Dynamic Hydrogen-Bonded Networks of Photosystem II

Federico Guerra

2017

A mio padre, che non perse mai curiosità e stupore per la meravigiosa
complessità del mondo.

To my father, who never lost a sense of wonder and curiosity about the
astonishing complexity of the world.

Abstract

Photosystem II is a large, membrane-embedded, protein complex that uses cofactor molecules and energy of absorbed light to split water molecules into electrons, molecular oxygen and protons. The protons, generated at the reaction site, are then transported to the lumen, covering distances larger than 20 Å across the protein interior. The identity of titrable amino acid groups that could participate in proton transfer, and the dynamics of water-mediated hydrogen-bonded networks that could serve as possible proton transfer pathways, are fundamental open questions.

As a first step to address these questions, we performed all-atom molecular dynamics simulations of wild type and mutant photosystem II embedded in a hydrated lipid bilayer. To this aim we first derived a new set of CHARMM force field parameters for three cofactors harbored in photosystem II interior: chlorophyll-a, pheophytin-a and plastoquinone-9. These parameters proved to accurately describe cofactors dynamics both in gas phase and in the protein matrix.

To facilitate efficient data analysis of hydrogen-bonded networks, we developed a data analysis tool for fast tracking of hydrogen-bonds. We detected, in wild type photosystem II, two water-mediated hydrogen-bonded networks connecting the manganese cluster to the lumen. These networks appear to be drastically perturbed by changes in protonation states and by single site mutation.

One of the networks we identified ends at the surface of the PsbO subunit. Even though this extrinsic protein subunit has been largely studied, it remained unclear the reason why the oxygen evolution rate drops when PsbO is removed from photosystem II. We performed simulations of photosystem II in the absence of PsbO and we observed that the region surrounding the manganese cluster has higher protein flexibility compared to the wild type, and an increased hydration level that associates with the release to the bulk of a chloride ion known to be required for oxygen evolution.

Zusammenfassung

Photosystem II ist ein großer integraler Membranproteinkomplex, der Kofaktormoleküle und Lichtenergie nutzt, um Wassermoleküle in Elektronen, molekularen Sauerstoff und Protonen zu spalten. Nachdem die Protonen im Reaktionszentrum erzeugt wurden, werden sie über bis zu 20 Å durch das Innere des Proteins zum Lumen transportiert. Die Identität der titrierbaren Aminosäuren, die am Protonentransfer beteiligt sein könnten, sowie die Dynamik der wassergesteuerten Wasserstoffbrückenbindungsnetzwerke (HBN) sind elementare bislang ungelöste Fragen.

Als ersten Schritt, um diesen Fragen nachzugehen, führten wir Ganzatom-Molekulardynamiksimulationen von Wildtyp- und mutiertem Photosystem II in einer hydrierten Lipidmembran durch. Zu diesem Zwecke leiteten wir einen neuen Satz von CHARMM-Kraftfeldparametern für drei Kofaktoren her, die im Inneren von Photosystem II zu finden sind: Chlorophyll a, Phäophytin a und Plastochinon-9. Die neuen Parameter haben bewiesen, dass sie in der Lage dazu sind, die Dynamik der Kofaktoren sowohl in der Gasphase als auch in der Proteinmatrix exakt zu beschreiben.

Wir entwickelten ein Datenanalysewerkzeug zum schnellen Verfolgen von Wasserstoffbrückenbindungen, um die effiziente Datenanalyse von HBN zu erleichtern. Im Wildtyp von Photosystem II machten wir zwei wassergesteuerte HBN ausfindig, die den Mangan-Komplex mit dem Lumen verbinden. Änderungen der Protonierungszustände sowie Mutationen scheinen diese Netzwerke stark zu beeinträchtigen.

Eines der von uns identifizierten Netzwerke endet an der Oberfläche der PsbO Untereinheit. Obwohl diese extrinsische Protein-Untereinheit intensiv untersucht wurde, bleibt es weiterhin unklar, warum die Sauerstoffgewinnungsrate sinkt, wenn PsbO von Photosystem II entfernt wird. Wir führten Simulationen ohne PsbO durch und konnten beobachten, dass der Bereich, der den Mangan-Komplex umgibt, im Vergleich zum Wildtyp eine größere Proteinflexibilität hat. Außerdem konnten wir eine erhöhte Hydration feststellen, die damit in Zusammenhang gebracht werden konnte, dass ein Chloridion, welches zur Sauerstoffgewinnung erforderlich ist, aus dem Proteininneren freigegeben wurde.

Publications

The present thesis is based on the following manuscripts, which have been published in peer-reviewed journals:

Federico Guerra, Suliman Adam, and Ana-Nicoleta Bondar. "Revised force-field parameters for chlorophyll-a, pheophytin-a and plastoquinone-9". In: Journal of Molecular Graphics and Modelling 58, 2015, pp. 30–39.
https://doi.org/10.1016/j.jmgm.2015.03.001

Federico Guerra, Malte Siemers, Christopher Mielack, and Ana-Nicoleta Bondar. "Dynamics of Long-Distance Hydrogen-Bond Networks in Photosystem II". In: Journal of Physical Chemistry B, 122 (17), 2018, pp. 4625-4641.
https://pubs.acs.org/doi/10.1021/acs.jpcb.8b00649

The following paper has been published during the time of my Ph.D as a side project on proton-coupled protein and water dynamics in molecular transporters:

Federico Guerra and Ana-Nicoleta Bondar. "Dynamics of the Plasma Membrane Proton Pump". In: The Journal of Membrane Biology 248 (3), 2015, pp. 443–453.
https://doi.org/10.1007/s00232-014-9732-2

Acknowledgements

Abbreviations list

CHARMM	chemistry at harvard macromolecular mechanics)
Chl-a	chlorophyll-a
DGDG	digalactosyl-diacyl-glycerol
H-bond	hydrogen-bond
HBN	hydrogen-bonded network
HF	Hartree-Fock
LJ	Lennard-Jones
MCCE	multi-conformer continuum electrostatics
MD	molecular dynamics
MGDG	distearoyl-monogalactosyl-diclyceride
MM	molecular mechanics
NVE	constant Number of atoms, constant Volume and Energy
NPT	constant Number of atoms, constant Pressure and Temperature
OEC	oxygen evolving complex
PDB	protein data bank
PG	dipalmitoyl-phosphatidyl- glycerol
Phe-a	pheophytin-a
Pl-9	plastoquinone-9
POPC	palmitololeoyl phosphatidylcholine
PSII	photosystem II
QM	quantum mechanics
SQDG	sulfoquinovosyl-diacylglycerols

List of Figures

List of Tables

xiii

Contents

Introduction

1.1 The photosystem II protein complex

Photosystem II (PSII) is a large protein complex embedded in the thylakoid membranes of plants, algae and cyanobacteria. This enzyme uses the energy of absorbed light to split water molecules into protons, electrons and molecular oxygen:

$$2H_2O + 4h\nu \rightarrow O_2 + 4e^- + 4H^+_{lumen} \tag{1.1}$$
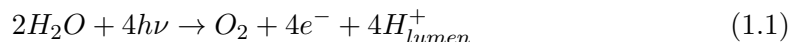
O_2 is then released in the environment whereas the other reaction products serve to advance the photosynthetic process. The electrons are used to reduce two plastoquinone molecules to plastoquinol and the protons are released on the luminal side of the thylakoid membrane generating an electrochemical gradient. The plastoquinol molecules and the generated transmembrane potential are then used to synthesize nicotinamide adenine dinucleotide phosphate (NADPH) and adenosine triphosphate (ATP), which are both needed for CO_2 fixation [7].

The water splitting reaction in PSII is initiated by the absorption of visible light by the many chlorophylls pigments harbored in the light-harvesting protein subunits: CP47 and CP43, Figure 1.1. The absorbed energy is then transfered to the reaction center, composed of the D1 and D2 subunits (Figure 1.1). Here is located a special pair of chlorophylls, P680, which acts as an exciton trap becoming a strong reducing agent upon excitation (P680*). P680* reduces a pheophytin (Phe) molecule, which again reduces in sequence two plastoquinone (PQ) molecules first PQ_A and then PQ_B. The oxidized P680 has a strong redox potential($> 1V$) and oxidizes a tyrosine (Y_Z) amino acid residue in the D1 subunit. The neutral radical $Y_Z^{\bullet+}$ extract an electron, and probably a proton, from the manganese cluster (Mn_4CaO_5) that binds the two substrate water molecules [7]. The work of Joliot and Kok [47, 57] revealed that four flashes of light are required to complete a catalytic cycle which was interpreted as
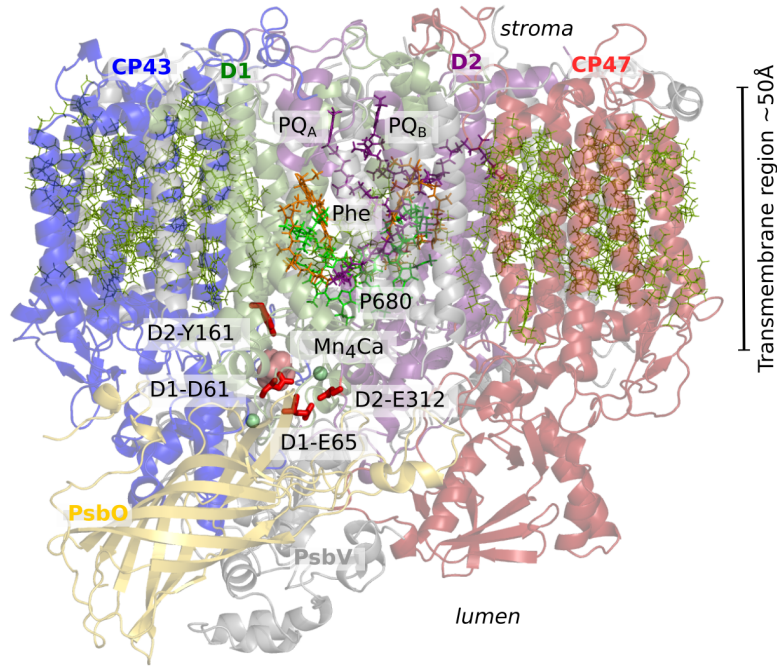
Figure 1.1: *Photosystem II and its components [37]. A molecular graphics representation of the PSII structure based on the coordinates of the 1.9 Å resolution crystal structure [115]. The PSII largest subunit are highlighted in different colors: D1 (green), D2 (purple), CP43 (blue), CP47 (red) and PsbO (yellow). Cofactors involved in the electron transfer process and amino acid residues possibly involved in proton transport are depicted using a licorice representation.*

consisting of five steps, Figure 1.2. The four light flashes produce the same amount of manganese cluster oxidizing equivalents which are used to oxidize the two substrate water molecules to form dioxigen. Finally, at each light flash, one extracted electron is transfered to PQ_A and PQ_B, following the path described above, forming doubly reduced PQs, which are then protonated to plastoquinol (PQH_2) and released into the lipid bilayer [7].

1.1.1 Open questions addressed in this work

Even though, in the last years, many progress have been made in understanding the mechanism regulating the catalytic cycle in PSII, as extensively reviewed in Refs. [7, 6, 123, 63, 23, 24], several questions are still to be answered. For instance, it is well known that during one catalytic cycle four protons are extracted by two substrate water molecules but it is yet not clear how these protons are then transfered from the
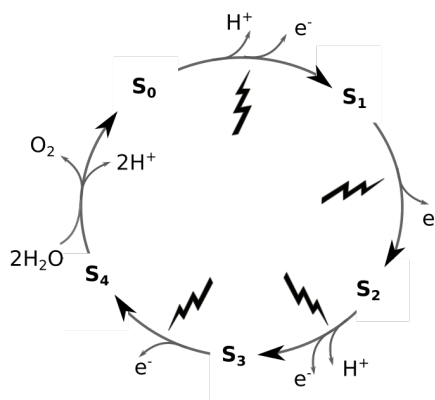
Figure 1.2: *A schematic representation of the S-state cycle. Upon absorption of four light photons (depicted as flashes) two water molecules are split in electrons, protons and molecular oxygen following a consecutive series of five intermediates: S_0, S_1, S_2, S_3 and S_4. During this cycle protons are released to the lumen and electrons are transfered to the plastoquinones (PQ_A , PQ_B).*

manganese cluster site to the lumen.

Proton transfer in biomolecules is a complicated process that involves titratable amino acid groups connected by movable water molecules [108]. Proton translocation across protein interior occurs in many enzymes [111]. For example, computer simulations of the influenza virus M2 channel [104], of the cytochrome c oxydase [61] and of the carbonic anhydrase [13] revealed proton conducting channels, constituted by chains of hydrogen-bonded (H-bonded) water molecules (water wires) and intermediate protonable sites (proton donors and acceptors). In these channels, proton do not diffuse as a localized charge but move from donor to acceptor groups along water wires via Grotthus mechanism [108, 1, 75, 56].

In this work we will refer to pairs of amino acid groups connected via water wires as water-mediated H-bonded amino acid groups.

In PSII, the identification of proton channels and the characterization of their dynamics are still open questions.

As a first step towards answering this question, we performed all atoms classical Molecular Dynamics (MD) simulations of PSII embedded in a hydrated lipid bilayer and we developed a data analysis tool to detect water mediated H-bonded networks (HBN), leading from the manganese cluster to the lumen, that could serve as possible proton exit pathways.

Based on observations of crystal structures, it has been proposed that proton exit pathways could end at the PsbO subunit surface, [82, 11] (the PsbO subunit is shown in

yellow in Figure 1.1). PsbO is the only extrinsic subunit of PSII present in all the oxygenic photosynthetic organisms and its removal shows to drastically alter the oxygen evolution rate of PSII [17]. Here, we simulated PSII in absence of PsbO to get insights in how the removal of this protein subunit perturbs PSII dynamics.
In this work all atom classical MD simulations have been performed using the CHARMM force field [70]. To set up PSII for MD simulations, it has been first necessary to derive new force field parameters to describe PSII a number of cofactors harbored in the protein complex interior.

In the next paragraphs, I will introduce what is known about proton transfer in PSII and about the role of the PsbO subunit and which CHARMM force field parameters do already exist for PSII cofactors. In the result and discussion section, I will present the contributions of this work to the field.

1.2 Proton pathways to the lumen in photosystem II

During one reaction cycle, PSII produces 4 protons, 4 electrons and 1 dioxygen molecule. Experimental investigations established the proton release sequence from the manganese cluster in the four S-state transitions, Figure 1.2, ($S_1 \rightarrow S_2$, $S_2 \rightarrow S_3$, $S_3 \rightarrow [S_4] \rightarrow S_0$ and $S_0 \rightarrow S_1$) to be around 0, 1, 2, 1 [59, 49]. Protons need to be transported to the protein luminal side in order to establish a trans-membrane proton gradient, which fuels the ATP-synthase for ATP production. For these reasons, the presence of pathways able to efficiently transport protons from the OEC to the lumen, covering a distance of more than 20 Å, have been proposed to exist in PSII [27, 82].

Even though, in the last years, great advancements have been achieved in the identification of these specialized pathways, it is still not fully clear how protons are transported to the lumen in PSII.

The first crystal structures of PSII from thermophilic cyanobacteria, resolved between the years 2000 and 2005 [128, 51, 27, 66], gave valuable information to identify channels within the protein complex. The 3.5 Å resolution crystal structure from Ferreira et al. [27] is the first providing not only information about cofactors relative position and orientation, but also details of their protein environments. Based on this structure and on the 3.0 [66] and 2.9 [38] Å resolution crystal structures, Murray and Barber [82] first, Ho and Styring [42] and Gabdulkhakov et al. [32] later, used search algorithms to find possible oxygen, water and proton exit pathways. Successively, the 1.9 Å resolution crystal structure [115] added more details on the structure of PSII and thus advancements in the identification of molecular oxygen, water and proton channels have been possible [11].

Murray and Barber, in the work published in Ref. [82], used the program CAVER [89] to locate channels (connected voids) that could possibly lead oxygen molecules and protons from the manganese cluster site to the lumen and water molecules in the opposite direction. They identified three channels (Table 1.1), of which, channel (i) is the most hydrophobic and thus assigned as being a possible oxygen exit pathways. The other two channels are both eligible as proton exit pathways.

Channel (ii) is about 35 Å long and it is, overall, polar. It starts at the C-terminal domain of the D1-subunit and continues at the large extrinsic loop of CP43. It is large enough to allow for water transport (channel radius larger than 1.4 Å), which does not exclude the possibility that it could work both as water and proton transport channel.

In contrast, channel (iii), Figure 1.3, is the narrowest and most polar, which make it more suitable for proton transport, i.e it is large enough to accommodate water molecules to form water wires connecting titratable groups, though, the function of water channel cannot be excluded. Channel (iii) involves residues of subunits D1 and D2 and reaches the surface of the PsbO subunit. The latter channel was postulated as being a proton exit pathway already by Ferreira et al. [27].

The 1.9 Å resolution crystal structure published by Umena et al. in 2011 [115] provided the position of more than 1000 water molecules per PSII monomer. This new piece of information is of paramount importance in the understanding of proton pathways. With the latter crystal structure it is, indeed, possible to upgrade the level of the pathways search from looking for void channels to compute H-bonding networks connecting titratable amino acid residues and water molecules. The authors proposed three HBNs as possible proton exit channels. The first network starts at the manganese cluster and continues at the interface of the D1, CP43 and PsbV subunits. Note that the first residues in this network is the Y_Z, one of the crucial residues in the oxygen evolving reaction, Figure 1.1. The other two networks lead to the PsbO and PsbU subunits passing through the Cl^-1 and Cl^-2 binding site respectively. The network leading to the PsbO surface involves the residues D1-D61, D1-E65 and D2-E312 showing a strong overlap with the previously discussed channel (iii) [82] and channels C,D from Ref. [32], as summarized in Table 1.1 and Figure 1.3. Bondar and Dau [11] further investigate the HBNs in the 1.9 Å crystal structure. In particular, they considered the D1-E65/D2-E312 pair as a possible proton-carrying carboxylate dyad, due to their close vicinity (2.6 Å). This dyad is part of a network connecting the manganese cluster to the PsbO surface (HBN3 from Ref. [11]), Table 1.1 and Figure 1.3.
Ogata et el. [87] and Vassieliev et al. [122], analyzed water permeated channels by means of MD simulations. This channels could serve as proton exit pathways and show great overlap with previously identified channels, Table 1.1.
A common feature to all of these studies is the presence of the amino acid residues D1-D61, D1-E65 and D2-E312 in highly hydrophilic channels commonly assigned as proton channels, [82, 42, 32, 115, 11, 122, 87]. The latter amino acid groups have been implicated in proton transport both in experimental and theoretical works[106, 99, 100]. Sproviero et al. investigated proton transfer in the manganese cluster region by means of QM simulations and proposed D1-D61 to be linked to the manganese cluster via a water molecule and to be the first element of a proton exit pathway [106].

The role of D1-D61, D1-E65 and D2-E312 and of other amino acid groups in their vicinity, has been experimentally investigated by Service et al. [99, 100] analyzing

Publication Ref.	Channel name as in Ref.	Channel type assigned by the authors	Titrable amino acid residues in the channel
Murray and Barber [82]	channel (i)	oxygen channel	D1-E189, D1-H190, D1-342
	channel (ii)	water channel	D1-E329, CP43-E83, CP43-E354, CP43-E413
	channel (iii)	proton channel	D1-D61, D1-D59,,D2-E312, D1-E65, PsbO-D222, PsbO-D224, PsbO-H228, PsbO-E229
Ho and Styring [42]	Narrow	proton/water channel	D1-D61, D1-E333
	Broad	proton/water channel	D1-D61, D1-E65, D1-Y161, D1-S169, D1-D170, D1-H332, D1-E333, D2-E312
	Large	water channel	D1-E329, D1-H332, D1-D342, CP43-E354, CP43-T412, CP43-E413
Gabulkhakov et al. [32]	A1, A2	oxygen channel	D1-D342, D1-H190, CP43-E221, CP43-S226, CP43-D360
	B1,B2	water channel	D1-E329, D1-D342, CP43-E413, CP43-H398, CP43-E83
	C, D	proton channel	D1-E333,D1-D61, D1-E65, D2-E312, D2-E310
	E, F	proton channel	D1-E333, CP43-E354, D1-S169, D1-D61, CP43-T335, PsbO-S192, PsbU-D126
	G	proton channel	D1-E333, D1-H332, D2-T316, D2-E323, D2-E302, D2-Y296, CP47-E364, CP47-S365
Umena et al. [115]	Figure 3b	proton channel	D1-Y161, D1-H190, D1-D319, D1-H304
	Figure 4b	proton channel	D1-E333, D1-D61, D1-E65, D2-E312, D2-E323,CP47-E364
	Figure 4c	proton channel	CP43-E354, D1-H337, PsbU-D96
Bondar and Dau [11]	HBN1	proton channel	D1-E189, D1-Y161, D1-H190, D1-D319, D2-E343, PsbV-Y136
	HBN2	proton channel	D1-D61, D1-D65, D2-E312, D2-Y313, D2-Y316
	HBN3	proton channel	D1-D61, D1-D65, D2-E312, D2-E310, D2-D308, D2-E307, PsbO-224
	HBN4	proton channel	D1-E65,D1-Y107, PsbO-Y107, D2-E312, D2-E310
Ogata et al.[87]	Path1	water channel	D1-Y161, D1-H190, D1-D319, D1-H304
	Path2-1 Path2-2	proton channel	D1-E333, D2-E323, D1-D61, D1-E65, D1-E302
	Path3	water channel	CP43-E354,D1-H337, PsbU-D96
Vassiliev et al. [122]	channel 1	water channel	D1-D65, D1-D61, D1-H332, D2-E312, D2-E310,PsbO-D224
	channel 2	water channel	D1-D61, D1-H337, PsbU-D96
	channel 3	water channel	D1-E333, D2-E323,CP47-S365
	channel 4	water channel	D1-E329, CP43-E413

Table 1.1: *Oxygen, water and proton channels proposed in the literature. Per each channel, a list of titratable amino acid residues contained in the channel is provided*

$S_{n+1} - S_n$ FTIR difference spectra. In the D1-D61A, D1-E65A, D2-E312A and D1-R334A mutants the efficiency of $S_2 \rightarrow S_3$ transition appears to be lower than in the wild type. Moreover, the $S_0 - S_3$ difference spectra applied to the latter mutants were practically devoid of features, this was interpreted either as a sensibly lower $S_3 \rightarrow S_0$ transition or as a fail of the PSII core complex to advance beyond the S_3 state. To further investigate the role of mutated amino acid residues, the authors placed special focus on a characteristic band at $1747cm^{-1}$. This frequency is associated with carboxylate groups $C = O$ double bond vibrations, thus, with changes in the protonation state or in the environment around a carboxylate group. In the wild type a negative peak is present at this frequency in the $S_2 - S_1$ difference spectra. This negative band either disappears or decreases sensibly its intensity in the D1-E65A, D1-E329Q, D2-E312A and D1-R334A mutants. However the D1-D61A mutation had no effect. These results were interpreted as indication that D1-E65, D1-E329, D2-E312 and D1-R334 are part of an extended HBN which includes an as yet unidentified carboxylate group (could be one of these four amino acid residues or a new one) that is responsible for the $1747cm^{-1}$ negative band in the wild type. Secondly, the disruption of this HBN via mutations could block the transmission of the perturbation at the unidentified carboxylate residue site, which is responsible for the negative band. Note that D1-E65 and D1-E329 are about 20 Å apart. For this reason water molecules are probably part of the HBN considered above. As a partial prove of that, they observed the same elimination of the $1747cm^{-1}$ negative band in the $S_2 - S_1$ difference spectra under partial dehydration [65]. Finally, they measured that upon D1-D61A, D1-E65A, D1-E329Q, D2-E312A and D1-R334A mutations, the O_2 evolving activity of the PSII core complexes drops respectively to the ~ 17, ~ 14, ~ 65, ~ 27 and $\sim 40\%$ of the wild type PSII. The central role of D1-R334 was confirmed by mutational studies conducted by Li and Burnap, where D1-R334E and D1-R334V mutants showed reduced O_2 evolution activity [64].

FTIR difference spectroscopy has also been employed by Nakamura et al. [84] to investigate the possibility of the Y_Z being involved in the proton transfer pathway. When Y_Z is oxidized by $P680^+$, it becomes a neutral $Y_Z^\bullet$ radical by releasing its phenolic proton. Measuring the $Y_Z - Y_Z^\bullet$ difference spectra, Nakamura et al. were able to associated a band to the $N - H$ stretching vibration of the protonated D1-His190, which is formed by accepting the phenolic proton from Y_Z and forms a strong H-bond with $Y_Z^\bullet$.
From the shape of the band it was also possible to attribute large polarizability to the proton at the $N_\tau - H$ of D1-His190, which implies that there are good chances for this

proton to hop to an H-bonded water molecule and to be subsequently transfered to the lumen via water wires by Grotthus mechanism [85]. This proton transfer scheme would be in good agreement with the one previously proposed by Klauss et al. [54] and with the proton exit pathway proposed by Umena et al. [115].
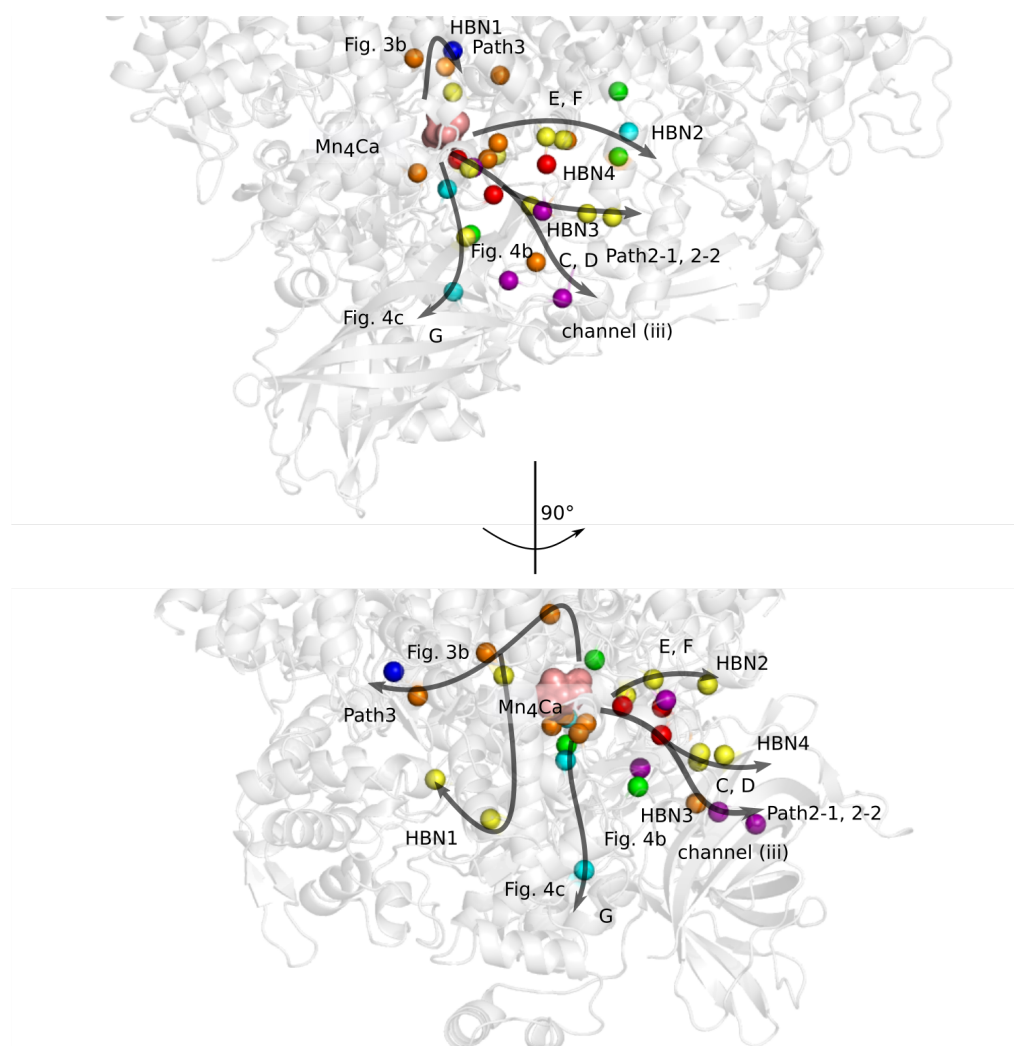
Amino acid groups common to all the proposed pathways
Amino acid groups common to at least two of the proposed pathways
Murray and Barber
Gabuldkakhov et al.
Umena et al
Bondar and Dau
Ogata et al.

Figure 1.3: *Proposed proton exit channels. Molecular graphics of PSII luminal side based on the coordinates of the 1.9 Å resolution crystal structure [115]. Arrows represent proposed proton exit channels and are named as in the respective publications (listed in Table 1.1). Colored spheres represent C_α atoms of titratable amino acid residues that constitute proposed proton exit channels. Amino acid groups common to all the proposed channels are colored red and amino acid groups common to at least two proposed proton transfer pathways are colored orange. Amino acid groups belonging to one single proton exit path have a different color per each Refs: Murray and Barber [82] (purple), Gabdulkhakov et al. [32] (green), Umena et al. [115] (cyan), Bondar and Dau [11] (yellow), Ogata et al. [87] (blue).*

10

1.3 Role of the PsbO subunit

The 26.5 KDa PsbO subunit is the largest extrinsic subunit of PSII. Located at the luminal side, PsbO is the only extrinsic subunit of PSII conserved in all plant, bacterial and algae organisms. The original name given to PsbO was: manganese-stabilizing protein. Under physiological condition, it was shown that the removal of PsbO destabilizes the manganese cluster. In particular, two of the four Mn ions are reduced and eventually lost [79]. It is possible to prevent the loss of two Mn ions using high concentrations of Ca^{2+} and Cl^- [14]. Nevertheless, even preventing the manganese cluster destabilization, the PsbO removal shows great impact on the oxygen evolution process.

Several studies have been conducted to characterize this effect. Flash oxygen yield experiments showed a retarded kinetics for O_2 release, related to a slower $S_2 \to S_3$ transition [80]. Moreover, it was possible to reveal, via a thermoluminescence study, that the $S_3 \to [S_4] \to S_0$ transition was inhibited after PsbO depletion [80]. The rate of oxygen evolution of the ΔPsbO varies in dependence of the Cl^- concentration and it can drop to about the 35% of the wild type [17].

PsbO binds to PSII during the last stages of the complex formation [95]. It mainly interacts with the D1 and D2 domains and it has been shown that mutations of Arg152 and Arg162 drastically affect PsbO binding affinity in *thermophilic S. elongatus* [81]. These arginine amino acid residues are only two of many charged amino acid groups distributed on PsbO surface. In particular a high number of carboxylate residues contribute to an overall negative charge of the PsbO, i.e. considering all amino acid residues in their standard protonation state, *T. vulcanus* PSbO has a total charge of -10 [68]. Even though not many amino acid residues appear to be conserved among PsbO subunits in different organisms, the total negative charge of the subunit appear to be a common feature in all the oxygenic photosynthetic organisms [116]. Due to its highly negatively charged surface, it has been proposed that PsbO could work as proton antenna [102, 10] - namely: a cluster of carboxylate groups located close to each other, which can attract and trap protons [97].
Finally, it has been observed that Cl^- ions are not efficiently retained in the vicinity of the manganese cluster in absence of the PsbO subunit [92, 19]. Even though the role of Cl^- ions is not fully understood, they appear to play a pivotal role in the oxygen evolution process.

1.3.1 Effect of Cl^- ions depletion on photosystem II function

The 1.9 Å crystal structure showed for the first time the location of two chloride ions in the second coordination shell of the manganeses cluster, Figure 1.4.
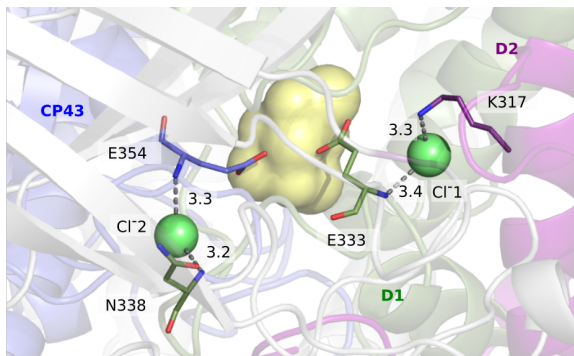
Figure 1.4: *Location of Cl^-1 and Cl^-2 in the 1.9 Å resolution crystal structure [37].*

The presence of two anion binding sites in the vicinity of the reaction center was previously revealed by Kawakami and coworkers [52]. In their experiments, they substituted Cl^- with Br^- and I^-. While PSII activity is not strongly affected by Br^- substitution, the oxygen evolution is completely abolished by replacing Cl^- with I^-. Flash-induced UV absorbance experiments revealed the importance of chloride cofactors along the oxygen evolutions steps [127]. The absence of Cl^- does not affect the UV absorbance spectra in the $S_1 \rightarrow S_2$ transition. On the other hand, the $S_2 \rightarrow S_3$ transition appears to be blocked under Cl^- depletion. [127]. Molecular Dynamics and Monte Carlo simulations gave us insights about the possible molecular mechanism that induces this strong effects on the reaction cycle upon removal of Cl^-. Rivalta and coworkers [94] conducted combined MD and MC simulations of the region surrounding the manganese cluster with and without Cl^-1. They observed that the depletion of Cl^-1 induces the formation of D1-D61/D2-K317 salt bridge, stabilizing a configuration in which D1-D61 cannot participate in the H-bonding network leading protons from the manganese cluster to the lumen. On the contrary, in presence of Cl^-1, D1-D61 side chain is free to form transient H-bonds with water molecules. The same effect of Cl^-1 removal was observed in multi-conformer continuum electrostatics (MCCE) [105] calculations performed by Amin and coworkers [3]. These calculation also aimed to compute the reduction potential of the $S_1 \rightarrow S_2$ and $S_2 \rightarrow S_3$ transitions. Both potentials show to increase in absence of Cl^-1. The authors explain this result with the impossibility of transferring protons to the lumen due to the D1-D61/D2-K317 salt

bridged formation.

In this work we investigated the possible role of the PsbO protein subunit by means of all-atom MD simulations. By simulating PSII in the absence of PsbO and comparing this simulation with the wild type PSII, we aimed to observe how does the removal of PsbO perturb the rest of the protein complex. In particular we were interested in understanding how does PsbO contribute to stabilize the region surrounding the manganese cluster and how does the absence of PsbO correlate with lower Cl^- binding affinity.

1.4 CHARMM force field parameters

All the simulations presented in this work were performed using the CHARMM empirical force field (Chemical at HARvard Macromolecular Mechanics) [16, 15, 70]. CHARMM is an all atoms, Class I additive force field [60] widely used to perform macromolecular simulations.

Equation 1.2 shows the potential energy function used in CHARMM:

$$
\begin{aligned}
U(\vec{R}) = {} & \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 \\
& + \sum_{\text{Urey-Bradley}} K_{UB}(S - S_0)^2 \\
& + \sum_{\text{dihedrals}} K_\varphi(1 + \cos(n\varphi - \delta)) + \sum_{\text{impropers}} K_\omega(\omega - \omega_0)^2 \\
& + \sum_{\text{non-bonded pairs}} \left\{ \varepsilon_{ij}^{min} \left[\left(\frac{R_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^{min}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon r_{ij}} \right\} \\
& + \sum_{\text{residues}} U_{\text{CMAP}}(\varphi, \psi)
\end{aligned}
\tag{1.2}
$$

The potential energy, $U(\vec{R})$, is a sum over individual terms defining atom pairs bonded and non-bonded interactions as a function of their coordinates. Bonded terms include bond stretching vibrations (b), valence angle bending motions (θ), Urey-Bradley coupling term for the 1-3 interactions (S), dihedral angle torsions (φ) with the backbone torsional correction (CMAP) [71, 72] and the improper angles bending (ω). The b_0, θ_0, S_0 and ω_0 parameters indicate equilibrium positions and the K parameters are the respective force constants. All the bonded terms are harmonic potentials, with the only exception of the dihedral angle torsion which is described by a sinusoidal function, where n is the periodicity of the angle and δ the phase shift. Non bonded terms include Coulombic interactions between fixed point charges and Lennard-Jones (LJ) potential used to describe van der Walls interactions.

In order to run simulations we need parameters to describe each bonded and non bonded interaction in our system. A general set up for a simulation of a membrane protein requires the following set of parameters:

- CHARMM36 protein parameters [9]

- CHARMM36 lipid parameters [53]

- TIP3P water model [48]

- Ions parameters [8]

- parameters for cofactors molecules

Parameters for cofactors molecules have to be derived in many cases and added to force field. In particular PSII harbors a large number of cofactor molecules. The most recent crystal structure [115] revealed the presence of more than 70 cofactor molecules for each PSII monomer. In this work we revised and derived CHARMM force field parameters for chlorophyll-*a* (Chl-*a*), pheophytin-*a* (Phe-*a*) and plastoquinone-9 (Pl-9), Figure 3.1.

Chl-*a* molecules ($C_{55}H_{72}MgN_4O_2$) count for about the half of the total number of cofactor molecules in PSII, 35 Chl-*a* per monomer. These chromophores are located within the transmembrane region of the protein complex (mostly within the light harvesting subunits) where they participate in energy transfer and electron transfer reactions [18, 86].
Two Phe-*a* molecules are located in the protein reaction center and are involved in electron transfer reactions [2, 55]. Phe-*a* is structurally similar to Chl-*a*, but the magnesium ion at the center of the chlorin ring is replaced by two hydrogen atoms covalently bound to the nitrogen atoms of the chlorin ring ($C_{55}H_{74}N_4O_2$, Figure 3.5). In the reaction center are also located two copies of Pl-9 ($C_{53}H_{80}O_2$), which are involved in proton and electron transfer reactions [98].
These three cofactor molecules tightly interact with elements of the protein matrix. It is, thus necessary to have set of parameters able to well characterize these interactions and the intrinsic dynamics of Chl-*a*, Phe-*a* and Pl-9.
Foloppe et al. [30, 29] developed CHARMM force-field parameters for analogs of bacteriochlorophyll and bacteriopheophytin. Bonded parameters were assembled by analogy with the CHARMM parametrization of heme [5] and with high-resolution crystal structure data. Partial atomic charges were derived via Mulliken population analysis using the semi-empirical AM1 (Austin Model 1 [45]) or PM3 (Parametrized Model number 3 [107]) for bacteriopheophytin and bacteriochlorophyll, respectively. Another set of partial charges have been derived for bacteriochlorophyll-*b* and bacteripheophytin-*b* using PM3 and a CHELPG-type (CHarges from ELectrostatic Potential using Grid based method) population analysis [93]. More recently, partial atomic charges for bacteriochlorophyll were computed performing electrostatic potential (ESP) calculations using the 6-31** basis set [21] and by means of DFT calculations [44].
The atomic partial charges reported in Refs. [93, 21, 44] are all the result of *ab initio*

or semi empirical calculations, and not of a standard CHARMM parametrization procedure [121]. Thus, these parameters are not consistent with the CHARMM additive force field. In particular, the aforementioned sets of atomic partial charges show great difference in the partial charge assigned to the magnesium (Mg) ion at the center of Chl-a chlorin ring and to its coordinating nitrogen atoms.

The Mg ion has a charge of $0.10e$ in the PM3 computations on bacteriochlorophyll-b [93], $0.93e$ lower than the charge obtained in the ESP computation on the bacteriochlorophyll-a from Ref. [21], and $0.822e$ lower than in Ref. [44]; values for nitrogen atoms range from $-0.535e$ to $-0.367e$ in the DFT calculations perfomed in Ref. [44], as compared to $-0.73e$ to $0.28e$ in Ref [21]. A good description of the atomic partial charges is necessary to well characterize cofactor molecules dynamics and their interactions with the protein environment. In this work we derived a new set of atomic partial charges for Chl-a, Phe-a and Pl-9, following the standard CHARMM procedure [121], largely discussed in the method section of this thesis.

Methods and methodological developments

2.1 Deriving CHARMM force field parameters

As a first step towards setting up photosystem II for MD simulation, we derived CHARMM force field parameters of a number of PSII cofactors molecules.

Even though the procedure to derive parameters for MM (classical Molecular Mechanics) simulations varies with the force field used, the parametrization philosophy is common to every force field: perform MM calculations and tune the force field parameters with the aim of reproducing *ab initio* quantum mechanical (QM) or experimental data [125, 121]. On the other hand, CHARMM is an additive all atoms force field, where transferability of parameters is guaranteed via consistency of the parametrization protocol.

For this reason the parametrization protocol applied in this work follows the standard algorithm published in Ref. [121]. We focused our attention on partial charges optimization and refinement of equilibrium bonds and angle positions parameters. In the next sections I will present the algorithm used and the cofactors' structural models employed in our parametrization .

2.1.1 Structural models of the cofactors

We extracted the initial coordinates of chlorophyll-a (Chl-a), pheophytin-a (Phe-a) and plastoquinone-9 (Pl-9) cofactor molecules from the crystal structure [115] of photosystem II, PDB ID: 3ARC, chain A. The PDB residue number were 604, 609 and 709 for Chl-a, Phe-a and Pl-9 molecules respectively. For computational efficiency, we parametrized the target molecules core obtained by removing all side chains. The latter were then replaced with hydrogen atoms using the software Avogadro version 1.0.3 [40] to locate their initial coordinates. The procedure of dissecting the target molecule is common in case of parametrization of large molecules, as in the case of the parametrization of the heme prosthetic group of *cytochrome c* [5].

2.1.2   Partial charges optimization

The CHARMM force field partial charge optimization is performed reproducing water-target molecule interaction energies computed at QM level at each hydrogen bonding site of the molecule. This calculations were performed in gas phase using the TIP3P water model and the Hartree-Fock (HF) with the 6-31G(d) QM method [20] . Even though higher level of QM theory could lead to more accurate interaction energies, the usage of the HF method is necessary to maintain compatibility with the CHARMM additive force field. In case of target molecules with zero net charges, as Chl-*a*, Phe-*a* and Pl-9, the QM dipole moment serves as further guideline for partial charges optimization. The force field dipole moment should overestimate the QM gas phase dipole moment by 20 to 50% to be relevant for the bulk phase [121]. In this work we performed all the QM calculations using Gaussian 09 [31].

For each target molecule, we first performed a QM geometry optimization of the core structure using the MP2 method [20] with the 6-31G(d) basis set and default convergence criteria. Thereafter, we used the Force Field Toolkit of VMD [43, 77] to generate Gaussian input files for HF/6-31G* computations of the interaction energies between the MP2-optimized target molecule core and water in TIP3P geometry. The HF/6-31G(d) geometry optimizations of the target molecule core/water complexes were performed independently for each hydrogen-bond donor and acceptor atoms (see Table 3.1 and Figure 2.1). For the oxygen atom acceptor site of the Chl-*a* and Phe-*a* cores we calculated water interaction energies for the linear geometry, and for an interaction pose whereby the water OH bond interacts with a lone electron pair of the acceptor oxygen atom (Figure 2.1). The interaction energy was computed by subtracting the energy of the target molecule and the energy of a water molecule in the TIP3P configuration to the energy of the system composed by the target molecule and the interacting water molecule:

$$\Delta E = E_{target\ molecule\ +\ water} - (E_{target\ molecule} + E_{water}) \tag{2.1}$$

The lowest-energy value of the interaction energy from the HF/6-31G(d) computations, scaled by 1.16 factor [73], is the target QM interaction energy, ΔE_{QM}. The optimized interaction distance is labeled R_{QM}. To describe bulk properties and for compatibility with CHARMM, in the case of neutral polar compounds one scales the HF water interaction energies by 1.16 [73].

We then optimized the partial charges by fitting the CHARMM computed water-
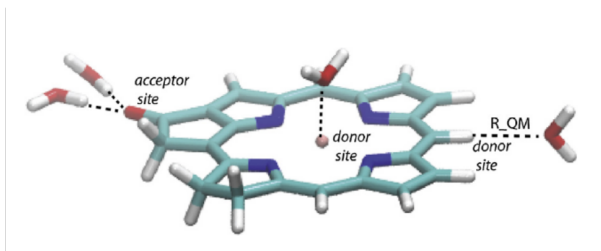
Figure 2.1: *Illustration of interaction geometries from HF/6-31G(d) computations of water interaction energies of the Chl-a core [35]. The geometry of the Chl-a core is optimized with MP2/6-31G(d), and the water molecule has TIP3P geometry [48]. The location of the each water molecule relative to the Chl-a core corresponds to the lowest-energy value for ΔE_{QM}. Computations were performed separately for each interaction site. The dotted lines indicate the directions along which the water molecule is constrained during the energy optimization. For the oxygen acceptor site we show the linear interaction pose, and a lone pair interaction. R_{QM} is the interaction distance.*

target molecule interaction energies, ΔE_{MM}, to ΔE_{QM}. The MP2-optimized geometry was used to compute a first estimate of the partial charges by performing a Merz-Kollmann population analysis [103]. To compute the CHARMM water interaction energies between the target molecule and water we made use of tools from the Force Field Development web page of the MacKerell Lab, and of the CGenFF protocol from Ref. [121]. For each interaction site, we thus performed an energy optimization of the complex by using as a degree of freedom the distance between the water oxygen atom and the target molecule; we varied this distance in steps of 0.01 Å, from 1.6 Å to 4.5 Å. The minimum energy value is the MM interaction energy, ΔE_{MM}, and the corresponding interaction distance is R MM. In the next step, we compared ΔE_{MM} and ΔE_{MM}. If ΔE_{MM} and ΔE_{QM} differed by more than 0.2 kcal/mol, we iteratively adjusted the partial charges and recalculated ΔE_{MM}. In adjusting the partial charges we considered local symmetry and the total charge of the molecule, and imposed for the partial charges of the methyl hydrogen and carbon atoms the standard CHARMM values. Finally, we compared the QM with the MM dipole moment.

2.1.3 Refinement of equilibrium positions

In our MM computations on the Chl-*a* core we used for starting bonded parameters values from Refs. [21, 30]. We adjusted several reference bond lengths and reference valence angles to better match the MP2-optimized geometries of the target molecule. In CGenFF, the accepted criteria for the deviations between the

QM and MM values for bond lengths and valence angles are 0.03 Å and 3°, re-

spectively [121]. MM geometry optimizations were performed to a root-mean squared energy gradient of 10^{-5} kcal/mol $\cdot$ Å. We used an infinite cutoff and set ϵ to 1.0.

The Phe-*a* core contains two more hydrogen atoms relative to the Chl-*a* core: HNB and HND (Figure 3.5). For the Phe-*a* core we thus used the same bonded parameters as for the Chl-*a* core, to which we added bonded parameters for HNB and HND. In the case of the Pl-9 core, all bonded parameters needed for the MM computations were taken from CGenFF and used without further changes.

2.2 Setting up MD simulations of photosystem II embedded in a lipid bilayer

We produced equilibrated trajectories of a full PSII protein complex monomer, embedded in a hydrated POPC lipid bilayer. The setup includes all the 19 protein subunits resolved in the 1.9 Å crystal structure [115] and all the protein cofactors molecules. The unit cell counts about $\sim 600K$ atoms. We run simulations of PSII in monomeric form, and not in the dimeric configuration in which it assembles in crystals [115, 38]. The main reason for this choice is related to the high computational effort necessary to compute a large system as PSII. Furthermore it is known that PSII works as monomer and its natural supra-molecular organization is still a matter of debate [126, 112]

2.2.1 Coordinates file preparation

Work done in collaboration with Christopher Mielack.

We used as starting coordinates chain A of the crystal structure of *T. Vulcanus* PSII at 1.9 Å resolution [115] (PDB ID: 3WU2).

This structure contains three lipid molecules presenting missing heavy atoms: one DGDG, one PG and one SQDG molecule. The DGDG lipid molecule (PDB resid: DGD 406) shows nine missing heavy atoms from the galactosyl group (PDB atom names: 2E C3E C4E C5E C6E O2E O3E O4E O5E). To rebuilt the missing heavy atoms we extracted the coordinates of a full resolved DGDG lipid molecule from the crystal structure (PDB resid: DGD 102). We then used the VMD fit function to compute a transformation matrix with which we could transfer the coordinates of the fully resolved lipid, DGD 102, at the position of the unresolved one, DGD 406, overlapping a selection of atoms present in both the molecules (PDB atom names: O6D C1D C2D C3D C4D C5D C6D). At this point we saved the translated coordinates of the missing atoms selection of DGD 102 in the DGD 406 coordinates files. We choose the DGD 102 molecule for this procedure because among the fully resolved DGDG lipids is the one that, upon translation, offers a reconstructed DGD 406 with the lowest overlap with atoms from the surrounding protein matrix, minimizing in this way possible strong steric interactions. The same procedure was applied to the remaining two incomplete lipids (PDB resid: LHG 410, SQD 407).

Although the structure by Guskov et al. [38] (PDB entry: 4V62) is at a lower resolution

of 2.9 Å, it does contain cofactors not resolved in the Umena et al. structure. To add these cofactors to our structure we superposed the two structures and copied the coordinates from the missing cofactors. Following this procedure we added a betacarotene molecule at the PsbJ subunit (XBCR 102, in 4V62: BCR 102), an SQDG molecule at the CP47-D2 subunits interface (SQD 409) and a PG lipid molecule (LHG 800, in 4V62: LHG 521) at the plastoquinone pocket site, which in both the crystal structures shows a water-accessible cavity. In order to have a hydrophobic plastoquinone pocket in our model, we filled the cavity with LHG lipid molecules. We took the coordinates from LHG 521 from the Guskov et al crystal structure [38] and add two more spatially shifted copies to fill the opening. We finally computed an energy minimization of the newly introduced lipids within a constrained protein matrix to optimize their steric interactions.

Finally we removed all the detergent molecules and the following residues that are located at the monomer-monomer interface and that in our set up would then be located at the protein-lipid bilayer interface: BCR 103, 620; CLA 602, DGD 406, SQD 103, 418, RRX.

The last step towards having a full set of coordinates for the PSII protein complex was to build the hydrogen atoms and we did that by using the CHARMM hbuild command. We considered all the titratable amino acid residues in their standard but PsbO-Asp102 [11]. Change in protonation state of the D1-Glu65/D2-Glu312 pair are discussed in section 2.2.3

Unit cell set up

The OPM server [67] was used to compute the correct orientation of the PSII monomer, with bound cofactors and special lipids, relative to the membrane normal. The membrane aligned apo-protein structure was then uploaded to the web-based software CHARMM-GUI [46, 62] in order to construct the simulation unit cell.

The apo-protein was inserted in a squared lipid bilayer with 200 Å size, composed of 910 POPC molecules. About 135000 water molecules were then added to solvate the system maintaining the water layer's thickness of at least 20 Å on the top and on the bottom of the protein. Finally the total charge of the unit cell was neutralize by placing 43 potassium ions following using a Montecarlo approach. The total number of atoms in the simulation system is $\sim$580000.

2.2.2   Force field description

We used the CHARMM [16] force field parameters for protein amino acid residues [69] with the CMAP correction [72], POPC lipids [53] and the TIP3P model to describe water molecules [48]. For chlorophyll-a, pheophytin-a and plastoquinone-9, we employed parameters previously derived by us [35], for a detailed description of how these parameters have been derived, go to section 3.1.

To describe the Chlorophyll-a coordination bonds with histidine and asparagine side chains we used the heam-histidine patch from Ref. [5], the same patch was applied twice for the heam doubly coordinated by histidines. We used the ParamChem web software [118] to assign force-field parameters for special lipids of photosystem II (DGDG, PG, MGDG, and SQDG), and for the bicarbonate and betacarotene molecules. To maintain the structural integrity of the complex composed of bicarbonate, iron, and four histidines moieties at the stroma side of PSII, we modeled coordination bonds by introducing bonds with associated force constants of 100 $kcal/mol \cdot \text{Å}^2$ between iron, the bicarbonate carboxylate oxygen atoms, and the histidine $N\epsilon$ atoms. To impose these constrains during our simulations we used the *extraBonds* NAMD function.

To preserve the internal geometry of the manganese cluster, we introduced bond and angle parameters for all the elements of the cluster. We assigned a value of 500 $kcal/mol \cdot \text{Å}^2$ to both bonds and angles force constants. For what concerns equilibrium position parameters, we relied on the coordinates form the 'radiation damage free' structure obtained by Suga et al. [110] by means of femtosecond X-ray pulses technique. Based on this structure, the authors tentatively assigned the following valences of the manganese atoms: Mn1(III), Mn2(IV), Mn3(IV) and Mn4(III). We assigned this valences as partial charges for the manganese ions and a charge of -2 e to oxygen atoms and $+2$ e to the calcium ion. The total charge of our manganese cluster model adds up to $+6$ e and due to the -6 e charge contribution from the amino acid residues coordinating the cluster, we have a neutral manganese cluster coordination pocket. Our model of the manganese cluster does not allow for geometrical variations along the simulation, due to the high force constants defined for bonds and angles. It will behave as a rigid body, which can be translated and rotated in space.

2.2.3   Simulations performed

We performed five simulations of a total length of 148 ns each, summarized in Table 2.1. Three simulations of the wild type PSII in which we changed the protonation state of the D1-Glu65/D2-Glu312 pair (Sim1-3); one single site mutation simulation (Sim4) and one simulation in which we depleted the PsbO subunit (Sim5). To prepare

Sim4 we made the D1-Glu65Ala mutation by using the 1.9 Å crystal structure [115] and replacing the side chain of D1-Glu65 by Ala. For setting up Sim5, we used the same protocol described in section 2.2.1, but we removed the coordinates of the PsbO subunit from the PDB file that we then feed to CHARMM-GUI.

Simulation	Description	*length (ns)*
Sim1	D1-Glu65 protonated	148
Sim2	D1-Glu312 protonated	148
Sim3	D1-Glu65/D2-Glu312 unprotonated	148
Sim4	D1-Glu65Ala mutant	148
Sim5	PsbO depleted	148

Table 2.1: *List of all simulation performed.*

2.2.4   Molecular dynamics simulations

The protocol described below was used to perform independently each one of our simulations. We computed all-atom molecular dynamics simulations using NAMD [50, 90] in the NPT ensemble ($T = 300K, P = 1\,bar$). We used a Langevin dynamics scheme and a Nose-Hoover Langevin piston [26, 76] to simulate constant pressure dynamics. We computed non bonding interactions using the smooth particle mesh Ewald summation for the Coulomb interactions [22, 25] and a switch distance with cutoff at 12 Å for the Van der Walls interactions. The non bonded interactions were computed with the reversible multiple time step algorithm [34, 114], i.e. short range every 2 steps and long range every 4 steps. All bonds involving hydrogen atoms were constrained using the SHAKE algorithm [96]. We first minimized the system and then, starting from the minimized coordinates, we performed an equilibration phase. The latter involved a slow release of weak harmonic constraints following this scheme: during the first 50 ps we imposed harmonic constraints of 3 $kcal/mol \cdot Å^2$ to the protein's backbone and to the cofactors and of 2 $kcal/mol \cdot Å^2$ to protein sidechains, during the next 100 ps we assigned constraints of 1.5 3 $kcal/mol \cdot Å^2$ to the protein's backbone and the cofactors and switched of constraints for sidechains. Finally we run 1 ns without harmonic constraints. During the whole equilibration we used the constant temperature control algorithm and we switched off the multiple time step algorithm. In what follows we will not consider the equilibration phase as part of the production run.

2.3 Developing new analysis tools for trajectories exploration

Due to the large size of the systems studied (the simulation unit cell counts $\sim$ 600K atoms), it was necessary to develop efficient data analysis tools able to characterize PSII properties along the time of our simulations. Our focus was placed on HBNs detection. In particular we developed two tools: one devoted to track stable direct HBNs, i.e. amino acid residues side-chain/side-chain or side-chain/backbone, and one capable to identify water-mediated HBNs of titratable amino acids residues. We developed these tools using the python programming language and the DCD-reader and universe object from the MDAnalysis library [33, 78] to get atoms coordinates. For the computations we used the numpy [124] and scipy [88] packages.

In the next sections I will explain in greater details the rational of the algorithms employed.

2.3.1 Using graph theory for H-bonded networks analysis

Work done in collaboration with Malte Siemers.

The 1.9 Å crystal structure of PSII [115] contains 2791 amino acid residues which are interconnected via 744 side-chain/side-chain or side-chain/backbone H-bonds. In this work, two amino acid residues are considered to be H-bonded to each other if the heavy atom-heavy atom distance (distance between donor and acceptor atom) is below 3.5 Å and if the angle defined by donor, hydrogen and acceptor atoms ranges between 120° and 180° [113]. The distance search is done employing a KDtree algorithm [74] (scipy.spatial.cKDTree), which scales as $\mathcal{O}(n \log(n))$, where n is the number of acceptor and donor atoms

At each trajectory frame the aforementioned distance and angle geometrical criterion is tested for all the possible pair of H-bond acceptor and donor in the system, excluding backbone-backbone H-bonds. If the condition are met for a specific pair of amino acid residues at a specific trajectory frame, then we consider the two residues to be H-bonded at that frame.

We are interested in stable H-bond. Thus, we consider an occupancy rate for each H-bonded pair of residues, which is computed as the percent of frames occupied by H-bonds along the analyzed trajectory. We set an occupancy rate cutoff of 60% to

define stable H-bonds.

Our data analysis tool is not simply able to compute each of the hundreds of H-bonds present in PSII in every trajectory frame and their occupancy rates, but it also stores the results in a way that highly facilitates the comparison between HBNs of different simulations.

We used the networkx python library [39] to generate graphs of stable H-bonds. Amino acid residues involved in H-bonds are the graph's nodes and stable H-bonds are the edges connecting them. This mathematical representation allow us to easily compare HBN of two distinct simulations, by evaluating their correspondent sets of edges. The intersection between the two sets contains H-bonds conserved in the two simulations,while the complementary sets contain H-bond present in one of the two simulations but not in the other. To better describe the complex H-bonded networks of PSII, we created a subsets of edges to differentiate between inter-subunit H-bond and intra-subunit H-bonds.

Finally, we generate a graphical output of this HBNs comparison, Figure A.3. The nodes are the amino acid residues involved in the networks, they are labeled by resid and colored depending on the subunit they belong. Black edges represent H-bonds conserved in the two simulations, red edges represent H-bonds present in simulation one but not in simulation two and the opposite is true for green edges. Thicker edges depict inter-subunit H-bonds. This graphical representation give us a clear picture of the difference in HBNs of two different simulation.

2.3.2 Detecting dynamic water-mediated H-bonded networks

Work done in collaboration with Malte Siemers.

We developed a tool to detect dynamic water-mediated HBNs in proteins. Part of the methods employed in the direct HBN analysis were also applied to the development of this tool. Nevertheless, in this case, we have do deal with a dynamic selection of water molecules and not only with a static selection of amino acid residues. Our system counts $\sim$130k water molecules, which access and leave, in the *ps* time scale, regions in the close vicinity of the protein, e.g. regions in which they could form water-mediated HBNs.

Our algorithm, schematically described in Figure 2.2 and 2.3, enters a loop over all frames in the dcd files. At each step, it first reads the coordinates for all atoms in the given protein residues selection (in this work we selected donor and acceptor atoms of titratable amino acid residues side chains that can participate in proton transfer: glutamate, aspartate, histidine, tyrosine, threonine and serine) as well as the coordinates of all oxygen belonging to water. Relying on this set of coordinates two kd-trees [74] (scipy.spatial.cKDTree) are build in order to find elements in H-bonding distance to each other. One contains all water molecules and the other contains the atoms of the selection. The water tree is cropped to contain only water molecules in a sphere around the selection in order to speed up the calculations. This sphere has its center at the center of coordinates of the selection and its radius is the distance from the center to the farthest atom in the selection plus 18 Å, the maximum length of a file of 5 H-bonded water molecules.

These two trees are then queried for three sets of pairs of atoms in H-Bond cut-off distance of 3.5 Å: a set of water pairs $\{w_i\}$, a set of protein residue pairs from the given selection, $\{s_j\}$, and a set containing water-residue pairs, $\{w_i, s_j\}$, Figure 2.2 (A). (Note that, as in the direct H-bonding algorithm, for all pairs the donor-hydrogen-acceptor angle condition is tested.)

For all water-water H-bonds an undirected, unweighted graph is build. The graph is a dictionary that contains for every node all its adjacent nodes. Every donor/acceptor that is hydrogen bonded to at least one water molecule is then sequentially added to the graph which is then queried for all shortest paths from this source to all possible targets, Figure 2.2 (B), with a maximal number of water molecules in the wire set to 5. All reachable targets are then referenced against the list of pairs of water-residue hydrogen bonds to find water-mediated H-bonds, Figure 2.2 (C). The source is removed afterwards. A final check is done to ensure no pair of residues connected by a water wire is connected by a direct HB.

All matches are then saved within a dictionary connecting the residue pair with the number of waters in the shortest path as well as a hash to identify a wire made of certain individual water molecules for further analysis.

Finally, we use an error function to smooth transitions from one specific water wire configuration to the next one. If a shortest paths between two residues is broken or increased in length the algorithm checks if the starting shortest path is present in a second graph that is built using larger cut-offs for H-bonding distances (3.8 Å) and angles (70°), which describes a weak H-bond [113]. This graph can only be queried for 10 frames in a row.
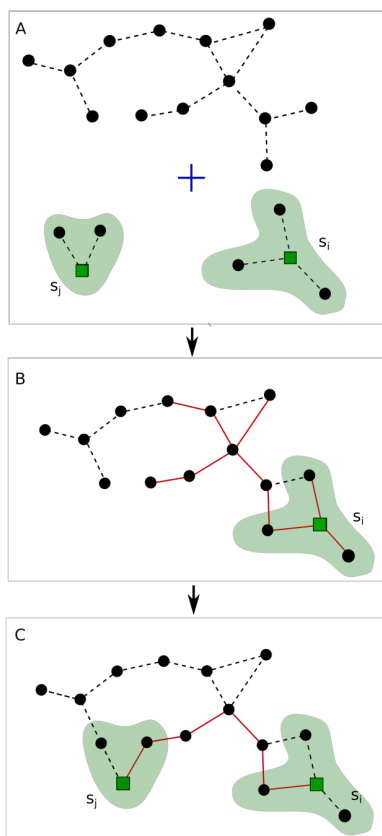
Figure 2.2: *Detecting dynamic water-mediated HBNs [37]. Graphical representation of the algorithm used to identify water wires. Black dots and green squares represent water molecules and side chains respectively. H-bonds are depicted as dotted lines. (A) top: graph of H-bonded water molecules; down: side chains $\{s_i\}$ and $\{s_j\}$ and relative H-bonded water molecules. (B) the graph is queried for shortest water wires (red lines paths) connected to the side chain s_i. (C) a water wire connecting s_i and s_j (red lines path) is detected.*

2.3.3 Data filtering procedure for water-mediated H-bonds

To describe water wires stability we did not use the same definition of occupancy rates as for the direct H-bonds. We imposed, in this case, a stronger condition for stability. We considered as occupied only those frames in which the water wire is exactly the same, e.g. same water index and same order in the wire, within an interval of minimum two consecutive frames. In Sim1-5 coordinates were printed every 10 *ps*, this implies that our filtering procedure selects only water wires that remain equal for at least 10 *ps*. More specifically, we are not able to determine what happens within the 10 *ps* interval, but we know if at the interval extremes the water wire connecting a specific pair of residues is identical or not. If it is identical we have good reasons to consider that the

① Input:
fixed selection (S) of
atoms from protein

② Identify a dynamic selection of water
molecules (W) within a radius (R) of
the center of geometry of S.

③ Compute all hydrogen bonds (HB) involving
elements of S {s$_i$} and elements of W {w$_i$}.

w$_i$ - w$_j$

w$_i$ - s$_i$

s$_i$ - s$_j$

④ Build a graph
with w$_i$ as nodes
and HB as edges

⑤ Generate a list (list 1)
of w$_i$ that are hydrogen
bonded to s$_i$

⑥ Generate a list
(list 2) of s$_i$-s$_j$
with direct HB

⑦ **Generate water wires:** One after the other, each s$_i$ is
added to the graph. All shortest paths starting at s$_i$ are
calculated. If the end of a path is in list 1, the path is
treated as the shortest path between the corresponding
residues s$_i$ and s$_j$. Then s$_i$ is removed from the graph.

⑧ A final check is done to be sure that s$_i$ and s$_j$ are not
directly HBN by compairing the pair with list 2.

Figure 2.3: *Schematic representation of the water-mediated H-bonded network
detection algorithm. Steps 2-8 are repeated for every frame in the trajectory of
the simulation*

wire either did not change within the interval or was subject to small and transient
changes.

2.3.4 Explore *fs* time scale water wires dynamics

Water molecules form and break H-bonds in the sub-picosecond time scale [58]. This fast dynamics do not emerge in Sim1-5 due to the fact that we used a coordinate saving step of 10 *ps* (because of data storage related issues). Nevertheless, we are interested in proving that what we computed in the *ps* time scale, holds true in the sub-picosecond time scale. For this reason we run a series of 100 *ps* long *NVE* (constant Number of atoms, constant Volume and Energy) simulations with a 10 *fs* coordinates writing step. We run these simulations starting from five equally spaced snapshots from the last 20 *ns* of Sim1 (at 128, 132, 136, 140 and 144 *ns*). We performed on these simulations the same water-mediated H-bonds analysis described above, but we computed water wires occupancy rates using new parameters for the filtering procedure. We considered as occupied just those frames belonging to interval of minimum size equal to 100 in which the water wire connecting two amino acid residues remained unchanged. This means that we now consider intervals of 1 *ps* (the reason for this choice is explained in the result chapter, Section 3.3.5). Finally we averaged the occupancy values computed in each simulation to get the final water wire occupancy rates.

Results and discussion

3.1 Deriving CHARMM force field parameters for photosystem II cofactors

Part of the results and material presented in this chapter has been published in Ref. [35].

We revised and derived new CHARMM force field parameters for the following PSII cofactors: chlorophyll-*a*, pheophytin-*a* and plastoquinone-*9*, Figure 3.1. This new set of parameters proved to well describe the geometry of the cofactors in vacuum and their non bonding interactions. Further tests on cofactors dynamics within the protein matrix confirms the reliability of our parameters to conduct MD simulation of the PSII protein complex.

Molecule	Atoms used as interaction sites	
Chl-a core	NA, NB, NC, ND, OBD, H2A1, H2A2, H2B, H2C, H2D, H3A1, H3A2, H3B, H3C, HBD1, HBD2, HBH, HCH, HDH, MG	Fig. 3.2
Phe-a core	NA, NB, NC, ND, OBD, H2A1, H2A2, H2B, H2C, H2D, H3A1, H3A2, H3B, H3C, HBD1, HBD2, HBH, HCH, HDH	Fig. 4A
Pl-9	O9, O10, H11, H12, H13, H15, H16, H18	Fig. 5A

Table 3.1: *Atoms selected for computations of water interaction energies [35].*

3.1.1 Parametrization of chlorophyll-*a*

In order to obtain a set of partial charges for Chl-*a* core we computed HF and CHARMM water/Chl-*a* core interaction energies for 21 different sites summarized in Table 3.1. We
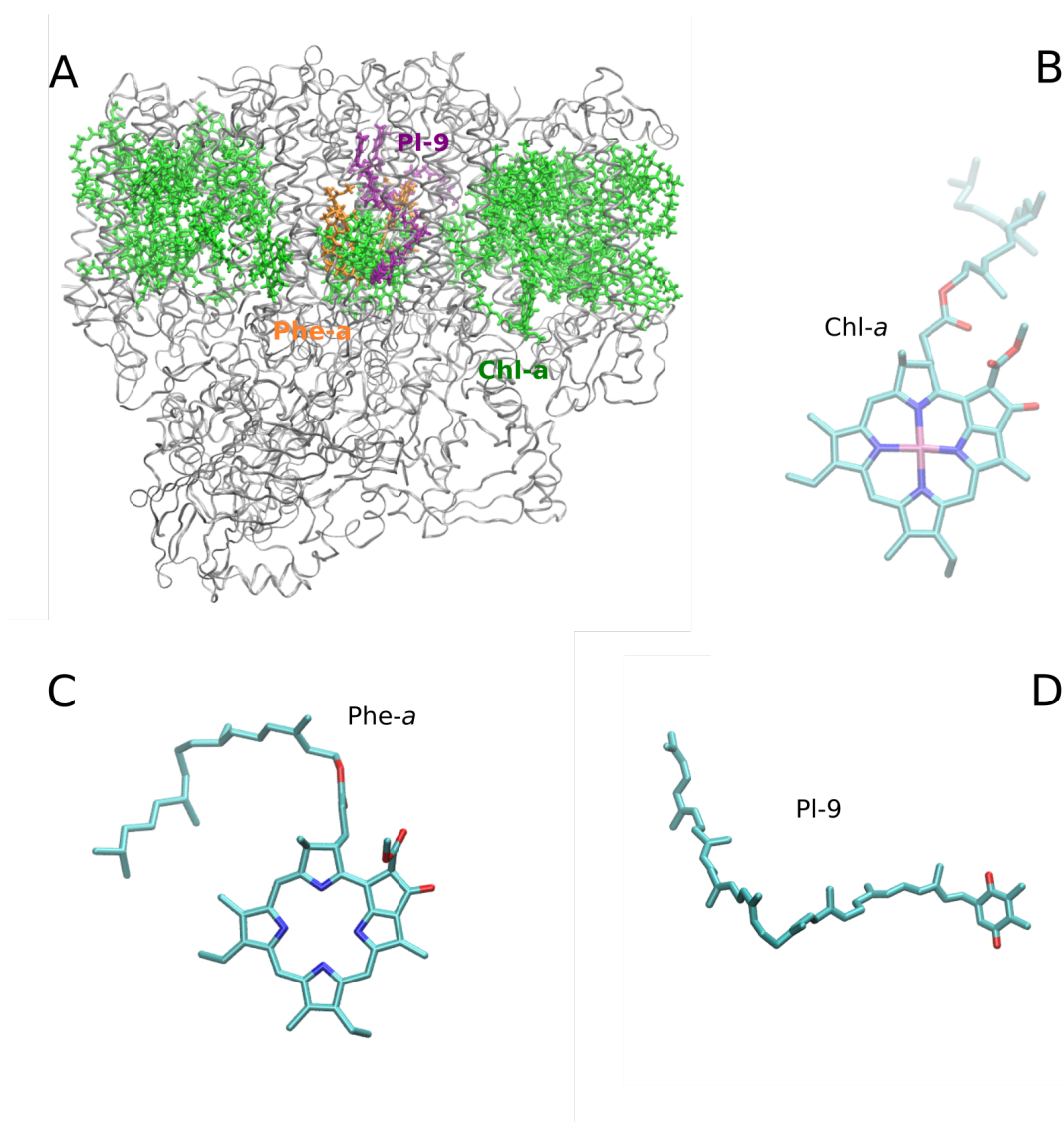
Figure 3.1: *The Chl-a, Phe-a and Pl-9 cofactors in PSII [35]. (A) Molecular graphics of monomer A of PSII from the crystal structure of Ref. [115] (PDB ID: 3WU2).The protein is shown as silver ribbons, and the selected cofactors are shown as licorice colored green (Chl-a), orange (Phe-a) and purple (Pl-9). (B, C and D) Close view of molecules of Chl-a (panel B), Phe-a (panel C) and Pl-9 (panel D).*

performed these calculations using the MP2-optimized structure of Chl-a core. The latter structure is depicted in Figure 3.2 (A) and Figure 3.3 with atoms labeled according to the CHARMM atom names and to the CHARMM atom types respectively. To minimize the energy difference between the HF and the CHARMM calculations we adjusted the partial charges of all the atoms considering the following constraint: the

34

partial charges on atoms CHB, CHC, CHD, HBH, HCH and HDH were kept fixed to -0.10e for the carbon atoms, and 0.10e for hydrogen atoms , similar to the CHARMM description of heme and to the set of bacteriopheophytin partial charges from Ref. [30]. The full list of converged interaction energies is contained in Table 3.2. We observe that the HF interaction energies values (ΔE_{QM}) range from -1.73 *kcal/mol*, at the HBD2 site, to -11.82 *kcal/mol*, at MG site. This range of values is in good agreement with the one of the CHARMM interaction energies computed with our partial charges (ΔE_{MM}): -1.32 *kcal/mol* and -11.92 *kcal/mol* at the HBD2 and MG sites respectively. The mean absolute difference between ΔE_{QM} and ΔE_{MM} is 0.28 *kcal/mol*. We observe the largest deviation from the HF interaction energies at the HCH donor site, where we compute a difference between ΔE_{QM} and ΔE_{MM} of 0.51 *kcal/mol*. We were able to reproduce at best the interaction energy at the NA site where ΔE_{QM} and ΔE_{MM} agree to -0.03 *kcal/mol*. At the sites of highest interest in term of possible protein-Chl-*a* interactions we obtained the following energy differences: -0.1 *kcal/mol* at the MG site and -0.27 *kcal/mol* at the OBD site.
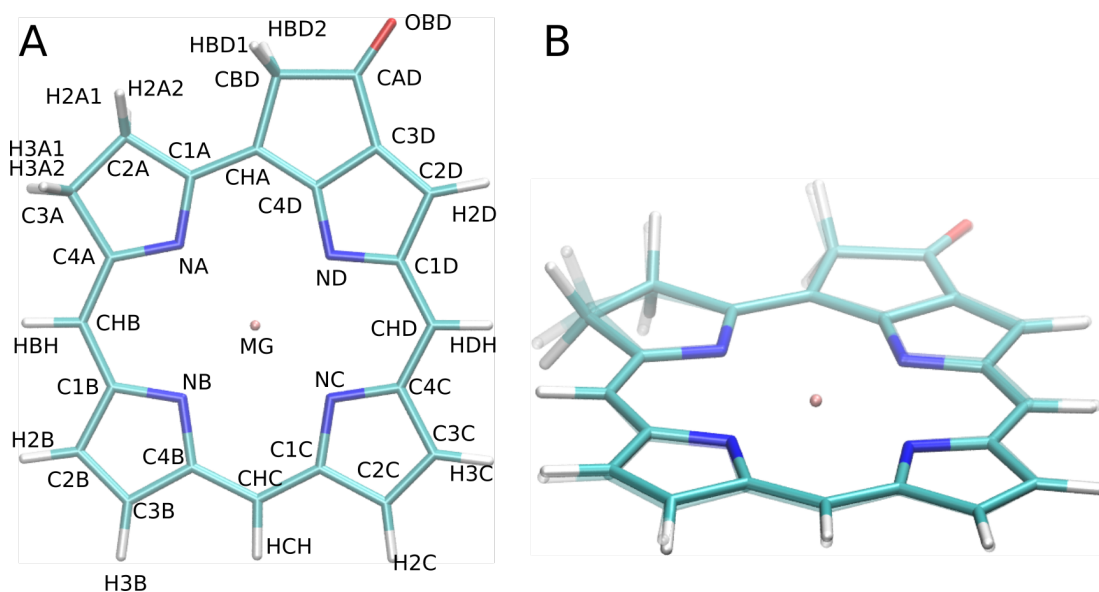
Figure 3.2: *The Chl-a core [35]. Panel(A) Molecular graphics of the MP2-optimized Chl-a core with atoms labeled according to the CHARMM atom names. Carbon atoms are colored cyan, oxygen - red, and hydrogen - white. (B) Overlap of the Chl-a core structure optimized with MP2/6-31G(d) (transparent) vs. MM (full colors). For the MM optimization of the Chl-a core, we controlled the planarity of the bond involving atoms H2B, H2C, H3B, H3C, and H2D by using improper angles. These definitions of the improper angles are not needed for the full Chl-a molecule, where the H atoms are replaced by sidechains*

Interaction site[a]	$\mathbf{\Delta E_{QM}}$[b]	$\mathbf{\Delta E_{MM}}$[c]	$\mathbf{\Delta E_{MM} - \Delta E_{QM}}$	$\mathbf{R_{QM}}$[b]	$\mathbf{R_{MM}}$[c]	$\mathbf{R_{MM} - R_{QM}}$
NA	-2.44	-2.47	-0.03	2.44	2.09	-0.35
NB	-3.64	-3.83	-0.18	2.37	2.03	-0.34
NC	-3.54	-4.03	-0.48	2.39	2.04	-0.35
ND	-3.35	-3.76	-0.41	2.4	2.09	-0.31
OBD	-5.90	-6.17	-0.27	2.06	1.8	-0.26
OBD_p	-5.48	-5.09	0.39	2.06	1.8	-0.26
H2A1	-3.10	-3.19	-0.09	2.53	2.51	-0.02
H2A2	-2.64	-2.90	-0.26	2.53	2.51	-0.02
H2B	-2.34	-2.50	-0.15	2.5	2.52	0.02
H2C	-2.38	-2.07	0.31	2.49	2.55	0.06
H2D	-2.47	-2.22	0.24	2.4	2.45	0.05
H3A1	-2.36	-2.64	-0.28	2.57	2.54	-0.03
H3A2	-2.44	-2.87	-0.43	2.57	2.54	-0.03
H3B	-2.61	-2.54	0.07	2.46	2.51	0.05
H3C	-2.50	-2.26	0.24	2.48	2.53	0.05
HBD1	-1.82	-1.33	0.49	2.56	2.55	-0.01
HBD2	-1.73	-1.32	0.42	2.59	2.57	-0.02
HBH	-2.49	-2.91	-0.42	2.55	2.56	0.01
HCH	-2.95	-2.44	0.52	2.45	2.57	0.12
HDH	-3.15	-3.10	0.06	2.42	2.78	0.36
MG	-11.82	-11.92	-0.10	2.26	2.18	-0.08

Table 3.2: *Water interaction energies (kcal/mol) and interaction distances (Å) computed for the Chl-a core [35].*

[a] *The interaction sites are depicted in Figure 2.1. OPD_p is an interaction site of OPD*
[b] *HF/6-31G(d) values. ΔE_{QM} values were scaled by 1.16.*
[c] *Energy and distance values from MM computations.*

Our parameters show to describe well not only the water interaction energies but also the associated minimum interaction energy distances. Overall, there is good agreement between interaction distances computed with HF (R_{QM}) and with CHARMM (R_{MM}). In particular the (R_{MM}) computed at the magnesium ion site, 2.18 Å, is close to the R_{QM}, 2.26 Å, and to the average water-magnesium distance found in the crystal structure (PDB ID: 3WU2, chain A) [115]. The latter contains seven chlorophyll-*a*, which coordinate one water molecule at the manganese site, with an average magnesium/water oxygen distance of 2.1 ± 0.02 Å. The crystal structure of the chlorophyllide-*a* dehydrated gives us another example of magnesium/water oxygen distance, which in this case measures 2.035 Å [18]

Following optimization of the partial charges we refined bonds and valence angles equilibrium positions parameters. We started by performing a CHARMM minimization

Figure 3.3: *Chl-a atom types with assigned partial charges: (A) core, (B) with sidechains [35].*

of Chl-*a* core in gas phase, using our new derived partial charges reported in Figure 3.3 and bonded parameters based on the bacteriochlorophyll-*a* from Ref. [30]. To improve

the agreement with the MP2 optimized structure, we introduced two new atom types, CPA3 and CPA4, and adjusted several equilibrium positions parameters as summarized in Table 3.3 and 3.4. The newly introduced atom types conserve the same force constants of atom type CPAN but have different bond equilibrium positions, which were optimized to reach a better agreement with the MP2 optimized structure. We assigned the same atom type, CPA2, to C1D and C4C.

Finally, we refined bonded parameters and the optimized set of partial charge to test the agreement between CHARMM and MP2 Chl-*a* core gas phase minimized structures. We computed very low absolute average deviation values both for bond lengths and valence angles in the MM-optimized geometry and the corresponding QM values: 0.01 Å for the bond lengths and 1.12 Å for the valence angles. We observed that both the MM and QM-optimized geometries of the Chl-*a* core show a planar chlorin ring, and the distances between the magnesium ion and the four adjacent nitrogen atoms are in good agreement between the two structures, Table 3.5.

The direction of the dipole moment computed with our new set of partial charges points in the same direction of the one computed with HF, with a deviation of only 4°. It does, however, overestimate the HF dipole value by 78.2%. Nevertheless, a force field overestimation of the dipole moment in gas phase, up to the 50% of the QM computed dipole moment, is required to be relevant in the bulk phase [121]. Due to the complexity of the Chl-*a* molecule and the large number of computed interaction sites, we consider that the good description of the dipole moment direction and of the water interaction energy, Table 3.2 is satisfactory.

To further probe the quality of our parameters, we performed a potential energy scan for a water interaction at the magnesium ion site. For this calculation we fixed the positions of Chl-*a* core and allowed, as only degree of freedom, the motion of the water molecule along the normal of the chlorin ring passing through the magnesium ion coordinates, Figure 3.4. Under this conditions we computed the scan both with the CHARMM force field (MM) and with HF/6-31G(d) (QM). We found a good agreement between MM and QM profiles Figure 3.4. In contrast, a coarse estimation using partial charges reported for bacteriochlorophyll models in Refs. [21] indicate an underestimation of the minimum magnesium/water interaction energy of $\sim 3kcal/mol$ (In Table 3.6: a comparison between partial charges for MG, NA, NB, NC and ND, derived by us and the ones computed in Refs. [21, 44]).

We prepared a CHARMM topology file for the complete Chl-*a* molecule starting from the partial charges optimized for the Chl-*a* core. To this aim we replaced specific

Bond/angle	Ref. value,b_0	MP2 value[a]	MM value[a]
Bonds			
CPBN-CPAN	1.437	1.430	1.426
CPBN-CPA2	1.462	1.462	1.455
CPBN-CPA3	1.422	1.420	1.404
CPBN-CPA4	1.467	1.468	1.456
CPBN-CPBN	1.365	1.371	1.363
CPM3-CPAN	1.424	1.439	1.410
CPMN-CPA2	1.400	1.405	1.393
CPM2-CPAN	1.404	1.405	1.417
CPM3-CPA4	1.371	1.371	1.357
NPH4-CPA3	1.385	1.364	1.367
NPH3-CPA4	1.390	1.390	1.375
NPH4-CPA2	1.360	1.365	1.346
NPH1-MG	2.184	2.183	2.167
NPH2-MG	2.028	2.028	2.010
C2-CPBN	1.474	1.481	1.482
Valence angles			
CPA2-CPMN-CPA2	124.827	124.8	127.723
CPAN-NPH2-CPAN	107.35	107.3	108.541
MG-NPH4-CPA2	132.662	132.7	130.650
NPH1-CPAN-CT2 N	110.348	111.4	111.902
NPH2-CPAN-CPMN	124.578	124.6	126.087

Table 3.3: *Reference bond lengths (Å) and valence angles (°) refined for the Chl-a core [35]. The atom chemical types used here are illustrated in Figure 3.3*

[a] *Some MM and MP2 bond lengths were averaged as summarized in Table 3.4*

hydrogen atoms with larger chemical groups, and adjusted the partial charges accordingly to Ref. [121]: summing the hydrogen's partial charge to the partial charge of its covalently bound carbon atom. For example, C2B has a partial charge of -0.36e in the Chl-a core and of -0.12e in the complete Chl-a molecule because we summed in the Chl-a core the partial charge of H2B, 0.24e, which is replaced by CTN in the complete Chl-a molecule.

Bond/angle, atom chemical type[a]	Bonds/angles averaged, atom name[b]
CPBN-CPAN	C1B-C2B; C3B-C4B
CPBN-CPA2	C3C-C4C; C1D-C2D
CPBN-CPA3	C3D-C4D
CPBN-CPBN	C2B-C3B; C2C-C3C; C2D-C3D
CPM3-CPAN	C4B-CHC
CPMN-CPA2	CHD-C1D; CHD-C4C
CPM2-CPAN	CHA-C1A
CMP3-CPA4	CHC-C1C
NPH4-CPA3	ND-C4D
NPH3-CPA4	NC-C1C
NPH4-CPA2	ND-C1D
CPBN-CPA4	C2C-C1C
NPH1-MG	NA-MG
NPH2-MG	NB-MG
C2-CPBN	CAD-C3D
CPA2-CPMN-CPA2	C1D-CHD-C4C
CPAN-NPH2-CPAN	C1B-NB-C4B
MG-NPH4-CPA2	MG-ND-C1D
NPH1-CPAN-CT2N	NA-C1A-C2A, NA-C4A-C3A
NPH2-CPAN-CPMN	NB-C1B-CHB

Table 3.4: *Bonds and valence angles: atom chemical types vs. atom names definitions [35]. Bonds and valence angles described by the reference bond stretching and angle bending parameters summarized in Table 3.3.*

[a] *See labels in Figure 3.3*
[b] *See labels in Figure 3.2 (A)*
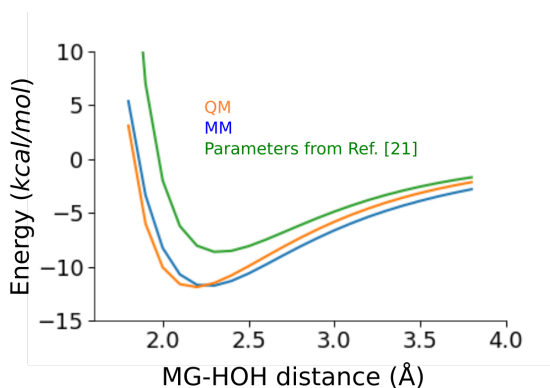
Figure 3.4: *Energy scan for the water interaction energy at the Chl-a manganese site [35]. The scan was performed constraining the Chl-a atoms and moving the water molecule along the normal of the chlorin ring. The same calculation was repeated using HF/6-31G(d) (QM) (orange line), our CHARMM parameter (blue line) and parameters from Refs. [21] (green line)*

	MP2	MM
Distances (Å)		
MG-NA	2.18	2.17
MG-NB	2.03	2.01
MG-NC	2.06	2.06
MG-ND	2.03	2.04
Angles (°)		
NA-MG-NC	177.0	178.2
NB-MG-ND	178.2	179.2

Table 3.5: *Comparison between Chl-a core structures optimized with MP2 vs. MM [35].*

Atom type	Partial atomic charge		
	Current set[a]	**Set A**[b]	**Set B**[c]
MG	1.17	1.03	0.922
NA	-0.58	-0.33	-0.424
NB	-0.6	-0.73	-0.535
NC	-0.59	-0.28	-0.367
ND	-0.51	-0.51	-0.473

Table 3.6: *Comparison of selected partial atomic charges of Chl-a and bacteriochlorophyll [35].*

[a] *Values from the refined force-field parameters reported here for Chl-a. The full set of atomic partial charges is reported in Fig. 6.*
[b] *Data reported in Ref. [21] for bacteriochlorophyll-a.*
[c] *Data reported in Ref. [44] for bacteriochlorophyll-a.*

3.1.2 Parametrization of pheophytin-*a*

We repeated for Phe-*a* the same parametrization procedure adopted for Chl-*a*. Firstly, we iteratively adjusted the partial charges of the Phe-*a* core. The water interaction energies are summarized in Table 3.7 and the CHARMM atoms names used are depicted in Figure 3.5 (A). We obtained an overall good agreement between ΔE_{QM} and ΔE_{MM} and between ΔR_{QM} and ΔR_{MM}. The mean absolute difference between the MM and QM calculations computed at each interaction sites of the Phe-*a* is 0.29 *kcal/mol*. NB and NC interaction site show the best agreement with the QM calculation of the water interaction energies: the absolute MM/QM difference is in the range 0.03-0.07 *kcal/mol*. Water interaction energies are extremely well characterized also

at the carbonyl oxygen, OBD atom, where the absolute MM/QM difference is in the range 0.08-0.19 *kcal/mol* for the two interaction orientations.
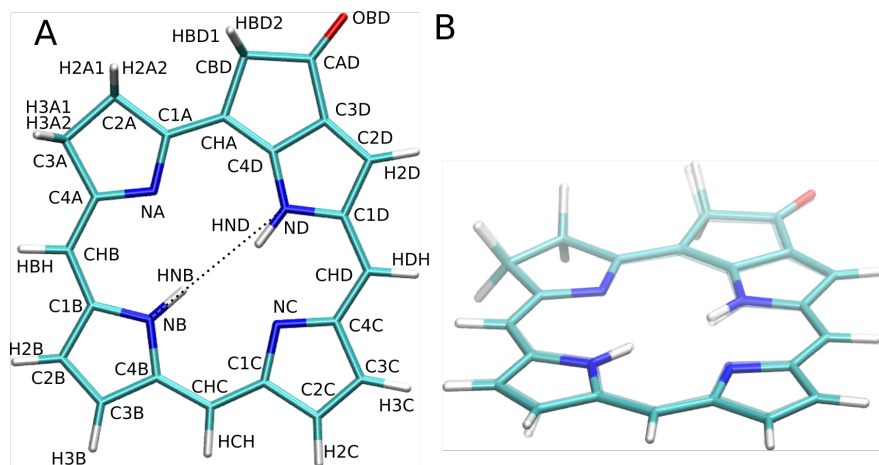
Figure 3.5: *The Phe-a core. Panel A: MP2 optimized structure of the Phe-a core with the used CHARMM atoms names [35]. Panel B: overlap between the structure of the Phe-a core optimized with MP2/6-31G(d) (transparent) vs. MM (full colors).*

The MP2-optimized geometry of the Phe-*a* core reveals that the NB-HNB and ND-HND bonds are coplanar with the chlorin ring, but are not co-linear with each other, Figure 3.5 (A), Table 3.8. The angles C1-NB-HNB and C4B-NB-HNB are almost identical, 125.1° and 124.2° respectively, meaning that the bond NB-HNB is almost along the bisector of the C1-NB-C4 valence angle. In contrast, the angles C1D-ND-C4D and C4D-ND-HND differ of about 10°, they measure 120.1° and 130.5° respectively, and the bond ND-HND slightly orients towards NC. To reproduce this feature with MM, we adjusted the reference valence angles (θ_0) for CPAN-NPH2-CPAN to 120.5° and for CPA3-NPH4-CPA2 to 130.5°. All other bonded parameters involving HNB and HND were taken from the CGenFF. The MM-optimized structure agrees within 3° for the valence angle involving HNB and HND with the QM-optimized structure, Table 3.8. In section A.7 is reported the full set of partial charges derived for Phe-*a*

3.1.3   Parametrization of plastoquinone-*9*

In this work we conducted a careful partial charge parametrization of the Pl-*9* core. Bonded parameters were computed using the paramchem web based software [119, 120], and no further modification have been performed.

The atom names assigned to the Pl-*9* core and the corresponding CHARMM atom types are depicted in Figure 3.6 (A) and (B) respectively. The results of the water

Interaction site[a]	ΔE_{QM}[b]	ΔE_{MM}[c]	$\Delta E_{MM} - \Delta E_{QM}$	R_{QM}[b]	R_{MM}[c]	$R_{MM} - R_{QM}$
NA	-3.00	-3.40	-0.40	2.44	2.17	0.27
NB	-2.05	-2.12	-0.07	2.37	2.18	0.19
NC	-3.34	-3.31	0.03	2.39	2.05	0.34
ND	-1.88	-2.37	-0.49	2.40	2.29	0.11
OBD	-5.54	-5.35	0.20	2.06	1.80	0.26
OBD$_m$	-4.37	-4.45	-0.08	2.00	1.83	0.17
H2A1	-2.72	-2.39	0.33	2.53	2.54	0.01
H2A	-2.72	-2.46	0.26	2.53	2.54	0.01
H2B	-2.61	-2.65	-0.03	2.50	2.52	0.02
H2C	-2.22	-1.85	0.38	2.49	2.56	0.07
H2D	-2.66	-2.27	0.39	2.40	2.34	0.06
H3A1	-2.14	-2.27	-0.14	2.57	2.57	0.00
H3A2	-2.14	-2.29	-0.15	2.57	2.57	0.00
H3B	-2.89	-2.53	0.35	2.46	2.52	0.06
H3C	-2.30	-2.03	0.27	2.48	2.54	0.06
HBD1	-1.88	-1.34	0.54	2.56	2.54	0.02
HBD2	-1.88	-1.35	0.53	2.59	2.54	0.05
HBH	-2.54	-3.02	-0.48	2.55	2.54	0.01
HCH	-3.10	-2.47	0.63	2.45	2.56	0.11
HDH	-3.16	-3.24	-0.09	2.42	2.70	0.28

Table 3.7: *Water interaction energies (kcal/mol) and interaction distances (Å) computed for the Phe-a core [35].*

[a] *The interaction sites are depicted in Figure 2.1. OPD$_m$ is an interaction site of OPD*
[b] *HF/6-31G(d) values. ΔE_{QM} values were scaled by 1.16.*
[c] *Energy and distance values from MM computations.*

interactions calculations are summarized in Table 3.9. We adjusted the partial charges considering standard constraints for partial charges of methyl groups: -0.27e for carbon and 0.09e for hydrogen atoms. With this constraints we obtained a set of partial charges that overall well describes the water interactions. In particular, difference values between ΔE_{QM} and ΔE_{MM} are within 0.2 $kcal/mol$ for the oxygen acceptor sites (O9, O10) and for the donor sites H11 and H12. The CHARMM geometry optimization performed with our partial charges overlap with a good agreement to the MP2-optimized structure, Figure 3.6 (C).

3.1.4 Testing the new force field parameters within the protein matrix

To further prove the quality of our force field parameters, we tested how well they describe cofactors dynamics within the protein matrix. We used as a test system the

	MP2	**MM**
Distances (Å)		
NA-NC	4.22	4.35
NB-ND	4.08	4.23
Angles (°)		
NB-HNB-ND	177.4	178.7
NB-HND-ND	161.7	170.0
C1B-NB-HND	125.1	126.9
C4B-NB-HNB	124.2	121.8
C1D-ND-HND	120.1	121.8
C4D-ND-HND	130.5	127.1
Improper angles (°)		
C1D-HND-C4D-NDa	0.0	0.0
C1B-C4B-HNB-NBa	0.0	0.1

Table 3.8: *Comparison between Phe-a core structures optimized with MP2 vs. MM [35].*

Interaction site	$\Delta E_{QM}{}^a$	$\Delta E_{MM}{}^b$	$\Delta E_{MM} - \Delta E_{QM}$	$R_{QM}{}^b$	$R_{MM}{}^c$	$R_{MM} - R_{QM}$
O9	-4.14	-4.29	-0.15	2.12	1.75	-0.37
O10	-4.13	-4.28	-0.15	2.12	1.75	-0.37
H11	-2.71	-2.53	0.18	2.37	2.48	0.11
H12	-2.70	-2.53	0.17	2.37	2.48	0.11
H13	-2.31	-1.49	0.82	2.66	2.70	0.04
H15	-1.97	-1.29	0.68	2.59	2.69	0.10
H16	-2.31	-1.49	0.82	2.66	2.70	0.04
H18	-1.97	-1.29	0.68	2.59	2.69	0.10

Table 3.9: *Water interaction energies (kcal/mol) and interaction distances (Å) computed for the Pl-9 core [35].*

[a] *HF/6-31G(d) values. ΔE_{QM} values were scaled by 1.16. Since the atomic partial charges of the carbon and hydrogen atoms of the methyl groups are set to the standard CHARMM values, we did not perform computations for atoms H14 and H17*
[b] *Energy and distance values from MM computations.*

wild type PSII with D1-E65 protonated, embedded in a hydrated POPC lipid bilayer (Sim1). Along 150 *ns* of trajectory, Chl-*a*, Phe-*a* and Pl-*9* cofactors show an overall good behavior and no anomalies in their dynamics or in their interactions with the protein matrix.
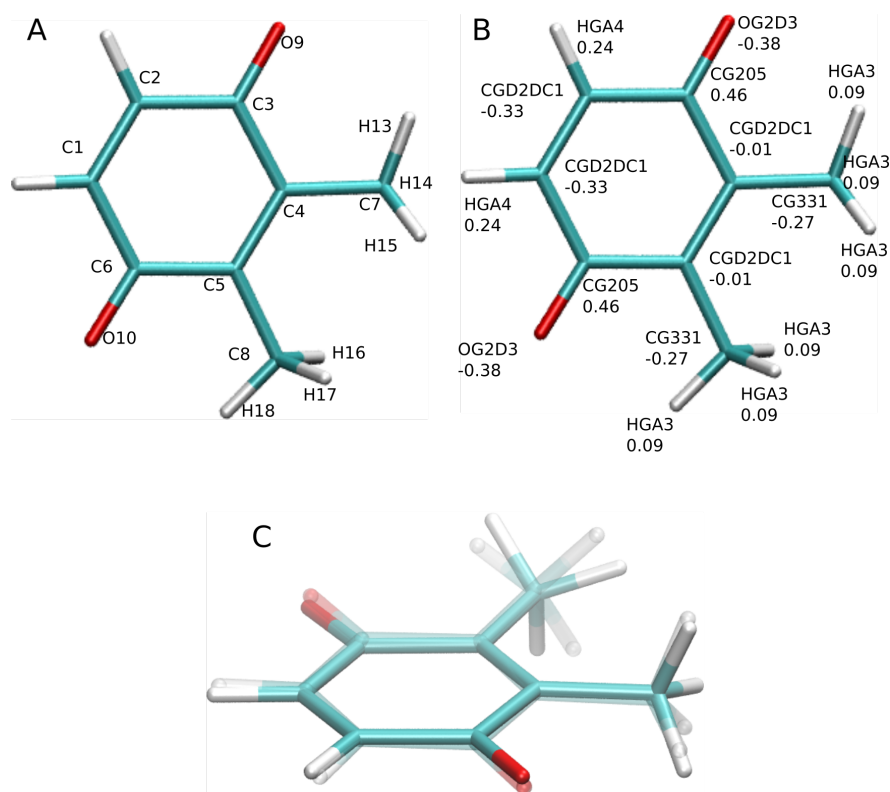
Figure 3.6: *The Pl-9 core [35]. Panel A,B: MP2 optimized structure of the Pl-9 core with the used CHARMM atoms names (A) and types (B). Panel C: overlap between the structure of the Pl-9 core optimized with MP2/6-31G(d) (transparent) vs. MM (full colors).*

In particular, the PSII crystal structure Ref. [115] presents seven Chl-*a* molecules coordinating one water molecule at the magnesium ion site (note that in our set-up we removed one of these Chl-*a* molecules, see section 2.2.1) . We computed the distances between magnesium ion and its coordinating crystal water oxygen along the whole trajectory for the six Chl-*a* molecules coordinating a water molecule present both in the crystal structure and in our system. In Figure 3.7 you can observe the distribution of all the computed distances. None of the Chl-*a* molecules looses its coordinating water molecule along the trajectory. The average distance value is 2.07 Å with a standard deviation of 0.08 Å. This result is in excellent agreement with the distance values found in the crystal structure [115] where the average value is 2.07 Å.

3.1.5 Concluding remarks

We revised the CHARMM force field description of the Chl-a, Phe-a and Pl-9 cofactors [35]. We derived new sets of partial charges by fitting MM water-target molecule inter-
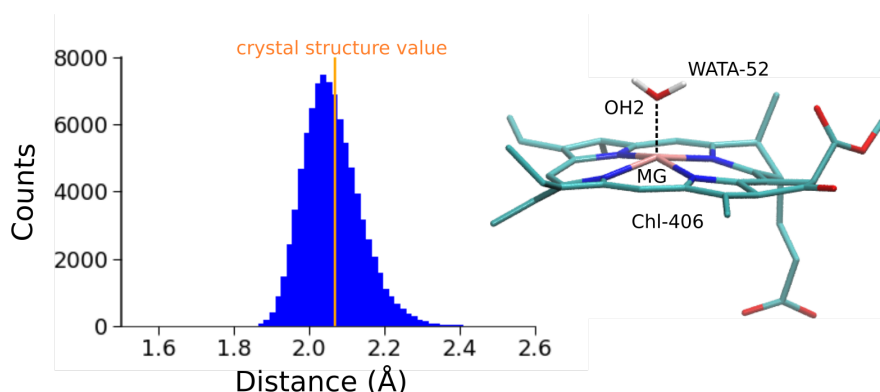
Figure 3.7: *Chl-a/water interactions in the protein environment [37]. Left panel: histogram of the distances between the Chl-a magnesium ions and their coordinating water from crystal structure along the whole trajectory of Sim1. The calculation was performed on all the Chl-a molecule which in the crystal structure are shown to coordinate a water molecule. The orange vertical line represent the average water/Chl-a Mg ion distance in the crystal structure [115]. Note that we did not impose any coordination bond, the interaction is fully described electrostatically. Right panel: snapshot from last 10 ns of Sim1 representing a Chl-a molecule and its coordinating water.*

action energies at each H-bonding site to values from HF/6-31G(d) computations. In the case of Chl-a we gave particular attention at describing accurate water interactions at the magnesium ion and at the oxygen acceptor sites. Chlorophylls, in PSII, do, indeed, show to interact either with protein residues or with water molecules at the these sites. We revised bonded parameters by Foloppe et al., [30] to better describe the MP2 optimized structures of Chl-a and Phe-a.

Finally, we tested our parameters into the PSII protein matrix, along 150 *ns* MD simulations. Chl-a, Phe-a and Pl-9 show to well preserve their structure. In particular, our parameters show to optimally describe the Chl-a coordination bond with a water molecule at the magnesium site.

3.2 Direct H-bonded networks in photosystem II

Part of the results and material presented in this chapter has been published in Ref. [37].

A PSII monomer counts about 3000 amino acid residues. The hundreds of dynamic H-bonds connecting them play a role in shaping the structure and the dynamics of the protein complex. We employed a data analysis tool developed by us to detect HBNs in our simulations and to compare this networks among different simulations or with the crystal structure's HBN.

We computed 744 H-bonds connecting amino acid residues, excluding backbone-backbone H-bonds, in the crystal structure [115]. Due to thermal fluctuation the total number of stable (present in more of the 60% of the equilibrated trajectory) H-bonds in the wild type PSII (Sim1) drops to about 647. Nevertheless, the crystal structure remains a valid experimental reference to test how well our simulation set-ups are able to describe PSII HBNs in the S1 catalytic state. In particular, as a first step towards understanding proton transfer in PSII, we are interested in understanding how HBNs reacts to changes in targeted carboxylate groups protonation states. We were, indeed, able to show in a previous work [36], that change in titratable amino acids protonation state can lead to drastic change in the HBN of the plasma membrane proton pump both locally (in the region surrounding the protonation site) and in the whole protein system. A strong coupling between titratable groups protonation state and protein/water dynamics was also observed in the lactose permease [4] and in proton transporters [117]. The protonation states of carboxylate groups in PSII is still unclear. As a first step towards probing the likely protonation states and the response of PSII to a change in the protonation state, we performed three independent simulations of PSII that are distinguished by the protonation state of D1-E65 or D2-E312: Sim1 (D1- E65 protonated) Sim2 (D2-E312 protonated) Sim3 (D1-E65 and D2-E312 unprotonated). These two carboxylate groups were chosen for tests of protonation-dependent dynamics because, in the crystal structure [115], they are within 2.5 Å distance, suggesting that this carboxylate pair might be protonated [11]. A model whereby the D1-E65/D2-E312 pair is protonated is compatible with FTIR spectroscopy data suggesting that these two carboxylates are part of an extended HBN that gives rise to a protonated carboxylate vibrational mode [100]. On the other hand, the close vicinity of the positively charged

D1-R334, leaves open the possibility that both glutamate residues are unprotonated. Determining the correct protonation state of the D1-E65/D2-E312 glutamate pair is necessary to further study other properties of the system, e.g. proton exit pathways.

3.2.1 Local dynamics of photosystem II depends on the protonation state

Analysis of Sim1, Sim2 and Sim3 indicates that interactions in the local environment of D1-E65 and D2-E312 depend significantly on the protonation state of these two groups. We compared the local HBN of Sim1-3 with the one found in the starting crystal structure, Figure 3.8 (A).

To this aim we computed all the possible H-bond partners for the following residues: D1-E65, D2-E312, D1-R334, D1-N335, D1-D59, D1-R64, PsbO-R152.

This residues forms a total of 13 H-bonds in the crystal structure. Sim1 shows to conserve the most of these H-bonds, 11, compared to the other two simulations: 8 conserved H-bonds in Sim2 and 7 in Sim3, Figure 3.8. In particular Sim1 shows to better conserve the inter domain H-bonds at the PsbO-D1 subunit interface and the geometry of the D1-E65, D1-R334, D1-N335 and D2-E312 H-bonded cluster, Figure 3.8. The D1-E65, D2-E312 direct H-bond is transiently lost: the two residues side chains sample H-bonding distances during the trajectory, Figure 3.9, but they do not steadily meet our H-bond angle criteria. On the contrary in Sim2 D1-E65, D2-E312 are H-bonded but the larger cluster including D1-R334 and D1-N335 is perturbed. D1-E65 forms a salt bridge with D1-R334 and breaks the H-bond with D1-N335.

Sim3 shows the largest deviation from the crystal structure at the glutamate pair site: D1-E65 side chain samples a large rotation and forms a salt bridge with PsbO-R152 while D2-E312 forms a salt bridge with D1-R334. It is therefore unlikely that none of the two glutamate residues is protonated in the starting crystal structure.

3.2.2 Effect of protonation state on the inter-subunit H-bonded network

To measure the effect of change in protonation state of a single residue on the HBN of the whole protein, we conducted a comprehensive analysis of the inter subunit H-bonds of core subunits: D1, D2, CP43, CP47 and extrinsic subunits: PsbO, PsbU and PsbV, Table 3.10. We computed the total number of inter-subunit stable H-bonds formed by amino acid residues of core and extrinsic subunits among them selves and with amino
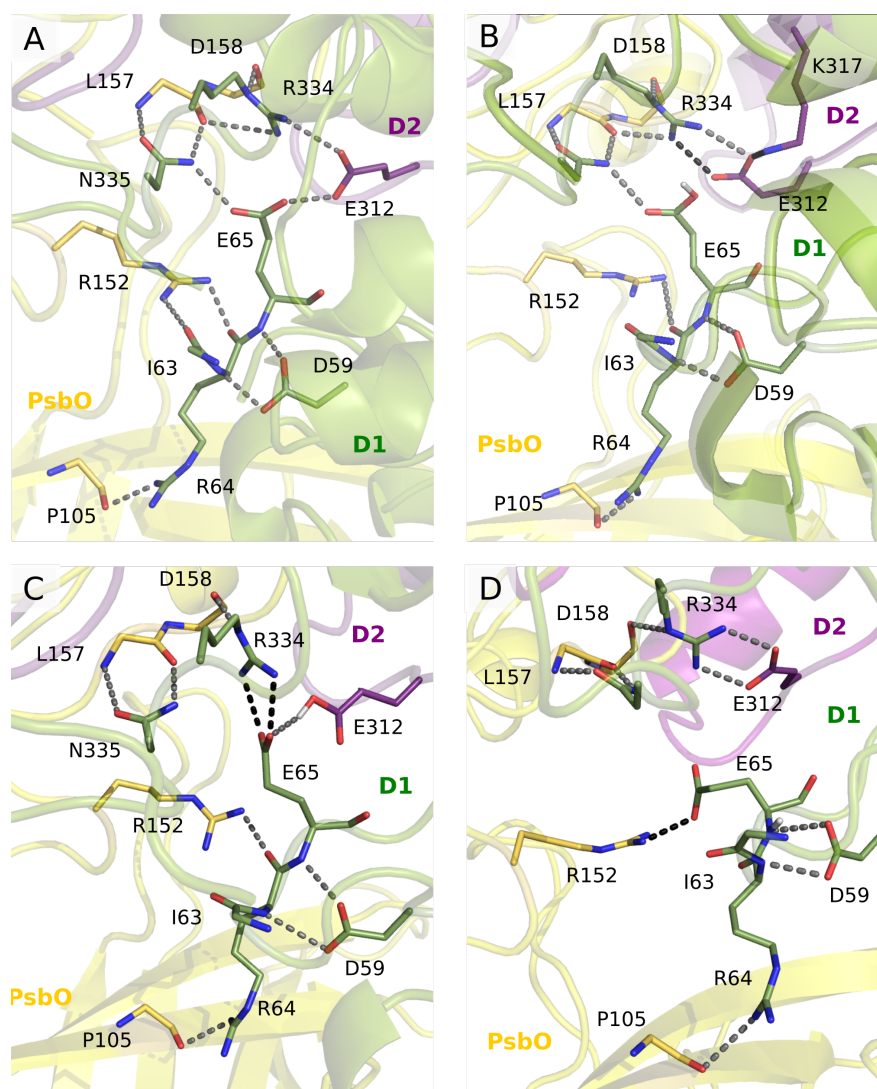
Figure 3.8: *Molecular graphics representing the HBN at the D1-E65/D2-E312 glutamate pair site [37], in the crystal structure (A) and in the wild type simulations with different protonation state of the D1-E65/D2-E312 glutamate pair: D1-E65 protonated (Sim1) (panel B), D2-E312 protonated (Sim2) (panel C) and both glutamate amino acid residues unprotonated (Sim3), (panel D). The H-bonds are depicted by dotted lines colored gray if the H-bond is conserved with the crystal structure and black if the H-bond is not present in the crystal structure. The H-bond are represented only if they are present for more than 60% of the analyzed trajectory, e.g the last 48 ns of the trajectories.*

acid residues belonging to other subunits of the protein complex. Even though the protonation state used in Sim3 showed to significantly alter the local HBN at the D1-E65/D2-E312 glutamate pair site, it does not show the same dramatic effect on the inter-subunit HBN.

49
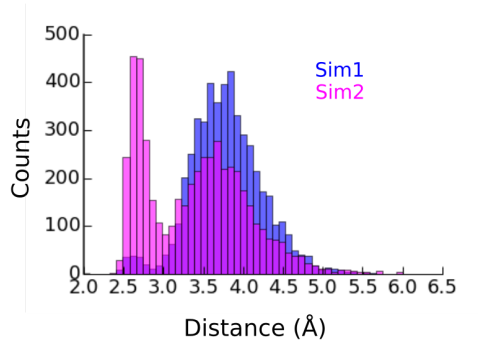
Figure 3.9: *Histogram of the distance between carboxylate oxygens for the D1-E65/D2-E312 glutamate pair along the whole trajectory*

Inter-subunit H-bonds

Simulation	Total	D1	D2	CP43	CP47	PsbO	PsbU	PsbV
Crystal [115]	196	76	81	51	30	30	13	14
Sim1	160	69	66	44	21	21	8	11
Sim2	133	63	59	29	13	13	7	9
Sim3	152	73	68	35	18	18	7	11

Table 3.10: *PSII inter-subunit H-bonds. Number of inter-subunit H-bonds for core (D1, D2, CP43 and CP47) and extrinsic (PsbO, PsbU and PsbV) subunits, computed in the crystal structure and in Sim1-3*

Sim1 is the simulation that better conserves the total amount of inter subunits H-bonds with respect to the crystal structure, wheres Sim2 shows the worst agreement. In particular, the PsbO, D2 and CP43 subunit appear to be the most affected by change in protonation state of D2-E312. In Sim2 these subunits loose 8, 7 and 15 inter-subunit H-bonds respectively with respect to Sim1. Interestingly, in Sim2, CP47 subunits forms 5 inter-subunit H-bonds less both with D1 and D2 and the PsbO subunit looses all the inter-subunits H-bonds that binds it with the D2 subunit. In Sim1, PsbO is connected to D2 via two H-bonds between PsbO-R162 and the backbone oxygens of D2-A305 and D2-A306, whereas in Sim2 there is no stable H-bond interconnecting the 2 subunits.

These analyses suggest that the protonation state of the D1-E65/D2-E312 glutamate pair used in Sim1 is the one that better describes the local and inter-subunits HBN. Therefore, we propose that in the crystal structure, D1-E65 is likely protonated. Pursuant to the considerations above, we consider Sim1, with D1-E65 protonated, as reference Sim for this work.

3.2.3 Concluding remarks

The D1-E65 and D2-E312 glutamate amino acid residues are found in the crystal structure [115] to be in a H-bonding distance. For this reason, they have been proposed as possible proton carrying dyad [11]. To get insights into the most likely protonation state of this glutamate pair, we performed three simulations: D1-E65 protonated, D2-E312 unprotonated and both glutamate residues unprotonated. From our calculation the D1-E65 protonation appears to be the most likely protonation state for the glutamate pair. It, indeed, conserves at best the crystal structure local HBN displayed in Figure 3.8 and the inter-subunit HBN. For this reason we considered this protonation state for the wild type protein in our work.

3.3 Proton exit pathways in photosystem II

Part of the results and material presented in this chapter has been published in Ref.
[37].

The mechanism of proton transfer in PSII is still a key open question. As a first step
towards answering this question we analyzed water mediated HBNs in PSII. We are
interested in finding stable networks that connect titratable amino acid residues side
chains, starting from the first coordination shell of the manganese cluster and ending
at the bulk water exposed regions at the luminal side of the protein complex. These
networks of titratable residues connected via water wires could may serve as proton
exit pathways.
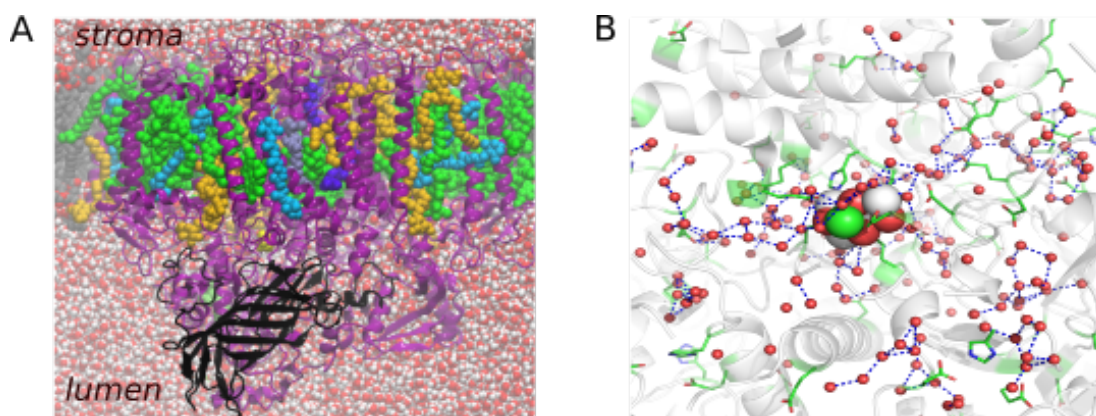
Figure 3.10: *PSII in a hydrated lipid membrane environment [37]. (A) Molecular
graphics of PSII from Sim1. The PsbO subunit is colored black, other protein
subunits are colored purple. Heavy atoms of cofactor molecules are depicted as van
der Waals spheres using the following color code: chlorophyll – green, betacarotene
– cyan, and special lipids - yellow. (B) Water-mediated hydrogen-bond networks
in PSII. We show a view from the stroma side of the manganese cluster region
from the crystal structure (chain A, PDB ID:3WU2, [115]). Blue dotted lines
represent hydrogen bonding distances between water oxygen atoms, shown as red
dots, and selected titratable side chains, shown in green.*

From the observation of the crystal structure several possible proton exit pathways
have been proposed [82, 42, 11, 32], as largely discussed in section 1.2. You can observe
in Figure 3.10 (B), how complex and highly interconnected are water-mediated HBNs
in PSII. Figure 3.10 (B) shows a static picture from the crystal structure coordinates
[115]; however, these networks are the result of the combined dynamics of the protein

matrix (*ps* time scale) and the water dynamics (*fs* time scale). It is, indeed necessary, to have a dynamic representation of the system to derive a more realistic characterization of water mediated HBNs. This is not an easy task due to the large number of possible hydrogen bonding partners of titratable amino acid residues present in PSII and to the vast amount of water molecules ($\sim$130000), present in our setup, that can be transiently part of the water-mediated HBN. We used our data analysis tool and the data filtering procedure described in the method section, to identify stable water wires in the PSII wild type and mutants MD trajectories. For our analysis we considered the side chains of the following titratable amino acid residues to be possible water mediated H-bonding partner: aspartate, glutamate, histidine, serine, threonine, and tyrosine for a total of 657 amino acids residues. A proton transfer path can include both direct and water-mediated sidechain-sidechain H-bonds. For this reason also direct sidechain-sidechain H-bonds, between residues belonging to our selection, have been included in this analysis.

3.3.1 Water dynamics in the manganese cluster region

The 1.9 Å resolution crystal structure [115] provided the position of more than 1300 water molecules per PSII monomer (1457 water molecules in monomer 1 and 1338 in monomer 2). The possible role of these water molecules is deeply discussed in Ref. [65]. In monomer 1, the 63% of the crystal waters are located on the protein surface. The 80% of the remaining 545 internal water molecules are located at the luminal side of PSII, among the extrinsic proteins PsbO, PsbU and PsbV and the loops of the core subunits, D1, D2, CP43 and CP47 [65]. The dehydrated environment and the cryogenic temperature at which crystals were prepared, can strongly affect water dynamics. We simulated PSII in a hydration state and at room temperature. We are interested in observing how do crystal waters behave in such environment.

Here I present an analysis of the dynamics of the water molecules present in the crystal structure along the trajectory of the wild type PSII (Sim1). The analysis is focused on the region surrounding the manganese cluster. It aims to understand at which extent the X-ray resolved water molecules do stay close to their initial positions (crystal structure positions) during 150 ns simulation, thus may playing a structural role within PSII, and, on the contrary, at which extent do they exchange with bulk water molecules.

Taking a radius of 20 Å from the manganese cluster, we define a region that extends to D1, D2 and PsbO subunits interface at the luminal side (region 1). This region

contains 185 crystal waters. If we look at a smaller region (region 2), considering a radius of 10 Å around the manganese cluster, which extends to the D1-E65/D2-E312 glutamate pair position, the number of crystal waters drops to 42. We monitored water dynamics in these two regions along the whole trajectory lengths of Sim1. The average number of water molecules found in region 1 and region 2 is: 192 ± 6 and 45 ± 2 respectively. This means that the hydration of the region surrounding the manganese cluster does not show drastic changes due to the simulation environment. An interesting question to ask is whether the crystal water molecules populating region 1 and region 2 do exchange with bulk water along the trajectory or not. In Figure 3.11 you can observe how the number of crystal waters in region 1 and 2 varies over time.
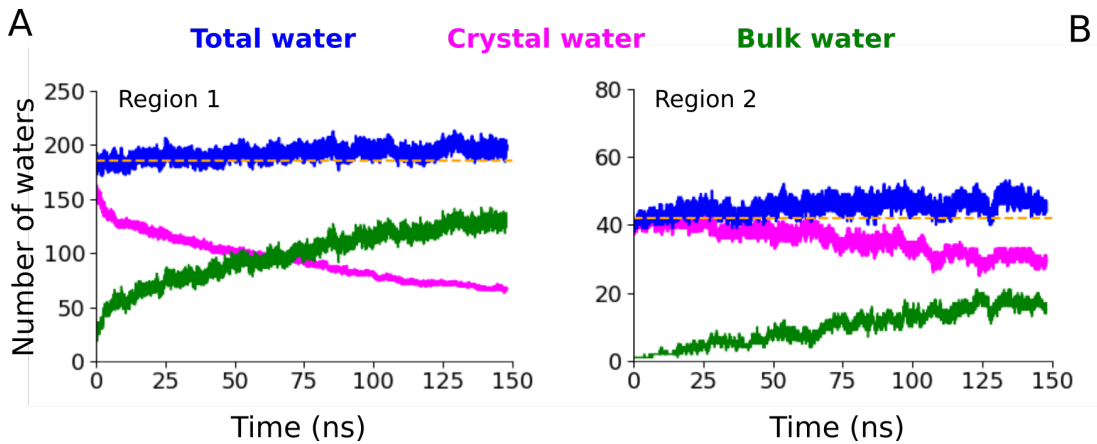
Figure 3.11: *Water dynamics at the OEC site. Number of water molecules over the whole lengths of Sim1 (wild type with D1-E65 protonated), within 2 specific regions: (A) region 1: volume defined by considering a radius of 20 Å from the manganese cluster; (B) region 2: volume defined by considering a radius of 10 Å from the manganese cluster. We computed the population of water molecules in these regions for the following sets: water molecules present in the crystal structure [115] (Crystal water), magenta line; water molecules added by us to hydrate the system (Bulk water), green line; the sum of the latter two sets (Total water), blue line. The orange dotted lines represent the number of water molecules found in the crystal structure in the 2 regions.*

The number of crystal waters in region 1 decreases with time reaching a plateau in the last 20 ns of the trajectory. The average number of crystal waters in this interval is 69 ± 4. This means that about 120 crystal water left region 1 within the time of our simulation. These water molecules have been replaced by water molecules coming from the bulk. In contrast, if we reduce the radius delimiting the region to 10 Å from the manganese cluster (region 2) we do not observe the same large exchange of water molecules with the bulk. Crystal water shows, indeed, to be more caged in this region,

which is not surprising due to the fact that these water molecules are more buried in the protein interior. Nevertheless, about 10 water molecules leave region 2 and are substituted by bulk waters during the time of the simulation. The average number of crystal water populating region 2 in the last 20 ns of the trajectory amounts to 30 ± 2. From this calculation we can infer that the most of the crystal water in region 2 could cover a structural role in PSII, for instance forming stable water-mediate HBNs.

Finally, in the crystal structure four water molecules coordinate the manganese cluster two at the Mn4 site (W1 and W2, following the nomenclature of Ref. [115]) and two at the calcium ion site (W3 and W4). Of these four crystal waters, we found that only W2 is still coordinating Mn4 towards the end of Sim1. The other 3 coordinating water molecules have been replaced by 2 other crystal waters and one water molecule coming from the bulk (Note that our simulation setup well conserves the water coordination at the Mn4 and calcium ion site).

3.3.2 Dynamic water-mediated H-bonded networks

A first inspection of the wild type PSII (Sim1) showed us a large and complex network of water mediated H-bonds connecting side chains of titratable amino acids groups. We computed 1201 water wires and direct H-bonds being present in at least 5% of the frames in the analyzed trajectory. In Figure 3.12 (A) we can observe how water wires are distributed. There are no water wires forming in the intra membrane part of the protein, whereas the most of the water wires are located on the bulk water exposed protein surface.

The latter result is a direct consequence of the topology of PSII surface. It, indeed, offers in large quantity the two elements necessary to form water wires, i.e. amino acids titratable groups side chains and water molecules. On the other hand we are interested in stable water wires. On the bulk water/protein interface, water wires are constantly broken and reformed with high frequencies. For this reason we filtered our results in order to get rid of transient and unstable water wires.

To this aim we computed a new occupancy rate by which a water wire is accounted as occupied any time that it is present for at least two consecutive frame without changing in its conformation, i.e. it has to be constituted by the same water molecules, following the same order in at least two consecutive frames. Using this data filtering procedure (described in greater details in section 2.3.3) the number of water wires sensibly reduced to about one third of the value computed for the unfiltered occupancy rate. You can observe in Figure 3.12 (B) that the most of the surface water wires
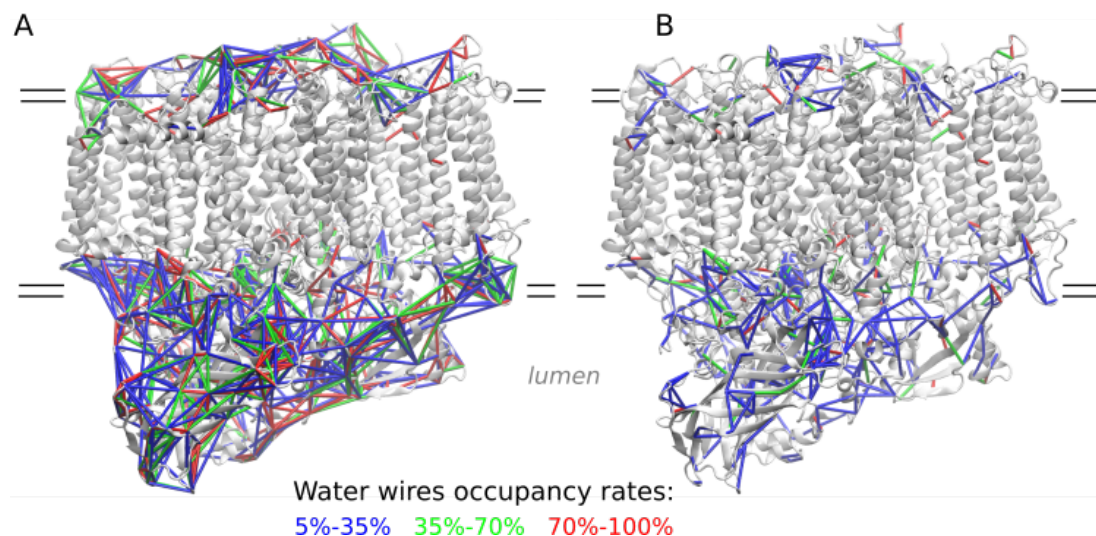
Figure 3.12: *Graphical representation of the water mediated HBN computed with our method (section 2.3.2) on the last 50 ns of the equilibrated trajectory. Water wires and direct H-bonds are depicted as colored lines connecting the C_α atoms of pairs of titratable amino acid groups. (A) The occupancy rate is computed as an occupancy frequency count of occupied water wires along the analyzed trajectory frames. (B) The occupancy rate is computed with our filtering method described in section 2.3.3*

have been filtered out, whereas the wires located in the protein interior still show an appreciable occupancy rate. Note in particular the high density of water wires present at the PsbO interface with the D1 and D2 protein subunits. Inspecting the 3D graphical representation of the water mediated HBN shown in Figure 3.12 (B) we were able to search for possible proton exit pathways.

In our inspection we considered only water wires with occupancy rates higher than 5% as reported in Figure 3.12 (B). We identified in Sim1 two stable water mediated HBNs connecting the manganese cluster to the luminal side of PSII protein complex, Figure 3.13, that could serve as possible proton exit pathways. In the next sections I will present in greater details the two HBNs that we called: PsbO HBN and CP43 HBN.

3.3.3 The PsbO H-bonding network

The PsbO HBN is constituted by four water wires which connect the first coordination shell of the manganese cluster to the surface of the PsbO subunit passing through the D1 and D2 subunits, Figure 3.14 (A), covering a distance of about 23 Å. It follows a
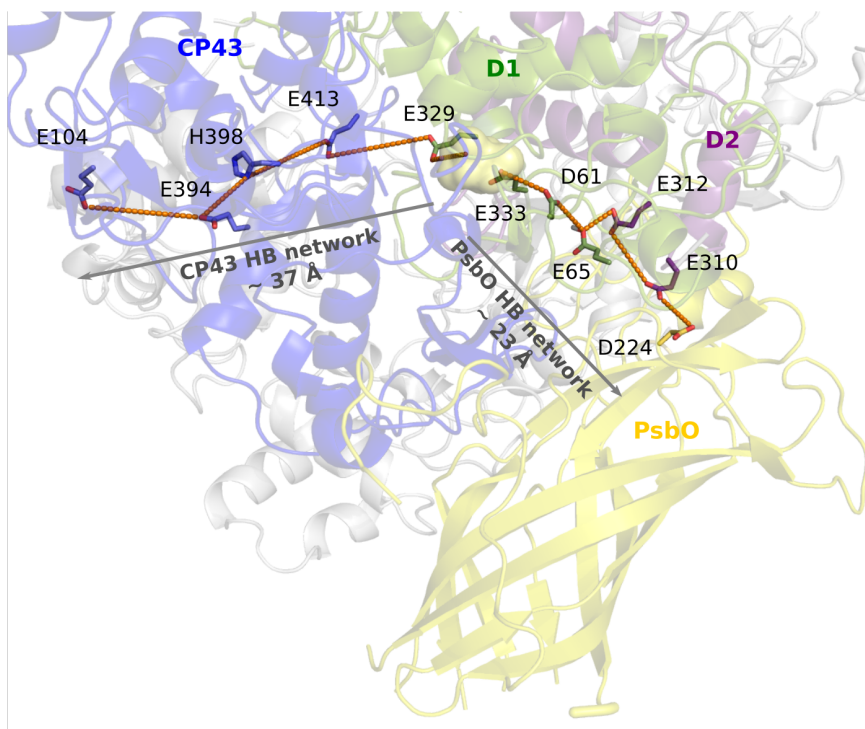
Figure 3.13: *We identified two HBNs connecting the manganese cluster to the lumen in the wild type PSII. The figure illustrates a coordinate snapshot from Sim1 (wild-type PSII with protonated D1-E65). Orange dotted lines represent water wires connecting pairs of titratable groups. One network connects the manganese cluster to the PsbO surface, spanning the D1 and D2 subunits (PsbO HBN). A second network extends mainly along the CP43 protein subunit (CP43 HBN).*

detailed description of the water wires constituting the network and their correspondent occupancy values, these data are summarized in Table A.1. The first wire connects the manganese cluster with the side chain of the D1-D61 amino acid residue. This connection can either be direct: a water wire connects Mn4 to D1-D61 with an occupancy rate 19.96%, or via an amino acid residue from the first coordination shell of the manganese cluster: D1-D170 with a rate of 26.58 % and D1-E333 which shows the highest occupancy rate: 67.94 %, and we therefore named it as Wire 1 in PsbO HBN, Figure 3.14. Wire 2 links D1-D61 to D1-E65 with an occupancy rate of 94.17%. D1-E65 samples H-bond distances with D2-E312 along the trajectory, Figure 3.9. D2-E312 is connected to D2-E310 via Wire 3 with an occupancy rate of 50.06%. Note that, at this stage, D2-E312 is already connected with the D222 and D224 amino acid residues, located on the water exposed surface of the PsbO subunit, via water wires that count lower occupancy rates with respect to Wire 3: 33.33 % and 44.98 % respectively. Finally, Wire 4 connects D2-E310 to the PsbO subunit surface at the PsbO-D224 site with an

occupancy rate of 95.63%. D2-E310 is the amino acid residue, in PsbO HBN, that shows the highest connectivity with surrounding residues, it builds a total of 5 water wires. D2-E310 is, indeed, linked to the PsbO surface also via D222 (62.94 %) and D223 (14.29 %). Moreover it is strongly connected to the adjacent D2-D308 (82.67 %), which in turn links to the PsbO-D224 (46.75 %). The high occupancy rates computed for the water wires constituting PsbO HBN, as well as the large amount of water wires found to connect D2-E312 and D2-E310 to residues located on the surface of PsbO, are observations that support the possibility that PsbO HBN could actually serve as pathway to transport protons from the manganese cluster to the lumen during the catalytic cycle.

PsbO HBN shows a remarkable overlap with residues constituting previously proposed proton channels, summarized in Table 1.1 and Figure 1.3. In particular, all the residues
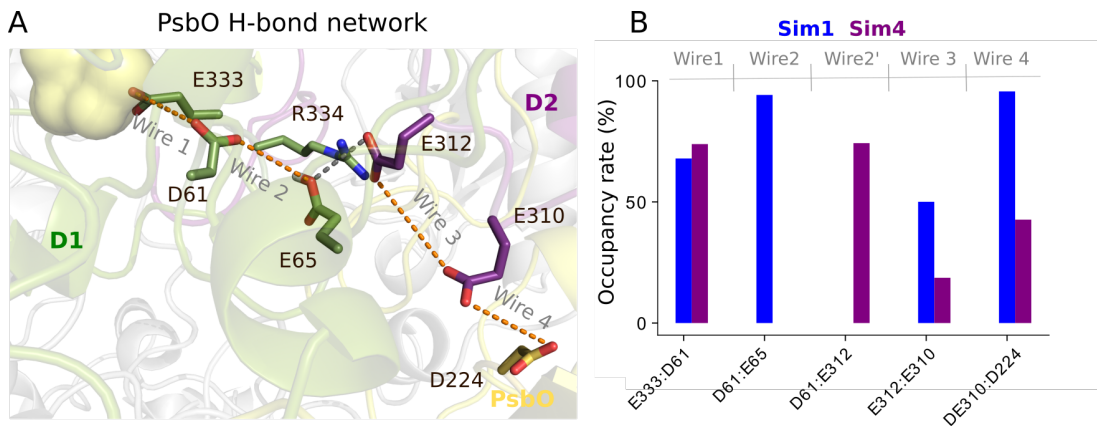
Figure 3.14: *The PsbO HBN [37]. (A) Coordinate snapshot from Sim1, wild-type PSII with protonated D1-E65 simulation. Orange dotted lines represent water wires connecting two titratable groups. A gray dotted line represent transient H-Bond distance between D1-D65 and D2-E312 siedechains. (B) Bar plot showing occupancy rates of water wires constituting PsbO HBN for the wild type with D1-E65 protonated (Sim1) and the D1-E65A mutant simulations (Sim4).*

involved in Wire 1 - Wire 4 are also part of channel (iii) proposed by Murray and Barber [82] and, with the exception of PsbO-D224, of the C channels by Gabdulkhakov et al. [32], Table 3.11. The latter channels were observed in the static 3.5 Å [27] and 2.9 Å [38] resolution crystal structures respectively. The resolution of these crystal structures is not high enough to reveal the position of water molecules in the protein interior. Thus, the search of proton channels in these structures was limited to detection of connected voids surrounded prevalently by hydrophilic and charged amino acid residues. On the contrary the 1.9 Å crystal structure [115], also used as starting coordinates for this work, displays the position of several water molecules. This

allowed for a search of proton transfer pathways that better resemble the biophysical properties of proton transport in protein. Protons move, via Grotthus mechanism [108, 1, 75, 56] , on water-mediated HBNs connecting side chains of titratable amino acid residues. Analyzing water-mediated HBNs in the Umena et al. crystal structure [115], Bondar and Dau discussed on the possible central role of the D1-E65/D2-E312 dyad in the proton transport process [11]. This glutamate pair has been, indeed, found to be part of a water-mediated HBN extending from D1-D61 to PsbO-D224, which is in great agreement with our results. Moreover, D1-D61, D1-E65 and D2-E312 have been proved experimentally [100, 99] to be part of an extended HBN including D1-E329 and D1-R334, involved in proton transfer. Site mutations of these amino acid residues show to slow down the catalytic cycle and to drastically reduce the oxygen evolution rate [100, 99] of PSII.

3.3.4 The CP43 H-bonding network

CP43 HBN links the manganese cluster to the CP43 subunit water exposed surface, Figure 3.15 (A), covering a distance of about 37 Å. It counts five water wires that show an overall lower occupancy rate compared to PsbO HBN, Figure 3.15 (B). D1-E329 is connected via water wires to the manganese cluster at three sites: Mn1 (Wire 1) , Mn2 and Ca with occupancy rate values of 27.25 %, 7.06 and 6.05 respectively. D1-E329 is then stability connected to CP43-E413 (Wire 2: 73.56 %). CP43-E413 and CP43-H398 are more than 10 Å apart, a water wire connecting them (Wire 3: 12.88 %) can contain between 3 and 5 water molecules, that could explain the low occupancy rate. The same argument would hold for Wire 5 connecting CP43-E394 and CP43-E104 (10.02 %).

Note that D1-E329 establishes a water-mediated H-bond with D1-H190 (12.60 %), which is constantly H-bonded to D1-Y161 (Y_Z). Both D1-H190 and D1-Y161 are linked to D1-E189, an element of the coordination shell of the manganese cluster, with occupancy rates: 31.06 % and 96.04 % respectively. This opens the possibility for an alternative proton pathway that starts at the D1-H190/Y_Z pair and then connects to the CP43 HBN via D1-E329.

The observation that Y_Z and D1-H190 are part of the CP43 HBN is in good agreement with experimental evidences suggesting that these two amino acid residues, which play a crucial role in the water oxidation process, could participate in a proton exit pathway [84, 83, 54].
Several amino acid residues belonging to the CP43 HBN were proposed to be part of
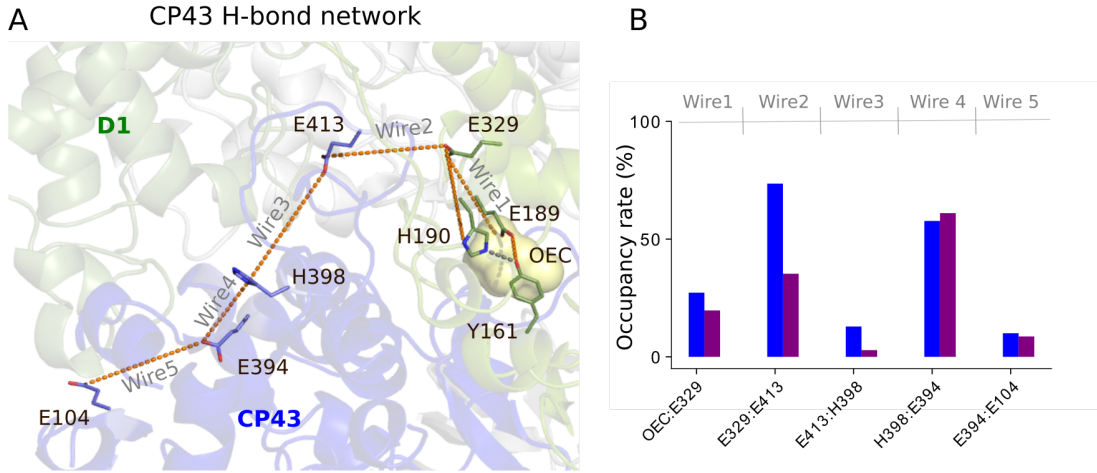
Figure 3.15: *The CP43 HBN [37]. (A) Coordinate snapshot from Sim1, wild-type PSII with protonated D1-E65 simulation (Sim1). Orange dotted lines represent water wires connecting two titratable groups. A gray dotted line represent a constant H-Bond between D1-Y161 and D1-H190 siedechains. (B) Bar plot showing occupancy rates of water wires constituting CP43 HBN for wild type with D1-E65 protonated (Sim1) and the D1-E65A mutant simulations (Sim4).*

a water channels in previous studies, see Table 3.11 and Table 1.1. Our simulations showed that water can access the manganese cluster following many pathways that finally merge at a large cavity surrounding the cluster. CP43-E413 faces this cavity and this could may be a contributing factor for the low occupancy rate computed for Wire 3 of CP43 HBN, since water in a highly hydrated cavity are more dynamic than in a dehydrated protein environment [28].

3.3.5 Water-mediated H-bonded networks dynamics in the *fs* time domain

Water molecules brake and reform H-bonds in the *fs* time scale [58, 28, 41]. This aspect need to be considered when discussing water wires stability. In the previous analysis we showed results obtained after applying our data filtering procedure on trajectory frames distant 10 *ps* from each other. Within this interval of time water molecules are able to sample large dynamics [101]. It is then necessary to verify the behavior of the previously identified water-mediated HBN in the *fs* time domain. To this aim we produced five independent short (100 *ps* long) *NVE* simulations this time with a coordinates printing frequency of 10 *fs*. We started these set of simulation from five equidistant coordinate snapshots of the wild type (Sim1) equilibrated trajectory (for

This work	PsbO HBN	CP43 HBN
Murray and Barber [82]	iii	ii
Ho and Styring [42]	("broad")	"large"
Gabdulkhakov et al. [32]	C, (D)	B1, (B2)
Umena et al. [115]	Fig. 4b	-
Vassiliev et al. [122]	1	4A, (4B)
Ogata et al. [87]	Path2	-

Table 3.11: *Proton or water channels found in other works which show an overlap with the water-mediated HBNs found in this study. The nomenclature used in this table is the original one used in the cited publications. Each column contains channels which are constituted of common amino acid residues, brackets indicate a low overlap. For more details on channels location in PSII and on their constituting amino acid residues, please see Table 1.1 and Figure 1.3.*

more details about the setup, please read section 2.3.4).
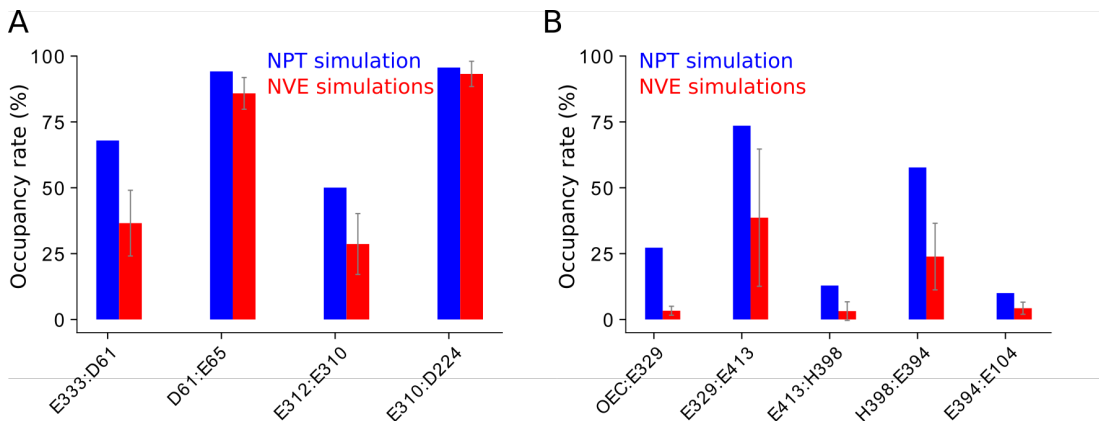
Figure 3.16: *Bar plot showing a comparison between the occupancy rates of water wires constituting PsbO (A) and CP43 (B) HBNs computed in the NPT vs. NVE simulations of the wild type PSII. Occupancy rates from the NPT simulation were computed over the last 48 ns of Sim1 (colored blue), whereas occupancy rates for the NVE simulations were computed as an average value of the five 100 ps long simulations (colored red).*

For this analysis we used different data filtering parameters respect to the *NPT*

simulations. We considered as occupied by a certain water wire only those frames in which a water wire is present without showing any change for at least 100 consecutive frames for a total time window of 1 *ps*. This choice is justified by the fact, that it has been computed in Ref. [91] that short periods of time, in the order of 1 *ps*, are enough for a proton to be transported along a stable water wire of the size considered in this work (not larger than 5 water molecules).

Finally, the water wires occupancy rates were averaged over the five trajectories.

We computed one water-mediated HBN that leads from the manganese cluster to the lumen, which corresponds to the PsbO HBN computed in the *NPT* wild type simulations. The occupancy rates of Wire 2 and Wire 4 are in great agreement with the *NPT* simulation whereas Wire 1 and Wire 3 show lower rates and a large standard deviation, indicating a large difference in the occupancy rates for these two wires in the five *NVE* trajectories, Figure 3.16. The presence of the PsbO HBN in the short *NVE* trajectories verifies the stability of this water-mediated HBN in the *fs* time scale.

From the analysis of the *NVE* simulations we observed that the PsbO HBN is the only HBN leading from the manganese cluster to the lumen, CP43 HBN is, in this case, broken (occupancy rate below 5%) at the connection between CP43-E413 and CP43-H398. The occupancy rate for this water wire in the *NPT* simulation is low (12.88%). Thus, the probability that this wire is occupied in one of the 100 *ps NVE* trajectories is extremely low. Due to the overall low occupancy rates computed in the *NPT* simulation for the water wires constituting CP43 HBN, we need to produce more short *NVE* trajectories starting at different points of the equilibrated Sim1 trajectory to have a set large enough to describe this water-mediated HBN.

3.3.6 Effect of the D1-E65A site mutations on the PsbO and CP43 H-bonded networks

D1-E65A mutation have been proved experimentally to slow down the catalytic cycle and to drop the O_2 evolution to 14 % of the wild type PSII [99, 100]. For this reason we are interested in simulating this site mutation and assess any eventual effect it could have on the water-mediated H-bonding networks found in the wild type.

Analysis of Sim4 shows that the D1-E65A single site mutations largely perturbs both PsbO and CP43 HBNs. Moreover we did not find any other alternative water-mediated HBN connecting the manganese cluster to the lumen.

Although with overall lower water wires occupancy rates, the D1-E65A mutant shows the same PsbO HBN found in Sim1; with the obvious difference that Wire 2 does not involve the D1-E65 side chain due to the fact that we mutated it into an alanine residue. However we computed a new water wire, Wire 2', which directly connects D1-D61 to D2-E312. We observed a drop of more than 30% and 50% in the water wire occupancy rates for Wire 3 and Wire 4 respectively, Figure 3.14 (B). This observation could be explained by differences between Sim1 and Sim4 in the local HBNs at the D2-E310 site.
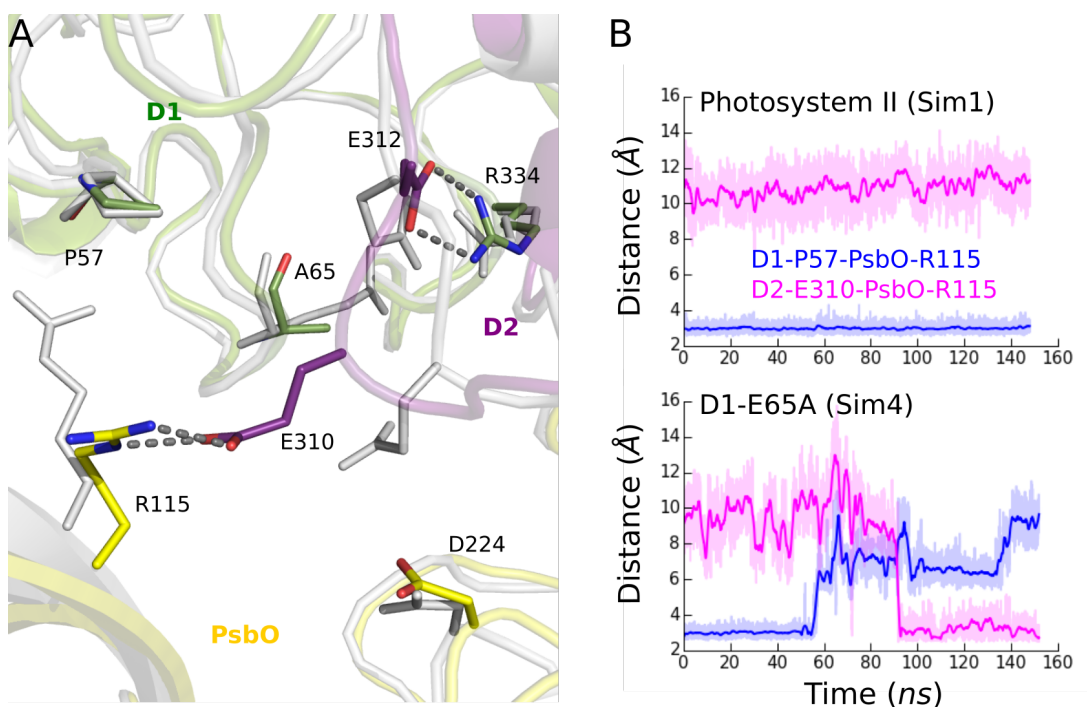
Figure 3.17: *Structural rearrangements could lead to lower occupancy rates in the water wires constituting the PsbO HBN in the D1-E65A mutant (Sim4) [37]. (A) Overlap between the Sim1 (white) and Sim4 (colored) simulations: close view at the D2-E312, PsbO-R115 sites in Sim4. Gray dotted lines represent H-bonds. PsbO-R115 exchange H-bond partner in the two simulation: D1-P57 in Sim1 and D2-E310 in Sim4. (B) Distance time series between PsbO-R115 sidechain and D1-P57 backbone oxygen (blue line), and between PsbO-R115 and D2-E310 sidechains (magenta line) along the whole trajectories: upper panel Sim1, lower panel Sim4. We used a Savitzky-Golay filter (window size = 301, order = 2) to smooth the distance data (thick lines).*

In Sim1 D2-E310 is not involved in any stable H-bond and its side chain is thus free to orient in a rotamer conformation that can favor the formation of Wire 3 and Wire 4, Figure 3.17 (A). Moreover, PsbO-R115 forms a steady H-bond with the backbone

oxygen of D1-P57. The latter H-bond is broken in Sim4 and PsbO-R115 forms a salt bridge with the D2-E310 side chain, Figure 3.17. This new H-bond pattern induces a strong reorientation of D2-E310 side chain. We conclude that the D1-E65A sensibly reduces the stability of PsbO HBN possibly by perturbing the local HBN at the D2-E310 site located about 10 Å away from the mutation site.

The D1-E65A simulation conserves all the water wires constituting CP43 HBN in Sim1 except for wire 3, which shows an occupancy rate lower than our cutoff of 5 %.
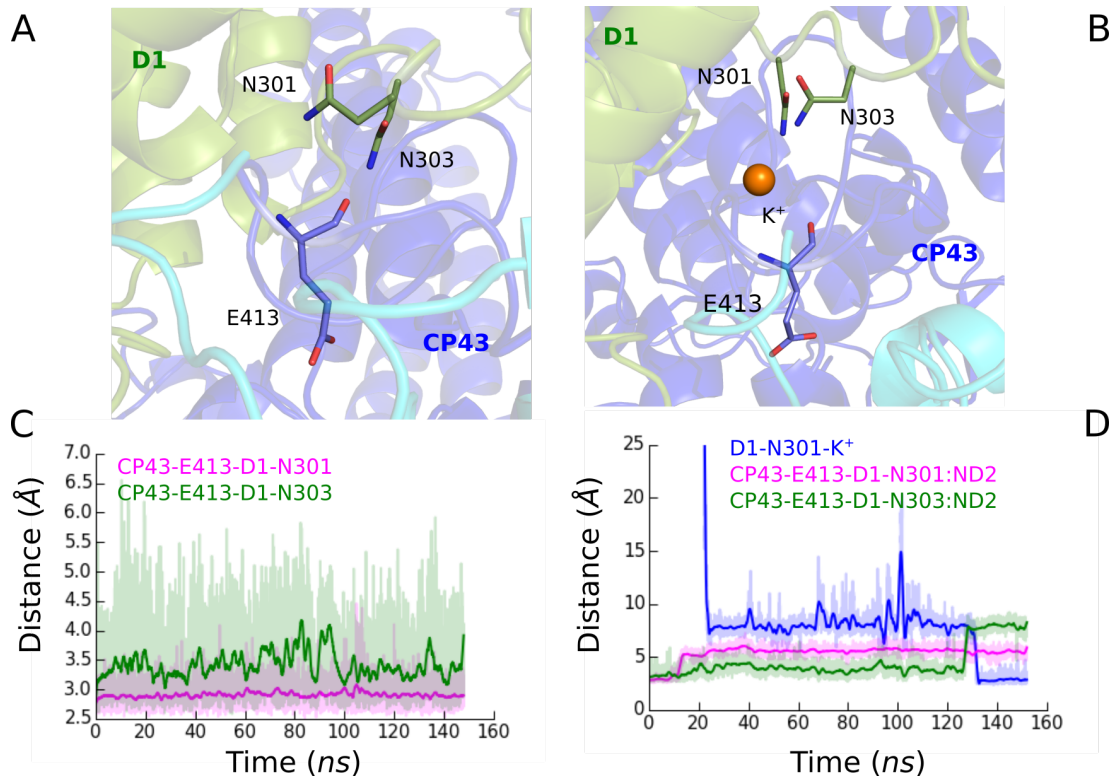
Figure 3.18: *Structural rearrangements could lead to lower occupancy rate in the water wires constituting the CP43 HBN in the D1-E65A mutant (Sim4). (A) Close view at the CP43-E413, D1-N301, D1-N303 site; left panel: Sim1, right panel: Sim4. Note that in Sim4 a Potassium ion reaches the protein interior from the solution; it is depicted as an orange VdW sphere. (B) Distance time series between Cp43-E413 backbone oxygen and D1-N301, D1-N303 sidechains (magenta and green line respectively); left panel: Sim1, right panel: Sim4. Note that the right panel includes the time series for the D1-N301 sidechain, potassium ion distance (blue line). We used a Savitzky-Golay filter (window size = 301, order = 2) to smooth the distance data (thick lines).*

Indeed, the path is broken at the CP43-E413 site. A possible explanation for this

behavior is the observed entrance of a potassium ion coming from the bulk solution into the protein interior. The potassium ion is coordinated by D1-N301 and D1-N303 which in Sim1 are H-bonded to the backbone oxygen of CP43-E413. The ion access the protein after about 20 *ns* of production run. This event immediately follows the loss of the H-bond between D1-N301 (located 35 A away from the mutation site) side chain and the CP43-E413 backbone oxygen, Figure 3.18. After 130 *ns* we observe the loss of the second H-bond formed by the backbone oxygen of CP43-E413, with the side chain of D1-N303. This event corresponds to a deeper penetration of the potassium ion in the protein interior. The ion is at this point steadily coordinated by D1-N301 and D1-N303 side chains, Figure 3.18. Sim4 is the only trajectory in which we observed a potassium ion accessing the protein interior from the bulk solution.

The strong perturbation induced by the D1-E65A mutation on PsbO and CP43 HBNs fits well with the experimental findings of Service et al. [99, 100]. Indeed, in their FTIR experiments, the catalytic cycle appears to slow down at the $S_2 \rightarrow S_3$ transition. During this transition one proton is transfered from the manganese cluster to the lumen [23]. If the PsbO and CP43 HBNs will be proven to be possible proton exit pathways, then we could infer from our simulation that D1-E65A would affect the proton transfer mechanism of PSII.

3.3.7   Concluding remarks

Crystal structures of PSII [128, 51, 27, 66, 38, 115], gave us invaluable insights into the protein function and architecture, but they have the limit of offering a static picture that does not account for protein breathing motions and water dynamics. With our simulations we were able to add value to the already proposed proton paths and HBNs, by characterizing their dynamics.

In particular we found two stable water-mediated HBNs connecting the manganese cluster with the lumen. The most stable of these two networks, PsbO HBN, spans the D1 and D2 domains before reaching the surface of the PsbO extrinsic subunit. Several amino acid residues participating in this network were previously proposed as being part of a possible proton exit pathway [82, 32, 42, 115, 11]. Specifically the D1-E65/D2-E312 pair was suggested as possible proton carrying dyad [11].
The other network, CP43 HBN, partially overlaps with elements of water channels proposed in previous studies [82, 42, 32, 122]. This network shows low occupancy rate

for Wire 3 connecting CP43-E413 to CP43-H398, this could depend on the fact that these amino acid residues are located in a large water cavity. We found that Y_Z and D1-H190 are part of the CP43 HBN. This result is in agreement with the proposal from Nakamura et al. [84, 83] and Klauss et al. [54] of Y_Z and D1-H190 being part of a proton exit pathway.

Finally we tested in silico the effect of the D1-E65A mutation which showed to strongly perturb PsbO and CP43 networks by reducing occupancy rates of several water wires constituting the network. The same mutation was proved experimentally to drastically influence PSII function by slowing down the catalytic cycle, by reducing the O_2 evolution to the 14 % of the wild type and by perturbing proton transfer events [99, 100].

3.4 PsbO absence perturbs the manganese cluster region

Part of the results and material presented in this chapter has been published in Ref. [37].

PsbO is a 26.5 KDa extrinsic protein subunit of PSII. Due to its highly conservation in all mesophiles and thermophiles and to its β-barrel shape it has been largely studied in the past years. Nevertheless, the role of the PsbO subunit is still poorly understood. There is experimental evidence that the removal of the PsbO subunit from the protein complex has a strong impact on its oxygen evolution rate, which can drop to the 35% of the wild type PSII [17]. Moreover, it has been shown that in the PsbO depleted PSII, the $S_2 \rightarrow S_3$ transition is slowed down [80]. Finally, due to the high concentration of charged residues on its surface, PsbO has been recently suggested to work as a proton antenna [68]. We performed simulations of the PSII protein complex removing the PsbO subunit (ΔPsbO) in order to get insights into its functional role by comparison with simulations of the wild type PSII. We were able to quantify the following effects brought by PsbO removal: the region surrounding the manganese cluster appears to be more flexible and more water accessible compared to the wild type PSII, moreover one of the two chloride ions (Cl^-1) located in the second coordination shell of the manganese cluster, necessary for the oxygen evolution [127, 52], is released to the bulk within the first nanoseconds of the (ΔPsbO) simulation.

3.4.1 Effects on hydration and protein flexibility

The removal of PsbO subunit shows a strong effect both on the hydration and the flexibility of the protein region surrounding of the manganese cluster. We computed the Root Mean Squared Fluctuations (RMSF) for all the protein atoms located within 10 Å of the manganese cluster both in the wild type and in the ΔPsbO simulations. In Figure 3.19 (A) you can see the result of this calculation displayed in histogram form. Each histogram represent the RMSF distribution for all the atom in the selection for each simulation respectively. You can observe a shift toward higher RMSF values in the ΔPsbO simulation. This means that the residues surrounding the manganese cluster are more dynamic in the PsbO-depleted PSII.

 As a first step to measure the impact of PsbO depletion on the hydration state of PSII, we computed the amount of water molecules within a hydrogen bonding distance
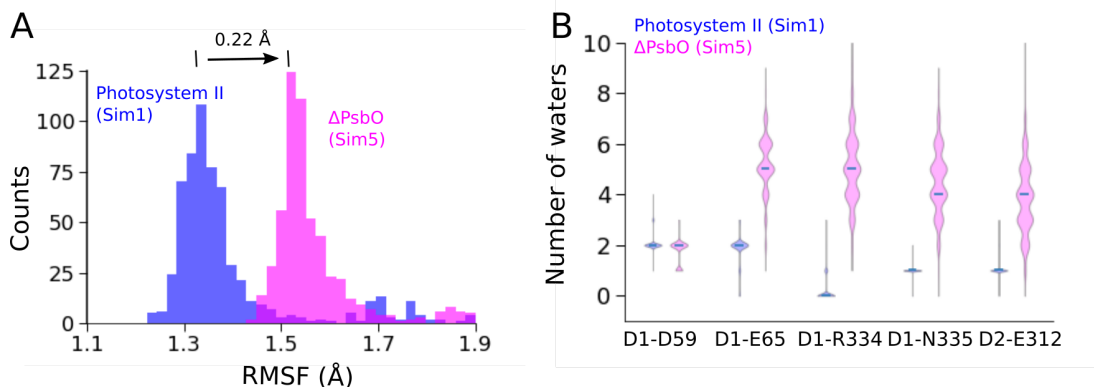
Figure 3.19: *Effects of PsbO depletion on protein hydration and flexibility [37]. In both panel A and B: wild type results are colored blue and ΔPsbO magenta. (A) RMSF for all protein atoms within 10 Å of the manganese cluster. (B) Violin plots representing the distribution, over the last 24 ns of the trajectories, of the number of water molecules within 3.5 Å of the sidechains of the following residues: D1-D59, D1-E65, D1-R334, D1-N335, D2-R334.*

of selected amino acid residues involved in the HBN displayed in Figure 3.8. The result of this analysis are displayed in Figure 3.19 (B). The violin plots show how the hydration level of the selected amino acid side chains stays constant in the wild type PSII, whereas oscillates spanning large intervals of values in the ΔPsbO simulation. Moreover we observe that PsbO depletion correlates with an overall increase in the amount of water molecules present in the first hydration shell of the selected residues. This results indicate that the hydration of the analyzed region is higher in the ΔPsbO simulation.

It was shown by experiments that PsbO depletion destabilizes the manganese cluster leading to the leakage of two of the four manganese ions [79]. Our results could provide a molecular interpretation of the experimental observations, since higher flexibility and hydration level in the region of the manganese cluster can be two factors contributing to the cluster destabilization.

3.4.2 One chloride ion leaves the second coordination shell of the manganese cluster

The 1.9 Å crystal structure [115] revealed the presence of two chloride ion cofactor, Cl^-1 and C^l-2, located about 7 Å aways from the manganese cluster and coordinated by D2-K317 and CP43-E354 respectively , Figure 1.4. We know from experiments that chloride depleted PSII is not able to advance above the $S_2 \to S_3$ [127] and that replacing Cl^- with I^- kills the protein's photo-cycle. In the ΔPsbO simulation we observe

the leaking of Cl^-1 to the bulk. The release event occurs within the first 7 ns of the trajectory. It takes 3.5 ns for the chloride ion to reach the protein surface, and other 3.5 ns to be released to the bulk, Figure 3.20. In the equilibrated wild type simulation, Cl^-1 is steadily coordinated by D2-K317 and D1-R334.
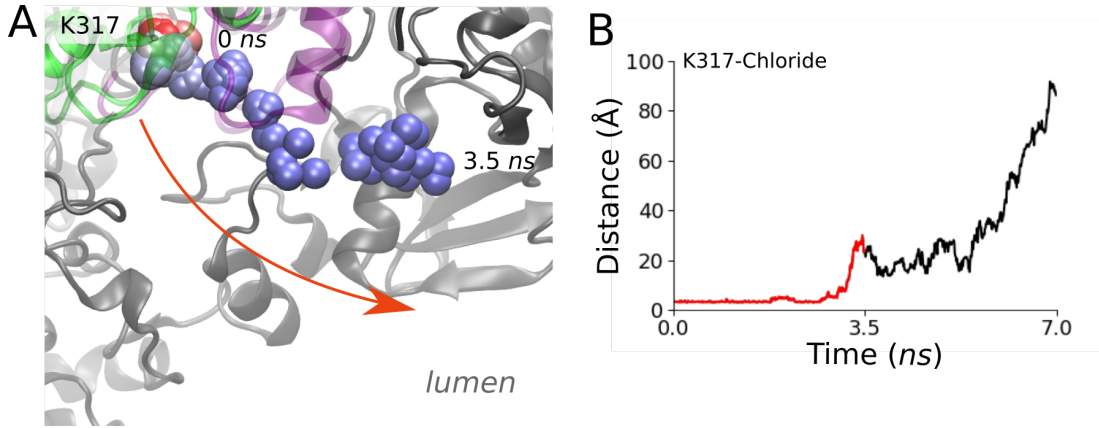
Figure 3.20: Cl^-1 *leaves the second coordination shell of the manganese cluster [37]. (A) Close view at the chloride ion binding site from Sim5. Positions of the chloride ion are shown simultaneously from 0 ns to 3.5 ns with a frequency of 10 ps. A RWB (red, white, blue) color gradient scheme allow to follow in time the chloride ion location. At time 0 ns (red VdW sphere) the chloride ion is coordinated by D2-K317, whereas at time 3.5 ns (blue VdW sphere) the ion has left its coordination pocket and faces the bulk water at the luminal side of the protein. (B) Time series of the D2-K317/chloride ion distance, during the first 10 ns of Sim5. In red are highlighted the first 3.5 ns of the trajectory, the same interval depicted in the molecular graphics of panel (A).*

In the ΔPsbO simulation Cl^-1 is coordinated by D2-K317 and D1-R334 side chains, by the D1-E333 backbone nitrogen and by 2 water molecules during the first 1.5 ns of the production run. During this interval of time the number of water molecules in the first hydration shell of Cl^-1 increases from 2 to 4, with a consequent break of the D1-E333 coordination bond. Within the successive ns another water adds to Cl^-1 hydration shell breaking the coordination bond with D2-K317. Finally the chloride ion is fully hydrated (8 water molecules in the first hydration shell), breaks the last coordination bond with D1-E334 and reaches the bulk at the protein's luminal side. This cascade of events could be correlated with the higher hydration and flexibility showed by ΔPsbO in the region of the manganese cluster.

This result is consistent with the experimental work from Popelkova et al. [92],

in which it is shown that the steady state Cl^- affinity of PsbO-depleted PSII is low compared to the wild type PSII.

3.4.3 Concluding remarks

We investigated the functional role of the PsbO extrinsic subunit by comparing PSII in absence of PsbO (ΔPsbO) with wild type PSII simulations. Our simulations revealed that PsbO absence increases protein flexibility in the region surrounding the manganese cluster and the hydration level of several amino acid residues in the second and third coordination shell of the manganese cluster. These results may correlate with the experimental evidence that the PsbO extrinsic subunit stabilizes the manganese cluster [79]. Indeed, a more flexible and hydrated region could favor the experimental observation of manganese ions leaking to the bulk [79] upon PsbO removal.

The higher hydration level observed in the ΔPsbO simulation couples with the release to the bulk of one of the two chloride ions in the second coordination shell of the manganese cluster (Cl^-1), which leaves its coordination pocket at the D1-K317 site within the first 7 ns of the trajectory. This result is consistent with the experimental finding of lower steady state Cl^- affinity in PsbO-depleted PSII in comparison with the wild type [92]. We were able to determine the path that conducts Cl^-1 from the protein interior to the bulk. The event of Cl^-1 release is of high importance in the context of PSII function. It is, indeed, known that removal of chloride ions blocks PSII photo cycle and consequently the oxygen evolution [127].

Conclusions and perspectives

During each reaction cycle, PSII transports four protons from the manganese cluster to the thylakoid lumen. In biomolecules, protons are transfered via Grotthus mechanism along water wires connecting titratable amino acid residues. In PSII, the identity of these amino acid groups and the dynamics of water-mediated HBNs involved in the proton transfer process, are still open questions. Here, we approached these questions by means of all atom classical MD simulations.

To this aim we first derived a new set of CHARMM force field parameters for three types of PSII cofactors: chlorophyll-a, pheophytin-a and plastoquinone-9 [35], which proved to well describe cofactors dynamics both in gas phase and in the protein matrix. We then employed these parameters to perform simulations of wild type and mutant PSII embedded in a hydrated lipid bilayer.

To facilitate efficient data analysis of HBNs, we developed a data analysis tool for fast tracking of H-bonds, that allows for detection of all water wires linking all the possible pairs of titratable amino acid groups at each simulation time step. We detected, in wild type photosystem II, two water-mediated HBNs connecting the manganese cluster to the lumen, that we named PsbO and CP43 HBNs. PsbO HBN reaches the surface of the PsbO subunit spanning the D1 and D2 subunits, Figure 3.14 (A) and CP43 HBN links the manganese cluster to the CP43 subunit water exposed surface, Figure 3.15 (A).

Possible proton exit channels have been previously proposed based either on crystal structure inspection [82, 32, 42, 115, 11] or on analysis of hydration dynamics in short (10 ns) MD simulations [87], as summarized in Figure 1.3 and Table 1.1. Amino acid groups, belonging to some of these channels, overlap with elements of the water-mediated HBNs computed in this work, Figure 4.1. In particular D1-D61, D1-E65 and D2-E312 constitute the core of PsbO HBN and are common elements to all of the previously proposed proton channels. These elements have also been proven experimentally to be involved in proton transfer [99, 100]. On the contrary, none of the the previously proposed proton exit pathways contain amino acid residues of the CP43

73

subunit. CP43-E413 and CP43-H398 of the CP43 HBN are, instead, part of water channels proposed by Murray and Barber [82], Ho and Styring [42], Gabdulkhakov et al. [32] and Vassiliev et al. [122].

By simulating the D1-Glu65Ala mutation we observed perturbations in the occupancy rate values of the water wires constituting the two water-mediated HBNs. The perturbation showed to be stronger in the PsbO HBN than in the CP43 HBN, in particular at the D2-Glu312 and D2-Glu310 sites. The D1-Glu65Ala mutation proved, experimentally, to perturb proton transfer and drastically reduce the O_2 evolution rate of PSII [99]. This experimental evidence could be related to the perturbation of the proton exit pathway observed in our simulations.

A second aspect analyzed in this work is the role of the PsbO extrinsic protein subunit of PSII. PsbO is the only extrinsic subunit conserved in all the photosynthetic organisms. There are experimental evidence that the removal of the PsbO subunit from the protein complex has a strong impact on its O_2 evolution rate, which can drop to the 35% of the wild type PSII [17]. Nevertheless, the function of PsbO is not yet fully understood. To get insights in this respect, we produced a simulation of PSII in absence of PsbO and we compared it with the wild type PSII simulation.

We observed increased protein flexibility (Figure 3.19 (A)) and higher hydration (Figure 3.19 (B)) in the region surrounding the manganese cluster when PsbO is absent. These results might correlate with the experimental evidence that the PsbO extrinsic subunit stabilizes the manganese cluster [79]. Indeed, a more flexible and hydrated region could favor the experimental observation of manganese ions leaking to the bulk upon PsbO removal [79]. Higher water accessibility to the manganese cluster region also associate to the release to the bulk of a chloride ion (Figure 3.20) known to be required for oxygen evolution [127, 52]. This result is in good agreement with the experimental finding of lower steady state Cl^- affinity in PsbO-depleted PSII in comparison with the wild type [92]. We were able to determine the path that conducts the chloride ion from its coordination site at the D2-K317 to the bulk.
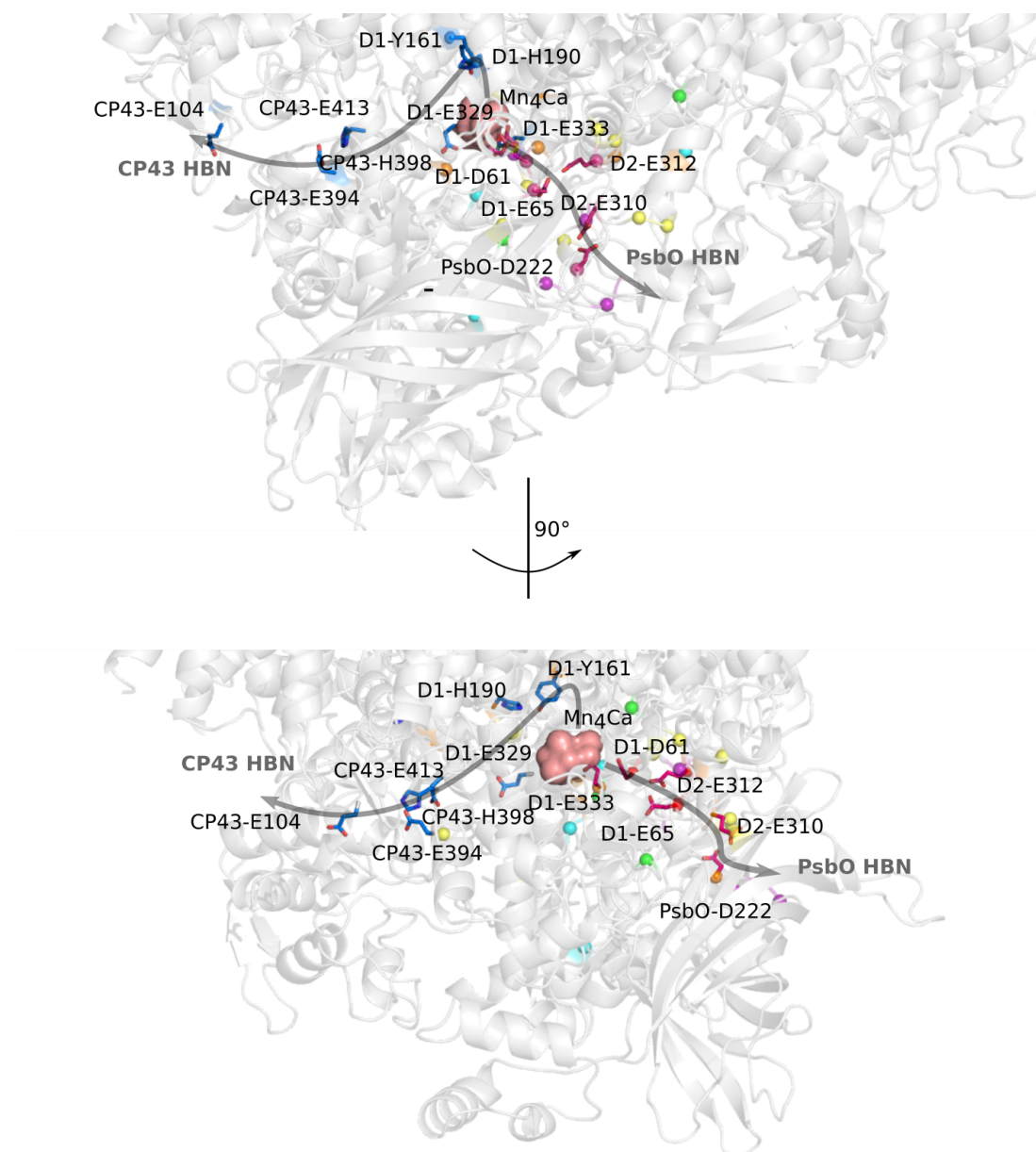
Figure 4.1: *PsbO and CP43 HBNs overlap with previously proposed proton exit channels. Molecular graphics of PSII luminal side based on the coordinates of the 1.9 Å resolution crystal structure [115]. Amino acid groups belonging to PsbO and CP43 HBN are represented in red and blue licorice respectively. Colored spheres represent C_α atoms of titratable amino acid residues that constitute proposed proton exit channels shown in Figure 1.3. Amino acid groups common to all the proposed channels are colored red and amino acid groups common to at least two proposed proton transfer pathways are colored orange. Amino acid groups belonging to one single proton exit path have a different color per each Refs: Murray and Barber [82] (purple), Gabdulkhakov et al. [32] (green), Umena et al. [115] (cyan), Bondar and Dau [11] (yellow), Ogata et al. [87] (blue).*

This work provide foundation to further investigate proton exit pathways in PSII. The two stable water-mediated HBNs connecting the manganese cluster to the lumen, revealed by our simulations, define a specific protein region on which we could focus future calculations. To assess if these two HBNs are suitable for proton transfer it is required to account for the presence of a hydrated extra proton and compute proton transfer energy barriers along the water-mediated HBNs steps. To perform this calculations, a higher level of theory and computation, with respect to all atom classical MD simulations, has to be employed. Examples are provided by studies on proton transfer in cytochrome c oxydase [61] or in bacteriorhodopsin [12].

Our analysis was performed using as starting coordinates the crystal structure of PSII in the S_1 state [115]. A recent study explored the response of PSII structure to changes in the catalytic state [109]. In particular changes around the PQ_B and OEC areas have been observed. These changes could perturb the PsbO and CP43 HBNs. Repeating our analysis starting from coordinates of PSII in different S-states, would tell us more about how much the PsbO and CP43 HBNs are conserved throughout the photocycle.

Appendix

A.1 Trajectories' RMSD

We computed the RMSD (Root Mean Square Deviation) for the C_α atoms of our simulations, Figure A.1. In the wild type (Sim1-3) equilibrated trajectories, RMSDs computed for the whole protein C_α atoms, converge at values between 1.5 and 2 Å. These small values could be correlated with the fact that a large part of photosystem II consists of hydrophobic transmembrane helices harboring a chlorophyll molecules. Both the lipid bilayer and the chlorophyll cofactors are likely to constrain protein motions.

The D1-E65A mutant (Sim4) and the $\Delta PsbO$ (Sim5) simulations show somewhat higher RMSD values. In particular the $\Delta PsbO$ simulation converges at about 2.5 Å. These higher values likely originate from the induced perturbations, site mutation and PsbO absence respectively. The perturbed systems equilibrated in a different configuration with respect to the wild type protein.

D1-E65A mutant appears to be the less converged among the five simulations performed. The RMSD value computed for the CP43 subunit increases towards the end of the trajectory affecting the total RMSD. This simulation should be prolonged to access higher convergence.
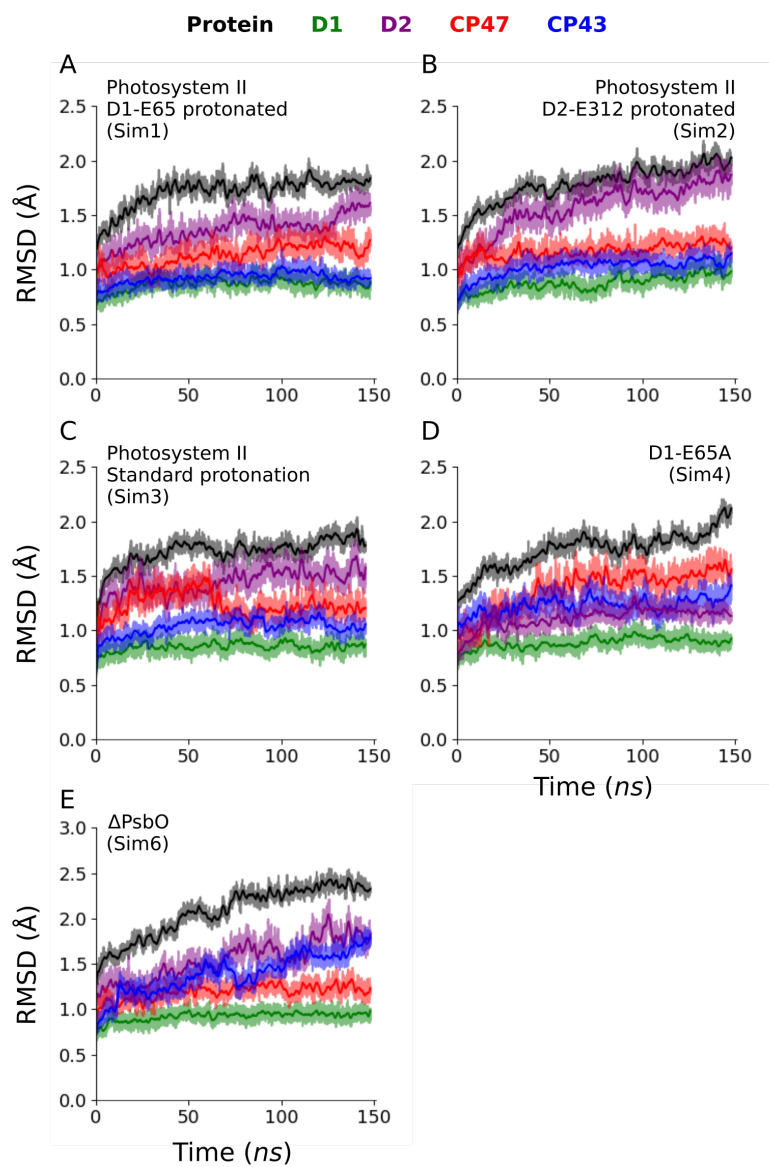
Figure A.1: *RMSD profiles of photosystem II C_α atoms for the whole protein (black line) and the core subunits: D1 (green), D2 (purple), CP43(blue) and CP47 (red). We used a Savitzky-Golay filter (window size = 301, order = 2) to smooth the RMSD data (thick lines).*

A.2 OEC coordination shell

Our wild type simulation (Sim1) shows a stable manganese cluster coordination shell along the whole equilibrated trajectory, Figure A.2. All the amino acid groups found in the crystal structure [115] to be coordinating the manganese cluster, maintain their coordination throughout the whole trajectory. The C_α-C_α distances of coordinating amino acid residues located on opposite sites of the manganese cluster shows displacements of less than 1 Å with respect to the crystal structure configuration A.2 (B), indicating that our simulations are able to describe a stable manganese cluster region.
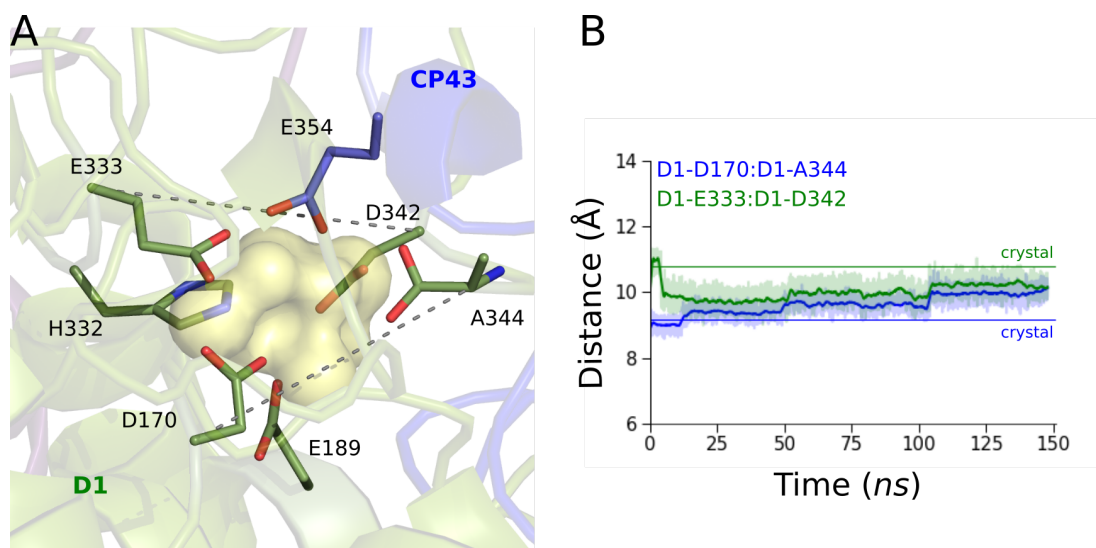
Figure A.2: *Stability of the OEC coordination shell. Panel A: snapshot from the last 10 ns of Sim1 representing the coordinating residues (licorice) of the manganese cluster (yellow surface). Gray dashed lines mark the C_α-C_α distances depicted as time series in panel B. We used a Savitzky-Golay filter (window size = 301, order = 2) to smooth the distance data (thick lines).*

A.3   Inter-subunit H-bond difference graph
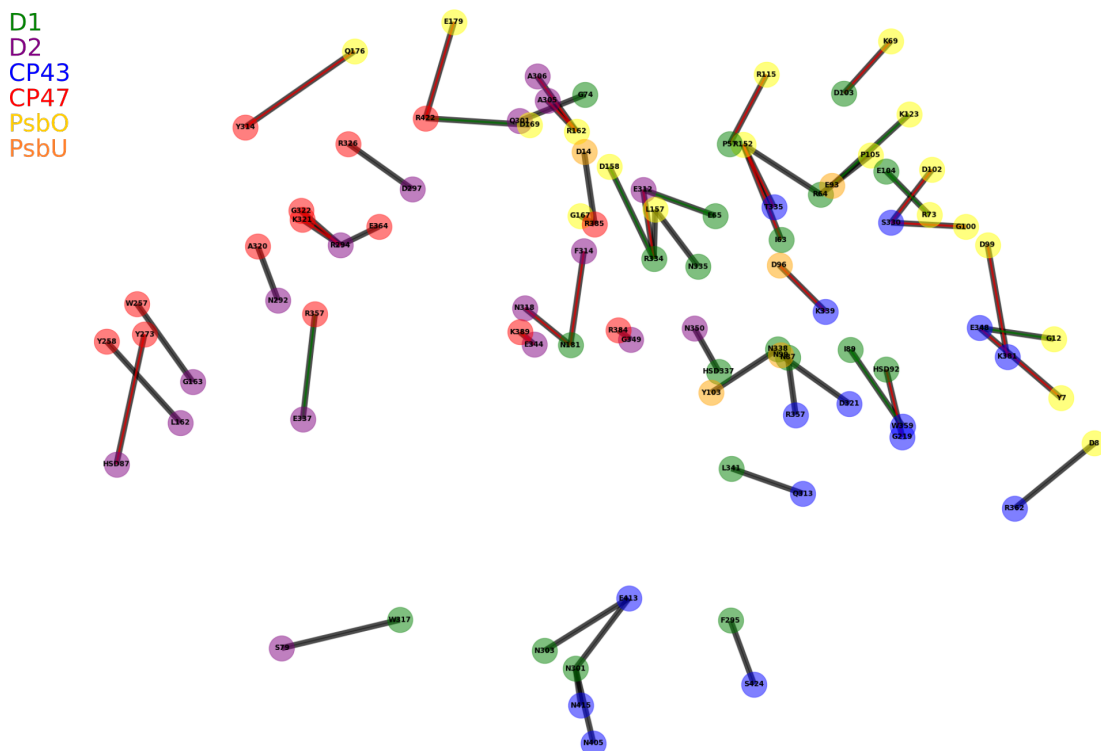
Figure A.3: *Visualization of the inter-subunit H-bond difference graph of Sim1 vs. Sim2. Nodes are amino acid groups involved in stable H-bonds. We define an H-bond as been stable when it is present in more than 60% of the analyzed trajectory. Black edges represent H-bonds conserved in the two simulations. Red edges represent bonds present in Sim1 but not in Sim2 and, on the contrary, green edges represent H-bonds present in Sim2 but not in Sim1.*

A.4 Reproducibility of the water-mediated H-bond networks analysis

We produced repeated simulations for the wild type (Sim1') and the D1-E65A mutant (Sim4'). The lengths of these simulations is 100 *ns*. Even though 50 *ns* shorter than their correspondent Sim1 and Sim4, we used these two simulations to prove the reproducibility of the water-mediated H-bond network analysis. Also in this case we performed the analysis on the last 48 *ns* of the trajectory.

	Water wires occupancy rates (%)			
PsbO **H-bond network**	**Wild Type** **(Sim 1)**	**Repeated** **Sim1**	**D1-E65A** **(Sim4)**	**Repeated** **Sim4**
Wire 1				
D1-E333—D1-D61	67.94	25.98	73.88	37.03
D1-D170—D1-D61	26.58	63.08	32.17	10.81
OEC-MN4—D1-D61	19.96	67.44	-	-
Wire 2/Wire2'				
D1-D61—D1-E65	94.17	79.06	-	-
D1-D61—D2-E312	-	80.56	74.27	73.92
Wire 3				
D2-E312—D2-310	50.06	54.46	18.69	20.58
D2-E312—PsbO-D222	33.33	48.75	13.29	14.48
D2-E312—PsbO-D224	44.98	43.79	15.44	19.79
Wire 4				
D2-E310—PsbO-D224	95.63	81.02	42.69	40.33
D2-E310—PsbO-D222	62.94	87.69	25.52	32.19
D2-E310—PsbO-D223	14.29	4.48	-	-
D2-E310—D2-D308	82.67	83.79	96.33	86.98
Wire 5				
D2-D308—PsbO-D224	46.75	72.90	53.73	21.21

Table A.1: Occupancy rates for the water wires constituting the PsbO HBN, discussed in section 3.3.3. All the possible wires connecting the manganese cluster to the lumen for Sim1, Sim1', Sim4 and Sim4' are listed in the table.

The repeated simulations confirmed the existence of only two water-mediated H-

bond networks connecting the manganese cluster to the lumen which correspond to the PsbO and CP43 HBNs found in Sim1, discussed in sections 3.3.3 and 3.3.4. The occupancy rates for the water wires constituting the PsbO HBN are in great agreement with Sim1 and Sim4 respectively, Table A.1. Sim1' shows significant differences from Sim1 only in the D1-D61 partner in Wire1 and in the occupancy rate of Wire2', which in Sim1' has high occupancy rates. Nevertheless, and considering the different lengths of the repeated simulations, the overall agreement in the water wires occupancy rates is remarkable.

A.5 Hydration dynamics of functionally important amino acid residues
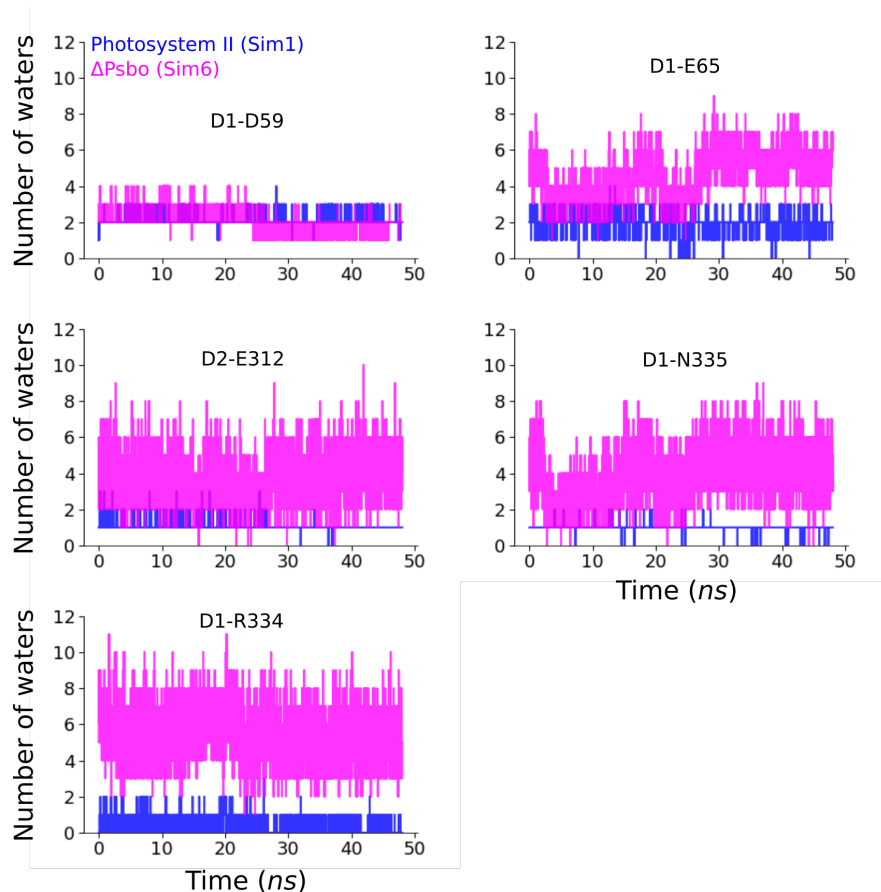
Figure A.4: *Time series of the number of water molecules computed within H-bond distance (3.5 Å) of specific residues sidechains donor or acceptor atoms, over the last 50 ns of the trajectory. The time series for D1-D59, D1-E65 and D1-N335 show to converge just on the last 24 ns.*

A.6 PsbO absence perturbs the wild type H-bond network at the D1-E65/D2-E312 pair site

We observed marked rearrangements of interactions involving protein groups which, in wild-type photosystem II (Sim1) participate in the PsbO H-bonded network (Figures 3.8). D2-R334, which in wild-type photosystem II interacts with PsbO-D158, reorients
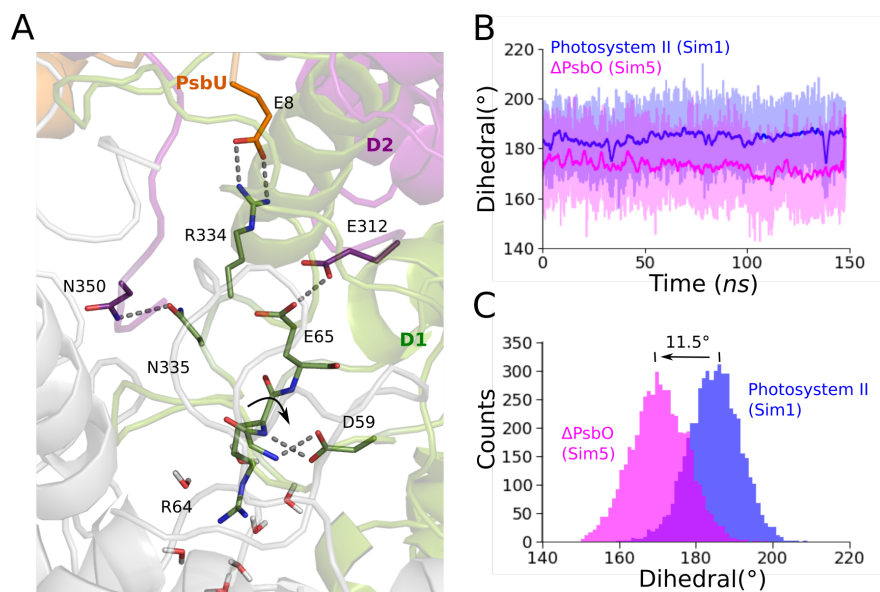
Figure A.5: *PsbO absence perturbs the wild type HBN at the D1-E65/D2-E312 glutamate pair site. (A) Molecular graphics of the HBN at the D1-E65/D2-E312 glutamate pair site depicted in Figure 3.8, snapshot from Sim5. (B-C) Dihedral torsion for the bond indicated by an arrow in figure 6C, computed in Sim1 and Sim6. Time series (panel B) and histogram (panel C) We used a Savitzky-Golay filter (window size = 301, order = 2) to smooth the dihedral data (thick lines)..*

and interacts with PsbU-E8, Figure A.5 (A), such that the interaction between D1-E65 and D2-E312 is strengthened. The altered interaction between D1-E65 and D2-E312 is accompanied by changes in the local structure and dynamics of the protein segment that holds D1-E65: In wild-type photosystem II, the backbone carbonyl groups of D1-E65 and D1-I63 H-bond to PsbO-R152, whereas their amide groups H-bond to D1-D59 A.5 (A). In ΔPsbO, absence of the two R152/backbone H-bonds associates with a ∼ 11° clockwise rotation of the protein backbone towards D2-D59. Pursuant to these considerations, we suggest that PsbO-R152 participates in a sidechain/backbone structural motif that helps control the local structure and dynamics at the D1-E65 site,

which is thought to be particularly important for proton transfer.

A.7   Topology files for Chl-a, Phe-a and Pl-9

I append here the topology files for chlorophyll-*a*, pheophytin-*a* and plastoquinone-9 derived in this work. The topology files include parameters for atoms masses, atoms partial charges, improper angles and a declaration of the bond defining the geometry of the molecule. All the other parameters required to run classical MD simulations of the these molecules are included in separate files (prm files), which I did not append here due to their large size. In case you are interested in using these files, please contact us at the addresses written in the header of the topology files, we will be glad to send them to you.

```
* Topology File for chlorophyll-a
*
! When using this top/par files please cite:
! 1. the original CHARMM papers for protein and Heme
!    K. Kuczera, Kuriyan J. and Karplus M,  Temperature Dependence of the
Structure and Dynamics of Myoglobin,
!    J. Mol. Biol. (1990) 213, 351-373
! 2. A. Damjanovic, I. Kosztin, U. Kleinekathöfer, K. Schulten, Excitons in a
photosynthetic light-harvesting system: a
!    combined molecular dynamics, quantum chemistry, and polaron model study,
Phys. Rev. E 65 (2002) 031919.
! 3. N. Foloppe, J. Breton, J.C. Smith, Potential energy function for
photosynthetic reaction centre chromophores: energy
!    minimisations of a crystalline bacteriophytin A analog, in: J. Brenton, A.
Vermeglio (Eds.), The Photosynthetic
!    Bacterial Reaction Center II, Plenum Press, New York, 1992.
! 4. N. Foloppe, M. Ferrand, J. Breton, J.C. Smith, Structural model of the
photosynthetic reaction center of rhodobacter
!    capsulatus, Proteins: Struct. Funct. Gen.22 (1995) 226–244.
! 5. Guerra F., Adam S. and Bondar A-N.,
!    Revised force-field parameters for chlorophyll-a, pheophytin-a and
plastoquinone-9,
!    Journal of Molecular Graphics and Modelling, 58 (2015) 30-39.
!
!  to take care of unwanted dihedral angles, the user may want to add an IC
table and a patch
!  similar to FHEM in toppar_all36_prot_heme.str
!
!  Please report bugs to Ana-Nicoleta Bondar, nbondar@zedat.fu-berlin.de, and
Federico Guerra: fguerra@zedat.fu-berlin.de
!
37  1


MASS  1  HAN    1.00800 ! nonpolar H
MASS  2  CTN   12.01100 ! tetrahedral C
MASS  3  CPAN  12.01100 ! heme alpha-C
MASS  4  CPBN  12.01100 ! heme beta-C
MASS  5  CPMN  12.01100 ! heme meso-C
MASS  6  CT2N  12.01100 ! tetrahedral C for reduced pyrrols
MASS  7  C2    12.01100 ! polar C for ketone in bpheo ring E
MASS  8  CPM2  12.01100 ! carbone meso de bpheo lie au cycle 5
MASS  9  CPM3  12.01100 ! carbone meso de bpheo lie au cycle 5
MASS  10 CPA2  12.01100 ! ring D of bpheo
MASS  11 CPA3  12.01100 ! ring D of bpheo
MASS  12 CPA4  12.01100 ! ring D of bpheo
MASS  13 NPH1  14.00700 ! heme pyrrole N
MASS  14 NPH2  14.00700 ! heme pyrrole N
MASS  15 NPH3  14.00700 ! heme pyrrole N
MASS  16 NPH4  14.00700 ! heme pyrrole N
MASS  17 OK    15.99940 ! ketone oxygen
MASS  18 MG    24.30500 !
MASS  19 CT2   12.01100 ! carbon of methylene group (-CH2-)
MASS  20 CT3   12.01100 ! carbon of methyl group (-CH3)
MASS  21 HA2    1.00800 ! alkane, CH2, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS  22 HA3    1.00800 ! alkane, CH3, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS  23 CE1   12.01100 ! for alkene; RHC=CR
MASS  24 CE2   12.01100 ! for alkene; H2C=CR
MASS  25 HE1    1.00800 ! for alkene; RHC=CR
MASS  26 HE2    1.00800 ! for alkene; H2C=CR
MASS  27 HA1    1.00800 ! alkane, CH, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS  28 CT1   12.01100 ! aliphatic sp3 C for CH
```

```
MASS  29 C      12.01100 ! MASS 145  CL   12.011000 C ! carbonyl C (acetic
acid/methyl acetate)
MASS  30 OS    15.99940 ! MASS 157  OSL  15.999400 O ! ester oxygen
MASS  31 OB    15.99940 ! MASS 155  OBL  15.999400 O ! acetic acid carboxyl
oxygen (e. to protein OB)


! -----------------------------------------------------------------

DEFAult FIRST none LAST none
AUTOgenerate ANGLes DIHEdrals


RESI CLA            0.00! 4-liganded chlorophyll-a
GROUP
ATOM MG     MG      1.17 !
ATOM CHA    CPM2    0.00 !                          HED1  HED2 HED3
ATOM NA     NPH1   -0.58 !                            \    |   /
ATOM C1A    CPAN    0.38 !                             CED
ATOM C2A    CT2N   -0.43 !                              |
ATOM H2A    HAN     0.33 !                             O2D   O1D
ATOM C3A    CT2N   -0.38 !                              \\ //
ATOM H3A    HAN     0.28 !                              CGD
ATOM C4A    CPAN    0.40 !        O2A   O1A             |
ATOM NB     NPH2   -0.60 !         \\ //                |  HBD
ATOM C1B    CPAN    0.26 !         CGA                  | /
ATOM C2B    CPBN   -0.12 !          |                  _CBD_
ATOM C3B    CPBN   -0.11 !     HBA1--CBA--HBA2        /      \
ATOM C4B    CPAN    0.26 !          |                /        \
ATOM NC     NPH3   -0.59 !     HAA1--CAA-HAA2  __ CHA       CAD--OBD
ATOM C1C    CPA4    0.29 !          |        /      \        /
ATOM C2C    CPBN   -0.10 !     H2A--C2A---C1A       C4D---C3D
ATOM C3C    CPBN   -0.10 !          |      |          |      |
ATOM C4C    CPA2    0.26 ! HMA1\    |      |          |      |      /HMD1
ATOM ND     NPH4   -0.51 ! HMA2-CMA--C3A     NA      ND    C2D--CMD-HMD2
ATOM C1D    CPA2    0.34 ! HMA3/    /  \   / \      / \   /      \HMD3
ATOM C2D    CPBN   -0.27 !        H3A   C4A   \    /   C1D
ATOM C3D    CPBN   -0.11 !               /     \  /      \
ATOM C4D    CPA3    0.15 !       HBH--CHB        MG        CHD--HDH
ATOM CAD    C2      0.47 !              \      /   \       /
ATOM OBD    OK     -0.44 !               C1B  /     \  C4C    HAC1 HCA2
ATOM CBD    CTN    -0.55 ! HMB1\        /  \ /       \ /  \    | /
ATOM HBD    HAN     0.30 ! HMB2-CMB--C2B    NB       NC    C3C--CAC
GROUP                    ! HMB3/      |      |        |      |     \   /HBC1
ATOM CHB    CPMN   -0.10 !            |      |        |      |      CBC-HBC2
ATOM HBH    HAN     0.10 !          C3B---C4B      C1C---C2C        \HBC3
GROUP                    !            |       \_CHC_/      |
ATOM CHC    CPM3   -0.10 !           CAB         |        CMC--HMC3
ATOM HCH    HAN     0.10 !          // \        HC       / |
GROUP                    !         CBB  HAB           HMC1  HMC2
ATOM CHD    CPMN   -0.10 !         /  \
ATOM HDH    HAN     0.10 !      HBB1   HBB2
GROUP !GLUP              !
ATOM CAA    CT2    -0.18 !
ATOM HAA1   HA2     0.09 !
ATOM HAA2   HA2     0.09
ATOM CBA    CT2    -0.21
ATOM HBA1   HA2     0.09
ATOM HBA2   HA2     0.09
ATOM CGA    C       0.75
ATOM O1A    OB     -0.17
ATOM O2A    OS     -0.55    !here attach the alkyl chain
GROUP
ATOM CAB    CE1    -0.15    ! as for ATOM CAC   CE1    -0.15 from heme
ATOM HAB    HE1     0.15    ! ATOM HAC   HE1     0.15 from heme
```

```
GROUP
ATOM CBB   CE2    -0.42   ! ATOM CBB   CE2     -0.42 in heme
ATOM HBB1  HE2     0.21   ! ATOM HBB1  HE2      0.21 in heme
ATOM HBB2  HE2     0.21   ! ATOM HBB2  HE2      0.21 in heme


!carboxyl+methyl

GROUP
ATOM CGD   C       0.72
ATOM O1D   OB     -0.55
ATOM O2D   OS     -0.17
ATOM CED   CT3    -0.27
ATOM HED1  HA3     0.09
ATOM HED2  HA3     0.09
ATOM HED3  HA3     0.09


!C2H-methyl

GROUP
ATOM CAC   CT2    -0.18
ATOM HAC1  HA2     0.09
ATOM HAC2  HA2     0.09
GROUP
ATOM CBC   CT3    -0.27
ATOM HBC1  HA3     0.09
ATOM HBC2  HA3     0.09
ATOM HBC3  HA3     0.09


!Methyl groups
GROUP
ATOM CMA   CT3    -0.27
ATOM HMA1  HA3     0.09
ATOM HMA2  HA3     0.09
ATOM HMA3  HA3     0.09


GROUP
ATOM CMB   CT3    -0.27
ATOM HMB1  HA3     0.09
ATOM HMB2  HA3     0.09
ATOM HMB3  HA3     0.09


GROUP
ATOM CMC   CT3    -0.27
ATOM HMC1  HA3     0.09
ATOM HMC2  HA3     0.09
ATOM HMC3  HA3     0.09


GROUP
ATOM CMD   CT3    -0.27
ATOM HMD1  HA3     0.09
ATOM HMD2  HA3     0.09
ATOM HMD3  HA3     0.09


!alkyl chain

GROUP
ATOM C1    CT2    -0.18
ATOM HP11  HA2     0.09
ATOM HP12  HA2     0.09
GROUP
ATOM C2    CE1    -0.09
ATOM  HP2  HE1     0.09
ATOM C3    CE1    -0.00
GROUP
```

```
ATOM C4    CT3     -0.27
ATOM HP41  HA3      0.09
ATOM HP42  HA3      0.09
ATOM HP43  HA3      0.09
GROUP
ATOM C5    CT2     -0.18
ATOM HP51  HA2      0.09
ATOM HP52  HA2      0.09
GROUP
ATOM C6    CT2     -0.18
ATOM HP61  HA2      0.09
ATOM HP62  HA2      0.09
GROUP
ATOM C7    CT2     -0.18
ATOM HP71  HA2      0.09
ATOM HP72  HA2      0.09
GROUP
ATOM C8    CT1     -0.09
ATOM HP8   HA1      0.09
GROUP
ATOM C9    CT3     -0.27
ATOM HP91  HA3      0.09
ATOM HP92  HA3      0.09
ATOM HP93  HA3      0.09
GROUP
ATOM C10   CT2     -0.18
ATOM H101  HA2      0.09
ATOM H102  HA2      0.09
GROUP
ATOM C11   CT2     -0.18
ATOM H111  HA2      0.09
ATOM H112  HA2      0.09
GROUP
ATOM C12   CT2     -0.18
ATOM H121  HA2      0.09
ATOM H122  HA2      0.09
GROUP
ATOM C13   CT1     -0.09
ATOM HP13  HA1      0.09
GROUP
ATOM C14   CT3     -0.27
ATOM H141  HA3      0.09
ATOM H142  HA3      0.09
ATOM H143  HA3      0.09
GROUP
ATOM C15   CT2     -0.18
ATOM H151  HA2      0.09
ATOM H152  HA2      0.09
GROUP
ATOM C16   CT2     -0.18
ATOM H161  HA2      0.09
ATOM H162  HA2      0.09
GROUP
ATOM C17   CT2     -0.18
ATOM H171  HA2      0.09
ATOM H172  HA2      0.09
GROUP
ATOM C18   CT1     -0.09
ATOM HP18  HA1      0.09
GROUP
ATOM C19   CT3     -0.27
ATOM H191  HA3      0.09
ATOM H192  HA3      0.09
ATOM H193  HA3      0.09
```

```
      GROUP
      ATOM C20  CT3    -0.27
      ATOM H201  HA3    0.09
      ATOM H202  HA3    0.09
      ATOM H203  HA3    0.09



         ! For porphyrin ring :
      BOND NA    MG       NB  MG  NC  MG  ND  MG
      !     -for pyrrol A:
      BOND NA       C1A  C1A  C2A  C2A  C3A  C3A  C4A
      BOND C2A  H2A   C3A  H3A
      BOND C4A   NA
      !     -for pyrrol B:
      BOND NB       C1B  C1B  C2B   C2B  C3B  C3B  C4B
      BOND C4B   NB
      !     -for pyrrol C:
      BOND NC       C1C  C1C  C2C   C2C  C3C  C3C  C4C
      BOND C4C   NC
      !     -for pyrrol D:
      BOND ND       C1D  C1D  C2D   C2D  C3D  C3D  C4D
        BOND C4D   ND
      !     -for methylene bridges:
      BOND C4A  CHB   CHB  C1B   C4B  CHC  CHC  C1C
      BOND C4C  CHD    CHD  C1D   C4D  CHA  CHA  C1A
      BOND CHB  HBH    CHC  HCH   CHD  HDH
      !     -for ring E and its substituants:
      BOND C3D  CAD  CAD  OBD   CBD CHA
      BOND CBD  CAD   CBD HBD




      !     -for pyrrol A substituants:

      BOND C3A  CMA      CMA  HMA1   CMA  HMA2   CMA  HMA3
      BOND C2A  CAA   CAA  HAA1      CAA  HAA2
      BOND CAA  CBA    CBA  HBA1      CBA  HBA2
      BOND CBA  CGA    CGA  O1A       CGA  O2A
      BOND O2A  C1     C1  HP11       C1  HP12

      !     -Phytyl chain:

      BOND C1   C2      C2  HP2     C2  C3
      BOND C3   C4      C4  HP41   C4  HP42    C4    HP43
      BOND C3   C5      C5  HP51   C5  HP52
      BOND C5   C6      C6  HP61   C6  HP62
      BOND C6   C7      C7  HP71   C7  HP72
      BOND C7   C8      C8  HP8
      BOND C8   C9      C9  HP91   C9  HP92    C9  HP93
      BOND C8   C10     C10 H101   C10  H102
      BOND C10  C11     C11 H111   C11  H112
      BOND C11  C12     C12 H121   C12  H122
      BOND C12  C13     C13  HP13
      BOND C13  C14     C14  H141   C14  H142   C14  H143
      BOND C13  C15     C15  H151   C15  H152
      BOND C15  C16     C16  H161   C16  H162
      BOND C16  C17     C17  H171   C17  H172
      BOND C17  C18     C18  HP18
      BOND C18  C19     C19  H191   C19  H192   C19  H193
      BOND C18  C20     C20  H201   C20  H202   C20  H203

      !     -For pyrrol B substituants:
```

```
BOND C2B  CMB       CMB  HMB1   CMB  HMB2   CMB  HMB3
BOND C3B  CAB       CAB  HAB    CAB  CBB
BOND CBB  HBB1   CBB  HBB2


!     -for pyrrol C substituants:

!BOND C2C  H2C      C3C  H3C
BOND C2C  CMC       CMC  HMC1   CMC  HMC2   CMC  HMC3
BOND C3C  CAC    CAC  HAC1   CAC  HAC2
BOND CAC  CBC    CBC  HBC1   CBC  HBC2   CBC  HBC3


!     -for pyrrol D substituants

BOND C2D  CMD       CMD  HMD1   CMD  HMD2   CMD  HMD3



!        -for ring E substituants:
BOND CBD CGD CGD O1D CGD O2D O2D CED
BOND CED HED1 CED HED2 CED HED3




!     IMPROPRES :

!     type CPAN NPHN     CPAN CPBN : 2 enleves pour CT2
IMPR   C4A NA C1A C2A       C1A NA C4A C3A
IMPR   C4B NB C1B C2B       C1B NB C4B C3B
IMPR   C4C NC C1C C2C       C1C NC C4C C3C
IMPR   C4D ND C1D C2D       C1D ND C4D C3D


!    type NPHN CPAN CPBN CPBN : 2 enleves pour CT2
IMPR   NB C1B C2B C3B       NB C4B C3B C2B
IMPR   ND C1D C2D C3D       ND C4D C3D C2D


!     type NPHN CPAN CPMN CPA
IMPR   NA C1A CHA C4D       NA C4A CHB C1B
IMPR   NB C1B CHB C4A       NB C4B CHC C1C
IMPR   NC C1C CHC C4B       NC C4C CHD C1D
IMPR   ND C1D CHD C4C       ND C4D CHA C1A


!     type CPMN CPAN CPAN HA
IMPR   CHB C1B C4A HBH      CHC C1C C4B HCH       CHD C1D C4C HDH


!     type CPAN CPBN CPMN NPH
IMPR   C1A C2A CHA NA       C4A C3A CHB NA
IMPR   C1B C2B CHB NB       C4B C3B CHC NB
IMPR   C1C C2C CHC NC       C4C C3C CHD NC
IMPR   C1D C2D CHD ND       C4D C3D CHA ND


!     type NPHN CPAN CPAN MG
IMPR   NA C1A C4A MG
IMPR   NB C1B C4B MG
IMPR   NC C1C C4C MG
IMPR   ND C1D C4D MG


ACCEPTOR OBD CAD

END
```

```
* Topology File for pheophytin-a
*
! When using this top/par files please cite:
! 1. the original CHARMM papers for protein and Heme
!    K. Kuczera, Kuriyan J. and Karplus M,  Temperature Dependence of the
Structure and Dynamics of Myoglobin,
!    J. Mol. Biol. (1990) 213, 351-373
! 2. A. Damjanovic, I. Kosztin, U. Kleinekathöfer, K. Schulten, Excitons in a
photosynthetic light-harvesting system: a
!    combined molecular dynamics, quantum chemistry, and polaron model study,
Phys. Rev. E 65 (2002) 031919.
! 3. N. Foloppe, J. Breton, J.C. Smith, Potential energy function for
photosynthetic reaction centre chromophores: energy
!    minimisations of a crystalline bacteriophytin A analog, in: J. Brenton, A.
Vermeglio (Eds.), The Photosynthetic
!    Bacterial Reaction Center II, Plenum Press, New York, 1992.
! 4. N. Foloppe, M. Ferrand, J. Breton, J.C. Smith, Structural model of the
photosynthetic reaction center of rhodobacter
!    capsulatus, Proteins: Struct. Funct. Gen.22 (1995) 226–244.
! 5. Guerra F., Adam S. and Bondar A-N.,
!    Revised force-field parameters for chlorophyll-a, pheophytin-a and
plastoquinone-9,
!    Journal of Molecular Graphics and Modelling, 58 (2015) 30-39.
!
!
!  Please report bugs to Ana-Nicoleta Bondar, nbondar@zedat.fu-berlin.de, and
Federico Guerra: fguerra@zedat.fu-berlin.de
!
37  1


MASS  1     HAN    1.00800 ! nonpolar H
MASS  2     CTN   12.01100 ! tetrahedral C
MASS  3     CPAN  12.01100 ! heme alpha-C
MASS  4     CPBN  12.01100 ! heme beta-C
MASS  5     CPMN  12.01100 ! heme meso-C
MASS  6     CT2N  12.01100 ! tetrahedral C for reduced pyrrols
MASS  7     C2    12.01100 ! polar C for ketone in bpheo ring E
MASS  8     CPM2  12.01100 ! carbone meso de bpheo lie au cycle 5
MASS  9     CPM3  12.01100 ! carbone meso de bpheo lie au cycle 5
MASS  10    CPA2  12.01100 ! ring D of bpheo
MASS  11    CPA3  12.01100 ! ring D of bpheo
MASS  12    CPA4  12.01100 ! ring D of bpheo
MASS  13    NPH1  14.00700 ! heme pyrrole N
MASS  14    NPH2  14.00700 ! heme pyrrole N
MASS  15    NPH3  14.00700 ! heme pyrrole N
MASS  16    NPH4  14.00700 ! heme pyrrole N
MASS  17    OK    15.99940 ! ketone oxygen
MASS  18    CT2   12.01100 ! carbon of methylene group (-CH2-)
MASS  19    CT3   12.01100 ! carbon of methyl group (-CH3)
MASS  20    HA2    1.00800 ! alkane, CH2, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS  21    HA3    1.00800 ! alkane, CH3, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS  22    CE1   12.01100 ! for alkene; RHC=CR
MASS  23    CE2   12.01100 ! for alkene; H2C=CR
MASS  24    HE1    1.00800 ! for alkene; RHC=CR
MASS  25    HE2    1.00800 ! for alkene; H2C=CR
MASS  26    HA1    1.00800 ! alkane, CH, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS  27    CT1   12.01100 ! aliphatic sp3 C for CH
MASS  28    C     12.01100 ! MASS 145  CL   12.011000 C ! carbonyl C (acetic
acid/methyl acetate)
MASS  29    OS    15.99940 ! MASS 157  OSL  15.999400 O ! ester oxygen
MASS  30    OB    15.99940 ! MASS 155  OBL  15.999400 O ! acetic acid carboxyl
```

```
oxygen (e. to protein OB)

! --------------------------------------------------------------

DEFAult FIRST none LAST none
AUTOgenerate ANGLes DIHEdrals



RESI PHO             0.00 !Pheophytin
GROUP
ATOM CHA   CPM2    0.00 !                              HED1  HED2 HED3
ATOM CHB   CPMN   -0.10 !                                 \   |  /
ATOM CHC   CPM3   -0.10 !                                  CED
ATOM HCH   HAN     0.10 !                                   |
ATOM CHD   CPMN   -0.10 !                                  O2D  O1D
ATOM NA    NPH1   -0.31 !                                   \\ //
ATOM C1A   CPAN    0.20 !                                    CGD
ATOM C2A   CT2N   -0.45 !         O2A    O1A                  |
ATOM H2A   HAN     0.32 !          \\ //                      |   HBD
ATOM C3A   CT2N   -0.34 !           CGA                       | /
ATOM H3A   HAN     0.24 !            |                       _CBD_
ATOM C4A   CPAN    0.38 !       HBA1--CBA--HBA2            /       \
ATOM NB    NPH2   -0.59 !            |                   /          \
ATOM C1B   CPAN    0.25 !       HAA1--CAA-HAA2  ___ CHA          CAD--OBD
ATOM C2B   CPBN   -0.13 !            |          /      \          /
ATOM C3B   CPBN   -0.13 !        H2A-C2A---C1A          C4D---C3D
ATOM C4B   CPAN    0.30 !             |      |           |     |
ATOM NC    NPH3   -0.68 ! HMA1\       |      |           |     |       /HMD1
ATOM C1C   CPA4    0.28 ! HMA2-CMA--C3A     NA          ND     C2D--CMD-HMD2
ATOM C2C   CPBN   -0.11 ! HMA3/    /  \  /              / \   /      \HMD3
ATOM C3C   CPBN   -0.11 !        H3A   C4A            HND  C1D
ATOM C4C   CPA2    0.26 !             /                      \
ATOM ND    NPH4   -0.44 !      HBH--CHB                      CHD--HDH
ATOM C1D   CPA2    0.34 !            \                       /
ATOM C2D   CPBN   -0.27 !            C1B  HNB          C4C      HAC1 HCA2
ATOM C3D   CPBN   -0.11 ! HMB1\      /   \ /          /    \      | /
ATOM C4D   CPA3    0.10 ! HMB2-CMB--C2B     NB       NC     C3C--CAC
ATOM CAD   C2      0.44 ! HMB3/     |      |          |     |      \   /HBC1
ATOM OBD   OK     -0.38 !           |      |          |     |     CBC-HBC2
ATOM CBD   CTN    -0.56 !          C3B---C4B        C1C---C2C        \HBC3
ATOM HBD   HAN     0.32 !           |       \_CHC_/            |
ATOM HND   HAN     0.59 !          CAB           |          CMC--HMC3
ATOM HDH   HAN     0.10 !         //  \          HC        / |
ATOM HNB   HAN     0.59 !        CBB  HAB             HMC1  HMC2
ATOM HBH   HAN     0.10 !        /   \
GROUP!GLUP partial charge      HBB1  HBB2
ATOM CAA   CT2    -0.18 !
ATOM HAA1  HA2     0.09 !
ATOM HAA2  HA2     0.09
ATOM CBA   CT2    -0.21
ATOM HBA1  HA2     0.09
ATOM HBA2  HA2     0.09
ATOM CGA   C       0.75
ATOM O1A   OB     -0.17
ATOM O2A   OS     -0.55     !here attach the alkyl chain
GROUP
ATOM CAB   CE1    -0.15   ! as for ATOM CAC   CE1    -0.15 from heme
ATOM HAB   HE1     0.15   ! ATOM HAC   HE1     0.15 from heme
GROUP
ATOM CBB   CE2    -0.42   ! ATOM CBB   CE2    -0.42 in heme
ATOM HBB1  HE2     0.21   ! ATOM HBB1 HE2     0.21 in heme
ATOM HBB2  HE2     0.21   ! ATOM HBB2 HE2     0.21 in heme
```

```
!carboxyl+methyl

GROUP
ATOM CGD  C       0.72
ATOM O1D  OB     -0.55
ATOM O2D  OS     -0.17
ATOM CED  CT3    -0.27
ATOM HED1 HA3     0.09
ATOM HED2 HA3     0.09
ATOM HED3 HA3     0.09


!C2H-methyl

GROUP
ATOM CAC  CT2    -0.18
ATOM HAC1 HA2     0.09
ATOM HAC2 HA2     0.09
GROUP
ATOM CBC  CT3    -0.27
ATOM HBC1 HA3     0.09
ATOM HBC2 HA3     0.09
ATOM HBC3 HA3     0.09


!Methyl groups
GROUP
ATOM CMA  CT3    -0.27
ATOM HMA1 HA3     0.09
ATOM HMA2 HA3     0.09
ATOM HMA3 HA3     0.09


GROUP
ATOM CMB  CT3    -0.27
ATOM HMB1 HA3     0.09
ATOM HMB2 HA3     0.09
ATOM HMB3 HA3     0.09


GROUP
ATOM CMC  CT3    -0.27
ATOM HMC1 HA3     0.09
ATOM HMC2 HA3     0.09
ATOM HMC3 HA3     0.09


GROUP
ATOM CMD  CT3    -0.27
ATOM HMD1 HA3     0.09
ATOM HMD2 HA3     0.09
ATOM HMD3 HA3     0.09


!alkyl chain

GROUP
ATOM C1   CT2    -0.18
ATOM HP11 HA2     0.09
ATOM HP12 HA2     0.09
GROUP
ATOM C2   CE1    -0.09
ATOM  HP2 HE1     0.09
ATOM C3   CE1    -0.00
GROUP
ATOM C4   CT3    -0.27
ATOM HP41 HA3     0.09
ATOM HP42 HA3     0.09
ATOM HP43 HA3     0.09
GROUP
```

```
ATOM C5    CT2    -0.18
ATOM HP51  HA2     0.09
ATOM HP52  HA2     0.09
GROUP
ATOM C6    CT2    -0.18
ATOM HP61  HA2     0.09
ATOM HP62  HA2     0.09
GROUP
ATOM C7    CT2    -0.18
ATOM HP71  HA2     0.09
ATOM HP72  HA2     0.09
GROUP
ATOM C8    CT1    -0.09
ATOM HP8   HA1     0.09
GROUP
ATOM C9    CT3    -0.27
ATOM HP91  HA3     0.09
ATOM HP92  HA3     0.09
ATOM HP93  HA3     0.09
GROUP
ATOM C10   CT2    -0.18
ATOM H101  HA2     0.09
ATOM H102  HA2     0.09
GROUP
ATOM C11   CT2    -0.18
ATOM H111  HA2     0.09
ATOM H112  HA2     0.09
GROUP
ATOM C12   CT2    -0.18
ATOM H121  HA2     0.09
ATOM H122  HA2     0.09
GROUP
ATOM C13   CT1    -0.09
ATOM HP13  HA1     0.09
GROUP
ATOM C14   CT3    -0.27
ATOM H141  HA3     0.09
ATOM H142  HA3     0.09
ATOM H143  HA3     0.09
GROUP
ATOM C15   CT2    -0.18
ATOM H151  HA2     0.09
ATOM H152  HA2     0.09
GROUP
ATOM C16   CT2    -0.18
ATOM H161  HA2     0.09
ATOM H162  HA2     0.09
GROUP
ATOM C17   CT2    -0.18
ATOM H171  HA2     0.09
ATOM H172  HA2     0.09
GROUP
ATOM C18   CT1    -0.09
ATOM HP18  HA1     0.09
GROUP
ATOM C19   CT3    -0.27
ATOM H191  HA3     0.09
ATOM H192  HA3     0.09
ATOM H193  HA3     0.09
GROUP
ATOM C20   CT3    -0.27
ATOM H201  HA3     0.09
ATOM H202  HA3     0.09
ATOM H203  HA3     0.09
```

```
! For porphyrin ring :
BOND  NB  HNB   ND   HND


!     -for pyrrol A:
BOND NA        C1A  C1A  C2A  C2A  C3A  C3A  C4A
BOND C2A  H2A  C3A  H3A
BOND C4A  NA
!     -for pyrrol B:
BOND NB        C1B  C1B  C2B   C2B  C3B  C3B  C4B
BOND C4B  NB
!     -for pyrrol C:
BOND NC        C1C  C1C  C2C   C2C  C3C  C3C  C4C
BOND C4C  NC
!     -for pyrrol D:
BOND ND        C1D  C1D  C2D   C2D  C3D  C3D  C4D
BOND C4D  ND
!     -for methylene bridges:
BOND C4A  CHB   CHB  C1B   C4B  CHC  CHC  C1C
BOND C4C  CHD   CHD  C1D   C4D  CHA  CHA  C1A
BOND CHB  HBH   CHC  HCH   CHD  HDH
!     -for ring E and its substituants:
BOND C3D  CAD  CAD  OBD   CBD CHA
BOND CBD  CAD  CBD  HBD



!     -for pyrrol A substituants:

BOND C3A  CMA     CMA  HMA1  CMA  HMA2   CMA  HMA3
BOND C2A  CAA     CAA  HAA1  CAA  HAA2
BOND CAA  CBA     CBA  HBA1  CBA  HBA2
BOND CBA  CGA     CGA  O1A   CGA  O2A
BOND O2A  C1      C1   HP11  C1   HP12




!     -for pyrrol B substituants:
BOND C2B CMB
BOND CMB HMB1 CMB HMB2 CMB HMB3
BOND C3B CAB  CAB HAB
BOND CAB CBB  CBB HBB1 CBB HBB2



!     -for pyrrol C substituants:
BOND C2C CMC
BOND CMC HMC1 CMC HMC2 CMC HMC3
BOND C3C CAC  CAC HAC1 CAC HAC2
BOND CAC CBC
BOND CBC HBC1 CBC HBC2 CBC HBC3



!     -for pyrrol D substituants:
BOND C2D CMD
BOND CMD HMD1 CMD HMD2 CMD HMD3



!     -for ring E substituants:
BOND CBD CGD CGD O1D CGD O2D
BOND O2D CED
BOND CED HED1 CED HED2 CED HED3



!    -Phytyl chain:
```

```
BOND C1   C2        C2  HP2     C2  C3
BOND C3   C4        C4  HP41    C4  HP42   C4    HP43
BOND C3   C5        C5  HP51    C5  HP52
BOND C5   C6        C6  HP61    C6  HP62
BOND C6   C7        C7  HP71    C7  HP72
BOND C7   C8        C8  HP8
BOND C8   C9        C9  HP91    C9  HP92    C9  HP93
BOND C8   C10       C10 H101    C10 H102
BOND C10  C11       C11 H111    C11 H112
BOND C11  C12       C12 H121    C12 H122
BOND C12  C13       C13  HP13
BOND C13  C14       C14  H141   C14  H142   C14  H143
BOND C13  C15       C15  H151   C15  H152
BOND C15  C16       C16  H161   C16  H162
BOND C16  C17       C17  H171   C17  H172
BOND C17  C18       C18  HP18
BOND C18  C19       C19  H191   C19  H192   C19  H193
BOND C18  C20       C20  H201   C20  H202   C20  H203


!     ANGLES IMPROPRES :
!     type CPAN NPHN     CPAN CPBN : 2 enleves pour CT2

IMPR   C4A NA C1A C2A       C1A NA C4A C3A
IMPR   C4B NB C1B C2B       C1B NB C4B C3B
IMPR   C4C NC C1C C2C       C1C NC C4C C3C
IMPR   C4D ND C1D C2D       C1D ND C4D C3D


!     type NPHN CPAN CPBN CPBN : 2 enleves pour CT2

IMPR   NB C1B C2B C3B       NB C4B C3B C2B
IMPR   ND C1D C2D C3D       ND C4D C3D C2D


!     type NPHN CPAN CPMN CPA

IMPR   NA C1A CHA C4D       NA C4A CHB C1B
IMPR   NB C1B CHB C4A       NB C4B CHC C1C
IMPR   NC C1C CHC C4B       NC C4C CHD C1D
IMPR   ND C1D CHD C4C       ND C4D CHA C1A


!     type CPMN CPAN CPAN HA

IMPR   CHB C1B C4A HBH      CHC C1C C4B HCH       CHD C1D C4C HDH


!     type CPAN CPBN CPMN NPH

IMPR   C1A C2A CHA NA       C4A C3A CHB NA
IMPR   C1B C2B CHB NB       C4B C3B CHC NB
IMPR   C1C C2C CHC NC       C4C C3C CHD NC
IMPR   C1D C2D CHD ND       C4D C3D CHA ND


!     type NPHN CPAN CPAN MG
IMPR   NB C1B C4B HNB
IMPR   ND C1D C4D HND

ACCEPTOR OBD CAD
END
```

```
* Topology File for plastoquinone-9 !
*
! When using this topology file please cite:
! Guerra F., Adam S. and Bondar A-N.,
! Revised force-field parameters for chlorophyll-a, pheophytin-a and
plastoquinone-9,
! Journal of Molecular Graphics and Modelling, 58 (2015) 30-39.
!
!
! Please report bugs to Ana-Nicoleta Bondar, nbondar@zedat.fu-berlin.de, and
Federico Guerra: fguerra@zedat.fu-berlin.de
!
37 1


MASS 132 CG2DC1 12.01100 ! conjugated alkenes, R2C=CR2
MASS 133 CG2O5  12.01100 ! carbonyl C: ketones
MASS 134 CG331  12.01100 ! aliphatic C for methyl group (-CH3)
MASS 135 OG2D3  15.99940 ! carbonyl O: ketones
MASS 136 HGA3    1.00800 ! alphatic proton, CH3
MASS 137 HGA4    1.00800 ! alkene proton; RHC=
MASS 138 CT2    12.01100 ! carbon of methylene group (-CH2-)
MASS 139 CT3    12.01100 ! carbon of methyl group (-CH3)
MASS 140 HA2     1.00800 ! alkane, CH2, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS 141 HA3     1.00800 ! alkane, CH3, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS 142 CE1    12.01100 ! for alkene; RHC=CR
MASS 143 HA1     1.00800 ! alkane, CH, new LJ params (see
toppar_all22_prot_aliphatic_c27.str)
MASS 144 CT1    12.01100 ! aliphatic sp3 C for CH



DEFAult FIRST none LAST none !
AUTOgenerate ANGLes DIHEdrals !

RESI PL9 0.000 !
GROUP
ATOM C2  CG2DC1 -0.33 !
ATOM C3  CG2DC1 -0.09 ! here attach the alkyl chain
ATOM C4  CG2O5   0.46 !
ATOM C5  CG2DC1  0.01 !                       O1
ATOM C6  CG2DC1  0.01 !                      //
ATOM C1  CG2O5   0.46 !             C3---C4        H13
ATOM C52 CG331  -0.27 !            //      \      /
ATOM C53 CG331  -0.27 !        H11--C2      C5--C52-H14
ATOM O1  OG2D3  -0.38 !            \      //      \
ATOM O2  OG2D3  -0.38 !             C1---C6        H15
ATOM H11 HGA4    0.24 !            //      \
ATOM H13 HGA3    0.09 !           O2        C53-H16
ATOM H14 HGA3    0.09 !                     / \
ATOM H15 HGA3    0.09 !                  H18   H17
ATOM H16 HGA3    0.09 !
ATOM H17 HGA3    0.09 !
ATOM H18 HGA3    0.09 !
!
!
!
!alkyl chain
GROUP
ATOM CP1  CT2 -0.18 !
ATOM HP11 HA2  0.09 !
ATOM HP12 HA2  0.09 !
GROUP !
ATOM CP2  CT1 -0.09 !
```
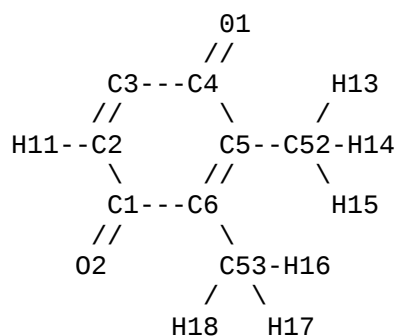
```
ATOM HP2  HA1  0.09 !
GROUP
ATOM CP3  CE1  0.00 !
GROUP
ATOM CP4  CT3 -0.27 !
ATOM HP41 HA3  0.09 !
ATOM HP42 HA3  0.09 !
ATOM HP43 HA3  0.09 !
GROUP
ATOM CP5  CT2 -0.18 !
ATOM HP51 HA2  0.09 !
ATOM HP52 HA2  0.09 !
GROUP
ATOM CP6  CT2 -0.18 !
ATOM HP61 HA2  0.09 !
ATOM HP62 HA2  0.09 !
GROUP
ATOM CP7  CT2 -0.18 !
ATOM HP71 HA2  0.09 !
ATOM HP72 HA2  0.09 !
GROUP
ATOM CP8  CT1 -0.09 !
ATOM HP8  HA1  0.09 !
GROUP
ATOM CP9  CT3 -0.27 !
ATOM HP91 HA3  0.09 !
ATOM HP92 HA3  0.09 !
ATOM HP93 HA3  0.09 !
GROUP
ATOM CP10 CT2 -0.18 !
ATOM H101 HA2  0.09 !
ATOM H102 HA2  0.09 !
GROUP
ATOM CP11 CT2 -0.18 !
ATOM H111 HA2  0.09 !
ATOM H112 HA2  0.09 !
GROUP
ATOM CP12 CT2 -0.18 !
ATOM H121 HA2  0.09 !
ATOM H122 HA2  0.09 !
GROUP
ATOM CP13 CT1 -0.09 !
ATOM HP13 HA1  0.09 !
GROUP
ATOM CP14 CT3 -0.27 !
ATOM H141 HA3  0.09 !
ATOM H142 HA3  0.09 !
ATOM H143 HA3  0.09 !
GROUP
ATOM CP15 CT2 -0.18 !
ATOM H151 HA2  0.09 !
ATOM H152 HA2  0.09 !
GROUP
ATOM CP16 CT2 -0.18 !
ATOM H161 HA2  0.09 !
ATOM H162 HA2  0.09 !
GROUP
ATOM CP17 CT2 -0.18 !
ATOM H171 HA2  0.09 !
ATOM H172 HA2  0.09 !
GROUP
ATOM CP18 CT1 -0.09 !
ATOM HP18 HA1  0.09 !
GROUP
```

```
ATOM CP19 CT3 -0.27 !
ATOM H191 HA3  0.09 !
ATOM H192 HA3  0.09 !
ATOM H193 HA3  0.09 !
GROUP
ATOM CP20 CT2 -0.18 !
ATOM H201 HA2  0.09 !
ATOM H202 HA2  0.09 !
GROUP
ATOM CP21 CT2 -0.18 !
ATOM H211 HA2  0.09 !
ATOM H212 HA2  0.09 !
GROUP
ATOM CP22 CT2 -0.18 !
ATOM H221 HA2  0.09 !
ATOM H222 HA2  0.09 !
GROUP
ATOM CP23 CT1 -0.09 !
ATOM HP23 HA1  0.09 !
GROUP
ATOM CP24 CT3 -0.27 !
ATOM H241 HA3  0.09 !
ATOM H242 HA3  0.09 !
ATOM H243 HA3  0.09 !
GROUP
ATOM CP25 CT2 -0.18 !
ATOM H251 HA2  0.09 !
ATOM H252 HA2  0.09 !
GROUP
ATOM CP26 CT2 -0.18 !
ATOM H261 HA2  0.09 !
ATOM H262 HA2  0.09 !
GROUP
ATOM CP27 CT2 -0.18 !
ATOM H271 HA2  0.09 !
ATOM H272 HA2  0.09 !
GROUP
ATOM CP28 CT1 -0.09 !
ATOM HP28 HA1  0.09 !
GROUP
ATOM CP29 CT3 -0.27 !
ATOM H291 HA3  0.09 !
ATOM H292 HA3  0.09 !
ATOM H293 HA3  0.09 !
GROUP
ATOM CP30 CT2 -0.18 !
ATOM H301 HA2  0.09 !
ATOM H302 HA2  0.09 !
GROUP
ATOM CP31 CT2 -0.18 !
ATOM H311 HA2  0.09 !
ATOM H312 HA2  0.09 !
GROUP
ATOM CP32 CT2 -0.18 !
ATOM H321 HA2  0.09 !
ATOM H322 HA2  0.09 !
GROUP
ATOM CP33 CT1 -0.09 !
ATOM HP33 HA1  0.09 !
GROUP
ATOM CP34 CT3 -0.27 !
ATOM H341 HA3  0.09 !
ATOM H342 HA3  0.09 !
ATOM H343 HA3  0.09 !
```

```
GROUP
ATOM CP35 CT2 -0.18 !
ATOM H351 HA2  0.09 !
ATOM H352 HA2  0.09 !
GROUP
ATOM CP36 CT2 -0.18 !
ATOM H361 HA2  0.09 !
ATOM H362 HA2  0.09 !
GROUP
ATOM CP37 CT2 -0.18 !
ATOM H371 HA2  0.09 !
ATOM H372 HA2  0.09 !
GROUP
ATOM CP38 CT1 -0.09 !
ATOM HP38 HA1  0.09 !
GROUP
ATOM CP39 CT3 -0.27 !
ATOM H391 HA3  0.09 !
ATOM H392 HA3  0.09 !
ATOM H393 HA3  0.09 !
GROUP
ATOM CP40 CT2 -0.18 !
ATOM H401 HA2  0.09 !
ATOM H402 HA2  0.09 !
GROUP
ATOM CP41 CT2 -0.18 !
ATOM H411 HA2  0.09 !
ATOM H412 HA2  0.09 !
GROUP
ATOM CP42 CT2 -0.18 !
ATOM H421 HA2  0.09 !
ATOM H422 HA2  0.09 !
GROUP
ATOM CP43 CT1 -0.09 !
ATOM H431 HA1  0.09 !
GROUP
ATOM CP44 CT3 -0.27 !
ATOM H441 HA3  0.09 !
ATOM H442 HA3  0.09 !
ATOM H443 HA3  0.09 !
GROUP
ATOM CP45 CT3 -0.27 !
ATOM H451 HA3  0.09 !
ATOM H452 HA3  0.09 !
ATOM H453 HA3  0.09 !


BOND C52 C5 !
BOND C53 C6 !
BOND C5 C6 !
BOND C5 C4 !
BOND O1 C4 !
BOND C6 C1 !
BOND C4 C3 !
BOND C1 O2 !
BOND C1 C2 !
BOND C3 C2 !
BOND C2 H11 !
!BOND C3    H12
BOND C52 H13 !
BOND C52 H14 !
BOND C52 H15 !
BOND C53 H16 !
```

```
BOND C53 H17 !
BOND C53 H18 !
BOND C3 CP1 CP1 HP11 CP1 HP12 !
!
!      -alkyl chain:
BOND CP1 CP2 CP2 HP2 CP2 CP3 !
BOND CP3 CP4 CP4 HP41 CP4 HP42 CP4 HP43 !
BOND CP3 CP5 CP5 HP51 CP5 HP52 !
BOND CP5 CP6 CP6 HP61 CP6 HP62 !
BOND CP6 CP7 CP7 HP71 CP7 HP72 !
BOND CP7 CP8 CP8 HP8 !
BOND CP8 CP9 CP9 HP91 CP9 HP92 CP9 HP93 !
BOND CP8 CP10 CP10 H101 CP10 H102 !
BOND CP10 CP11 CP11 H111 CP11 H112 !
BOND CP11 CP12 CP12 H121 CP12 H122 !
BOND CP12 CP13 CP13 HP13 !
BOND CP13 CP14 CP14 H141 CP14 H142 CP14 H143 !
BOND CP13 CP15 CP15 H151 CP15 H152 !
BOND CP15 CP16 CP16 H161 CP16 H162 !
BOND CP16 CP17 CP17 H171 CP17 H172 !
BOND CP17 CP18 CP18 HP18 !
BOND CP18 CP19 CP19 H191 CP19 H192 CP19 H193 !
BOND CP18 CP20 CP20 H201 CP20 H202 !
BOND CP20 CP21 CP21 H211 CP21 H212 !
BOND CP21 CP22 CP22 H221 CP22 H222 !
BOND CP22 CP23 CP23 HP23 !
BOND CP23 CP24 CP24 H241 CP24 H242 CP24 H243 !
BOND CP23 CP25 CP25 H251 CP25 H252 !
BOND CP25 CP26 CP26 H261 CP26 H262 !
BOND CP26 CP27 CP27 H271 CP27 H272 !
BOND CP27 CP28 CP28 HP28 !
BOND CP28 CP29 CP29 H291 CP29 H292 CP29 H293 !
BOND CP28 CP30 CP30 H301 CP30 H302 !
BOND CP30 CP31 CP31 H311 CP31 H312 !
BOND CP31 CP32 CP32 H321 CP32 H322 !
BOND CP32 CP33 CP33 HP33 !
BOND CP33 CP34 CP34 H341 CP34 H342 CP34 H343 !
BOND CP33 CP35 CP35 H351 CP35 H352 !
BOND CP35 CP36 CP36 H361 CP36 H362 !
BOND CP36 CP37 CP37 H371 CP37 H372 !
BOND CP37 CP38 CP38 HP38 !
BOND CP38 CP39 CP39 H391 CP39 H392 CP39 H393 !
BOND CP38 CP40 CP40 H401 CP40 H402 !
BOND CP40 CP41 CP41 H411 CP41 H412 !
BOND CP41 CP42 CP42 H421 CP42 H422 !
BOND CP42 CP43 CP43 H431 !
BOND CP43 CP44 CP44 H441 CP44 H442 CP44 H443 !
BOND CP43 CP45 CP45 H451 CP45 H452 CP45 H453 !
!
!
!
!
!
IMPR C4 C5 C3 O1 !
IMPR C1 C6 C2 O2 !
END !
```

Bibliography

[1] Noam Agmon. "The Grotthuss mechanism". In: *Chemical Physics Letters* 244.5 (1995), pp. 456–462.

[2] Suleyman I. Allakhverdiev et al. "Redox potential of pheophytin a in photosystem II of two cyanobacteria having the different special pair chlorophylls". In: *Proceedings of the National Academy of Sciences* 107.8 (2010), pp. 3924–3929.

[3] Muhamed Amin et al. "Effect of Chloride Depletion on the Magnetic Properties and the Redox Leveling of the Oxygen-Evolving Complex in Photosystem II". In: *The Journal of Physical Chemistry B* 120.18 (2016), pp. 4243–4248.

[4] Magnus Andersson et al. "Proton-Coupled Dynamics in Lactose Permease". In: *Structure* 20.11 (2012), pp. 1893–1904.

[5] Felix Autenrieth et al. "Classical force field parameters for the heme prosthetic group of cytochrome c". In: *Journal of Computational Chemistry* 25.13 (2004).

[6] James Barber. "Photosystem II: the engine of life". In: *Quarterly Reviews of Biophysics* 36.1 (Feb. 2003), pp. 71–89.

[7] James Barber. "'Photosystem II: the water splitting enzyme of photosynthesis and the origin of oxygen in our atmosphere'". In: *Quarterly Reviews of Biophysics* 49 (2016).

[8] Dmitrii Beglov and Benoît Roux. "Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations". In: *The Journal of Chemical Physics* 100.12 (1994), pp. 9050–9063.

[9]   Robert B. Best et al. "Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone , and Side-Chain 1 and 2 Dihedral Angles". In: *Journal of Chemical Theory and Computation* 8.9 (2012), pp. 3257–3273.

[10]   Martin Bommer et al. "Crystallographic and Computational Analysis of the Barrel Part of the PsbO Protein of Photosystem II: Carboxylate–Water Clusters as Putative Proton Transfer Relays and Structural Switches". In: *Biochemistry* 55.33 (2016), pp. 4626–4635.

[11]   Ana-Nicoleta Bondar and Holger Dau. "Extended protein/water H-bond networks in photosynthetic water oxidation". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1817.8 (Aug. 2012), pp. 1177–1190.

[12]   Ana-Nicoleta Bondar et al. "Mechanism of Primary Proton Transfer in Bacteriorhodopsin". In: *Structure* 12.7 (2004), pp. 1281–1288.

[13]   Sonja Braun-Sand, Marek Strajbl, and Arieh Warshel. "Studies of Proton Translocations in Biological Systems: Simulating Proton Transport in Carbonic Anhydrase by EVB-Based Models". In: *Biophysical Journal* 87.4 (2004), pp. 2221–2239.

[14]   TM Bricker. "Oxygen evolution in the absence of the 33-kilodalton manganese-stabilizing protein." In: *Biochemistry* 31.19 (1992), pp. 4623–4628.

[15]   B. R. Brooks et al. "CHARMM: The biomolecular simulation program". In: *Journal of Computational Chemistry* 30.10 (2009).

[16]   Bernard R. Brooks et al. "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations". In: *Journal of Computational Chemistry* 4.2 (1983), pp. 187–217.

[17]   Robert L. Burnap et al. "Role of Disulfide Linkage and Putative Intermolecular Binding Residues in the Stability and Binding of the Extrinsic Manganese-Stabilizing Protein to the Photosystem II Reaction Center". In: *Biochemistry* 33 (1994), pp. 13712–13718.

[18]   Hsi-Chao Chow, Robert Serlin, and Charles E. Strouse. "Crystal and molecular structure and absolute configuration of ethyl chlorophyllide a-dihydrate. Model for the different spectral forms of chlorophyll a". In: *Journal of the American Chemical Society* 97.25 (1975), pp. 7230–7237.

[19]   Alan Commet et al. "pH Optimum of the Photosystem II H2O Oxidation Reaction: Effects of PsbO, the Manganese-Stabilizing Protein, Cl– Retention, and Deprotonation of a Component Required for O2 Evolution Activity". In: *Biochemistry* 51.18 (2012), pp. 3808–3818.

[20]   C. J Cramer. *Essentials of Computational Chemistry: Theories and Models. 2nd Edition.* Chichester: John Wiley & Sons, Ltd, 2010.

[21]   Ana Damjanović et al. "Excitons in a photosynthetic light-harvesting system: A combined molecular dynamics, quantum chemistry, and polaron model study". In: 65 (Apr. 2002), p. 031919.

[22]   Tom Darden, Darrin York, and Lee Pedersen. "Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems". In: *The Journal of Chemical Physics* 98.12 (1993), pp. 10089–10092.

[23]   Holger Dau and Michael Haumann. "Eight steps preceding O–O bond formation in oxygenic photosynthesis—A basic reaction cycle of the Photosystem II manganese complex". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1767.6 (June 2007), pp. 472–483.

[24]   Holger Dau and Michael Haumann. "The manganese complex of photosystem II in its reaction cycle—Basic framework and possible realization at the atomic level". In: *Coordination Chemistry Reviews* 252.3-4 (Feb. 2008), pp. 273–295.

[25]   Ulrich Essmann et al. "A smooth particle mesh Ewald method". In: *The Journal of Chemical Physics* 103.19 (1995), pp. 8577–8593.

[26]   Scott E. Feller et al. "Constant pressure molecular dynamics simulation: The Langevin piston method". In: *The Journal of Chemical Physics* 103.11 (1995), pp. 4613–4621.

[27]   Kristina N. Ferreira et al. "Architecture of the Photosynthetic Oxygen-Evolving Center". In: *Science* 303.5665 (2004), pp. 1831–1838.

[28]   Aoife C. Fogarty and Damien Laage. "Water Dynamics in Protein Hydration Shells: The Molecular Origins of the Dynamical Perturbation". In: *The Journal of Physical Chemistry B* 118.28 (2014), pp. 7715–7729.

[29]   Nicolas Foloppe, Jacques Breton, and Jeremy C. Smith. "Potential energy function for photosynthetic reaction centre chromophores: energy minimisations of a crystalline bacteriophytin A analog". In: (1992).

[30]   Nicolas Foloppe et al. "Structural model of the photosynthetic reaction center of Rhodobacter capsulatus". In: *Proteins: Structure, Function, and Bioinformatics* 22.3 (1995), pp. 226–244.

[31]   M. J. Frisch et al. *Gaussian09 Revision E.01*. Gaussian Inc. Wallingford CT 2009.

[32]   Azat Gabdulkhakov et al. "Probing the Accessibility of the Mn4Ca Cluster in Photosystem II: Channels Calculation, Noble Gas Derivatization, and Cocrystallization with DMSO". In: *Structure* 17.9 (2009), pp. 1223–1234.

[33]   Richard J. Gowers et al. "MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations". In: (2016), pp. 98–105.

[34]   H. Grubmüller et al. "Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-range Interactions". In: *Molecular Simulation* 6.1-3 (1991), pp. 121–142.

[35]   Federico Guerra, Suliman Adam, and Ana-Nicoleta Bondar. "Revised force-field parameters for chlorophyll-a, pheophytin-a and plastoquinone-9". In: *Journal of Molecular Graphics and Modelling* 58 (2015), pp. 30–39.

[36]   Federico Guerra and Ana-Nicoleta Bondar. "Dynamics of the Plasma Membrane Proton Pump". In: *The Journal of Membrane Biology* 248.3 (June 2015), pp. 443–453.

[37]   Federico Guerra et al. "Dynamics of Long-Distance Hydrogen-Bond Networks in Photosystem II". In: *The Journal of Physical Chemistry B* 122.17 (2018), pp. 4625–4641.

[38]   Albert Guskov et al. "Cyanobacterial photosystem II at 2.9 Å resolution and the role of quinones, lipids, channels and chloride". In: *Nature Structural & Molecular Biology* (2009).

[39]   Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. "Exploring network structure, dynamics, and function using NetworkX". In: *Proceedings of the 7th Python in Science Conference (SciPy2008)*. Pasadena, CA USA, 2008, pp. 11–15.

[40]   Marcus D. Hanwell et al. "Avogadro: an advanced semantic chemical editor, visualization, and analysis platform". In: *Journal of Cheminformatics* 4.1 (2012), p. 17.

[41]   Matthias Heyden and Douglas J. Tobias. "Spatial Dependence of Protein-Water Collective Hydrogen-Bond Dynamics". In: *Phys. Rev. Lett.* 111 (21 2013), p. 218101.

[42]   Felix M. Ho and Stenbjörn Styring. "Access channels and methanol binding site to the CaMn4 cluster in Photosystem II based on solvent accessibility simulations, with implications for substrate water access". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1777.2 (2008), pp. 140–153.

[43]   William Humphrey, Andrew Dalke, and Klaus Schulten. "VMD – Visual Molecular Dynamics". In: *Journal of Molecular Graphics* 14 (1996), pp. 33–38.

[44]   Hiroshi Ishikita et al. "Tuning electron transfer by ester-group of chlorophylls in bacterial photosynthetic reaction center". In: *FEBS Letters* 579.3 (2005), pp. 712–716.

[45]   Michael J. S. Dewar et al. "AM1 a New General Purpose Quantum Mechanical Molecular Model". In: *Journal of the American Chemical Society* 107 (Apr. 1985).

[46]   Sunhwan Jo et al. "CHARMM-GUI: A web-based graphical user interface for CHARMM". In: *Journal of Computational Chemistry* 29.11 (), pp. 1859–1865.

[47]   P. Joliot, G. Barbieri, and R. Chabaud. "Un nouveau modele des centres photochimiques du system II". In: *Photochemistry and Photobiology* 10.5 (1969), pp. 309–329.

[48]   William L. Jorgensen et al. "Comparison of simple potential functions for simulating liquid water". In: *The Journal of Chemical Physics* 79.2 (1983), pp. 926–935.

[49]   Wolfgang Junge et al. "Electrostatics and Proton Transfer in Photosynthetic Water Oxidation". In: *Philosophical Transactions: Biological Sciences* 357.1426 (2002), pp. 1407–1418.

[50]   Laxmikant Kalé et al. "NAMD2: Greater Scalability for Parallel Molecular Dynamics". In: *Journal of Computational Physics* 151.1 (1999), pp. 283–312.

[51]   Nobuo Kamiya and Jian-Ren Shen. "Crystal structure of oxygen-evolving photosystem II from Thermosynechococcus vulcanus at 3.7 Å resolution". In: *PNAS* 100.1 (2003), pp. 98–103.

[52]   K. Kawakami et al. "Location of chloride and its possible functions in oxygen-evolving photosystem II revealed by X-ray crystallography". In: *Proceedings of the National Academy of Sciences* 106.21 (May 2009), pp. 8567–8572.

[53]   Jeffery B. Klauda et al. "Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types". In: *The Journal of Physical Chemistry B* 114.23 (2010), pp. 7830–7843.

[54]   A. Klauss, M. Haumann, and H. Dau. "Alternating electron and proton transfer steps in photosynthetic water oxidation". In: *Proceedings of the National Academy of Sciences* 109.40 (Sept. 2012), pp. 16035–16040.

[55]   V.V. Klimov, A.V. Klevanik, and V.A. Shuvalov. "Reduction of pheophytin in the primary light reaction of photosystem II". In: *FEBS Letters* 82.2 (1977), pp. 183–186.

[56]   Chris Knight and Gregory A. Voth. "The Curious Case of the Hydrated Proton". In: *Accounts of Chemical Research* 45.1 (2012), pp. 101–109.

[57]   BESSEL KOK, BLISS FORBUSH, and MARION McGLOIN. "Cooperation of charges in photosyntetic O2 evolution- a linear four step mechanism". In: *Photochemistry and Photobiology* 11.6 (1970), pp. 457–475.

[58]   Damien Laage and James T. Hynes. "On the Molecular Mechanism of Water Reorientation". In: *The Journal of Physical Chemistry B* 112.45 (2008), pp. 14230–14242.

[59]   Jérôme Lavergne and Wolfgang Junge. "Proton release during the redox cycle of the water oxidase". In: *Photosynthesis Research* 38.3 (Jan. 1993), pp. 279–296.

[60]   Andrew R. Leach. *Molecular modelling : principles and applications*. 2nd ed. Pearson Prentice Hall, 2009.

[61]   Hyun Ju Lee et al. "Intricate Role of Water in Proton Transport through Cytochrome c Oxidase". In: *Journal of the American Chemical Society* 132.45 (2010), pp. 16225–16239.

[62]   Jumin Lee et al. "CHARMM-GUI Input Generator for NAMD, GRO-MACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field". In: *Journal of Chemical Theory and Computation* 12.1 (2016), pp. 405–413.

[63]   Xichen Li and Per E. M. Siegbahn. "Alternative mechanisms for O2 release and O-O bond formation in the oxygen evolving complex of photosystem II". In: *Phys. Chem. Chem. Phys.* 17 (2015), pp. 12168–12174.

[64]   Zhaoliang Li and Robert L. Burnap. "Mutations of basic arginine residue 334 in the D1 protein of Photosystem II lead to unusual S2 state properties in Synechocystis sp. PCC 6803". In: *Photosynthesis Research* 72.2 (May 2002), pp. 191–201.

[65]   Katrin Linke and Felix M. Ho. "Water in Photosystem II: Structural, functional and mechanistic considerations". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1837.1 (2014), pp. 14–32.

[66]   Bernhard Loll et al. "Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II". In: *Nature* 438.15 (2005), pp. 1040–1044.

[67]   Andrei L. Lomize, Irina D. Pogozheva, and Henry I Mosberg. "Anisotropic Solvent Model of the Lipid Bilayer. 2. Energetics of Insertion of Small Molecules, Peptides, and Proteins in Membranes". In: *Journal of Chemical Information and Modeling* 51.4 (2011), pp. 930–946.

[68]   Sebastian Lorch et al. "Dynamic Carboxylate/Water Networks on the Surface of the PsbO Subunit of Photosystem II". In: *The Journal of Physical Chemistry B* 119.37 (2015), pp. 12172–12181.

[69]   A. D. MacKerell et al. "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins". In: *The Journal of Physical Chemistry B* 102.18 (1998), pp. 3586–3616.

[70]   A. D. MacKerell et al. "CHARMM: The Energy Function and Its Parameterization with an Overview of the Program". In: vol. 1. The Encyclopedia of Computational Chemistry. 1998, pp. 271–277.

[71]   Alexander D. MacKerell, Michael Feig, and Charles L. Brooks. "Improved Treatment of the Protein Backbone in Empirical Force Fields". In: *Journal of the American Chemical Society* 126.3 (2004), pp. 698–699.

[72]   Alexander D. Mackerell, Michael Feig, and Charles L. Brooks. "Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations". In: *Journal of Computational Chemistry* 25.11 (2004), pp. 1400–1415.

[73]   Alexander D. MacKerell and Martin Karplus. "Importance of attractive van der Waals contribution in empirical energy function models for the heat of vaporization of polar liquids". In: *The Journal of Physical Chemistry* 95.26 (1991), pp. 10559–10560.

[74]   Songrit Maneewongvatana and David M. Mount. "Analysis of approximate nearest neighbor searching with clustered point sets". In: *CoRR* cs.CG/9901013 (1999). URL: `http://arxiv.org/abs/cs.CG/9901013`.

[75]   Omer Markovitch et al. "Special Pair Dance and Partner Selection: Elementary Steps in Proton Transport in Liquid Water". In: *The Journal of Physical Chemistry B* 112.31 (2008), pp. 9456–9466.

[76]   Glenn J. Martyna, Douglas J. Tobias, and Michael L. Klein. "Constant pressure molecular dynamics algorithms". In: *The Journal of Chemical Physics* 101.5 (1994), pp. 4177–4189.

[77]   Christopher G. Mayne et al. "Rapid parameterization of small molecules using the force field toolkit". In: *Journal of Computational Chemistry* 34.32 (2013), pp. 2757–2770.

[78]   Naveen Michaud-Agrawal et al. "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations". In: *Journal of Computational Chemistry* 32.10 (2011).

[79]   Mitsue Miyao and Norio Murata. "Effect of urea on Photosystem II particles. Evidence for an essential role of the 33 kilodalton polypeptide in photosynthetic oxygen evolution". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 765.3 (1984), pp. 253–257.

[80]   Mitsue Miyao et al. "Effect of the 33-kDa protein on the S-state transitions in photosynthetic oxygen evolution". In: *Biochimica et Biophysica Acta* 890 (1987), pp. 151–159.

[81] Akihiro Motoki et al. "A Domain of the Manganese-stabilizing Protein fromSynechococcus elongatus Involved in Functional Binding to Photosystem II". In: *Biological chemistry* 277.17 (2002), pp. 14747–14756.

[82] James W. Murray and James Barber. "Structural characteristics of channels and pathways in photosystem II including the identification of an oxygen channel". In: *Journal of Structural Biology* 159.2 (Aug. 2007), pp. 228–237.

[83] Shin Nakamura and Takumi Noguchi. "Infrared Determination of the Protonation State of a Key Histidine Residue in the Photosynthetic Water Oxidizing Center". In: *Journal of the American Chemical Society* 139.27 (2017), pp. 9364–9375.

[84] Shin Nakamura et al. "Fourier Transform Infrared Detection of a Polarizable Proton Trapped between Photooxidized Tyrosine YZ and a Coupled Histidine in Photosystem II: Relevance to the Proton Transfer Mechanism of Water Oxidation". In: *Biochemistry* 53.19 (2014), pp. 3131–3144.

[85] Takumi Noguchi. "Fourier transform infrared difference and time-resolved infrared detection of the electron and proton transfer dynamics in photosynthetic water oxidation". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1847.1 (2014), pp. 35–45.

[86] Vladimir I. Novoderezhkin, Jan P. Dekker, and Rienk van Grondelle. "Mixing of Exciton and Charge-Transfer States in Photosystem II Reaction Centers: Modelling of Stark Spectra with Modified Redfield Theory". In: *Biophysical Journal* 93.4 (2007), pp. 1293–1311.

[87] Koji Ogata et al. "All-Atom Molecular Dynamics Simulation of Photosystem II Embedded in Thylakoid Membrane". In: *Journal of the American Chemical Society* 135.42 (2013), pp. 15670–15673.

[88] Travis E. Oliphant. "Python for Scientific Computing". In: *Computing in Science & Engineering* 9.3 (2007), pp. 10–20.

[89] Martin Petřek et al. "CAVER: a new tool to explore routes from protein clefts, pockets and cavities". In: *BMC Bioinformatics* 7.1 (2006), p. 316.

[90] James C. Phillips et al. "Scalable molecular dynamics with NAMD". In: *Journal of Computational Chemistry* 26.16 (2005), pp. 1781–1802.

[91] R. Pomès and B. Roux. "Structure and dynamics of a proton wire: a theoretical study of H+ translocation along the single-file water chain in the gramicidin A channel". In: *Biophysical Journal* 71.1 (1996), pp. 19–39.

[92] Hana Popelkova et al. "Inorganic Cofactor Stabilization and Retention: The Unique Functions of the Two PsbO Subunits of Eukaryotic Photosystem II". In: *Biochemistry* 47.47 (2008), pp. 12593–12600.

[93] Björn Rabenstein, G. Matthias Ullmann, and E.-W. Knapp. "Calculation of protonation patterns in proteins with structural relaxation and molecular ensembles – application to the photosynthetic reaction center". In: *European Biophysics Journal* 27.6 (Sept. 1998), pp. 626–637.

[94] Ivan Rivalta et al. "Structural–Functional Role of Chloride in Photosystem II". In: *Biochemistry* 50.29 (2011), pp. 6312–6315.

[95] Anne Rokka et al. "Synthesis and assembly of thylakoid protein complexes: multiple assembly steps of photosystem II". In: *Biochemical Journal* 388.1 (2005), pp. 159–168.

[96] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman J.C Berendsen. "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes". In: *Journal of Computational Physics* 23.3 (1977), pp. 327–341.

[97] V. Sacks et al. "The dynamic feature of the proton collecting antenna of a protein surface". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1365.1 (1998), pp. 232–240.

[98] G. Schultz et al. "Synthesis of prenylquinones in chioropiasts". In: *Physiol. Plant* 64 (1985), pp. 123–129.

[99] Rachel J. Service, Warwick Hillier, and Richard J. Debus. "Evidence from FTIR Difference Spectroscopy of an Extensive Network of Hydrogen Bonds near the Oxygen-Evolving Mn4Ca Cluster of Photosystem II Involving D1-Glu65, D2-Glu312, and D1-Glu329". In: *Biochemistry* 49.31 (Aug. 2010), pp. 6655–6669.

[100] Rachel J. Service, Warwick Hillier, and Richard J. Debus. "Network of Hydrogen Bonds near the Oxygen-Evolving Mn4CaO5Cluster of Photosystem II Probed with FTIR Difference Spectroscopy". In: *Biochemistry* 53.6 (Feb. 2014), pp. 1001–1017.

[101] Sheh-Yi Sheu and Dah-Yen Yang. "Determination of Protein Surface Hydration Shell Free Energy of Water Motion: Theoretical Study and Molecular Dynamics Simulation". In: *The Journal of Physical Chemistry B* 114.49 (2010), pp. 16558–16566.

[102] Tatiana Shutova et al. "A cluster of carboxylic groups in PsbO protein is involved in proton transfer from the water oxidizing complex of Photosystem II". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1767.6 (2007), pp. 434–440.

[103] U. Chandra Singh and Peter A. Kollman. "An approach to computing electrostatic charges for molecules". In: *Journal of Computational Chemistry* 5.2 (1984), pp. 129–145.

[104] Alexander M. Smondyrev and Gregory A. Voth. "Molecular Dynamics Simulation of Proton Transport through the Influenza A Virus M2 Channel". In: *Biophysical Journal* 83.4 (2002), pp. 1987–1996.

[105] Yifan Song, Junjun Mao, and M. R. Gunner. "MCCE2: Improving protein pKa calculations with extensive side chain rotamer sampling". In: *Journal of Computational Chemistry* 30.14 (2009), pp. 2231–2247.

[106] Eduardo M. Sproviero et al. "Computational studies of the O2-evolving complex of photosystem II and biomimetic oxomanganese complexes". In: *Coordination Chemistry Reviews* 252.3 (2008), pp. 395–415.

[107] James J. P. Stewart. "Optimization of parameters for semiempirical methods I. Method". In: *Journal of Computational Chemistry* 10.2 (1989), pp. 209–220.

[108] Alexei A. Stuchebrukhov. "Mechanisms of proton transfer in proteins: Localized charge transfer versus delocalized soliton transfer". In: *Phys. Rev. E* 79 (3 Mar. 2009), p. 031927.

[109] Michihiro Suga et al. "Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL". In: *Nature* (2017).

[110] Michihiro Suga et al. "Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses". In: *Nature* (2015).

[111] Jessica M. J. Swanson et al. "Proton Solvation and Transport in Aqueous and Biomolecular Systems: Insights from Computer Simulations". In: *The Journal of Physical Chemistry B* 111.17 (2007), pp. 4300–4314.

[112]   Takeshi Takahashi et al. "Photosystem II Complex in Vivo is a Monomer".
        In: *The Journal of Biological Chenmistry* 284.23 (Apr. 2009), pp. 15598–
        15606.

[113]   Ivan Y Torshin, Irene T Weber, and Robert W Harrison. "Geometric cri-
        teria of hydrogen bonds in proteins and identification of "bifurcated" hy-
        drogen bonds." In: *Protein engineering* 15 5 (2002), pp. 359–63.

[114]   M. Tuckerman, B. J. Berne, and G. J. Martyna. "Reversible multiple
        time scale molecular dynamics". In: *The Journal of Chemical Physics* 97.3
        (1992), pp. 1990–2001.

[115]   Yasufumi Umena et al. "Crystal structure of oxygen-evolving photosystem
        II at a resolution of 1.9 Å." In: *Nature* (2011).

[116]   Coral del Val and Ana-Nicoleta Bondar. "Charged groups at binding inter-
        faces of the PsbO subunit of photosystem II: A combined bioinformatics
        and simulation study". In: *Biochimica et Biophysica Acta (BBA) - Bioen-
        ergetics* 1858.6 (2017), pp. 432–441.

[117]   Coral del Val, Luiza Bondar, and Ana-Nicoleta Bondar. "Coupling be-
        tween inter-helical hydrogen bonding and water dynamics in a proton
        transporter". In: *Journal of structural biology* 186.1 (Apr. 2014).

[118]   K. Vanommeslaeghe and A. D. MacKerell. "Automation of the CHARMM
        General Force Field (CGenFF) I: Bond Perception and Atom Typing". In:
        Journal of Chemical Information and Modeling 52.12 (2012), pp. 3144–
        3154.

[119]   K. Vanommeslaeghe and A. D. MacKerell. "Automation of the CHARMM
        General Force Field (CGenFF) I: Bond Perception and Atom Typing". In:
        Journal of Chemical Information and Modeling 52.12 (2012), pp. 3144–
        3154.

[120]   K. Vanommeslaeghe, E. Prabhu Raman, and A. D. MacKerell. "Automa-
        tion of the CHARMM General Force Field (CGenFF) II: Assignment of
        Bonded Parameters and Partial Atomic Charges". In: *Journal of Chemical
        Information and Modeling* 52.12 (2012), pp. 3155–3168.

[121] K. Vanommeslaeghe et al. "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields". In: *Journal of Computational Chemistry* 31.4 (2010), pp. 671–690.

[122] Sergey Vassiliev, Tatiana Zaraiskaya, and Doug Bruce. "Exploring the energetics of water permeation in photosystem II by multiple steered molecular dynamics simulations". In: *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1817.9 (2012), pp. 1671–1678.

[123] David J. Vinyard and Gary W. Brudvig. "Progress Toward a Molecular Mechanism of Water Oxidation in Photosystem II". In: *Annual Review of Physical Chemistry* 68.1 (2017), pp. 101–116.

[124] Stéfan van der Walt, S. Chris Colbert, and Gaël Varoquaux. "The NumPy Array: A Structure for Efficient Numerical Computation". In: *Computing in Science & Engineering* 13.2 (2011), pp. 22–30.

[125] Junmei Wang et al. "Development and testing of a general amber force field". In: *Journal of Computational Chemistry* 25.9 (2004).

[126] Mai Watanabe et al. "Is the Photosystem II Complex a Monomer or a Dimer?" In: *Plant Cell Physiol.* 50.9 (2009), pp. 1674–1680.

[127] Hanna Wincencjusz, van Hans J. Gorkom, and Charles F. Yocum. "The Photosynthetic Oxygen Evolving Complex Requires Chloride for Its Redox State S2→S3 and S3→S0 Transitions But Not for S0→S1 or S1→S2 Transitions". In: *Biochemistry* 36.12 (1997), pp. 3663–3670.

[128] Athina Zouni et al. "Crystal structure of photosystem II from Synechococcus elongatus at 3.8 Å resolution". In: *Nature* 409.6821 (2001), pp. 739–743.

Selbstständigkeitserklärung

Hiermit bestätige ich, dass die vorliegende Doktorarbeit von mir allein und nur unter Nutzung der angegebenen Hilfsmittel verfasst wurde. Ferner bestätige ich, dass diese Arbeit nicht in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt wurde.

Berlin 7. Dezember 2017

_____

Federico Guerra