bonds during the QMMM simulation, it takes the proton back from pCA within a picosecond. The FFMD simulations show a fluctuating number of hydrogen bonds in the first tens of picoseconds, until the stable new protonation state is reached (with on average more than three hydrogen bonds to Glu46). This means that an intermediate state after 2 ps can be stable (three hydrogen bonds), while 10 ps later it can be unstable again. It is noteworthy that in these early stages after shifting the negative charge from pCA to Glu46, the protein stabilizes itself by internal rearrangement of hydrogen bonds. This makes it plausible that, indeed, the negatively charged Glu46 is one of the driving forces for the structural rearrangement that leads to pB, and therefore, that the chromophore is protonated before being solvated. Still, our computationally very expensive methods did not provide the vast number of trajectories one would like to substantiate this line of reasoning.

In our pG simulations, Arg52 stays in-plane with the chromophore ring; in this way, the electron density in the π-system of the ring stabilizes the positive charge on Arg52. In system F, the binding pocket in vacuum after isomerization, Arg52 can donate a hydrogen bond to the pCA carbonylic oxygen while staying in-plane with the ring. In the pR state, the dihedral angles in the chromophore tail fluctuate more than those in the pG state; by bending its tail a bit toward Arg52, the chromophore can form a very stable hydrogen bond that hardly ever breaks. The same applies to Glu46: as soon as a hydrogen bond with any of the Arg52 NH groups has formed, this very stable bond only breaks to be immediately replaced by a bond with another NH group of the same molecule. When restrained by the protein environment, other hydrogen-bond donors have to stabilize Glu46, although these have to break free from their original hydrogen-bond acceptors first. The FFMD simulation of the protonated chromophore shows a more rapid hydrogen-bond rearrangement than the QMMM simulations. Apparently this hydrogen-bond switching is more easily

accomplished within the Gromos96 force field than within the setup with the forces determined by quantum calculations. As DFT is more precise in describing hydrogen-bond interactions, this might indicate that the description of the potential-energy surface of the hydrogen-bond network by the force field is not sufficiently accurate. Other simulation studies of PYP with a protonated cis chromophore used the same force field;[3,8,22] the dynamics of the hydrogen bonds after protonation have not been studied before in a QMMM setup.

The simulations described above elucidate the factors that stabilize either of the two protonation states of the chromophore in PYP. A stable protonation product requires the accommodation of the negative charge on Glu46. In all pR simulations and X-ray structures, a very hydrophobic environment surrounds Glu46. The OH bond points toward the chromophore, but the other oxygen points into a "void" with the nearest amino acids at considerable distance. Hydrogen-bond-donating groups that can help to stabilize the negative charge are not readily available. Therefore, the stabilization has to be accomplished by hydrogen-bond rearrangements in- and outside the chromophore binding pocket.

4. Conclusions and Outlook

In this paper we report force field MD, CPMD, and QMMM simulation studies on the proton transfer in the photocycle of PYP. We have shown that the direct environment of the proton is highly important for this reaction. If the surrounding residues can adjust rapidly and without constraints to stabilize the new protonation state, frequent transfer occurs, as in our gas-phase models. CPMD simulations of the chromophore binding pocket in vacuum show proton transfer between Glu46 and pCA more than 10 times within 8 ps. The estimated barrier for this reaction is $2.5k_BT$. However, in QMMM simulations including the protein environment, protonation does not occur within accessible simulation times. The chromophore binding pocket cannot adjust its configuration to the new state fast enough, because the protein backbone restrains its movement. As the simulations of the binding pocket in vacuum do not include the full system, the estimated barrier is significantly lower than the one inside the protein. Still, this simulation gives us a good indication of which conditions lower this protonation barrier.

We argue that it is the stabilization of the charge on Glu46 after protonation that is important: both Glu46 oxygens need to accept hydrogen bonds to accommodate the charge in the hydrophobic protein environment. In the CPMD vacuum system, an ionic hydrogen bond from Arg52, plus the hydrogen bond with pCA, suffices. In our QMMM simulations, Arg52 is not flexible enough to form a stable hydrogen bond to Glu46 quickly. Only Thr50 switches its hydrogen bond from Tyr42 to Glu46 after manual protonation of the chromophore. FFMD simulations show that other hydrogen-bond donors switch from their original acceptor to Glu46 as well. When both oxygens accept together at least three hydrogen bonds, the new protonation state is stable. All hydrogen bonds are made available by internal rearrangement of protein hydrogen bonds. QMMM simulations proved that this new conformation is stable on a picosecond time scale. This is an indication that partial unfolding and solvation of the chromophore and Glu46 take place only after the proton transfer.

The question remains what happens first: is the hydrogen-bond rearrangement caused by the proton transfer, or do the hydrogen bonds have to rearrange (partly) before the proton hops over to pCA? And if the latter is true, in what order and how do the bonds rearrange? To answer these questions, further extensive dynamical quantum-chemical simulations are necessary, using techniques to stimulate the reactive process.

5. Appendix

In section 3.1 we have summarized the results of the validation of our methods. Here, we provide the results in detail. For the various systems and computational methods described in section 2, we investigated the structure of the chromophore and the hydrogen-bond network in the chromophore binding pocket. As a validation of our computational approach, we will first compare our calculated chromophore bond lengths to those measured in crystal structures and other computational studies.

5.1. Chromophore Bond Lengths. Our DFT-BLYP geometry optimizations of the isolated deprotonated trans and cis chromophore and the neutral cis chromophore yield bond lengths that are within 0.02 Å of the values obtained in DFT-BP calculations[10] and DFT-B3LYP and CASSCF calculations.[41] Only the distance from the phenolic oxygen to the nearest carbon atom shows some variation (see below). This indicates that our computational approach of the electronic structures is sufficiently accurate. Figures 3 and 4 compare calculated pCA bond lengths to experimental data for pG and pR crystal structures, respectively (see Figure 1 for an explanation of the atom numbers). The X-ray structures that define the gray area in the figures have temperatures ranging from 85 to 287 K; both high- and low-temperature structures contribute to the upper boundary. Our calculated bond-length values are all in or very close to the area that is defined by the experiments. The bonds in the ring (between C_2 to C_7) have a mirror axis from C_2 to C_5 in each simulated structure. This tendency is also found in the experiments. Only the C−C bonds near the phenolic oxygen (C_2 to C_3 and C_2 to C_7) show some variation, which can be attributed to the behavior of the C−O bond next to it. In the isolated geometry optimization, it is shorter and has a more double-bond-like character. The chromophore is more *p*-benzoquinone-like in this case, elongating especially the two bonds immediately neighboring the C−O bond. When the negative charge can be stabilized by hydrogen bonds, in the protein-embedded (system B/E) and binding-pocket (system C/F) simulations, the C−O bond is longer, more single bond-like, and the C−C bonds are shorter. This compares well with the experimental X-ray measurements. In the tail of the chromophore (C_5 to C_9), the variation in bond lengths clearly indicates alternating single and double bonds. Somewhat larger differences in length exist in the dynamical simulations of the cis chromophore than in the trans configuration. This suggests that the chromophore is less conjugated in the cis configuration; in the pR state, the chromophore is not as flat as in the pG state. We will address this at the end of this section in more detail.

5.2. Hydrogen-Bond Network. Next we will discuss the hydrogen-bond network around the chromophore. Here we will compare the calculated structures of the ground-state chromophore binding pocket in vacuum (system C) and embedded in the protein with those from literature experiments. We studied the protein-embedded system in the pG state employing both FFMD and QMMM simulations, referred to as systems A and B. The crystal structure used as a starting point shows Tyr42 and Glu46 donating a hydrogen bond to the phenolic oxygen of pCA and Thr50 donating a hydrogen bond to Tyr42. Table 2 shows the calculated time-averaged hydrogen-bond lengths (oxygen−oxygen distances) together with values obtained from X-ray and NMR experiments. The FFMD simulation at 300 K shows a dynamic process of hydrogen-bond breaking and (re-

)forming. In this simulation, the Thr50 switches its donated hydrogen bond between Tyr42 (primarily), Glu46, pCA and an acceptor outside the binding pocket. When Thr50 hydrogen bonds to pCA, the phenolic oxygen accepts three hydrogen bonds. Also the hydrogen bond between Glu46 and pCA occasionally breaks, with Glu46 moving out of the binding pocket. In a 1-ns simulation, the Glu46-pCA bond breaks about 10 times for a short period; only in four cases does Glu46 form a relatively stable new hydrogen bond temporarily. Glu46 is in its "native" hydrogen-bond state 85% of the time, Thr50 only less than 50%. In the 100 K FFMD simulation, the hydrogen-bond network remains in its initial state, as can be expected from a lower temperature system. In the QMMM simulations, Thr50 is predominantly hydrogen bonded to Tyr42, occasionally switching to a hydrogen-bond acceptor outside the binding pocket, similar to the FFMD simulation. Note that other hydrogen-bond dynamics observed in the FFMD simulation are beyond the time scale accessed in the QMMM simulation.

Our observations are qualitatively consistent with the NMR measurements,[11] which show the pCA−Glu46 and Tyr42−Thr50 hydrogen bonds existing in most frames, and the pCA−Tyr42 hydrogen bond being present in all cases. Note that the numbers in Table 2 for the NMR experiments are averages over 26 lowest-energy conformations, not time averages. The X-ray measurements suggest that in the protein crystal the hydrogen-bond network around the phenolic oxygen is rather static with the pCA−Glu46, pCA−Tyr42, and Tyr42−Thr50 hydrogen bonds all stable. This could be attributed to the suppressed motion of the protein backbone in the crystalline state compared to the solvated situation. A quantitative comparison shows that the calculated hydrogen-bond lengths to pCA are somewhat longer than the experimental values. Here we should note that the hydrogen bonds to pCA in the crystal structure are extremely short.[6] The NMR solution structure even shows a Tyr42-to-pCA oxygen−oxygen distance of 2.21 Å in one of its energy-minimized structures. In this case, the oxygen−hydrogen−oxygen angle is 82° and the distance from the hydrogen to the pCA-oxygen is 2.13 Å. In the employed force field, these distances yield high energies and are therefore very unlikely to be accessed in an FFMD simulation. The quantum-chemical simulations show larger hydrogen-bond lengths as well, although the difference with experiment is smaller. Still, they do give a good estimate of the length ratio of the hydrogen bonds: all simulations but one show correctly that the hydrogen bond from Tyr42 to pCA is the shorter one. Only the simulation with the QM/MM border in between the hydrogen-bond donor and the acceptor exhibits equal hydrogen-bond lengths.

5.3. Relaxation of the Chromophore after Isomerization. Ihee and Rajagopal[12] argue that after isomerization, the chromophore first adopts an intermediate planar cis configuration, I_{cp}, before relaxing into the nonplanar "cis wobble" pR structure, pR_{cw}. To address this, we determined the dihedral angles from our QMMM simulations of the protein-embedded chromophore binding pocket (systems E) and from our CPMD simulation of the binding pocket in vacuum (system F). We compared these angles to our PDB starting structure, 1OT9, that has a flat chromophore. Figure 5 shows a typical chromophore configuration from each of the simulations, where the dihedral angles are equal to the time-averaged values. Both quantum simulations (systems E and F) suggest that the curvature is determined by the position of Arg52. There is a stable hydrogen bond between one of the NH_2 groups of the Arg52 residue and the carbonylic oxygen of the chromophore. To maintain this, the oxygen has to stick out of plane.

In most crystal structures of the pR state (1S4R, 1UWP, 1TS7, 1S1Y) there is a cis wobble chromophore, but the dihedral angles vary. The exact dihedral values do not seem to be important, as long as they enable the hydrogen bonds from Glu46 and Tyr42 to pCA to stay intact after isomerization. The Arg52 hydrogen bond is not present in the various X-ray configurations. Before the isomerization, the chromophore oxygen points to the other side (often hydrogen bonding to the backbone NH group of Cys69) and is therefore not close enough to Arg52 to hydrogen bond. The hydrogen bond from Arg52 in pR only forms after relaxation within some picoseconds of simulation time.

Acknowledgment. We would like to thank the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research) for funding. L.G. thanks the Velux Stiftung for financial support and the Centro Fermi of Rome for computational resources. This work was sponsored by the Stichting Nationale Computerfaciliteiten (National Computing Facilities Foundation, NCF) for the use of supercomputer facilities.

References and Notes

(1) Meyer, T. E.; Yakali, E.; Cusanovich, M. A.; Tollin, G. *Biochemistry* **1987**, *26*, 418.
(2) Cusanovich, M. A.; Meyer, T. E. *Biochemistry* **2003**, *42*, 4759.
(3) Vreede, J.; Crielaard, W.; Hellingwerf, K. J.; Bolhuis, P. G. *Biophys. J.* **2005**, *88*, 3525.
(4) Hendriks, J.; Hoff, W. D.; Crielaard, W.; Hellingwerf, K. J. *J. Biol. Chem.* **1999**, *274*, 17655.
(5) Getzoff, E. D.; Gutwin, K. N.; Genick, U. K. *Nat. Struct. Biol.* **2003**, *10*, 663.
(6) Anderson, S.; Crosson, S.; Moffat, K. *Acta Crystallogr., Sect. D*: *Biol. Crystallogr.* **2004**, *60*, 1008.
(7) Groenhof, G.; Lensink, M. F.; Berendsen, H. J. C.; Snijders, J. G.; Mark, A. E. *Proteins* **2002**, *48*, 202.
(8) Groenhof, G.; Lensink, M. F.; Berendsen, H. J. C.; Mark, A. E. *Proteins* **2002**, *48*, 212.
(9) Groenhof, G. M.; Bouxin-Cademartory, M; Hess, B.; de Visser, S. P.; Berendsen, H. J. C.; Olivucci, M.; Mark, A. E.; Robb, M. A. *J. Am. Chem. Soc.* **2004**, *126*, 4228.
(10) Sergi, A.; Grüning, M.; Ferrario, M.; Buda, F. *J. Phys. Chem. B* **2001**, *105*, 4386.
(11) Dux, P.; Rubinstenn, G.; Vuister, G. W.; Boelens, R.; Mulder, F. A.; Hard, K.; Hoff, W. D.; Kroon, A. R.; Crielaard, W.; Hellingwerf, K. J.; Kaptein, R. *Biochemistry* **1998**, *37*, 12689.
(12) Ihee, H.; Rajagopal, S.; Srajer, V.; Pahl, R.; Anderson, S.; Schmidt, M.; Schotte, F.; Anfinrud, P. A.; Wulff, M.; Moffat, K. *PNAS* **2005**, *102*, 7145.
(13) Rajagopal, S.; Moffat, K. *PNAS* **2003**, *100*, 1649.
(14) Kort, R.; Hellingwerf, K. J.; Ravelli, R. B. G. *J. Biol. Chem.* **2004**, *279*, 26417.
(15) Perman, B.; Srajer, V.; Ren, Z.; Teng, T.; Pradervand, C.; Ursby, T.; Bourgeois, D.; Schotte, F.; Wulff, M.; Kort, R.; Hellingwerf, K. J.; Moffat, K. *Science* **1998**, *279*, 1946.
(16) Anderson, S.; Srajer, V.; Pahl, R.; Rajagopal, S.; Schotte, F.; Anfinrud, P. A.; Wulff, M.; Moffat, K. *Structure* **2004**, *12*, 1039.
(17) Xie, A.; Hoff, W. D.; Kroon, A. R.; Hellingwerf, K. J. *Biochemistry* **1996**, *35*, 14671.
(18) Imamoto, Y.; Mihara, K.; Hisatomi, O.; Kataoka, M.; Tokunaga, F.; Bojkova, N.; Yoshihara, K. *J. Biol. Chem.* **1997**, *272*, 12905.
(19) Röhrig, U.; Guidoni, L.; Laio, A.; Frank, I.; Röthlisberger, U. *J. Am. Chem. Soc.* **2004**, *126*, 15328.
(20) Thompson, M. J.; Bashford, D.; Noodleman, L.; Getzoff, E. D. *J. Am. Chem. Soc.* **2003**, *125*, 8186.
(21) Yoda, M.; Inoue, Y.; Sakurai, M. *J. Phys. Chem. B* **2003**, *107*, 14569.
(22) Antes, I.; Thiel, W.; Van Gunsteren, W. F. *Eur. Biophys. J.* **2002**, *31*, 504.
(23) Genick, U. K.; Borgstahl, G. E. O.; Ng, K.; Ren, Z.; Pradervand, C.; Burke, P. M.; Srajer, V.; Teng, T.; Schildkamp, W.; McRee, D. E.; Moffat, K.; Getzoff, E. D. *Science* **1997**, *275*, 1471.
(24) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.
(25) Laio, A.; VandeVondele, J.; Röthlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941.

(26) Laio, A.; VandeVondele, J.; Röthlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 7300.
(27) Lindahl, E.; Hess, B.; Van der Spoel, D. *J. Mol. Model* **2001**, *7*, 306.
(28) Van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular simulation*: *the GROMOS96 manual and user guide*; vdf Hochschulverlag AG an der ETH: Zürich, Switzerland; and Biomos B.V.: Zürich/Groningen, Switzerland, 1996.
(29) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; Van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596.
(30) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
(31) Hess, B.; Bekker, H.; Berendsen, H.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463.
(32) Miyamoto, S.; Kollman, P. *J. Comput. Chem.* **1992**, *13*, 952.
(33) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327.
(34) *CPMD* V3.9; Copyright IBM Corp 1990−2004; MPI für Festkörperforschung Stuttgart, 1997−2001.
(35) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(36) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(37) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993.
(38) http://lcbcpc21.epfl.ch.
(39) Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 2810.
(40) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
(41) Li, Q.-S.; Fang, W.-H. *Chem. Phys.* **2005**, *313*, 71.
(42) Frenkel, D.; Smit, B. *Understanding molecular simulation*, 2nd ed.; Academic Press: San Diego, CA, 2002.