

2 The MoD-QM/MM methodology for structural refinement 3 of photosystem II and other biological macromolecules

4 Eduardo M. Sproviero · Michael B. Newcomer ·
5 José A. Gascón · Enrique R. Batista ·
6 Gary W. Brudvig · Victor S. Batista

7 Received: 18 December 2008 / Accepted: 25 June 2009
8 © Springer Science+Business Media B.V. 2009

9 **Abstract** Quantum mechanics/molecular mechanics
10 (QM/MM) hybrid methods are currently the most powerful
11 computational tools for studies of structure/function rela-
12 tions and structural refinement of macromolecules (e.g.,
13 proteins and nucleic acids). These methods are highly
14 efficient, since they implement quantum chemistry tech-
15 niques for modeling only the small part of the system (QM
16 layer) that undergoes chemical modifications, charge
17 transfer, etc., under the influence of the surrounding envi-
18 ronment. The rest of the system (MM layer) is described in
19 terms of molecular mechanics force fields, assuming that
20 its influence on the QM layer can be roughly decomposed
21 in terms of electrostatic interactions and steric hindrance.
22 Common limitations of QM/MM methods include inaccura-
23 cies in the MM force fields, when polarization effects are
24 not explicitly considered, and the approximate treatment of
25 electrostatic interactions at the boundaries between QM
26 and MM layers. This article reviews recent advances in the
27 development of computational protocols that allow for
28 rigorous modeling of electrostatic interactions in extended

systems beyond the common limitations of QM/MM 29
hybrid methods. We focus on the moving-domain QM/MM 30
(MoD-QM/MM) methodology that partitions the system 31
into many molecular domains and obtains the electrostatic 32
and structural properties of the whole system from an 33
iterative self-consistent treatment of the constituent 34
molecular fragments. We illustrate the MoD-QM/MM 35
method as applied to the description of photosystem II as 36
well as in conjunction with the application of spectro- 37
scopically constrained QM/MM optimization methods, 38
based on high-resolution spectroscopic data (extended X- 39
ray absorption fine structure spectra, and exchange cou- 40
pling constants). 41

42
43 **Keywords** Quantum mechanics · Molecular mechanics ·
44 Photosystem II · EXAFS · MoD-QM/MM

45 Abbreviations

BS	Broken symmetry	46
DFT	Density functional theory	47
DNA	Deoxyribonucleic acid	48
EE	Electronic-embedding	49
ESP	Electrostatic potential	50
EXAFS	Extended X-ray absorption fine structure	51
FTIR	Fourier Transform Infrared Spectroscopy	52
MEP	Minimum energy path	53
MoD-QM/MM	Moving-domain QM/MM	54
OEC	Oxygen-evolving complex	55
P-EXAFS	Polarized extended X-ray absorption fine structure	56
PSII	Photosystem II	57
QM/MM	Quantum mechanics–molecular mechanics	58
SC-QM/MM	Spectroscopically constrained QM/MM	59

A1 E. M. Sproviero · M. B. Newcomer · J. A. Gascón ·
A2 G. W. Brudvig · V. S. Batista (✉)
A3 Department of Chemistry, Yale University, P. O. Box 208107,
A4 New Haven, CT, 06520-8107, USA
A5 e-mail: victor.batista@yale.edu

A6 *Present Address:*
A7 J. A. Gascón
A8 Department of Chemistry, University of Connecticut, Unit 3060,
A9 Storrs, CT 06269, USA

A10 E. R. Batista
A11 Theoretical Division, Los Alamos National Laboratory, Los
A12 Alamos, NM 87545, USA

Introduction

In recent years, the development of computer hardware has allowed computational chemistry methods to establish themselves as powerful tools for studies of structure/function relations and structural refinement of macromolecules (e.g., proteins and nucleic acids). In particular, quantum mechanics/molecular mechanics (QM/MM) hybrid methods have become the most popular approaches for modeling prosthetic groups, which undergo significant structural or electronic modifications when embedded in macromolecular structures (Maseras and Morokuma 1995; Vreven and Morokuma 2000a, b; Warshel and Levitt 1976). The main aspect, common to all QM/MM methods, is the partitioning of the system into two fragments: a small part (QM layer) that includes the prosthetic group as described by quantum chemistry methods and the rest of the system (MM layer), which is modeled in terms of molecular mechanics force fields, since its influence can be roughly decomposed in terms of steric hindrance and electrostatic interactions. This basic idea is at the heart of the Moving-Domain QM/MM (MoD-QM/MM) methodology and the applications studies reviewed in this article (Gascon et al. 2006a; Menikarachchi and Gascon 2008; Newcomer et al. 2009).

The various different versions of QM/MM methods differ in the way the energy is evaluated (e.g., additive and subtractive schemes are available); in the way, the boundary between QM and MM layers is treated (e.g., link-atoms, localized bond orbitals, and the pseudopotential approach); on the specific level of quantum chemistry and MM force field; and on how the electrostatic interactions between QM and MM layers are described, including mechanical, electronic, and polarized embedding schemes.

The quantum mechanics and molecular mechanics (QM/MM) methodologies require a starting structure, which is often obtained from X-ray diffraction models. However, the available X-ray diffraction structures might have only moderate resolution, or suffer from radiation damage due to the high doses of X-ray necessary for data collection. In addition, the potential-energy landscape of complex molecular systems is often so rugged that the obtained minimum energy structures are typically strongly correlated to the initial-guess configuration. Another major problem with traditional QM/MM modeling is the limited accuracy of the electrostatic potentials provided by molecular mechanics force fields. These limitations often introduce inaccuracies in the description of the interactions between the QM and MM layers in QM/MM hybrid models.

A structural-refinement scheme has been developed to circumvent the risk of obtaining minimum energy structures that are inconsistent with high-resolution spectroscopic data (Sproviero et al. 2008a). The method adjusts

the overall configuration to minimize the total QM/MM energy subject to the constraint of a minimum χ^2 deviation between the calculated and experimental properties of the system as described by an independent set of experimental data. This methodology has been successfully applied in conjunction with high-resolution polarized extended X-ray absorption fine structure (P-EXAFS) spectra to refine structural models of the Mn_4Ca active site of PSII (Sproviero et al. 2008a).

The MoD-QM/MM methodology has been developed to provide an accurate description of the electrostatic interactions between MM and QM layers in QM/MM hybrid models. The MoD-QM/MM approach partitions the system into many molecular domains and obtains the electrostatic and structural properties of the whole system from an iterative self-consistent treatment of the constituent molecular fragments (Gascon et al. 2006a; Menikarachchi and Gascon 2008; Newcomer et al. 2009). This involves a polarization scheme rendering a description of prosthetic groups at the polarized-embedding QM/MM level. Self-consistent optimized structures and electrostatic potential (ESP) atomic charges are obtained, modeling each fragment as a QM layer polarized by the distribution of atomic charges in the surrounding molecular fragments. The resulting computational task scales linearly with the size of the system, bypassing the enormous demands of memory and computational resources that would otherwise be required by 'brute-force' full QM calculations of the complete system. Computations for several benchmark systems, which allow for full QM calculations, have shown that the MoD-QM/MM method produces ab initio quality electrostatic potentials and optimized structures in quantitative agreement with full QM results (Gascon et al. 2006a; Menikarachchi and Gascon 2008; Newcomer et al. 2009).

The MoD-QM/MM approach is not limited to any specific level of quantum chemistry and has been applied in conjunction with Density Functional Theory (DFT) in studies of photosystem II and other macrobiological systems (Gascon et al. 2006a, b; Menikarachchi and Gascon 2008; Sproviero et al. 2006b, 2008d). The main advantage of DFT is its efficiency for the description of electron correlation at a computational cost comparable to single-determinant calculations, and the efficient description of metal centers and inorganic cores of metalloproteins in various possible spin-electronic states. MoD-QM/MM calculations have modeled charge redistribution, polarization effects and changes in the ligation scheme of metal centers induced by oxidation state transitions, bond breaking, and forming processes associated with chemical reactivity and proton transfer. In addition, the MoD-QM/MM method has provided a reliable description of electronic and magnetic properties of prosthetic groups, embedded in complex molecular environments, and has allowed for the application of

167 spectroscopically constrained QM/MM optimization meth-
 168 ods based on simulations of X-ray absorption spectra and
 169 direct comparisons with high-resolution spectroscopic data
 170 (e.g., FT-IR, NMR, EXAFS, and EPR; Sproviero et al.
 171 2008c, d).

172 The work reviewed in this article illustrates the MoD-
 173 QM/MM methodology as applied to the description of
 174 macromolecules of biological interest, including photo-
 175 system II, and in conjunction with spectroscopically con-
 176 strained QM/MM structural-refinement schemes based on
 177 simulations of X-ray absorption spectra and direct com-
 178 parisons with experimental data. The application studies
 179 show that the MoD-QM/MM methodology is well-suited to
 180 study equilibrium and transition states of biological sys-
 181 tems beyond the limitations of model structures obtained
 182 by X-ray diffraction or NMR spectroscopy, including
 183 reaction intermediates generated by water exchange, proton
 184 transfer, and the rearrangement of ligands during the cat-
 185 alytic cycle of water oxidation at the oxygen-evolving
 186 complex (OEC) of photosystem II (PSII; Sproviero et al.
 187 2006b, 2007, 2008b, d). This article is organized as fol-
 188 lows. First, we review the QM/MM methodology, includ-
 189 ing a discussion of its specific implementation to studies of
 190 high-valent metal complexes in metalloproteins. Then, we
 191 review the MoD-QM/MM method and a spectroscopically
 192 constrained (SC) QM/MM refinement scheme based on
 193 simulations of X-ray absorption spectra and direct com-
 194 parisons with experimental measurements.

195 QM/MM methodology

196 The MoD-QM/MM methodology (Gascon et al. 2006a;
 197 Menikarachchi and Gascon 2008; Newcomer et al. 2009;
 198 Sproviero et al. 2008e) has been applied in conjunction
 199 with the two-layer ONIOM electronic-embedding (EE)
 200 link-hydrogen atom method (Dapprich et al. 1999), as
 201 implemented in Gaussian 03 (Frisch et al. 2004). The
 202 ONIOM approach divides the system into two partitions
 203 (Fig. 1), including a reduced system X (the QM layer) and
 204 the rest of the system (region Y). The total energy E
 205 is obtained from three independent calculations:

$$E = E^{\text{MM},X+Y} + E^{\text{QM},X} - E^{\text{MM},X} \quad (1)$$

207 where $E^{\text{MM},X+Y}$ is the energy of the complete system
 208 computed at the molecular mechanics level of theory,
 209 while $E^{\text{QM},X}$ and $E^{\text{MM},X}$ correspond to the energy of
 210 the reduced system X computed at the QM and MM
 211 levels of theory, respectively. Electrostatic interactions
 212 between regions X and Y are included in the calculation
 213 of both $E^{\text{QM},X}$ and $E^{\text{MM},X}$ at the quantum mechanical
 214 and molecular mechanics levels, respectively. Thus, the
 215 electrostatic interactions computed at the MM level in $E^{\text{MM},X}$ and

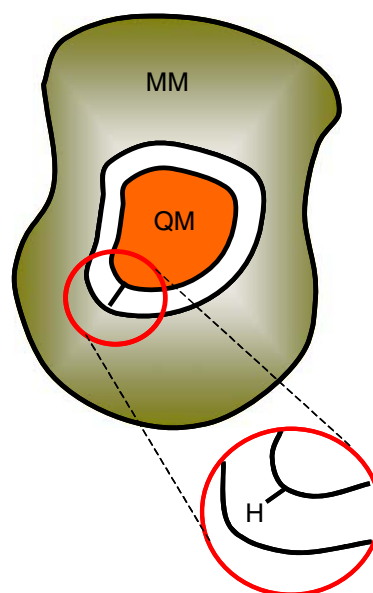


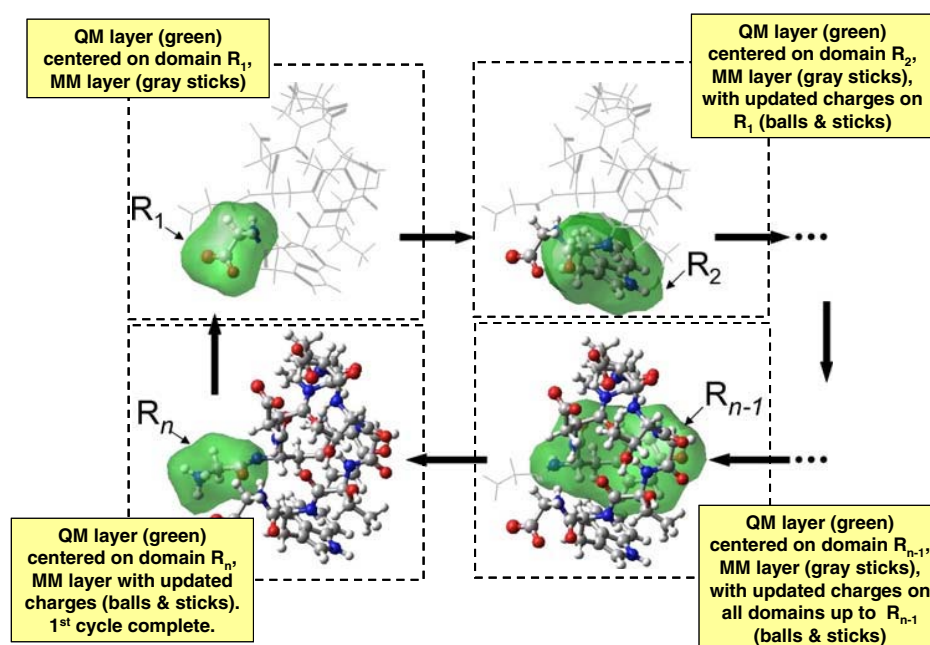
Fig. 1 QM/MM fragmentation of the system into a reduced QM layer (region X) and the surrounding molecular environment described as a MM layer (region Y)

$E^{\text{MM},\text{full}}$ cancel and the resulting DFT-QM/MM evaluation
 of the total energy involves a quantum mechanical
 description of the polarization of the reduced system, due
 to the electrostatic influence of the surrounding protein
 environment.

MoD-QM/MM methodology

Figure 2 illustrates the MoD-QM/MM computational pro-
 tocol where the system is partitioned into n molecular
 domains by using a simple space-domain decomposition
 scheme (Gascon et al. 2006a; Menikarachchi and Gascon
 2008; Newcomer et al. 2009). Electrostatic potential (ESP)
 atomic charges of the constituent molecular domains are
 sequentially computed by using the QM/MM hybrid
 methods (e.g., the two-layer ONIOM-EE scheme) consid-
 ering the updated distribution of atomic charges in the
 molecular domains previously modeled as QM layers in the
 cycle. The whole cycle of QM/MM calculations is iterated
 several times until obtaining a self-consistent point-charge
 model of the electrostatic potential that explicitly includes
 mutual polarization interactions between the constituent
 molecular domains. The underlying computational effort
 usually scales linearly with the size of the system, since the
 iterative scheme typically converges within a few iteration-
 cycles (i.e., 4 or 5 cycles). The method, thus, bypasses the
 enormous demands of memory and computational resour-
 ces that would be required by a 'brute-force' quantum
 chemistry calculation of the complete system.

Fig. 2 Schematic representation of the MoD-QM/MM method in terms of an iterative cycle of QM/MM calculations with QM layers (in green) defined on different molecular domains for different steps of the cycle (Gascon et al. 2006a). Balls and sticks represent molecular domains with ESP atomic charges updated in previous steps of the cycle



243 The MoD-QM/MM method typically corrects the
 244 atomic charges prescribed by popular molecular mechanics
 245 force fields (Cornell et al. 1995; Dykstra 1993) by only 10–
 246 20%. However, summing the corrections from multiple
 247 domains over the entire QM/MM interface often introduces
 248 significant corrections (10–15 kcal/mol) to the total inter-
 249 action between the QM and MM layers. These energy
 250 corrections are often critical since they are comparable to
 251 energy differences associated with different protonation
 252 states, hydrogen-bonding structures, or the energy splitting
 253 between high-spin and low-spin states of the metal cluster.

254 The accuracy of the resulting molecular electrostatic
 255 potential comes at the expense of transferability, since the
 256 atomic charges obtained for one configuration are, in
 257 principle, nontransferable to other configurations of the
 258 system. Therefore, while accurate, the computed electro-
 259 static potential is most useful for applications where the
 260 system does not involve large conformational changes,
 261 such as intermediate structures of the OEC of PSII where
 262 there are no significant temperature-dependent variations in
 263 structure, protonation state, or charge localization, as
 264 recently reported by studies of X-ray absorption at 20 K
 265 and room temperature (Haumann et al. 2005).

266 The resulting generalized computational protocol
 267 requires the space-domain fragmentation of the system into
 268 partially overlapping molecular domains and an iterative
 269 cycle of QM/MM calculations as depicted in Fig. 2. Each
 270 step in the cycle considers a different molecular domain as
 271 a QM layer and updates its distribution of ESP atomic

charges, after relaxing its geometry by energy minimiza-
 272 tion. The rest of the system is considered at the MM level,
 273 with an updated distribution of atomic charges for all
 274 molecular domains previously considered as QM layers in
 275 the iteration cycle.

276 The geometry optimization is performed with standard
 277 methods (Fan and Ziegler 1991; Versluis and Ziegler 1988),
 278 using the Newton–Raphson method and updating the sec-
 279 ond derivative (Hessian matrix) with the Broyden–
 280 Fletcher–Goldfarb–Shanno strategy (Schlegel 1987).
 281 Starting with the first molecular domain R_1 , its geometry is
 282 relaxed as influenced by the surrounding environment. The
 283 ESP charges of the QM layer in its minimum energy con-
 284 figuration are computed and updated for subsequent steps in
 285 the cycle (see Fig. 2). The relaxed configuration and ESP
 286 atomic charges of the next domain (R_2) are computed
 287 analogously, considering the updated charges and configu-
 288 ration of the previously considered molecular domain. The
 289 procedure is subsequently applied to all fragments and the
 290 entire computational cycle is iterated several times until
 291 reaching self-consistency in the description of the geometry
 292 and distribution of atomic charges as influenced by mutual
 293 polarization effects. The capabilities of the resulting pro-
 294 tocol have been recently demonstrated at the ONIOM-EE
 295 level, as applied to the description of benchmark models
 296 systems, which allowed for direct comparisons with full
 297 QM calculations as well as in the study of the structural
 298 characterization of the DNA *Oxytricha nova* guanine
 299 quadruplex (G4) (Newcomer et al. 2009).
 300

301 MoD-QM/MM structural refinement constrained with 302 high-resolution spectroscopic data

303 The typical energy landscape of macrobiological systems is
304 very rugged, with a manifold of configurations of compa-
305 rable energy (e.g., within the ~ 5 kcal/mol energy range of
306 error of DFT calculations). Therefore, it is rather difficult to
307 determine the most likely configuration of the system solely
308 from the energetic analysis of possible configurations. In
309 order to address these limitations, a spectroscopically con-
310 strained MoD-QM/MM structural-refinement scheme has
311 been introduced. The method allows one to assess the
312 likelihood of structures of similar energies in terms of
313 simulations of high-resolution spectroscopy (e.g., EXAFS
314 spectra) and direct comparisons with experimental data
315 (Sproviero et al. 2008c). These MoD-QM/MM methods
316 have been successfully applied to the development of
317 structural models of the OEC of PSII (Sproviero et al.
318 2008a) and other systems (Soderhjelm and Ryde 2006) and
319 proved particularly valuable to refine configuration of
320 prosthetic groups with multiple high-valent metal centers
321 beyond the current limitations of DFT-QM/MM methods.

322 The refinement algorithm starts with a reference con-
323 figuration of the system (e.g., the MoD-QM/MM minimum
324 energy structure) and iteratively adjusts its configuration to
325 minimize the mean-square deviation between the simulated
326 and experimental EXAFS spectra, subject to the constraint
327 of minimum displacements relative to the reference con-
328 figuration (see Fig. 3). Upon convergence, the resulting

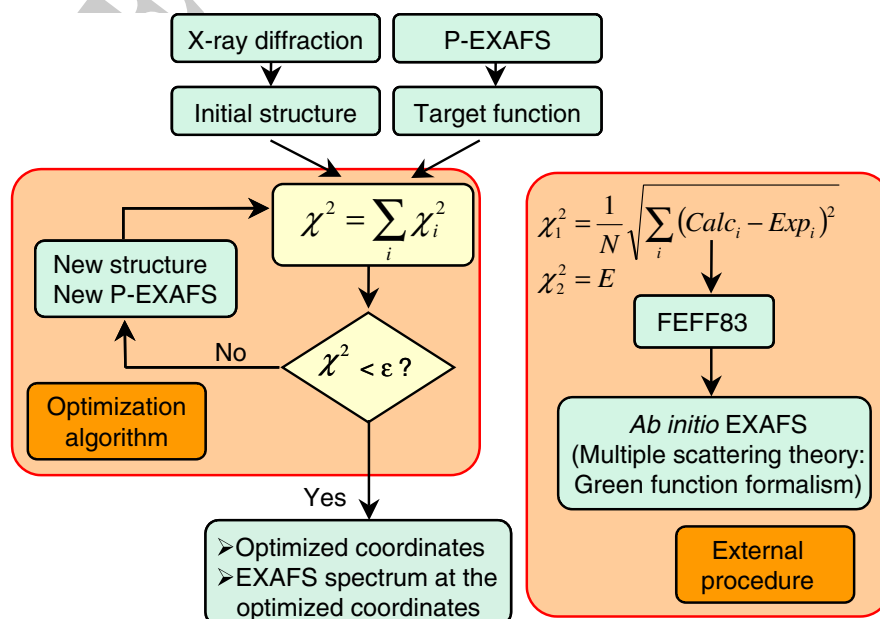
329 structures are often consistent with the reference MoD-
330 QM/MM minimum energy structures (with energies and
331 root-mean-square deviations within the range of error of
332 DFT-QM/MM minimization methods) and predict high-
333 resolution spectroscopic data in quantitative agreement
334 with experiments. The refinement scheme allows one to
335 narrow the range of possible structures through direct
336 comparisons with high-resolution spectroscopic data. As
337 mentioned earlier, this method is particularly valuable for
338 metalloproteins for which the most accurate and reliable
339 experimental information on the structure of the prosthetic
340 group metal clusters comes from high-resolution spec-
341 troscopy, e.g., EXAFS (Yano et al. 2006).

342 The MoD-QM/MM refinement protocol has been applied
343 in conjunction with simulations of EXAFS spectra for the
344 development of models of the OEC of PSII (Sproviero et al.
345 2008c). The EXAFS spectra were computed as a function of
346 the nuclear coordinates and compared to the experimental
347 spectra. The nuclear coordinates of the Mn cluster and
348 coordinated ligands were adjusted to minimize the mean-

349 square deviation $\chi^2 = \frac{1}{N} \sqrt{\sum_i^N (A_i^{\text{calc}} - A_i^{\text{exp}})^2}$, where

350 A_i^{calc} and A_i^{exp} correspond to the calculated and experimental
351 EXAFS amplitudes for a grid of electron kinetic energies
352 K_i with $i = 1, 2, \dots, N$. The χ^2 minimization is performed
353 subject to the constraints of minimum displacements of
354 nuclear positions relative to their coordinates in the refer-
355 ence MoD-QM/MM configuration. The minimization

Fig. 3 SC-QM/MM structural-refinement protocol based on simulations of X-ray absorption spectra with FEFF8 and direct comparisons with experimental data (Sproviero et al. 2008c)



algorithm implements a multidimensional variable metric method through the Broyden–Fletcher–Goldfarb–Shanno (BFGS) approach (Broyden 1970; Fletcher 1970; Goldfarb 1970; Shanno 1970), using the Cartesian coordinates x_i of the nuclei as input variables and an external simulation procedure that returns the corresponding spectrum as computed by the program FEFF8 (version 8.2; Ankudinov et al. 1998, 2002; Bouldin et al. 2001).

Ab initio simulations of EXAFS spectra involve solving the multiscattering problem associated with the photodetached electrons emitted upon X-ray absorption by the Mn centers. The quantum mechanical interference of outgoing photoelectrons with the backscattering waves from atoms surrounding the Mn ions, gives rise to oscillations of EXAFS intensities as a function of the kinetic energy (or momentum) of the photoelectron. The Fourier transform of these oscillations provide information on the Mn–Mn distances, the coordination of Mn ions, and changes in the Mn coordination sphere as determined by changes in the oxidation state of the metal centers. Calculations are typically carried out according to the Real Space Green function approach as implemented in the program FEFF8 (version 8.2; Ankudinov et al. 1998, 2002; Bouldin et al. 2001), using the theory of the oscillatory structure due to multiple-scattering originally proposed by Kronig (1931) and Kronig and Penney (1931) and worked out in detail by Sayers et al. (1971), Stern (1974) Lee and Pendry (1975), and by Ashley and Doniach (1975).

One of the advantages of the structural-refinement scheme based on EXAFS simulations and comparisons with experimental data is that EXAFS experiments require doses of X-ray radiation, which are substantially lower than those required by X-ray diffraction (XRD) measurements (Yano et al. 2005). In addition, EXAFS provides metal-to-metal and metal-to-ligand distances with high accuracy (~ 0.02 Å) and a resolution of ~ 0.1 Å. For polarized-EXAFS experiments on one-dimensionally oriented membranes or three-dimensionally oriented single crystals, the EXAFS amplitude is orientation-dependent and proportional to $\cos^2(\theta)$, where θ is the angle between the electric field vector of the polarized X-ray beam and the absorber back-scatterer vector. Therefore, the structural refinement can provide additional geometric information about the metal site in metallo-protein crystals (Flank et al. 1986; Scott et al. 1982; Yano et al. 2005).

403 Exchange coupling constants

404 The SC-QM/MM analysis of MoD-QM/MM structural
405 models can be based on calculations of magnetic properties

and direct comparisons with experimental data. Biomimetic oxomanganese complexes (Cooper and Calvin 1977; Cooper et al. 1978; Limburg et al. 1999; Manchanda et al. 1994; Ruettinger et al. 1999; Sarneski et al. 1990; Stebler et al. 1986; Thorp and Brudvig 1991) and models of the Mn_4Ca complex of PSII extracted from complete MoD-QM/MM models (Sproviero et al. 2006b) have been recently analyzed (Pantazis et al. 2009; Sproviero et al. 2006a). The exchange coupling constants J_{ij} between centers i and j were computed. Assuming that the centers interact according to the Heisenberg Hamiltonian:

$$H = - \sum_{i,j}^N J_{ij} \hat{S}_i \hat{S}_j, \quad (9)$$

where S_i is the spin corresponding to metal center i . In the weak-coupling limit, the spin states are essentially localized on individual metal centers and the exchange coupling constants are determined by splittings between energy levels as follows:

$$J_{ij} = \frac{E_{BS}^{(ij)} - E_{HS}^{(ij)}}{2S_i S_j + \min(S_i, S_j)}, \quad (10)$$

where $E_{BS}^{(ij)}$ and $E_{HS}^{(ij)}$ are the magnetic energy contributions from the pair of metal centers i and j (with $S_i < S_j$) in the broken-symmetry (BS) and high-spin (HS) configurations, respectively. The HS configuration corresponds to equally oriented spins while in the BS state the α and β densities are allowed to localize on different atomic centers, providing a multiconfigurational character to the spin state (Noodleman 1981; Noodleman and Case 1992; Noodleman and Davidson 1986; Noodleman et al. 1995).

The physical picture supporting this calculation is that in the weak-coupling limit, atomic d -orbitals remain singly occupied, and the spin states are essentially localized on individual metal centers. Under those conditions, the ground state of a multinuclear complex involves antiparallel coupling between spins localized on different centers. However, metal–metal interactions induce electronic delocalization over multiple centers, forming bonding and antibonding orbitals by either direct overlap of atomic orbitals or super-exchange. In the BS state, the α and β electronic densities can be localized on different centers. These relaxed symmetry requirements allow for substantial interactions mediated by d -orbitals of the metal centers and the p -orbitals of the ligands.

In order to perform ab initio calculations of J , it is convenient to define spin configurations that differ only in the orientation of the spin of one center with respect to the HS configuration. The energy difference associated with flipping a spin of center k with respect to the HS configuration is

$$\begin{aligned}
 E_{HS} - E_{SCK} &= \sum_{j \neq k}^N E_{HS}^{(jk)} - E_{BS}^{(jk)} \\
 &= - \sum_{j \neq k}^N J_{jk} (2S_j S_k + \min(S_j, S_k)), \quad (11)
 \end{aligned}$$

454 where $k = 1, \dots, N$. Considering that there are $N + 1$ rel-
 455 evant spin configurations, including one with the highest
 456 spin and N with a spin flipped in only one center, the
 457 coupling constants can be readily obtained, since the
 458 number of equations is equal or greater than the number of
 459 unknown J_s .

460 Applications

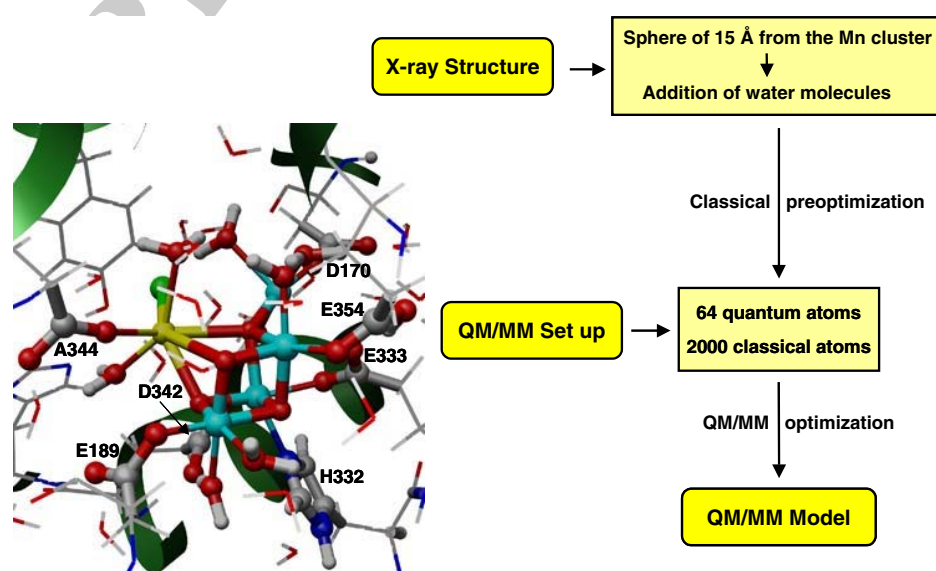
461 MOD-QM/M model of photosystem II

462 The moving-domain quantum mechanics/molecular
 463 mechanics (MoD-QM/MM) models of the OEC of PSII
 464 have been based on the 1S5L X-ray crystal structure of PSII
 465 from the cyanobacterium *Thermosynechococcus elongatus*
 466 1S5L (Ferreira et al. 2004; see Fig. 4). The models explicitly
 467 considered 1,987 atoms of PSII, including the proposed
 468 $\text{Mn}_3\text{CaO}_4\text{Mn}$ unit and all amino acid residues with α -car-
 469 bons within 15 Å from any atom in the OEC metal ion
 470 cluster. The coordination spheres of the Mn centers were
 471 completed by hydration, assuming a minimum displacement
 472 of the proteinaceous ligands from their crystallographic
 473 positions. The usual coordination of 5 or 6 ligands, for high-
 474 valent Mn ions with oxidation states III and IV, was also
 475 satisfied. In addition, the coordination of calcium (typically
 476 with 6–8 ligands) was satisfied by coordination of water
 477 molecules and negative counterions.

478 The MoD-QM/MM molecular structures were hydrated
 479 by including water molecules that could bind to polar
 480 amino acid residues, or existing water molecules in the
 481 model. The hydration process involved inserting the model
 482 system in a large box containing a thermal distribution of
 483 water molecules. After adding to the model, the water
 484 molecules that did not interfere sterically with the protein
 485 residues, the structure of hydrogen bonds was optimized by
 486 energy minimization. The hydration and relaxation pro-
 487 cesses were repeated several times until the number of
 488 water molecules in the system converged, typically with
 489 the addition of 85–100 water molecules. A few of the
 490 additional water molecules (up to six molecules) attached
 491 to metal centers in the cuboidal $\text{Mn}_3\text{CaO}_4\text{Mn}$ cluster,
 492 including two water molecules that were considered as
 493 potential substrate, water molecules responsible for O–O
 494 bond formation in the S_4 to S_0 transition.

495 The moving-domain quantum mechanics/molecular
 496 mechanics (MoD-QM/MM) computations of the OEC of
 497 PSII could only be efficiently performed, according to the
 498 ONIOM-EE level of theory, after obtaining high-quality
 499 initial-guess states of the QM layer. These correspond to
 500 spin-electronic states of the cluster of Mn ions based on
 501 ligand field theory as implemented in Jaguar 5.5 (2003). The
 502 electronic states of the Mn cluster involve antiferromagnetic
 503 couplings between manganese centers and are defined by
 504 broken-symmetry (BS) states, in which α and β densities are
 505 allowed to localize on different atomic fragments (Noodle-
 506 man 1981; Noodleman and Davidson 1986; Noodleman and
 507 Case 1992; Noodleman et al. 1995). Each fragment is simply
 508 a group of atoms that would be bonded together even if all
 509 bonds between transition metal atoms and non-transition
 510 metal atoms were broken. Typically, the system is broken
 511 into ligand fragments and transition metal fragments. The

Fig. 4 *Left:* Structural model of the OEC of PSII, including QM layer (balls and sticks) and MM layer (wires). *Right:* QM/MM setup of the OEC of PSII



512 BS approach, in conjunction with DFT, is the most accurate
513 approach available for these types of high-valent metal
514 clusters. BS-DFT also facilitates the ligand field analysis,
515 including metal $d \rightarrow d$, charge transfer (ligand \rightarrow metal or
516 metal \rightarrow ligand), and intervalence charge transfer
517 (metal \rightarrow metal or ligand \rightarrow ligand) transitions.

518 From a theoretical point of view, changes in the spin
519 state will affect the spatial distribution and energetics of
520 the molecular orbitals. As different electronic distributions
521 will give different relaxed geometries, having a proper
522 description of the spin state of the system allows a correct
523 description of the corresponding relaxed geometry (Lund-
524 berg and Siegbahn 2004).

525 Once the initial state was properly defined, the com-
526 bined approach exploits important capabilities of ONIOM
527 as implemented in Gaussian 03 (Frisch et al. 2004),
528 including both the link-hydrogen atom scheme for efficient
529 and flexible definitions of QM layers and the possibility of
530 modeling open-shell systems by performing unrestricted
531 DFT (e.g., UB3LYP) calculations in a variety of spin-
532 electronic states.

533 The QM layer of MoD-QM/MM calculations (region X)
534 included the Mn_3CaO_4 -Mn complex and the directly liga-
535 ted proteinaceous carboxylate groups of D1-D189, CP43-
536 E354, D1-A344, D1-E333, D1-D170, and D1-D342. In
537 addition, the imidazole ring of D1-H332 as well as bound
538 water molecules, hydroxide, and chloride ions were inclu-
539 ded in the QM layer. The rest of the system defined the MM
540 layer (region Y). The boundary between the QM and MM
541 layers was defined by cutting covalent bonds in the corre-
542 sponding amino acid residues (i.e., D1-D189, CP43-E354,
543 D1-A344, D1-E333, D1-D170, D1-D342, and D1-H332)
544 and completing the covalency of frontier atoms according to
545 the standard link-hydrogen atom scheme (see Fig. 1).

546 The efficiency of MoD-QM/MM calculations has been
547 optimized by using a combination of basis sets for the QM
548 layer (Sprovierio et al. 2006c). Such a combination of basis
549 sets has been validated through extensive benchmark cal-
550 culations on high-valent manganese complexes (Sprovierio
551 et al. 2006a). The molecular structure beyond the QM layer
552 is described by the Amber MM force field (Cornell et al.
553 1995).

554 Fully relaxed MoD-QM/MM model structures were
555 obtained by energy minimization of the QM layer, sur-
556 rounded by the MM layer of amino acid residues with
557 α -carbons within 15–20 Å from any atom in the OEC ion
558 cluster. The buffer shell is subjected to harmonic constraints
559 in order to preserve the natural shape of the system even in
560 the absence of the membrane. Upon convergence, the
561 resulting optimized structures are analyzed and assessed not
562 only in terms of the total energy of the system but also in
563 terms of the comparison of structural, electronic, and
564 mechanistic features that should be consistent with

565 experimental data. Such a comprehensive analysis is crucial
566 for these complex biological systems often with many possi-
567 ble configurations with similar energies (e.g., within the
568 ~ 5 kcal/mol range of error of DFT-QM/MM calculations).

569 SC-QM/MM refinement model of the OEC of PSII

570 The MoD-QM/MM minimum energy structure of the OEC
571 of PSII in the S_1 state was used as the initial reference
572 configuration for the EXAFS structural-refinement algo-
573 rithm based on P-EXAFS data from Yano et al. (2008). The
574 algorithm searched for a configuration with the smallest
575 possible displacement relative to the initial structure with
576 the constraint of a minimum χ^2 deviation between the
577 calculated and experimental EXAFS data. Upon conver-
578 gence of the structural-refinement procedure, the isotropic
579 and polarized-EXAFS spectra of the resulting refined (R)-
580 QM/MM model were found to be in quantitative agreement
581 with high-resolution EXAFS measurements (Haumann
582 et al. 2005; Yano et al. 2005; Figs. 5, 6) as well as with the
583 description of the protein environment suggested by XRD
584 models (Ferreira et al. 2004; Loll et al. 2005; Fig. 7).

585 The predicted metal–metal and metal–ligand distances
586 are consistent with EXAFS studies of frozen solutions of
587 PSII that have provided accurate distances for Mn–Mn, Mn–
588 Ca, and Mn/Ca–ligand vectors in the Mn_4Ca cluster of PSII
589 (Dau et al. 2001; Haumann et al. 2005; Robblee et al. 2001;
590 Yano et al. 2005). In addition, the R-QM/MM model pro-
591 vides insights on structural details currently unresolved by
592 XRD spectroscopy, including the spatial orientation of the
593 cluster, which has been inherited from the experimental
594 models (Yano et al. 2006), as well as the approximate
595 placement of the models within the protein environment,
596 which is achieved in this study (Sprovierio et al. 2008a).
597 However, it is important to emphasize that the R-QM/MM
598 model does not rule out other possible structures since it is
599 not only the result of an exhaustive refinement of all possible
600 structures of the system but also rather a local minimum
601 solution relative to the reference MoD-QM/MM structure.

602 The oxomanganese inorganic core of the R-QM/MM
603 model involves a cuboidal core Mn_3CaO_4 with a dangler
604 manganese ligated to a corner μ_4 -oxide ion. Dangler Mn
605 models were originally proposed (Britt et al. 2004; Britt
606 et al. 2000) within a set of 11 possible structural motifs
607 (Derose et al. 1994), suggested by solution EXAFS studies,
608 from which the dimer-of-dimers model was highly favored
609 (Britt et al. 2000, 2004; Yachandra 2002). Similar dangler
610 Mn models were also preferred by ^{55}Mn electronuclear
611 double resonance (ENDOR) studies that strongly disfa-
612 vored the dimer-of-dimers motif over models with a tri-
613 nuclear Mn core and a fourth Mn set off from the core by a
614 longer Mn–Mn internuclear distance (Britt et al. 2000).
615 However, recent polarized-EXAFS studies of single

Fig. 5 Comparison of the experimental polarized-EXAFS data (red) (Yano et al. 2006) and the corresponding calculated spectra (blue) obtained with the MoD-QM/MM (top) and R-QM/MM (bottom) models (Sproviero et al. 2008c)

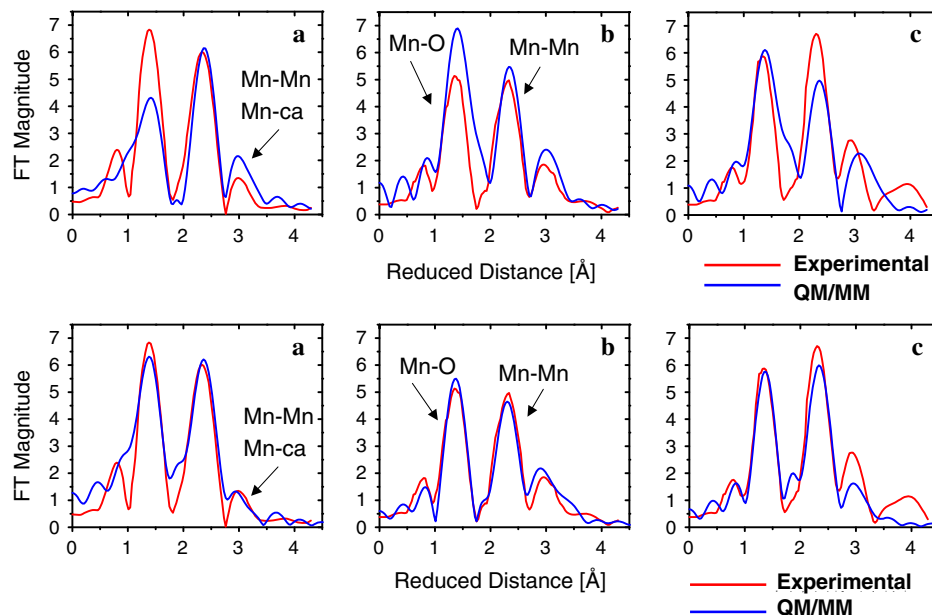
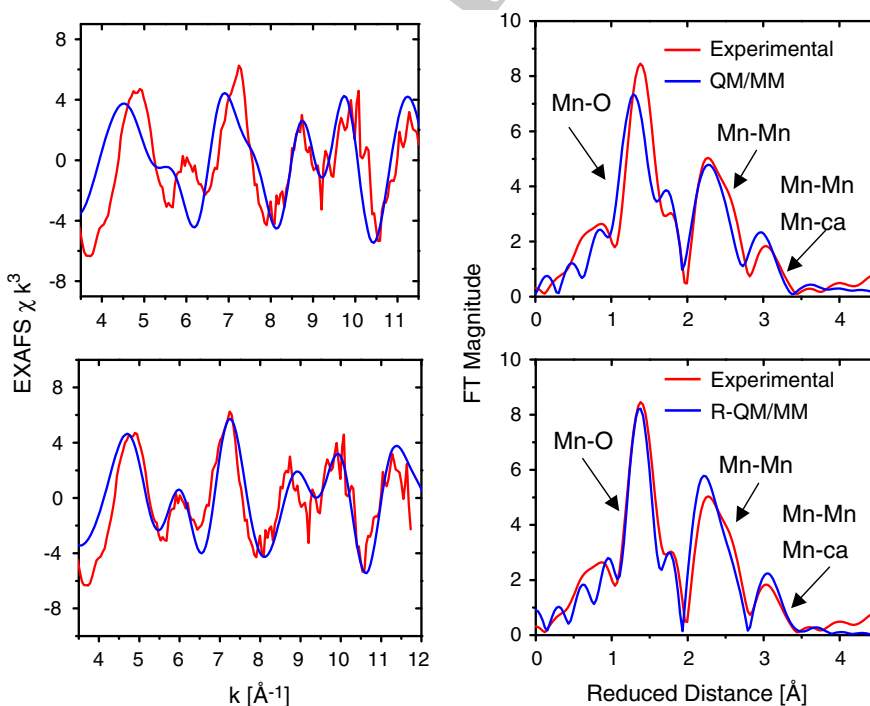


Fig. 6 Comparison of the experimental (red) (Dau et al. 2004) isotropic EXAFS spectrum in k-space (left) and FT EXAFS magnitude (right) and the calculated spectra (blue) obtained with the QM/MM (top) and R-QM/MM (bottom) models (Sproviero et al. 2008c)



616 crystals have disfavored dangler models (Yano et al. 2006)
 617 and have suggested instead four alternative models (models
 618 I, II, IIa, and III (Yano et al. 2006)) for which the simulated
 619 EXAFS spectra (in the absence of ligands) match the
 620 experimental data. These empirical models, however,
 621 remain uncertain since structural differences between some
 622 of them (e.g., models I and II) are large. Furthermore,
 623 placing any of the four models into the XRD structures of
 624 PSII results in unsatisfactory metal–ligand distances,

coordination numbers, and geometries. The underlying
 inconsistencies might be due to the coordinate error in the
 X-ray crystal models, or due to the intrinsic limitations of
 polarized-EXAFS models built by neglecting the contribu-
 tions of electron scattering paths from the ligands (i.e.,
 assuming that the EXAFS amplitudes result solely from
 scattering paths associated with metal centers and oxo-
 bridges in the inorganic core). These issues should be
 addressed in the development of empirical models since

625
 626
 627
 628
 629
 630
 631
 632
 633

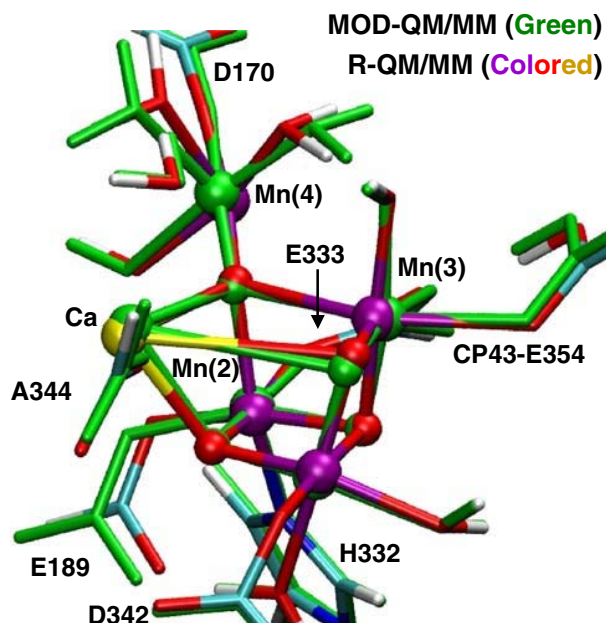


Fig. 7 Superposition of the refined (R)-QM/MM (*colored*) and MoD-QM/MM (*green*) reference structure of the OEC of PSII (Sproviero et al. 2008c). The R-QM/MM model reproduces the experimental isotropic (Dau et al. 2004) and polarized (Yano et al. 2006) EXAFS spectra and is consistent with XRD models (Ferreira et al. 2004; Loll et al. 2005)

Table 1 Oxidation numbers, Mulliken spin population analyses (*in parenthesis*), and ab initio exchange coupling constants J_{ij} computed at the DFT B3LYP level with the TZV basis set for the high-valent multinuclear oxomanganese complexes **1**, **2**, and **3** introduced in the text (see Fig. 8)

Complex	Mn(1)	Mn(2)	Mn(3)	J_{12} (cm^{-1})	J_{23} (cm^{-1})
1	III(+3.8)	IV(-2.6)	-	298(300) ^a	-
2	III(+3.9)	IV(-2.6)	-	295(268–300) ^a	-
3	IV(-1.2)	IV(-1.2)	IV(+2.8)	211(182) ^a	70(98) ^a

^a Values in parenthesis indicate experimental values of coupling constants (Cooper and Calvin 1977; Cooper et al. 1978)

These results serve to validate the analyzed model structure and its electronic description, as applied to the description of magnetic properties of oxomanganese complexes with pre-defined spin-electronic states. The reported agreement between calculated and experimental exchange coupling constants was consistent with anti-ferromagnetic couplings of high-valent metal centers in high-spin states. The analysis of the electronic structures show that the most important exchange interactions between Mn centers involved symmetric super-exchange pathways $J_{xz/xz}$ and $J_{yz/yz}$ mediated by the overlap of orbitals d_{xz} and d_{yz} of Mn with out-of-the-plane p_{π} orbitals of the μ -oxo-bridges (Sproviero et al. 2006a). These results are particularly relevant to the description of active sites of metalloproteins with common prosthetic groups and provide insight on the origin of electronic delocalization over multiple centers (Pantazis et al. 2009).

calculations of polarized-EXAFS spectra suggest that the contributions of backscattering from the ligands could introduce amplitude corrections of the order of 10–15% (Sproviero et al. 2008c).

Exchange coupling constants

Benchmark calculations of exchange coupling constants involved the analysis of di-, tri- and tetra-nuclear Mn complexes (Sproviero et al. 2006a), previously characterized both chemically and spectroscopically (Cooper and Calvin 1977; Cooper et al. 1978; see Table 1). These include the di- μ -oxo bridged dimers **1**: $[\text{Mn}^{\text{IV}}\text{Mn}^{\text{III}}(\mu\text{-O})_2(\text{H}_2\text{O})_2(\text{terpy})_2]^{3+}$ (terpy = 2,2':6,2''-terpyridine) and **2**: $[\text{Mn}^{\text{III}}\text{Mn}^{\text{IV}}(\mu\text{-O})_2(\text{phen})_2]^{3+}$ (phen = 1,10-phenanthroline), and the trimer **3**: $[\text{Mn}_3\text{O}_4(\text{bpy})_4(\text{OH}_2)_2]^{4+}$ (bpy = 2,2'-bipyridine), shown in Fig. 8. These synthetic complexes have been previously characterized by electrochemical and high-resolution EPR and X-ray diffraction methods (Cooper and Calvin 1977; Cooper et al. 1978).

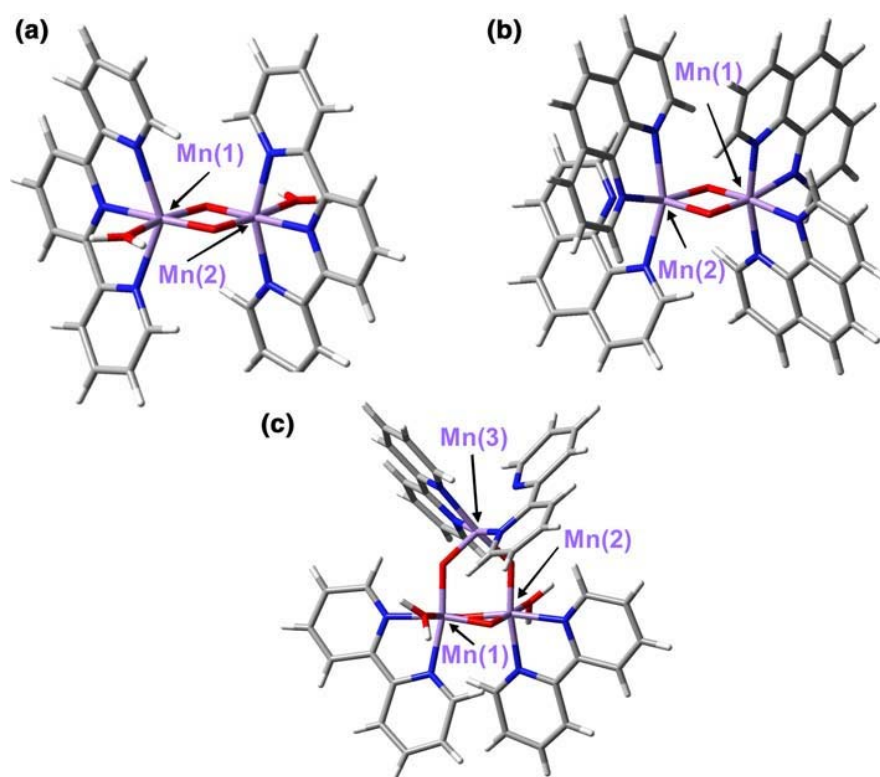
Table 1 compares exchange coupling constants for synthetic oxomanganese complexes **1–3** to readily available experimental data (Cooper and Calvin 1977; Cooper et al. 1978; Sarneski et al. 1990; Stebler et al. 1986). Exchange coupling constants were obtained at the DFT/B3LYP level and show very good agreement with experimental data.

Water-exchange rates

The MoD-QM/MM methodology has been recently applied to develop structural models of transition-state structures the OEC of PSII generated by water exchange, in an effort to provide insights on the nature of water-exchange rates, and rigorous validations of the structural models by direct comparisons of water-exchange rates with time-resolved mass spectrometry (MS) measurements (Sproviero et al. 2008e).

Time-resolved MS studies have determined different water-exchange rates (k_{ex}) of two substrate water molecules of the OEC with bulk ^{18}O -labeled water. The more slowly exchanging water (W_{slow}) was found to be associated with Ca^{2+} , implying that the fast-exchanging water (W_{fast}) would be bound to a manganese ion (Hendry and Wydrzynski 2002; Hillier and Wydrzynski 2008; Messinger et al. 1995). These findings were rather surprising since high-valent manganese ions (e.g., Mn^{3+} or Mn^{4+}) were expected to

Fig. 8 Molecular structures of (a) complex 1, $[\text{Mn}^{\text{IV}}\text{Mn}^{\text{III}}(\mu\text{-O})_2(\text{H}_2\text{O})_2(\text{terpy})_2]^{3+}$ (terpy = 2,2':6,2''-terpyridine), (b) complex 2, $[\text{Mn}^{\text{III}}\text{Mn}^{\text{IV}}(\mu\text{-O})_2(\text{phen})_2]^{3+}$ (phen = 1,10-phenanthroline), and (c) complex 3, $[\text{Mn}_3\text{O}_4(\text{bpy})_4(\text{OH}_2)_2]^{4+}$ (bpy = 2,2'-bipyridine) (Sproviero et al. 2006a). Color key: red = oxygen, blue = nitrogen, gray = carbon, white = hydrogen, and purple = manganese



693 bind water more strongly than Ca^{2+} . In addition, it was
694 observed that the exchange rate of W_{slow} (k_{ex} slow)
695 increased by two orders of magnitude upon S_1 to S_2 oxi-
696 dation, with $k_{\text{ex}}(S_1) = 0.02 \text{ s}^{-1}$ and $k_{\text{ex}}(S_2) = 2.0 \text{ s}^{-1}$
697 (Hendry and Wydrzynski 2003; Hillier and Wydrzynski
698 2004). These rates correspond approximately to activation
699 barriers of 20 and 17 kcal/mol, respectively. This faster
700 exchange of W_{slow} in the S_2 state was also intriguing since it
701 implied that the oxidation of a manganese center should
702 indirectly affect the exchange rate of a calcium-bound water
703 molecule.

704 The QM/MM structural models allowed for a quantita-
705 tive analysis of specific metal–water interactions and
706 potential-energy profiles associated along minimum energy
707 paths (MEPs) for water detachment and exchange
708 (Sproviero et al. 2008e). The MEPs were found by energy
709 minimization with respect to nuclear and electronic coordi-
710 nates, while progressively detaching substrate water
711 molecules from their corresponding coordination metal
712 centers. Since elongation of the metal–oxygen bond is the
713 primary step in water exchange, and presumably rate
714 determining in this case, the resulting structural rear-
715 rangements provided insights on the detailed mechanisms
716 of water exchange.

717 The MoD-QM/MM models are consistent with coordi-
718 nation of substrate water molecules as terminal ligands, in
719 agreement with earlier proposals (Haumann and Junge

1999; Hoganson and Babcock 1997; McEvoy and Brudvig
2004; Messinger 2004; Schlodder and Witt 1999; Sproviero
et al. 2006b, 2007), although in contrast to other sug-
gested models that involve coordination as oxo-bridges
between Mn ions (Brudvig and Crabtree 1986; Messinger
2004; Nugent et al. 2001; Pecoraro et al. 1994; Robblee
et al. 2001; Yachandra et al. 1996). The energy barriers for
water exchange in the MoD-QM/MM models were esti-
mated to be 19.3, 16.6, 8.4, and 7.9 kcal/mol for exchange
of water ligands from $\text{Ca}^{2+}(S_1)$, $\text{Ca}^{2+}(S_2)$, $\text{Mn}(4)^{3+}(S_2)$,
and $\text{Mn}(4)^{3+}(S_1)$, respectively. Figures 9, 10, and 11 show
the initial and TS configurations along the MEPs for water
exchange from Ca^{2+} and Mn(4), as described by the DFT-
QM/MM structural models of the OEC, including a
detailed description of coordination bond lengths. In the
initial states (Fig. 9), substrate water molecules are shown
attached to the metal centers forming hydrogen bonds with
the exchanging water molecules. In the TS configurations
(Figs. 10, 11), the coordination bonds of the substrate
water molecules are stretched and the substituting water
molecules are displaced relative to their original configu-
rations. Further displacement beyond the transition state
leads to coordination of the exchanging water molecule to
the metal center. The underlying process involves converted
exchange with dissociative character (i.e., without
formation of any intermediate state of lower or higher
coordination number).

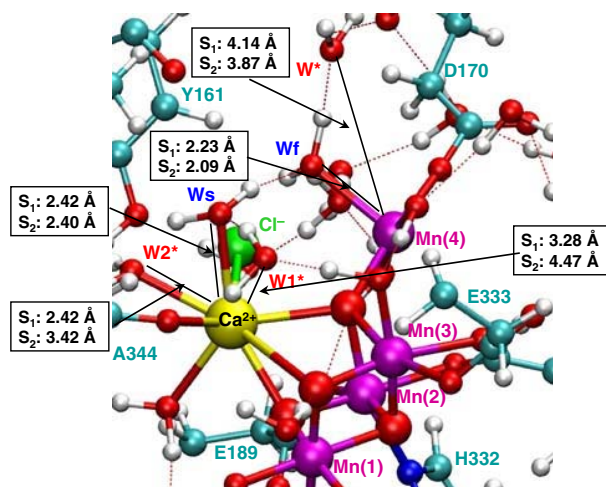


Fig. 9 Structural model of the OEC of PSII in the S_1 and S_2 states. Ligated and exchanging water molecules are labeled W_s (W_{slow}), W_f (W_{fast}) and W^* , respectively. All amino acid residues correspond to the D1 protein subunit, unless otherwise indicated (Sproviero et al. 2008e)

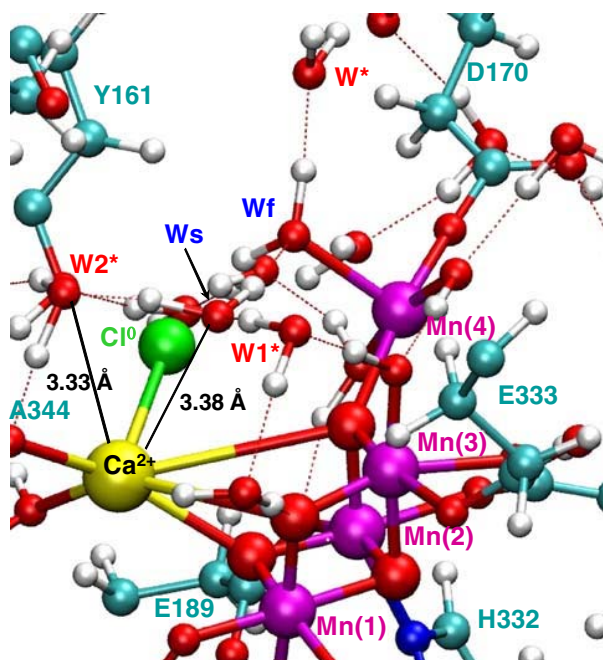


Fig. 11 Structural model of the OEC of PSII in the S_1 and S_2 states at the transition-state configurations for water exchange from Ca^{2+} . The water-exchange energy barriers in the S_1 and S_2 states are 19.3 and 16.6 kcal mol $^{-1}$, respectively. Substrate and exchanging water molecules are labeled W_s (W_{slow}) and W_1^* , respectively. All amino acid residues correspond to the D1 protein subunit, unless otherwise indicated (Sproviero et al. 2008e)

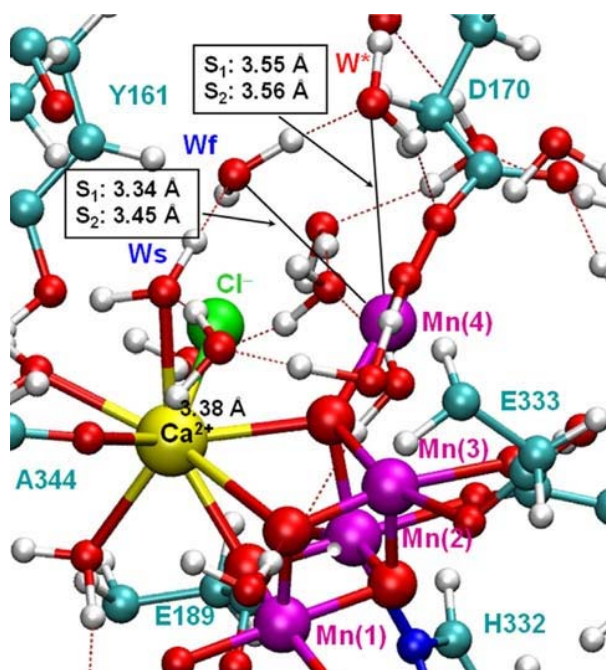


Fig. 10 Structural model of the OEC of PSII in the S_1 and S_2 states at the transition-state configurations for exchange of the fast-exchanging water from Mn(4). The water-exchange energy barriers in the S_1 and S_2 states are 7.9 and 8.4 kcal mol $^{-1}$, respectively. Substrate and exchanging water molecules are labeled W_f (W_{fast}) and W_s , respectively. All amino acid residues correspond to the D1 protein subunit, unless otherwise indicated (Sproviero et al. 2008e)

exchange mechanism gains dissociative character upon S_1/S_2 oxidation of the OEC. The electronic origin of these differences has been traced to the redistribution of charge induced by charge-transfer interactions with D1-A344 leading to a net reduction of the atomic charge of Ca^{2+} . The electrostatic interactions with negatively charged ligands are strengthening upon S_1 to S_2 oxidation, since the metal complex becomes positively charged. Most of the electronic density acquired by Ca^{2+} upon S_1/S_2 oxidation ($\Delta q = -0.21$) is transferred from the ligated carboxylate group of D1-A344. These results suggest that oxidation state transitions of the oxomanganese complex can indirectly regulate substrate water binding to the metal cluster by modulating charge-transfer interactions.

Water exchange in the coordination sphere of the dangling manganese is also predicted to have dissociative character in both the S_1 and the S_2 states, probably due to the reduced charge of the metal center in the presence of charge-transfer interactions between Mn and oxo-bridges. The calculated energy barriers for water exchange from Mn(4) (7.9 and 8.4 kcal/mol in the S_1 and S_2 states, respectively) are comparable to the barriers in hydrated Mn complexes, such as $[(H_2O)_2(OH)_2 Mn^{IV}(\mu-O)_2Mn^{IV}(-H_2O)_2(OH)_2]$ and $[Mn^{IV}(H_2O)_2(OH)_4]$, where the exchange of terminal water ligands requires 8.6–9.6 kcal/mol

The comparison of the coordination bond lengths of water ligands in TS configurations indicates that the coordination bonds are more stretched when water is exchange from Ca^{2+} in the S_2 state, suggesting that the

776 (Lundberg et al. 2003; Tsutsui et al. 1999). In contrast, the
777 DFT-QM/MM energy barriers for water exchange from
778 Ca^{2+} are much higher (19.3 and 16.6 kcal/mol for the OEC
779 in the S_1 and S_2 states, respectively). These larger barriers
780 stabilizing substrate water molecules bound to the catalytic
781 metal center for much longer times (ms) than typically
782 bound to similar complexes in solutions.

783 The analysis of MoD-QM/MM models suggests that the
784 higher energy barriers are mainly determined by the nature
785 of the interactions in the OEC-binding pocket and the
786 incomplete solvation of dissociated water ligands. These
787 calculations, thus, provided theoretical support for the
788 rather slow exchange rates and for the surprising experi-
789 mental finding that the slow-exchanging substrate water is
790 associated with calcium, while the fast-exchanging sub-
791 strate water is coordinated with manganese. Furthermore,
792 the QM/MM models provided a rationale for the opposite
793 effect of the S_1/S_2 transition on the two water-exchange
794 rate constants. It was concluded that the mechanism and
795 energetics of water binding to metal centers in the OEC is
796 not only simply correlated to the formal oxidation states of
797 the ions but also rather to their corresponding partial
798 charges, as modulated by charge-transfer interactions with
799 coordinated ligands.

800 Conclusions

801 A wide range of computational methods has been devel-
802 oped during the last few decades to model electrostatic
803 interactions in complex molecular systems, including the
804 influence of polarization effects on structural refinement
805 and the analysis of structure/function relations. However,
806 rigorous and practical approaches to obtain ab initio quality
807 electrostatic potentials have yet to be established. Several
808 techniques have emerged within the field of linear scaling
809 methods, based on the fragmentation of extended system
810 into small molecular domains and the subsequent calcula-
811 tion of the properties of the entire system from accurate
812 quantum chemistry calculations of the constituent frag-
813 ments. The recently developed MoD-QM/MM hybrid
814 method belongs to this general class of linear scaling
815 approaches, and is based on a simple space-domain
816 decomposition scheme and an iterative self-consistent
817 treatment of the constituent molecular fragments according
818 to state-of-the-art QM/MM methods. The resulting meth-
819 odology complements previous approaches (Dapprich et al.
820 1999) and extends the capabilities of current QM/MM
821 hybrid methods to the description of electrostatic interac-
822 tions that explicitly include mutual polarization effects in
823 both the QM and the MM layers.

824 The moving-domain quantum mechanics/molecular
825 mechanics (MoD-QM/MM) structural models have been

used as reference configurations in conjunction with the
development and application of SC-QM/MM structural-
refinement schemes. The SC-QM/MM refinement method-
ology allowed for the assessment of MoD-QM/MM
molecular structures of similar energy (e.g., within
 ~ 5 kcal mol $^{-1}$ from the minimum energy configuration) in
terms of direct comparisons with high-resolution spectro-
scopic data. The method has been developed and imple-
mented in terms of simulations of EXAFS spectra and the
minimization of χ^2 -deviations between simulated and
experimental spectra by adjustment of nuclear coordinates,
subject to the constraint of minimum displacements relative
to the reference MOD-QMM minimum energy configura-
tion. The same methodology could be implemented, more
generally, with other types of spectroscopic data (e.g.,
FTIR, NMR) whenever the experimental data are available
so long as the spectroscopic properties can be readily
computed. Such SC-QM/MM refinement methods are par-
ticularly valuable for the analysis of complex systems with
rugged potential-energy landscapes, where many configu-
rations have similar energy, including metalloproteins for
which the most accurate and reliable information on the
structure of prosthetic metal clusters comes from high-res-
olution spectroscopy.

The reviewed applications addressed fundamental
questions that for many years have remained largely elu-
sive to rigorous first-principles examinations, showing that
QM/MM hybrid methods where polarization effects are
explicitly considered are powerful tools of modern com-
putational chemistry when applied in conjunction with
high-resolution spectroscopy. It is, therefore, expected that
these QM/MM techniques will continue to make many
more important contributions in the characterization of
structure and reactivity of biological macromolecules.

Acknowledgements V. S. Batista acknowledges a generous allo-
cation of DOE supercomputer time from the National Energy
Research Scientific Computing Center (NERSC) and financial support
from the Grants NSF ECCS # 0404191, NSF CHE # 0345984, DOE
DE-FG02-07ER15909, NIH 2R01-GM043278 and the US-Israel BSF
R08164. G. W. Brudvig acknowledges support from NIH GM32715.

References

- Ankudinov AL, Ravel B, Rehr JJ, Conradson SD (1998) Real-space
multiple-scattering calculation and interpretation of X-ray-
absorption near-edge structure. *Phys Rev B* 58:7565–7576
- Ankudinov AL, Bouldin CE, Rehr JJ, Sims J, Hung H (2002) Parallel
calculation of electron multiple scattering using Lanczos algo-
rithms. *Phys Rev B* 65:104–107
- Ashley CA, Doniach S (1975) Theory of extended X-ray absorption-
edge fine-structure (EXAFS) in crystalline solids. *Phys Rev B*
11:1279–1288
- Bouldin C, Sims J, Hung H, Rehr JJ, Ankudinov AL (2001) Rapid
calculation of X-ray absorption near edge structure using parallel
computation. *X-Ray Spectrom* 30:431–434

- 879 Britt RD, Peloquin JM, Campbell KA (2000) Pulsed and parallel-
880 polarization EPR characterization of the photosystem II oxygen-
881 evolving complex. *Annu Rev Biophys Biomol Struct* 29:463–
882 495
- 883 Britt RD, Campbell KA, Peloquin JM, Gilchrist ML, Aznar CP, Dicus
884 MM, Robblee J, Messinger J (2004) Recent pulsed EPR studies
885 of the photosystem II oxygen-evolving complex: implications as
886 to water oxidation mechanisms. *Biochim Biophys Acta*
887 1655:158–171
- 888 Broyden CG (1970) Convergence of single-rank quasi-Newton
889 methods. *Math Comput* 24:365–382
- 890 Brudvig GW, Crabtree RH (1986) Mechanism for photosynthetic O-2
891 evolution. *Proc Natl Acad Sci USA* 83:4586–4588
- 892 Cooper SR, Calvin M (1977) Mixed-valence interactions in di- μ -oxo
893 bridged manganese complexes. *J Am Chem Soc* 99:6623–6630
- 894 Cooper SR, Dismukes GC, Klein MP, Calvin M (1978) Mixed-
895 valence interactions in di- μ -oxo bridged manganese complexes:
896 electron-paramagnetic resonance and magnetic-susceptibility
897 studies. *J Am Chem Soc* 100:7248–7252
- 898 Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson
899 DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995)
900 A second generation force-field for the simulation of proteins,
901 nucleic-acids, and organic-molecules. *J Am Chem Soc*
902 117:5179–5197
- 903 Dapprich S, Komaromi I, Byun KS, Morokuma K, Frisch MJ (1999)
904 A new ONIOM implementation in Gaussian98. Part I. The
905 calculation of energies, gradients, vibrational frequencies and
906 electric field derivatives. *THEOCHEM J Mol Struct* 461:1–21
- 907 Dau H, Iuzzolino L, Dittmer J (2001) The tetra-manganese complex
908 of photosystem II during its redox cycle: X-ray absorption results
909 and mechanistic implications. *Biochim Biophys Acta* 1503:24–
910 39
- 911 Dau H, Liebisch P, Haumann M (2004) The structure of the
912 manganese complex of photosystem II in its dark-stable S-1-
913 state-EXAFS results in relation to recent crystallographic data.
914 *Phys Chem Chem Phys* 6:4781–4792
- 915 Derose VJ, Mukerji I, Latimer MJ, Yachandra VK, Sauer K, Klein
916 MP (1994) Comparison of the manganese oxygen-evolving
917 complex in photosystem-II of Spinach and *Synechococcus* Sp
918 with multinuclear manganese model compounds by X-ray-
919 absorption spectroscopy. *J Am Chem Soc* 116:5239–5249
- 920 Dykstra CE (1993) Electrostatic interaction potentials in molecular-
921 force fields. *Chem Rev* 93:2339–2353
- 922 Fan LY, Ziegler T (1991) The influence of self-consistency on
923 nonlocal density functional calculations. *J Chem Phys* 94:6057–
924 6063
- 925 Ferreira KN, Iverson TM, Maghlaoui K, Barber J, Iwata S (2004)
926 Architecture of the photosynthetic oxygen-evolving center.
927 *Science* 303:1831–1838
- 928 Flank AM, Weininger M, Mortenson LE, Cramer SP (1986) Single-
929 crystal EXAFS of nitrogenase. *J Am Chem Soc* 108:1049–1055
- 930 Fletcher R (1970) A new approach to variable metric algorithms.
931 *Comput J* 13:317–322
- 932 Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA,
933 Cheeseman JR, Montgomery Jr JA, Vreven T, Kudin KN, Burant
934 JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B,
935 Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada
936 M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M,
937 Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox
938 JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J,
939 Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R,
940 Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA,
941 Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S,
942 Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD,
943 Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG,
944 Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A,
Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham
MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW,
Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004)
Gaussian 03, Revision B.04. Gaussian, Inc., Wallingford, CT
- Gascon JA, Leung SSF, Batista ER, Batista VS (2006a) A self-
consistent space-domain decomposition method for QM/MM
computations of protein electrostatic potentials. *J Chem Theory
Comput* 2:175–186
- Gascon JA, Sproviero EM, Batista VS (2006b) Computational studies
of the primary phototransduction event in visual rhodopsin. *Acc
Chem Res* 39:184–193
- Goldfarb D (1970) A family of variable-metric methods derived by
variational means. *Math Comput* 24:23–26
- Haumann M, Junge W (1999) Evidence for impaired hydrogen-
bonding of tyrosine Y-Z in calcium-depleted photosystem II.
Biochim Biophys Acta 1411:121–133
- Haumann M, Muller C, Liebisch P, Iuzzolino L, Dittmer J, Grabolle
M, Neisius T, Meyer-Klaucke W, Dau H (2005) Structural and
oxidation state changes of the photosystem II manganese
complex in four transitions of the water oxidation cycle (S0 -
>S1, S1 - >S2, S2 - >S3, and S3, S4 - >S0) characterized by X-
ray absorption spectroscopy at 20 K and room temperature.
Biochemistry 44:1894–1908
- Hendry G, Wydrzynski T (2002) The two substrate-water molecules
are already bound to the oxygen-evolving complex in the S-2
state of photosystem II. *Biochemistry* 41:13328–13334
- Hendry G, Wydrzynski T (2003) O-18 isotope exchange measure-
ments reveal that calcium is involved in the binding of one
substrate-water molecule to the oxygen-evolving complex in
photosystem II. *Biochemistry* 42:6209–6217
- Hillier W, Wydrzynski T (2004) Substrate water interactions within
the photosystem II oxygen evolving complex. *Phys Chem Chem
Phys* 6:4882–4889
- Hillier W, Wydrzynski T (2008) O-18-Water exchange in photosys-
tem II: substrate binding and intermediates of the water splitting
cycle. *Coord Chem Rev* 252:306–317
- Hoganson CW, Babcock GT (1997) A metalloradical mechanism for
the generation of oxygen from water in photosynthesis. *Science*
277:1953–1956
- Jaguar (2003) Jaguar 5.5. Schrodinger LLC, Portland, OR
- Kronig RD (1931) The quantum theory of dispersion in metallic
conductors II. In: *Proceedings of the Royal Society of London
Series A Containing Papers of a Mathematical and Physical
Character*, vol 133, pp 255–265
- Kronig RD, Penney WG (1931) Quantum mechanics of electrons on
crystal lattices. In: *Proceedings of the Royal Society of London
Series A Containing Papers of a Mathematical and Physical
Character*, vol 130, pp 499–513
- Lee PA, Pendry JB (1975) Theory of the extended X-ray absorption
fine structure. *Phys Rev B* 11:2795–2811
- Limburg J, Vrettos JS, Liable-Sands LM, Rheingold AL, Crabtree
RH, Brudvig GW (1999) A functional model for O–O bond
formation by the O₂-evolving complex in photosystem II.
Science 283:1524–1527
- Loll B, Kern J, Saenger W, Zouni A, Biesiadka J (2005) Towards
complete cofactor arrangement in the 3.0 Å resolution structure
of photosystem II. *Nature* 438:1040–1044
- Lundberg M, Siegbahn PEM (2004) Theoretical investigations of
structure and mechanism of the oxygen-evolving complex in
PSII. *Phys Chem Chem Phys* 6:4772–4780
- Lundberg M, Blomberg MRA, Siegbahn PEM (2003) Modeling water
exchange on monomeric and dimeric Mn centers. *Theor Chem
Acc* 110:130–143
- Manchanda R, Brudvig GW, Degala S, Crabtree RH (1994) Improved
syntheses and structure of (Mn(III)Mn(IV))(O)(2)(Phen)(4)
(ClO₄)(3)·2CH₃COOH·2H₂O. *Inorg Chem* 33:5157–5160

- 1011 Maseras F, Morokuma K (1995) IMOMM: a new integrated ab initio
1012 plus molecular mechanics geometry optimization scheme of
1013 equilibrium structures and transition-states. *J Comput Chem* 16:
1014 1170–1179
- 1015 McEvoy JP, Brudvig GW (2004) Structure-based mechanism of
1016 photosynthetic water oxidation. *Phys Chem Chem Phys* 6:4754–
1017 4763
- 1018 Menikarachi LC, Gascon JA (2008) Optimization of cutting
1019 schemes for the evaluation of molecular electrostatic potentials
1020 in proteins via moving-domain QM/MM. *J Mol Model* 14:479–
1021 487
- 1022 Messenger J (2004) Evaluation of different mechanistic proposals for
1023 water oxidation in photosynthesis on the basis of Mn_4O_4Ca
1024 structures for the catalytic site and spectroscopic data. *Phys*
1025 *Chem Chem Phys* 6:4764–4771
- 1026 Messenger J, Badger M, Wydrzynski T (1995) Detection of one
1027 slowly exchanging substrate water molecule in the S-3 state of
1028 photosystem-II. *Proc Natl Acad Sci USA* 92:3209–3213
- 1029 Newcomer MB, Ragain CM, Gascón JA, Batista ER, Strobel SA,
1030 Loria JP, Batista VS (2009) A self-consistent MOD-QM/MM
1031 structural refinement method: characterization of hydrogen
1032 bonding in the oxytricha nova G-quadruplex. *J Chem Theory*
1033 *Comput* (in press)
- 1034 Noodleman L (1981) Valence bond description of anti-ferromagnetic
1035 coupling in transition-metal dimers. *J Chem Phys* 74:5737–5743
- 1036 Noodleman L, Case DA (1992) Density functional theory of spin
1037 polarization and spin coupling in iron–sulfur clusters. *Adv Inorg*
1038 *Chem* 38:423–470
- 1039 Noodleman L, Davidson ER (1986) Ligand spin polarization and
1040 antiferromagnetic coupling in transition-metal dimers. *Chem*
1041 *Phys* 109:131–143
- 1042 Noodleman L, Peng CY, Case DA, Mouesca JM (1995) Orbital
1043 interactions, electron delocalization and spin coupling in iron–
1044 sulfur clusters. *Coord Chem Rev* 144:199–244
- 1045 Nugent JHA, Rich AM, Evans MCW (2001) Photosynthetic water
1046 oxidation: towards a mechanism. *Biochim Biophys Acta* 1503:
1047 138–146
- 1048 Pantazis DA, Orio M, Petrenko T, Zein S, Bill E, Lubitz W,
1049 Messenger J, Neese F (2009) A new quantum chemical approach
1050 to the magnetic properties of oligonuclear transition-metal
1051 complexes: application to a model for the tetranuclear manga-
1052 nese cluster of photosystem II. *Chem Eur J* 15:5108–5123
- 1053 Pecoraro VL, Baldwin MJ, Gelasco A (1994) Interaction of manga-
1054 nese with dioxygen and its reduced derivatives. *Chem Rev*
1055 94:807–826
- 1056 Robblee JH, Cinco RM, Yachandra VK (2001) X-ray spectroscopy-
1057 based structure of the Mn cluster and mechanism of photosyn-
1058 thetic oxygen evolution. *Biochim Biophys Acta* 1503:7–23
- 1059 Ruettinger WF, Ho DM, Dismukes GC (1999) Protonation and
1060 dehydration reactions of the $Mn_4O_4L_6$ cubane and synthesis and
1061 crystal structure of the oxidized cubane $Mn_4O_4L_6$ (+): a model
1062 for the photosynthetic water oxidizing complex. *Inorg Chem* 38:
1063 1036–1037
- 1064 Sarneski JE, Thorp HH, Brudvig GW, Crabtree RH, Schulte GK
1065 (1990) Assembly of high-valent oxomanganese clusters in
1066 aqueous solution. Redox equilibrium of water-stable $Mn_3O_4^{4+}$
1067 and $Mn_2O_2^{3+}$ complexes. *J Am Chem Soc* 112:7255–7260
- 1068 Sayers DE, Stern EA, Lytle FW (1971) New technique for
1069 investigating noncrystalline structures: Fourier analysis of
1070 extended X-ray absorption fine structure. *Phys Rev Lett* 27:
1071 1204–1207
- 1072 Schlegel HB (1987) Optimization of equilibrium geometries and
1073 transition structures. *Adv Chem Phys* 67:249–285
- 1074 Schlodder E, Witt HT (1999) Stoichiometry of proton release from
1075 the catalytic center in photosynthetic water oxidation:
reexamination by a glass electrode study at pH 5.5–7.2. *J Biol*
Chem 274:30387–30392
- Scott RA, Hahn JE, Doniach S, Freeman HC, Hodgson KO (1982)
Polarized X-ray absorption-spectra of oriented plastocyanin
single-crystals: investigation of methionine copper coordination.
J Am Chem Soc 104:5364–5369
- Shanno DF (1970) Conditioning of quasi-Newton methods for
function minimization. *Math Comput* 24:647–656
- Soderhjelm P, Ryde U (2006) Combined computational and crystal-
lographic study of the oxidised states of NiFe hydrogenase.
THEOCHEM J Mol Struct 770:199–219
- Sproviero EM, Gascon JA, McEvoy JP, Brudvig GW, Batista VS
(2006a) Characterization of synthetic oxomanganese complexes
and the inorganic core of the O_2 -evolving complex in photosys-
tem II: evaluation of the DFT/B3LYP level of theory. *J Inorg*
Biochem 100:786–800
- Sproviero EM, Gascon JA, McEvoy JP, Brudvig GW, Batista VS
(2006b) QM/MM models of the O_2 -evolving complex of
photosystem II. *J Chem Theory Comput* 2:1119–1134
- Sproviero EM, Gascon JA, McEvoy JP, Brudvig GW, Batista VS
(2006c) QM/MM models of the O_2 -evolving complex of
photosystem II. *J Chem Theory Comput* 2:1119–1134
- Sproviero EM, Gascon JA, McEvoy JP, Brudvig GW, Batista VS
(2007) Structural models of the oxygen-evolving complex of
photosystem II. *Curr Opin Struct Biol* 17:173–180
- Sproviero EM, Brudvig GW, Batista VS (2008a) A P-EXAFS
molecular refinement protocol. *J Phys Chem* (in press)
- Sproviero EM, Gascon JA, McEvoy JP, Brudvig GW, Batista VS
(2008b) Computational studies of the O_2 -evolving complex of
photosystem II and biomimetic oxomanganese complexes.
Coord Chem Rev 252:395–415
- Sproviero EM, Gascon JA, McEvoy JP, Brudvig GW, Batista VS
(2008c) A model of the oxygen evolving center of photosystem
II predicted by structural refinement based on EXAFS simula-
tions. *J Am Chem Soc* 130:6728–6730
- Sproviero EM, Gascon JA, McEvoy JP, Brudvig GW, Batista VS
(2008d) QM/MM study of the catalytic cycle of water splitting in
photosystem II. *J Am Chem Soc* 130:3428–3442
- Sproviero EM, Shinopoulos K, Gascon JA, McEvoy JP, Brudvig GW,
Batista VS (2008e) QM/MM computational studies of substrate
water binding to the oxygen evolving complex of Photosystem
II. *Phil Trans R Soc Lond B* 363:1149–1156
- Stebler M, Ludi A, Burgi HB (1986) (Phen) $2Mn-4(Mu-O)2Mn-3$ (Phen) 2 (PF $_6$) $_3CH_3CN$ and (Phen) $2Mn-4(Mu-O)2Mn-4$ (Phen) 2
(ClO $_4$) $_4CH_3CN$ (Phen = 1,10-Phenanthroline): crystal-structure
analyses at 100-K, interpretation of disorder, and optical,
magnetic, and electrochemical results. *Inorg Chem* 25:4743–4750
- Stern EA (1974) Theory of extended X-ray-absorption fine-structure.
Phys Rev B 10:3027–3037
- Thorp HH, Brudvig GW (1991) The physical inorganic-chemistry of
manganese relevant to photosynthetic oxygen evolution. *New J*
Chem 15:479–490
- Tsutsui Y, Wasada H, Funahashi S (1999) Reaction mechanism of
water exchange on di- and trivalent cations of the first transition
series and structural stability of seven-coordinate species. *J Mol*
Struct THEOCHEM 462:379–390
- Versluis L, Ziegler T (1988) The determination of molecular-
structures by density functional theory: the evaluation of
analytical energy gradients by numerical-integration. *J Chem*
Phys 88:322–328
- Vreven T, Morokuma K (2000a) On the application of the IMOMO
(integrated molecular orbital plus molecular orbital) method. *J*
Comput Chem 21:1419–1432
- Vreven T, Morokuma K (2000b) The ONIOM (our own N-layered
integrated molecular orbital plus molecular mechanics) method

- 1141 for the first singlet excited (S-1) state photoisomerization path of
 1142 a retinal protonated Schiff base. *J Chem Phys* 113:2969–2975
- 1143 Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions:
 1144 dielectric, electrostatic and steric stabilization of carbonium-ion
 1145 in reaction of lysozyme. *J Mol Biol* 103:227–249
- 1146 Yachandra VK (2002) Structure of the manganese complex in
 1147 photosystem II: insights from X-ray spectroscopy. *Phil Trans R*
 1148 *Soc Lond Ser B* 357:1347–1357
- 1149 Yachandra VK, Sauer K, Klein MP (1996) Manganese cluster in
 1150 photosynthesis: where plants oxidize water to dioxygen. *Chem*
 1151 *Rev* 96:2927–2950
- 1152 Yano J, Pushkar Y, Glatzel P, Lewis A, Sauer K, Messinger J,
 1153 Bergmann U, Yachandra V (2005) High-resolution Mn EXAFS
 of the oxygen-evolving complex in photosystem II: structural
 implications for the Mn₄Ca cluster. *J Am Chem Soc* 127:14974–
 14975
- Yano J, Kern J, Sauer K, Latimer MJ, Pushkar Y, Biesiadka J, Loll B,
 Saenger W, Messinger J, Zouni A, Yachandra VK (2006) Where
 water is oxidized to dioxygen: structure of the photosynthetic
 Mn₄Ca cluster. *Science* 314:821–825
- Yano J, Kern J, Pushkar Y, Sauer K, Glatzel P, Bergmann U,
 Messinger J, Zouni A, Yachandra VK (2008) High-resolution
 structure of the photosynthetic Mn₄Ca catalyst from X-ray
 spectroscopy. *Phil Trans R Soc Lond B* 363:1139–1147