# Computational Spectroscopy and Reaction Dynamics

Pierre-Andre Cazade, Stephan Lutz, Myung Won Lee, and Markus Meuwly*

*Abstract:* Physico- and bio-chemical processes on the femto- to picosecond time scale are ideally suited to be investigated with all-atom simulations. They include, amongst others, vibrational relaxation, ligand migration in sterically demanding environments (proteins, ices), or vibrational spectra. By comparing with experimental data, the results can be used to obtain an understanding of the *mechanisms* underlying the observations. Furthermore, most of these processes are sensitive to the intermolecular interactions. Therefore, detailed refinement of such interaction potentials is possible.

**Keywords:** Computational spectroscopy · Molecular dynamics simulations · Reaction dynamics

The classification of fast and ultrafast time scales in atoms and molecules depends to some extent on the process and property considered. For example, the electronic degrees of freedom change on considerably more rapid time scales than the nuclear motions. This is the basis of the Born-Oppenheimer approximation which is valid in many cases. In proteins, on the other hand, vibrations are much more rapid than changes in the secondary structure. Typical time scales for electronic motion in molecules are attoseconds, whereas vibrations are associated with femto- and pico-second time scales. In the following, ultrafast time scales in molecular systems – which are at the focus of the NCCR Molecular and Ultrafast Science and Technology (MUST) – will be associated with atto- to picosecond processes.

Understanding ultrafast processes in molecular systems requires multi-faceted approaches, including experiment, theory and computation. A typical example is the vibrational relaxation of a solvated diatomic molecule, *e.g.* cyanide ($CN^-$) in water.[1,2]

Although time scales of vibrational relaxation can be reliably measured using modern laser techniques, understanding the relaxation *mechanism* in molecular detail is much more difficult as it includes coupling and energy transfer to solvent modes. Another example is allostery which is the regulation of a protein by binding a small molecule at the allosteric site. An allosteric activator will enhance the activity of the protein. The most widely known but still incompletely understood allosteric protein is hemoglobin. Here, four binding sites for the ligand ($O_2$) exist. Binding of the first $O_2$ molecule increases the protein's affinity for binding the second $O_2$ molecule, *etc*. However, the molecular details underlying the enhanced affinity for subsequent $O_2$ ligands is still unexplained. Finally, one purpose of high-resolution spectroscopy is to understand the bonding pattern and geometrical details of a molecule by recording and analyzing spectroscopic signatures. As the size of the molecule grows, directly determining the structure of the molecule from the spectroscopic data becomes increasingly difficult. Even more difficult is the situation in proteins where spectroscopic features of small probe molecules such as CO or NO can be recorded but the splitting patterns of the vibrational spectra are difficult to interpret in structural terms. The Stark splitting is caused by the probe molecules interacting with the surrounding electric field and contains valuable information about structural features of the protein.

Under such circumstances atomistic simulations are valuable to complement and explain experimental results. They provide the link between the atomic motions and the experimental measurements. On the other hand, molecular dynamics simulations strongly depend on the force fields used and therefore much effort has

been spent over the past several years in refining and optimizing force fields for particular purposes. An alternative means to circumvent this problem is to use quantum mechanical methods to compute the intermolecular interactions. However, due to the computational effort involved in such calculations, it is often used for small systems or in conjunction with mixed quantum mechanics/molecular mechanics (QM/MM) calculations. It should be noted that ultrafast processes as defined above are ideal for molecular dynamics simulations as the processes of interest occur on the femto- to pico-second time scale.

## Vibrational Relaxation

Substantial progress has been made in observing and modeling bond relaxation after photo-excitation in small model systems and larger biomolecules.[3–9] Various approaches exist, perhaps the most intuitive of which uses non-equilibrium MD simulations to explicitly excite the bond(s) of interest and measure relaxation times.[3] By averaging over a number of non-equilibrium trajectories, the cooling time $T_1$ is obtained according to Eqn. (1)[10,11]

$$\frac{\overline{E_v(t)} - \overline{E_v(\infty)}}{\overline{E_v(0)} - \overline{E_v(\infty)}} = \exp(-t/T_1) \tag{1}$$

Here $E_v(t)$ is the vibrational energy at time '$t$', $T_1$ is the vibrational relaxation time and the overbar represents an average over non-equilibrium trajectories. $\overline{E_v(\infty)}$ is the thermal value[6] $k_B T$ (0.60 kcal/mol at 300 K), while $E_v(0)$ is the initial vibrational energy and $k_B$ the Bolzmann constant of the excited bond at $t = 0$.

Because relaxation involves coupling between different degrees of freedom,

*Correspondence:* Prof. Dr. M. Meuwly
Department of Chemistry
University of Basel
Klingelbergstrasse 80
CH-4056 Basel
Tel: +41 61 267 3821
E-mail: m.meuwly@unibas.ch

calculated relaxation times depend on the details with which interactions in the force field are described. Most commonly used force fields are based on harmonic functions for bonded interactions (bonds $r$ and valence angles $\theta$) and periodic functions for torsions $\phi$ to describe the total energy as a function of geometry.[12]

A number of studies analyzing the process of heme and CO ligand cooling,[4,5,8,13,14] as well as assigning the key vibrational modes of Fe in the MbCO vibrational spectrum[15] have been carried out. For an overview of the system see Fig. 1. Infrared pump-probe measurements of solvated MbCO, HbCO and a protoheme system find ultrafast (picosecond timescale) cooling rates for the vibrationally excited CO ligand.[13,14] The CO relaxation time of MbCO was found to be $T_1 = 17$ ps, significantly faster than is typical for metal carbonyls[16] such as $W(CO)_6$ in $CCl_4$ ($T_1 = 800 \pm 200$ ps). The relaxation time decreases with molecular complexity, as a result of the increased number of internal vibrational modes available for coupling. Relaxation times in heme proteins are therefore faster than in protoheme complexes (31 ps). Interaction with the solvent also influences the vibrational cooling rate, although to a lesser extent.[13] For MbCO a two-stage cooling process was predicted, with an initial fast cooling rate (decay constant 1–4 ps) that was found to be in good agreement with experimental values (time constant 2 ps).[8] The subsequent, slower cooling phase consists of two sub-components, a faster component with time constant 7.5 ps, followed by a slower component with time constant 20 ps. It is believed that the slow component arises from energy transfer from the heme to the surrounding solvent, while the very fast initial heme cooling rates arise from coupling within the molecule.[8]
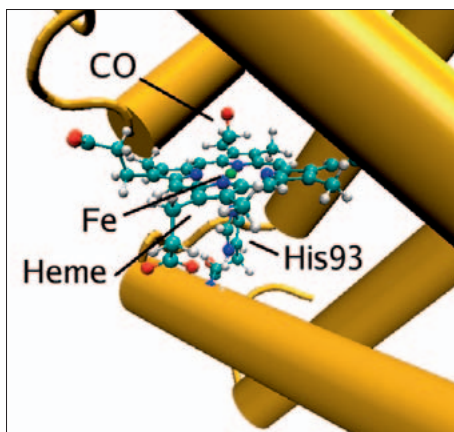


Fig. 1. Heme embedded in the Mb-protein structure. Heme, CO and His93 moieties, shown in ball-and-stick, are used to study energy relaxation.
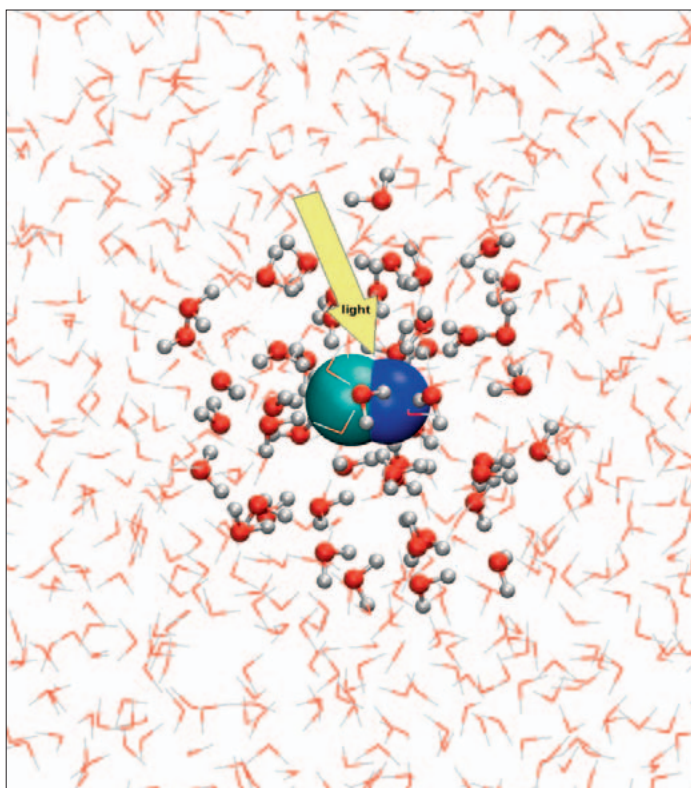


Fig. 2. The simulation system used for studying vibrational relaxation of CN⁻ in water. The cyanide ion is excited vibrationally by infrared light at $t = 0$. The C and N atoms of the cyanide ion are displayed in the van der Waals spheres. The atoms of the water molecules within 7 °A from CN⁻ are displayed as smaller spheres with the O–H bonds shown in cylinders. Other water molecules are depicted as lines.

MD simulations with accurate force fields for CO interacting with heme in MbCO demonstrated that anharmonic Morse potentials for both the C–O and the adjacent Fe–$C_{CO}$ bonds to allow anharmonic coupling are required for rapid (tens of ps) relaxation of the vibrationally excited CO ligand.[17,18] For the recommended force field parametrization which maintains all relevant vibrational frequencies within 45 cm⁻¹ of the experimental observations, $T_1 = 50$ ps (25 to 17 ps including quantum corrections[19]) is found compared to $T_1 = 432$ ps for the standard (harmonic) force field.[17] The simulations also establish that the cooling (relaxation) rate sensitively depends on the difference in the vibrational frequency between the coupled modes. Analysis of the trajectories reveals that the relaxation *mechanism* involves primarily coupling to local vibrational modes, rather than through non-bonded interactions. The Fe–$C_{CO}$ stretch and FeCO bending modes were identified as particularly important to vibrational cooling of CO on a realistic timescale. Introduction of Morse potentials to both the CO and Fe–$C_{CO}$ bonds improves coupling between these modes, and provides faster CO cooling times with a smaller impact on vibrational frequencies. Other local modes such as the heme doming motion were not found to significantly affect the cooling rate, although they may be important to dissipate energy away from the Fe–$C_{CO}$ bond once it has been transferred. Energy transfer from CO to Fe–$C_{CO}$ and the FeCO bending mode was identified as the rate determining step.[17]

More recently, we have also studied vibrational relaxation of simple ions in solution, in particular cyanide CN⁻ in water (see Fig. 2).[1,2,20] The intermolecular interactions are treated with distributed multipoles and relaxation was found to depend sensitively on the van der Waals parameters. The fundamental insight that can be gained from such simulations concerns the *mechanism* for vibrational relaxation, *i.e.* which modes of the surrounding solvent couple most strongly to the relaxing ion. For CN⁻ in water recent simulations suggest that energy flows primarily into the water-bending and librational modes.[20] It can even be shown that van der Waals parameters can be fitted quite reliably to experimental relaxation times in $H_2O$ and $D_2O$ and reproduce the observations very well. Compared to $28 \pm 7$ ps and $71 \pm 3$ ps for relaxation in $H_2O$ and in $D_2O$, respectively, computed relaxation times are $22 \pm 2$ ps and $68 \pm 11$ ps from simulations with van der Waals ranges increased by $\approx 10$ %.[2,20]

## Structural Origin of Spectroscopic Substates

Vibrational frequency shifts caused by external electric fields (Stark shifts) offer an increasingly important and attractive means to study the structure, electrostatics and dynamics of protein active sites.[21,22] The response of bond vibrations to changes in local electric fields can be accurately measured and used to probe heteroge-

neous chemical environments. However, while detailed spectroscopic data can be obtained it is usually difficult if not impossible to provide a structural interpretation of the spectroscopic features.

The bond between a ligand and the heme-iron atom in proteins including myoglobin (Mb) or neuroglobin (Ngb) is photolabile and can be broken by electronically exciting the Soret band.[23] This bond is of central importance, because it can be used to better understand reactivity, binding and rebinding dynamics in a well-defined biological context. In less than 100 fs after excitation, the Fe–CO bond is broken and the heme unit relaxes to a 5-coordinated structure that prevents immediate rebinding of CO.[23] Infrared (IR) spectroscopy has been useful to characterize protein and ligand dynamics.[24–30] By sensing the iron-bound CO vibrations in carbonmonoxy myoglobin (MbCO), several CO-stretching bands were observed. Vibrational spectra of reactants and products of CO dissociation contain signatures related to geometrical structures which can be directly associated with particular chemical bonding patterns and are therefore more readily interpreted in terms of chemical structure. Thus, an important aim of IR studies of proteins is the structural interpretation of such spectra. The association of spectroscopic features to chemical structure is particularly important if an understanding of dynamical aspects of such complicated systems is sought.

Force fields employed in many MD simulations allow extensive sampling, but are not sufficiently accurate to provide realistic response to local changes in electric field which is required for understanding vibrational spectroscopy.[31] Use of more detailed, multipolar charge models offers an attractive solution to more accurately describe the interaction of the probe molecule with the inhomogeneous electric fields in protein binding sites, while still allowing significant sampling of available phase space. In particular, it is possible to compare experimental and computed spectra and subsequently analyse the nuclear dynamics which provides a structural interpretation of the spectroscopic results. This has been successfully applied to the spectroscopy of CO in myoglobin, neuroglobin, and in ordered and disordered ices.[32–35]

An example for a structural interpretation of an IR band is presented in Fig. 3. Panel A shows a computed CO spectrum of photodissociated carbonmonoxyneuroglobin in the region of the CO stretching vibration. By analysing the underlying trajectory it is possible to relate typical conformations to the spectroscopic features. Characteristic conformations de-
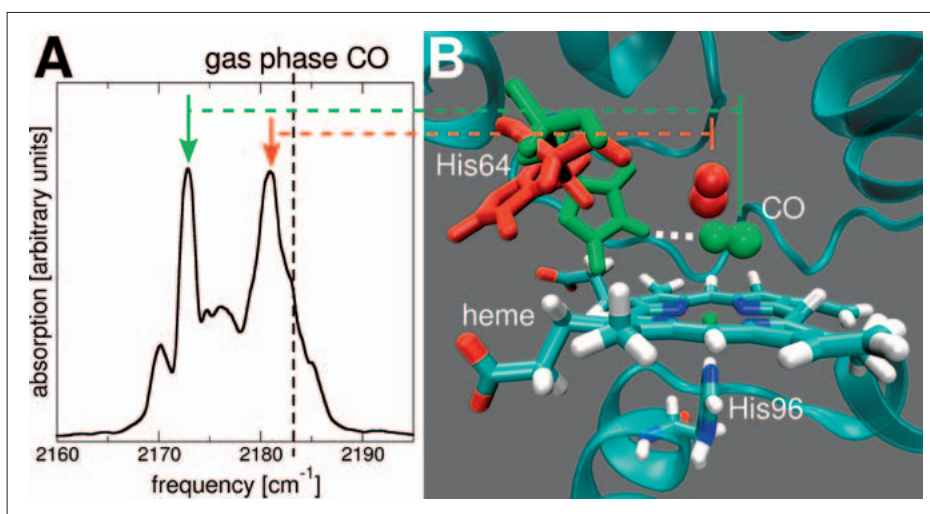


Fig. 3. Example for a structural identification of spectroscopic substrates. Panel A: Computed IR spectrum of the CO stretching vibration in photodissociated Ngb–CO showing two pronounced absorption peaks (indicated by the green and red arrows). Panel B: Structure of the heme active site in Ngb. The two superimposed structures in green and red correspond to conformations responsible for the two IR band marked in panel A.

scribing these peaks are shown in Fig. 3B. As one can see, the green peak in panel A, shifted by $\approx$−10 cm⁻¹ compared to the IR peak of CO in gas phase, derives from a strong hydrogen-bond interaction of the CO-oxygen atom with the protonated $N_\varepsilon$ atom of the His64 side chain. The peak marked in red, which is only shifted by −2 cm⁻¹, was found to be related to a structure having the His64 side chain rotated away from the active site and showing no significant interaction with CO.

## Reaction Networks in Proteins

Characterizing the long-time scale dynamics in complex systems exhibiting several states is a fundamental problem in chemistry and biology. In biological systems – *i.e.* proteins – overall rates for reactions can be measured with a wide range of methods whereas transition times between spatially connected, but isolated regions are much more difficult to investigate directly. However, such rates have great significance in understanding the reaction mechanisms and to distinguish rate-limiting from other, less important, reactive processes.

Using MD simulations it is now possible to extensively sample the ligand network of a protein and to construct transition networks which allow to investigating the long-time behaviour of the system. Such ligand networks – an example is shown in Fig. 4 – are coarse-grained representations of multi-dimensional free energy surfaces which are much more difficult to compute and to analyse in detail due to their inherent roughness.[36–38]

For the truncated hemoglobin (trHbN) of *Mycobacterium tuberculosis* such a ligand migration network for NO-diffusion has been built and analyzed from a transition network analysis.[39,40] This provides a detailed picture of probability flux between all accessible metastable states and of the 'logic' behind a ligand-migration network in space and time. Most of the transitions occur on the sub-nanosecond time scale which should make them accessible to time-resolved experiments together with spectroscopic techniques.[41–44] Also, it was found that harvesting trajectories from many different initial states leads to a kinetic model in close agreement with explicit long-time simulations. This makes TNA a relevant and robust technique for related problems such as small-molecule diffusion in disordered ices or in clathrates.[39]

[1] E. J. Heilweil, F. E. Doany, R. Moore, R. M. Hochstrasser, *J. Chem. Phys.* **1982**, *76*, 5632.
[2] P. Hamm, M. Lim, R. M. Hochstrasser, *J. Chem. Phys.* **1997**, *107*, 10523.
[3] R. M. Whitnell, K. R.Wilson, J. T. Hynes, *J. Phys. Chem.* **1990**, *94*, 8625.
[4] D. E. Sagnella, J. E. Straub, *Biophys. J.* **1999**, *77*, 70.
[5] D. E. Sagnella, J. E. Straub, *J. Phys. Chem. B* **2001**, *105*, 7057.
[6] L. Mu, J. E. Straub, *J. Phys. Chem. B* **2003**, *107*, 10634.
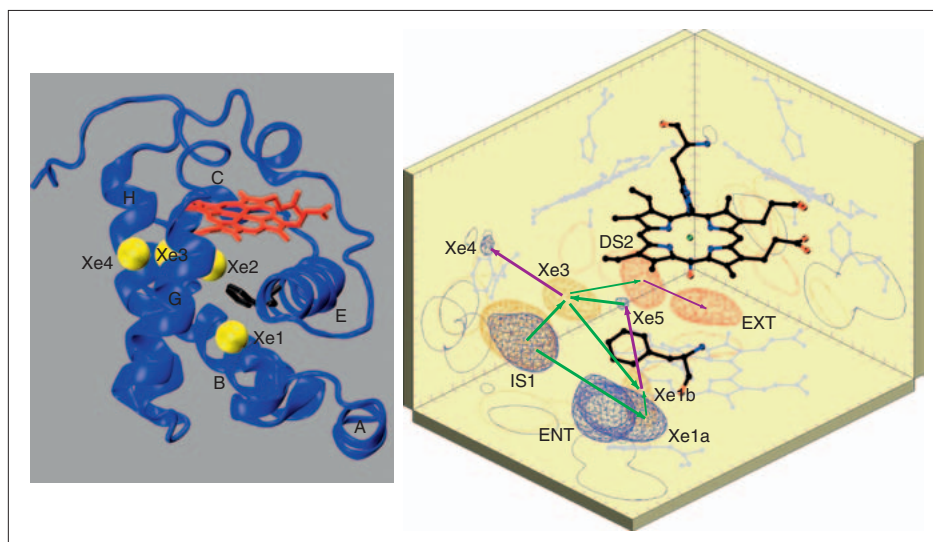[7] E. Henry, W. Eaton, R. M. Hochstrasser, *Proc. Natl. Acad. Sci.* **1986**, *83*, 8982.

Fig. 4. Schematic representation of the protein network (right panel). The density corresponds to probability distributions of the unbound O$_2$. Only densities higher than 10% of the maximum density are shown. Three trajectories corresponding to the three pocket colors are used to build the figure. The arrows show the main ligand migrations and were obtained by averaging over 15 trajectories. The heme, Phe62, His81 and NO residues are represented by spheres and sticks. On the lefthand side, the trHbN structure from experimental X-ray data is shown. The heme group is shown by red sticks and the Phe62 residue by black ones. The spheres are the four xenon docking sites identified experimentally.

[8]   Y. Mizutani, T. Kitagawa, *Chem. Rec.* **2001**, *1*, 258.
[9]   J. Danielsson, M. Meuwly, *J. Phys. Chem. B* **2007**, *111*, 218.
[10]  E. W. Montroll, K. E. Shuler, *J. Chem. Phys.* **1957**, *26*, 454.
[11]  H. A. Bethe, E. Teller, Ballistic Research Laboratory, **1941**, Report X-117.
[12]  A. D. MacKerell Jr., J. Wiorkiewicz-Kurczera, M. Karplus, *J. Am. Chem. Soc.* **1995**, *117*, 11946.
[13]  J. C. Owrutsky, M. Li, B. Locke, R. M. Hochstrasser, *J. Phys. Chem.* **1995**, *99*, 4842.
[14]  J. R. Hill, A. Tokmakoff, K. A. Peterson, B. Sauter, D. Zimdars, D. D. Dlott, M. D. Fayer, *J. Phys. Chem.* **1994**, *98*, 11213.
[15]  B. M. Leu, N. J. Silvernail, M. Z. Zgierski, G. R. A. Wyllie, M. K. Ellison, W. R. Scheidt, J. Zhao, W. Sturhahn, E. E. Alp, J. T. Sage, *Biophys. J.* **2007**, *92*, 3764.
[16]  E. J. Heilweil, R. R. Cavanagh, J. C. Stephenson, *Chem. Phys. Lett.* **1987**, *134*, 181.
[17]  M. Devereux, M. Meuwly, *J. Chem. Inf. Model.* **2010**, *50*, 349.
[18]  M. Devereux, M. Meuwly, *J. Phys. Chem. B* **2009**, *113*, 13061.
[19]  G. Stock, *Phys. Rev. Lett.* **2009**, *102*, 118301.
[20]  M. W. Lee, M. Meuwly, *J. Phys. Chem. A* **2011**, in press.
[21]  S. G. Boxer, *J. Phys. Chem. B* **2009**, *113*, 2972.
[22]  L. J. Webb, S. G. Boxer, *Biochemistry* **2008**, *47*, 1588.
[23]  J. W. Petrich, C. Poyart, J. L. Martin, *Biochem.* **1988**, *27*, 4049.
[24]  A. Ansari, J. Berendzen, D. Braunstein, B. R. Cowen, H. Frauenfelder, M. K. Hong, I. E. T. Iben, J. B. Johnson, P. Ormos, T. B. Sauke, R. Scholl, A. Schulte, P. J. Steinbach, J. Vittitow, R. D. Young, *Biophys. Chem.* **1987**, *26*, 337.
[25]  P. A. Anfinrud, C. Han, R. M. Hochstrasser, *Proc. Natl. Acad. Sci.* **1989**, *86*, 8387.
[26]  J. B. Johnson, D. C. Lamb, H. Frauenfelder, J. D. Müller, B. McMahon, G. U. Nienhaus, R. D. Young, *Biophys. J.* **1996**, *71*, 1563.
[27]  G. N. Phillips, M. L. Teodoro, T. Li, B. Smith, J. S. Olson, *J. Phys. Chem. B* **1999**, *103*, 8817.

[28]  D. Morikis, P. M. Champion, B. A. Springer, S. G. Sligar, *Biochem.* **1989**, *28*, 4791.
[29]  T. S. Li, M. L. Quillin, G. N. Phillips, J. S. Olson, *Biochem.* **1994**, *33*, 1433.
[30]  D. R. Nutt, M. Meuwly, *Proc. Natl. Acad. Sci.* **2004**, *101*, 5998.
[31]  N. Plattner, M. Meuwly, *Biophys. J.* **2008**, *94*, 2505.
[32]  N. Plattner, M. Meuwly, *Chem. Phys. Chem.* **2008**, *9*, 1271.
[33]  K. Nienhaus, S. Lutz, M. Meuwly, G. U. Nienhaus, *Chem. Phys. Chem.* **2010**, *11*, 119.
[34]  S. Lutz, K. Nienhaus, G. U. Nienhaus, M. Meuwly, *J. Phys. Chem. B* **2009**, *113*, 15334.
[35]  S. Lutz, M. Meuwly, *Faraday Discuss. Chem. Soc.* **2011**, in print.
[36]  P. Banushkina, M. Meuwly, *J. Phys. Chem. B* **2005**, *109*, 16911.
[37]  P. Banushkina, M. Meuwly, *J. Chem. Theo. Comp.* **2005**, *1*, 208.
[38]  P. Banushkina, M. Meuwly, *J. Chem. Phys.* **2007**, *127*, 135101.
[39]  S. Mishra, M. Meuwly, *Biophys. J.* **2010**, *99*, 3969.
[40]  S. Mishra, M. Meuwly, *Biophys. J.* **2009**, *96*, 2105.
[41]  V. Srajer, T. Y. Teng, T. Ursby, C. Pradervand, Z. Ren, S. Adachi, W. Schildkamp, D. Bourgeois, M. Wulff, K. Moffat, *Science* **1996**, *274*, 1726.
[42]  V. Srajer, Z. Ren, T. Y. Teng, M. Schmidt, T. Ursby, D. Bourgeois, C. Pradervand, W. Schildkamp, M. Wulff, K. Moffat, *Biochem.* **2001**, *40*, 13802.
[43]  F. Schotte, M. Lim, T. A. Jackson, A. V. Smirnov, J. Soman, J. S. Olson, G. N. Phillips Jr., M. Wulff, P. A. Anfinrud, *Science* **2003**, *300*, 1944.
[44]  M. Schmidt, K. Nienhaus, R. Pahl, A. Krasselt, S. Anderson, F. Parak, G. U. Nienhaus, V. Srajer, *Proc. Natl. Acad. Sci.* **2005**, *102*, 11704.