



**HAL**  
open science

# CpHMD-Then-QM/MM Identification of the Amino Acids Responsible for the Anabaena Sensory Rhodopsin pH-Dependent Electronic Absorption Spectrum

Elisa Pieri, Vincent Ledentu, Michael Sahlin, François Dehez, Massimo Olivucci, Nicolas Ferré

► **To cite this version:**

Elisa Pieri, Vincent Ledentu, Michael Sahlin, François Dehez, Massimo Olivucci, et al.. CpHMD-Then-QM/MM Identification of the Amino Acids Responsible for the Anabaena Sensory Rhodopsin pH-Dependent Electronic Absorption Spectrum. *Journal of Chemical Theory and Computation*, 2019, 15 (8), pp.4535-4546. 10.1021/acs.jctc.9b00221 . hal-02171700

**HAL Id: hal-02171700**

**<https://amu.hal.science/hal-02171700>**

Submitted on 3 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# CpHMD-then-QM/MM identification of the amino-acids responsible for the Anabaena Sensory Rhodopsin pH-dependent electronic absorption spectrum

Elisa Pieri,<sup>\*,†</sup> Vincent Ledentu,<sup>†</sup> Michael Sahlin,<sup>†</sup> François Dehez,<sup>‡</sup> Massimo Olivucci,<sup>¶,§</sup> and Nicolas Ferré<sup>\*,†</sup>

<sup>†</sup>*Aix-Marseille Univ, CNRS, Institut de Chimie Radicalaire, Marseille, France*

<sup>‡</sup>*Laboratoire de Physique et Chimie Thorique, UMR 7019, Facult des Sciences et Technique, Campus Aiguillettes, 54506 Vandoeuvre-les-Nancy, France*

<sup>¶</sup>*Department of Chemistry, Bowling Green State University, Bowling Green, Ohio 43403, United States*

<sup>§</sup>*Dipartimento di Biotecnologie, Chimica e Farmacia, Universit di Siena, via A. Moro 2, I-53100 Siena, Italy*

E-mail: elisa.pieri@univ-amu.fr; nicolas.ferre@univ-amu.fr

## Abstract

Anabaena Sensory Rhodopsin (ASR), a microbial photoactive protein featuring the retinal chromophore in two different conformations, exhibits a pH-dependent electronic absorption spectrum. Using the recently developed CpHMD-then-QM/MM multiscale protocol applied to ASR embedded in a membrane model, the pH-induced changes in its maximum absorption wavelength have been reproduced and analyzed. While the acidic

tiny red-shift is essentially correlated with the deprotonation of an aspartic acid located on the ASR extracellular side, the larger blue-shift experimentally reported at pH values larger than 5 involves a cluster of titrating residues sitting on the cytoplasmic side. The ASR pH-dependent spectrum is the consequence of the competitive stabilization of retinal ground and excited states by the protein electrostatic potential.

## 1 Introduction

The pH of the water phase is of crucial importance for the activity and stability of proteins, influencing a considerable amount of chemical properties, such as the tertiary structure,<sup>1</sup> the enzymatic activity<sup>2</sup> or the solubility.<sup>3</sup> It relies on the fact that, on average, 29% of protein sequence is composed by titratable (also often termed as ionizable) amino acids (i.e. ASP, GLU, HIS, LYS, CYS, TYR and ARG),<sup>4</sup> which can dynamically change their protonation state with the pH value. The loss/binding of a proton modifies the atomic charges on the side chains, triggering adjustments in the electrostatic interactions and conformational rearrangements. Consequently, the variation of the pH can induce major changes in the protein function. The understanding of the origin of such a pH-dependent behavior faces a great challenge: at a given pH, several protein residues can be in different protonation states (ie. they titrate) at the same time, making very difficult the identification of a single amino-acid as the main responsible for the pH-dependent property of interest. In other words, the protein has to be considered as a poly-acid macromolecule featuring tens of titrating sites which individual behaviors and their possible interactions (also allowed by water networks) shape the protein pH-dependency.<sup>5</sup> Nevertheless, the interpretation of the pH-dependency in terms of titrating sites is always appealing. As a matter of fact, such understanding opens the way to site-directed mutagenesis designed for altering the protein function in a controlled manner.<sup>6-11</sup>

In this work, we identify the titrating residues responsible for the pH-dependence of the visible-light absorption spectrum of Anabaena Sensory Rhodopsin (ASR), a transmem-

brane photochromic microbial protein whose photoactivity is carried out by the retinal chromophore. The retinal, present in the protonated Schiff base form, is linked to the lysine K210 and counterbalanced by the aspartate counterion D75 (Figure 1). Its predominant configuration in the ASR dark-adapted state is the *all-trans* (AT), while the *13-cis* (13C) configuration is the major one in the light-adapted state.<sup>12</sup> In this protein, the pH has the capacity of influencing the entire photochemistry, including absorption spectra, lifetimes and reaction dynamics.<sup>13,14</sup> In particular, a small red-shift (2 nm in both isomers<sup>14</sup>) in the maximum absorption wavelength ( $\lambda_{\max}$ ) has been experimentally found from pH 3.6 to 5.0, while a more consistent (6 nm for the *all-trans* isomer and 10 nm for the *13-cis* one) blue-shift is evidenced when the pH changes from 5.0 to 7.6 (Figure 2). The latter shift has been experimentally attributed putatively to the D217 deprotonation. While it is located about 15Å from retinal, this aspartic acid is known as the retinal proton acceptor during the ASR photocycle, thanks to an extended hydrogen bonding network present in ASR.<sup>15</sup> Accordingly, D217’s change of protonation probably perturbs this network, modifying the electrostatic potential experienced by retinal and ultimately the ASR  $\lambda_{\max}$ . The cause of the shift occurring at the more acidic pH remains unknown.

The elucidation of the molecular origin of such small  $\lambda_{\max}$  shifts is not always easy from the experimental point of view, and can take benefit from multiscale atomistic models rooted in theoretical chemistry: quantum mechanical (QM), molecular mechanical (MM) or hybrid QM/MM).<sup>16-26</sup> Most of the time, the pH effect is introduced thanks to very specific change of individual protonation states, without considering the relevance (probability) of such a chemical process at a given pH. In the present work, we will demonstrate that, indeed, the number of populated protonation states combinations (*microstates*) is always large. Accordingly, the corresponding ensemble has to be fully considered when small  $\lambda_{\max}$  shifts have to be reproduced and analyzed, as it is the case in the ASR absorption spectrum. Therefore, we have derived and used a molecular model which has been especially targeted towards the investigation of the ASR protonation microstates. For that purpose, we have

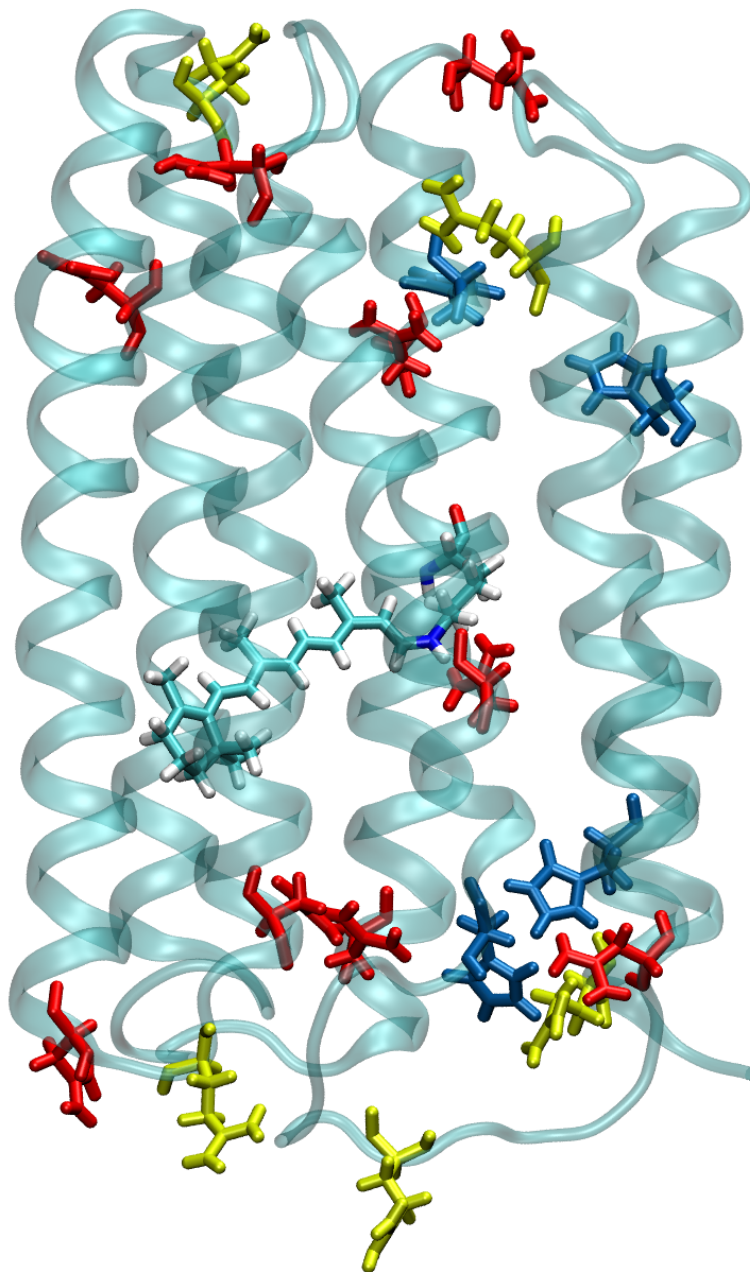


Figure 1: Schematic representation of the ASR model, featuring (i) retinal chromophore bonded to lysine 210 and (ii) 18 titratable residues between pH=3 and pH=8: aspartic acids (1-letter code D) in red; glutamic acids (1-letter code E) in yellow, histidines (1-letter code H) in blue.

chosen the Constant-pH Molecular Dynamics (CpHMD) approach which consists in sampling both the conformation and protonation state spaces.<sup>27,28</sup> With respect to usual molecular dynamics simulations, CpHMD ones include Monte Carlo-based sampling of the system

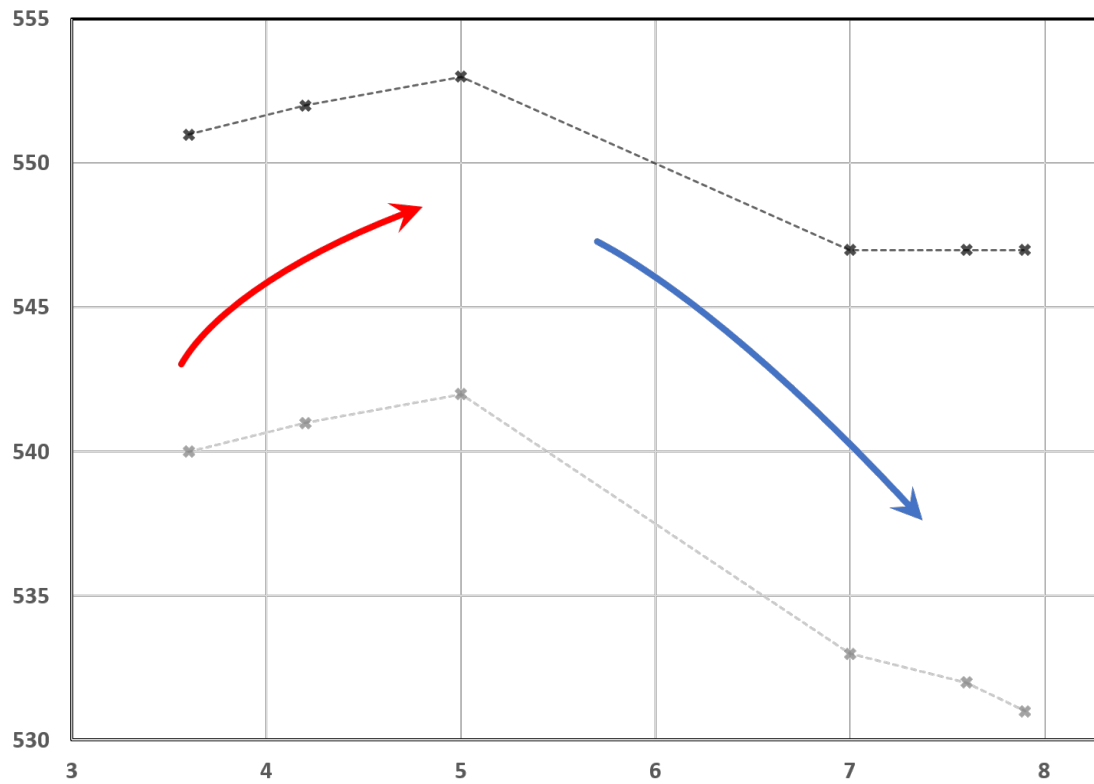


Figure 2: ASR experimental  $\lambda_{\max}$  values (in nm) between pH=3 and pH=8. Black (grey): AT retinal (13C retinal). Data taken from Rozin et al.<sup>14</sup> The red (blue) arrow indicates the pH region characterized by a  $\lambda_{\max}$  red-shift (blue-shift). Note the lack of data between pH=5 and pH=7.

protonation states, weighted for the semigrand canonical ensemble (constant temperature, volume and hydronium chemical potential). The nature of the opsin–retinal interactions responsible for the tuning of the chromophore photochemical properties like  $\lambda_{\max}$  has been studied for many years.<sup>16,29–35</sup> Within recent years, new models have emerged in which the direct electrostatic interactions are found to be less important than inductive, dispersive,<sup>36</sup> excitonic or charge transfer ones.<sup>37</sup> Nevertheless, in the context of the present study regarding pH-effects in ASR, in which most of the titratable amino-acids are not in the immediate vicinity of its chromophore (see Figure 1), we have based our model on a very qualitative, yet simple, QM/MM with electrostatic embedding approach. Such an approach has proven to be effective when one is looking for reproducing trends, eg due to rhodopsin mutations<sup>38</sup> or in other photoactive proteins.<sup>39</sup> The successful reproduction of the ASR  $\lambda_{\max}$  variations with the pH validates such an effective model.

## 2 Methods and computational details

From the microscopic point of view, the large number of interacting titratable sites translates to the existence of an ensemble of protonation microstates whose respective populations are pH-dependent.<sup>40,41</sup> Consequently, macroscopic  $pK_a$  values characterizing the acid-base properties of a macromolecule cannot be strictly attributed to independent and well-identified titrating amino-acid. Only the  $pK_{1/2}$ , ie the pH value at which a given titrating site is equally populated in its protonated and deprotonated forms, keeps a physically-ground meaning. Accordingly, a pH-dependent macroscopic property  $A$  can be evaluated as the average of individual contributions  $a_i$  for each member of the ensemble containing  $N$  possible microstates, weighted by their respective probabilities  $w_i$ :

$$A(\text{pH}) = \sum_{i=1}^N w_i(\text{pH}) a_i \tag{1}$$

Because the size of the ensemble grows exponentially with the number of titratable sites, it is very often assumed that a unique microstate dominates the sum in Equation 1, being then used as a template for further atomistic modeling, and eventually allowing for the identification of the microscopic origin of the protein pH-dependency. However, the sum in Equation 1 can be rewritten in a different way, grouping in a cluster all the microstates sharing the same  $a_i$  value ( $M$  being the number of clusters,  $M \leq N$ ):

$$A(\text{pH}) = \sum_j^M W_j(\text{pH}) a_j \quad (2)$$

in which  $W_j$  denotes the cumulated weight of a particular cluster,  $W_j = \sum_i w_i, \forall a_i = a_j$ . There is no mathematical evidence that the most probable cluster (ie the largest  $W_j$ ) coincides with the most probable microstate (ie the largest  $w_i$ ). As a consequence,  $A$  may significantly deviate from the value  $a_i$  characterizing the most probable protonation microstate. In order to mitigate the bias introduced by the selection of a single microstate, some of the authors have recently designed a multiscale modeling protocol, denoted as CpHMD-then-QM/MM, which heavily relies on the statistical sampling of both the protonation state and conformational spaces, together with large scale (thousands) of QM/MM  $\lambda_{\text{max}}$  evaluations. This approach, successfully applied to the modeling of a polypeptide dyad,<sup>42</sup> is summarized hereafter (all details are given in SuppInfo).

**System setup and CpHMD calculations.** The ASR initial structure has been retrieved from the 1XIO PDB entry,<sup>43</sup> a dimeric system in which both the retinal AT and the 13C forms are present. Having cleaned the structure of a monomer (see SuppInfo for the details), we have built two ASR models: (i) one solvated in water; (ii) one embedded in a membrane, modeled as a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) lipid bilayer. Using the Amber16 package,<sup>44</sup> explicit solvent CpHMD simulations<sup>27,45</sup> have been carried out using the ff14SB Amber and TIP3P forcefields,<sup>46</sup> complemented with parameters for retinal.<sup>47</sup> Enhanced sampling has been achieved using the Replica-Exchange technique for



pH values between 3.5 and 8.5, for a total simulation length of 20 to 30 ns (reduced sets of titrated residues) or 40 ns (full set of titrated residues) per pH value and per retinal isomer. For the sake of simplicity, individual  $pK_{1/2}$  values are simply called  $pK_a$  in the following and are fitted to the calculated deprotonated fraction  $x_i$  using the Hill equation:

$$pK_a(i) = \text{pH} - n \log \frac{x_i}{1 - x_i} \quad (3)$$

in which the fitted  $n$  value represents cooperative effects between residue  $i$  and the other titrating residues.

**Semi-empirical QM/MM calculations.** Using 20000 to 40000 equally-spaced snapshots (geometries and protonation microstate) extracted from CpHMD trajectories obtained at pH values between 3.5 and 8.5, we have carried out QM/MM vertical excitation energy calculations of retinal, in electrostatic interaction with the rest of the system. In order to keep reasonable the computational cost, we have used the PM7 semi-empirical Hamiltonian,<sup>48</sup> together with the Multi-Electron Configuration Interaction<sup>49</sup> approach implemented in Mopac.<sup>50</sup> The absorption spectra at different pH values have been generated using the Newton-X 2.0 package.<sup>51,52</sup>

**ARM calculations.** Automatic Rhodopsin Modeling (ARM)<sup>11,53</sup> calculations have been used to obtain accurate vertical excitation energies on a selected subset of protonation microstates. In short, ARM is a semi-automated protocol which starts from a PDB structure, builds a minimal rhodopsin model, samples the conformational space of the retinal cavity, before calculating vertical excitation energies using the high-level CASPT2 method.<sup>54</sup> In its current stage of development, it only deals with a single protonation microstate.

**Analysis tools.** In order to decipher the relations between calculated  $\lambda_{\text{max}}$  and the pH, we have mainly used two analysis approaches.

The first one is an improved version of the Minimal Electrostatic Model derived by some of us and already applied to the experimental ASR absorption spectrum.<sup>55</sup> It is based on the sound assumption that the  $\lambda_{\max}$  changes with the pH are governed by modified electrostatic interactions between the retinal chromophore and the protein protonation state changes. By fitting the  $\lambda_{\max}$  variations (actually, the related  $\Delta E_{max}$  variations) to the computed changes of the deprotonated fraction  $x_i$  (which depends on the pH) for each of the  $N_{tit}$  titrating residue  $i$  experiencing the pH-independent retinal transition electrostatic potential  $\phi_{ret}^{S_0 \rightarrow S_1}$  (only the 3 terms  $E_{full}$ ,  $\epsilon_{int}$  and  $\epsilon_{surf}$  are optimized, all the other ones are constants or parameters,  $c_i$ ,  $\phi_{ret}^{S_0 \rightarrow S_1}$  and  $r_i$  which are derived independently, see SuppInfo):

$$\Delta E_{max}(\text{pH}) = E_{full} + \sum_{i=1}^{N_{tit}} \left\{ x_i(\text{pH}) \frac{e}{4\pi\epsilon_0 [c_i\epsilon_{int} + (1 - c_i)\epsilon_{surf}]} \phi_{ret}^{S_0 \rightarrow S_1}(r_i) \right\} \quad (4)$$

the contribution  $\Delta E_i$  (corresponding to the term in between the braces) of each individual titrating residue  $i$  can be obtained at all the considered pH values.

The second analysis has been done using multilinear regression analysis of  $\lambda_{\max}$  as a function of the protonation states for all the considered independent titratable residues.<sup>56</sup> Ultimately, this analysis provides correlations between  $\lambda_{\max}$  and individual (de)protonation events.

### 3 Results

Before explicitly considering the very large ensemble of the ASR protonation microstates, we first report accurate  $\lambda_{\max}$  values for a selected set of microstates, in order to illustrate how it can be difficult to pick one or few ASR models as candidates for explaining the origin of its pH-dependency.

### 3.1 ARM calculations of some individual microstate $\lambda_{\max}$

When one builds an atomistic model of a protein, one of the main assumptions concerns the protonation state of each and every titratable residue. This is usually based on the comparison of each microscopic  $pK_a$  value with the targeted pH. It is then possible to roughly mimic a change of pH by modifying the protonation state of one or several titratable residues with respect to a reference microstate. We have applied this procedure to ASR  $\lambda_{\max}$  by means of ARM calculations (Figure 3). As expected, the largest  $\lambda_{\max}$  shift ( $> 120$  nm)

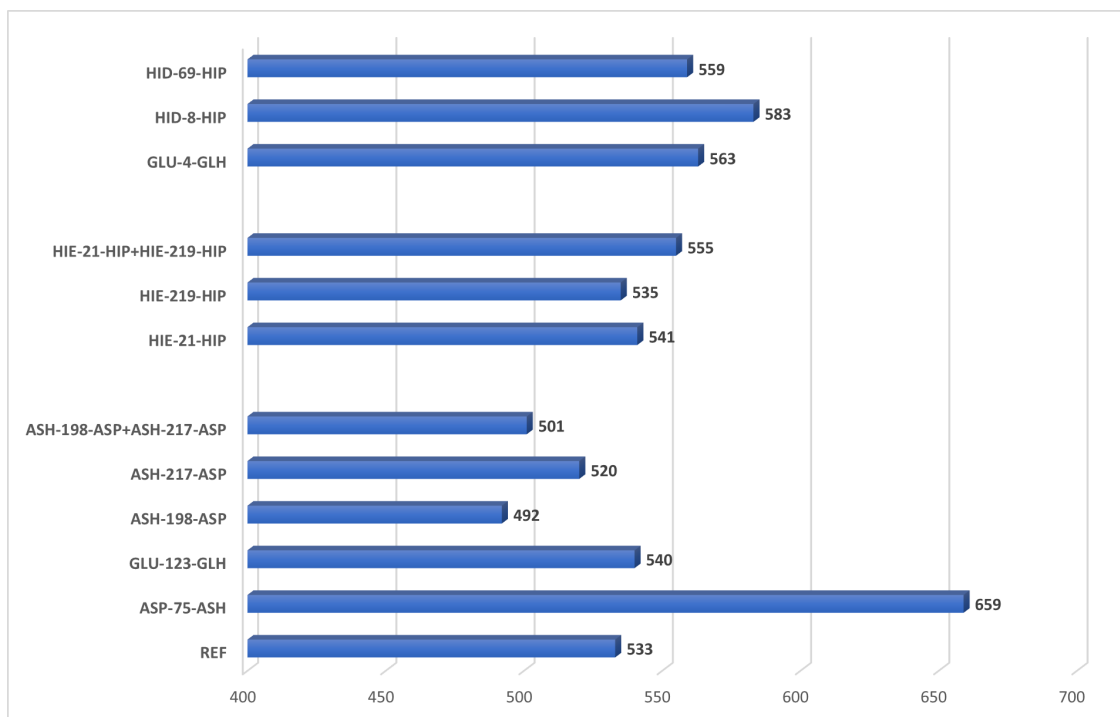


Figure 3:  $\lambda_{\max}$  values (in nm) of several ASR models, using the ARM protocol. The REF model corresponds to the most likely protonation microstate at pH=7. The 3-letter code for each amino-acid has been chosen to unambiguously denote each change of protonation state on the  $y$ -axis, as follows: initial protonation state–amino-acid number–final protonation state.

is due to protonation of retinal counterion, D75, keeping retinal as a protonated Schiff base. Because this shift is out of the range reported experimentally (max 10 nm), we will keep D75 as deprotonated in the following.

The deprotonation of D217 (15.4 Å from retinal) induces a blue-shift of 13 nm, in quali-

tative agreement with the experimental assignment.<sup>14</sup> Closer to retinal (14.2 Å), the deprotonation of D198 causes a much larger 41 nm  $\lambda_{\max}$  blue-shift. Simultaneous deprotonation of both D198 and D217 results in a blue-shift which is 32 nm with respect to the reference value, hence less than the one obtained by the single deprotonation of D198.

The smallest red-shift (2 nm) is due to protonation of H219. Accordingly, H219 could be a good candidate for explaining the ASR red-shift at acidic pH. This tiny effect is probably related to the large distance between H219 and retinal, more than 19 Å. The smaller distance between H21 and retinal (18 Å) results in a slightly larger red-shift (8 nm) upon H21 protonation. However, protonation of both H21 and H219 results in a 22 nm red-shift, larger than the ones resulting from each protonation taken separately.

In this configuration, H21 and H219 form an interacting pair in the ASR structure. Because of the large electrostatic repulsion between two close positively charged moieties, this doubly-protonated situation has very little chance to occur in the considered pH range. In other words, the corresponding protonation microstate is unlikely to be populated.

These ARM calculations result in accurate  $\lambda_{\max}$  values for each microstate. Accordingly, they can be used for predicting how  $\lambda_{\max}$  changes between two different microstates. However, they need to be complemented with the probability of such a process to take place at a given pH. For instance, it may occur that the large blue-shift caused by D198 deprotonation actually does not matter if this reaction is not likely to happen, because D198  $pK_a$  value could be significantly far from the pH range under study. Rephrased in a probabilistic language, they don't take into account the relative populations of D198 protonated and deprotonated forms at a given pH. The same remark applies to all the above-considered changes of protonation and this is the very reason why a correct ASR modeling protocol needs to include the determination of microstate populations. The latter are evaluated by means of CpHMD simulations.

### 3.2 Preliminary CpHMD-then-QM/MM study: effect of the membrane.

We have first applied the CpHMD-then-QM/MM protocol to an ASR model featuring 13C retinal at three pH values: 3.5, 5.5 and 7.5, using 6 equally-spaced 40 ns trajectories obtained between pH=3.5 and pH=8.5. Because the  $pK_a$  values of the 18 selected titrating residues (all aspartic and glutamic acids, all histidines) obtained from CpHMD simulations are known to be slowly converging with the system size and the number of titrating residues,<sup>45</sup> we have considered two models of ASR: in water and in a membrane solvated in water, using several 10, 20 or 40 ns windows (see the discussion in SuppInfo). As mentioned above, the retinal counterion D75 is always deprotonated. In all, our study demonstrates that most of the  $pK_a$  values are converged after 20 ns, with the exception of D57, H69, D98, D120 and D198 in water, but only D57 and D98 in the membrane (see Figure 4). Indeed, the titration curve of the latter residue shows a chaotic behavior which can be attributed to its peculiar location. It allows both a complete exposure to water and embedding in the membrane, so that the side chain can easily flip between the two opposite environments and experience completely different dielectric constants. As a matter of fact, the distance between the centers of mass of D98 and the closest phospholipid head changes sharply between 6Å and 18Å at all pH values (see SuppInfo). This behavior is responsible for the oscillating titration curve of D98, which would require prohibitively long simulations to achieve a good convergence. D57 belongs to a very flexible part of the protein, which was actually not resolved in the original 1XIO PDB structure. Accordingly, it is likely that its  $pK_a$  value would converge using a larger sampling (see SuppInfo for the time evolution of its  $pK_a$  value).

Even if the 40 ns trajectory length is relatively short to observe any major secondary or tertiary rearrangement, the ASR structural fluctuations are indeed smaller when the membrane is included into the model, as exemplified by the RMSD values reported in Table 1. The cavity in which the retinal chromophore lies keeps the same shape whatever the ASR environment. Its fluctuations, as well as those characterizing the retinal structure are

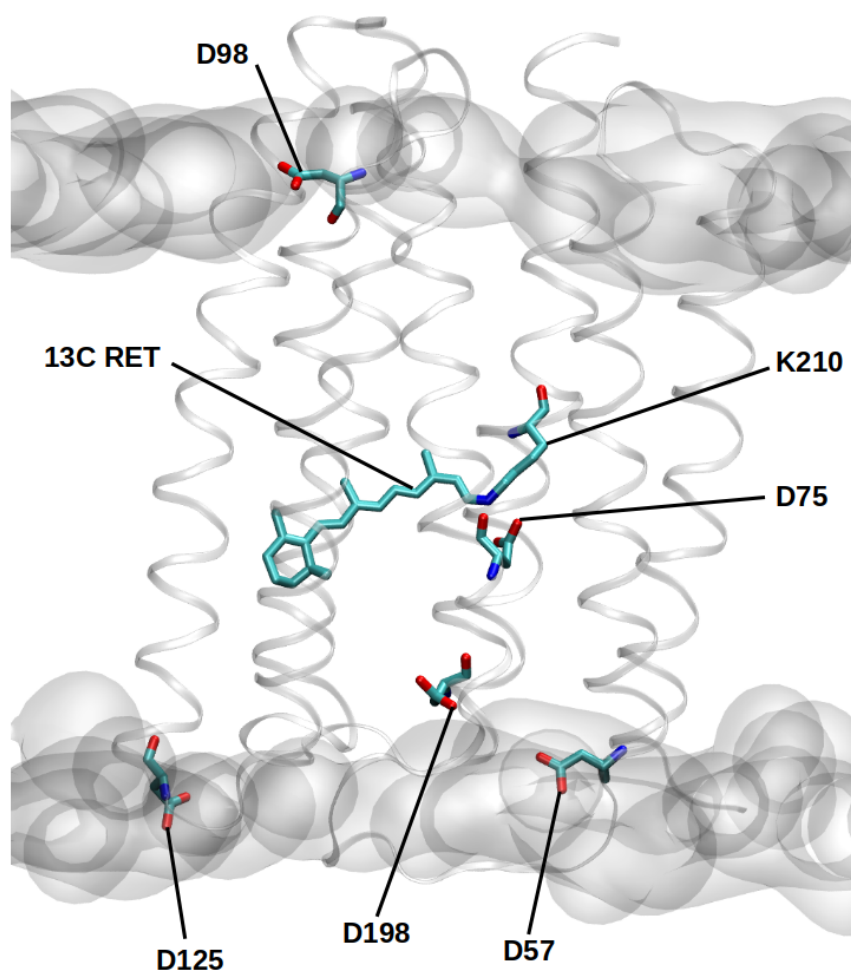


Figure 4: ASR structure with 13C retinal, bonded to K210. Retinal counterion is D75. Membrane phospholipid headgroups are represented as transparent surfaces. Other 4 titrating residues (D57, D98, D125, D198) are discussed in the text.

Table 1: Root-mean-square deviation (RMSD, in Å) with respect to the original crystallographic structure, without (Water column) and with the membrane (Membrane column), averaged over the 3 pH values 3.5, 5.5, 7.5 (individual pH RMSD values are available in SuppInfo).

RMSD	Water	Membrane
All	3.14	2.03
Backbone	2.36	1.27
Cavity	1.22	1.04
Retinal	1.03	1.03

reduced and independent of the pH value. The average distance between the centers of mass of the titrating residues is also independent of the pH (see SuppInfo). Moreover, the distances between titrating residues and retinal only change slightly, the maximum variations being observed for residues close or belonging to the reconstructed Q58 – Q66 missing loop (see SuppInfo). Accordingly, it is likely that the pH-dependence of ASR  $\lambda_{\max}$  is not primarily related to structural rearrangements involving the retinal cavity.

The importance of the membrane is striking when water-to-membrane  $pK_a$  shifts are considered (Table 2). Besides D57 and D98 which are not discussed for the reasons explained above, 2  $pK_a$  differences are much larger than 1 pH unit, corresponding to titrating D125 and D198. Actually, it turns out that D198 is always deprotonated in the membrane, while it can exist in water within its protonated form at acidic pH. The cause of this behavior can be found in the distances between D198 and the closest titratable sites (E4, H8, H69). For instance, the E4 – D198 distance is almost constant when ASR is embedded in the membrane ( $9.4 \text{ \AA} (\pm 0.3)$ ) while it fluctuates between  $9.8 \text{ \AA}$  and  $13.6 \text{ \AA}$  in its absence (see SuppInfo). The same behavior is found for H8, while it is reduced in the case of H69. In their protonated forms, these residues favor the deprotonation of D198, especially when they remain close to it. Eventually, D198  $pK_a$  becomes negative. Regarding D125, this residue is very close to the membrane in the acidic pH range (the distance between its carboxylate and the closest phospholipid head is  $5 \text{ \AA}$  as an average); the low dielectric constant of the membrane environment favors the neutral form of D125.

Table 2:  $pK_a$  values calculated fitting the Hill equation (3) using the deprotonated fractions retrieved from 40 ns long CpHMD trajectories (13C isomer) in water and in membrane.

Residue	Water		membrane	
	$pK_a$	$n$	$pK_a$	$n$
H8	7.07	0.63	6.15	0.88
H21	6.27	0.99	7.21	0.73
H69	5.43	0.75	6.05	0.83
H219	5.70	0.84	6.70	0.49
E4	4.89	0.78	5.27	0.60
E36	6.63	0.75	7.56	0.83
E62	4.83	0.86	4.72	0.83
E123	4.52	0.82	5.19	0.71
E160	4.77	0.81	4.92	0.90
D57	3.84	0.73	4.94	0.43
D75	<0	-	<0	-
D98	2.94	0.76	3.75	0.18
D120	5.30	0.69	2.05	0.36
D125	3.66	0.92	6.28	0.73
D166	<0	-	<0	-
D198	3.28	1.44	<0	-
D217	5.99	0.59	6.05	0.73
D226	4.71	0.84	4.51	0.88

Hill factors are also reported in Table 2. Ideally, for a mono-acid without any interaction with other titrating sites, the Hill factor  $n = 1$ .<sup>57</sup> If  $n$  is very far from 1, it probably means that the associated  $pK_a$  value has to be considered carefully, probably due to convergence issues (for instance D98 or D120). Besides this technical issue, it is expected that acidic sites are interacting with each other in a protein like ASR, eventually resulting in  $n$  values different from 1, as it is the case for the triad E4–H8–H69. This is also true for other titrating residues like H21–E36–E217–H219 which are close in space. Actually, in the latter case, H21 seems to behave as an independent titrating residue when ASR is not embedded in the membrane, resulting in a  $pK_a$  value closer to the reference one for histidine (6.5).

These  $pK_a$  changes with the nature of ASR surroundings have an impact on  $\lambda_{\max}$  values. In Figure 5 we have reported the QM/MM  $\lambda_{\max}$  values computed in both water and the membrane environments. When calculated in the membrane, ASR  $\lambda_{\max}$  values are closer to



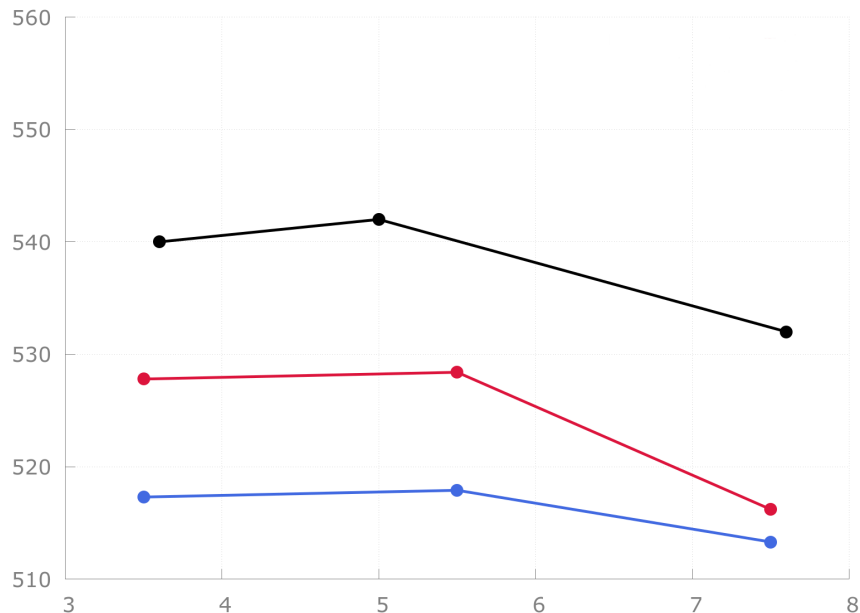


Figure 5: ASR (13C retinal) QM/MM (using 40000 structures extracted from 40 ns CpHMD trajectories) and experimental (black line)  $\lambda_{\max}$  values in nm as a function of the pH; the red line represents the calculations of ASR embedded in a membrane, while the blue line stands for the results of ASR in water.

the experimental ones. This result suggests an improved accuracy of our model when the protein is embedded in a lipid bilayer that resembles the physiological conditions. Between pH=3.5 and pH=5.5, we obtain a 0.6 nm red-shift, in agreement with the experimental 2 nm (note that the latter shift comes out of  $\lambda_{\max}$  values obtained at pH=5.0, not pH=5.5).<sup>14</sup> Between pH=5.5 and pH=7.5, our protocol predicts a 12.2 nm blue-shift, in excellent agreement with the experimentally reported 10 nm value. These two excellent results demonstrate that our CpHMD-then-QM/MM protocol is particularly well adapted to the study of ASR  $\lambda_{\max}$  pH-dependency. However, since we are titrating 18 residues at the same time, the analysis of its molecular origin is rather complex and cumbersome. Therefore, we decided to reduce the complexity of the problem by studying separately three smaller pH windows, each featuring a different  $\lambda_{\max}$  behavior: red-shifting in the acidic 3.0 to 4.5 pH range, showing a maximum in the intermediate 4.5 to 6.0 pH window, and blue-shifting in the neutral 6.0 to 7.5 pH window.

### 3.3 Refinements using three pH windows.

**CpHMD-then-QM/MM in each pH window.** In order to select a minimal number of titrating residues in each pH window, we have applied a two-step filtering procedure to our 13C ASR model:

1. A titratable residue is not titrated in a window if its deprotonated fraction within the pH range is lower than 5% or larger than 95%, according to our preliminary 40 ns CpHMD calculations.
2. Each list of titrating residues is then complemented with titratable residues whose protonation states show a large interaction with at least one of the members of the list.

All details are given in SuppInfo. The number and the list of titrating residues in each pH window are reported in Figure 6. Four residues (E4, D57, D98, E123) are included in all the three windows, due to their irregular or very flat titration curves (see SuppInfo). Some other ones belong to two consecutive pH windows (eg H69 which is titrating between pH=4.5 and 7.5).

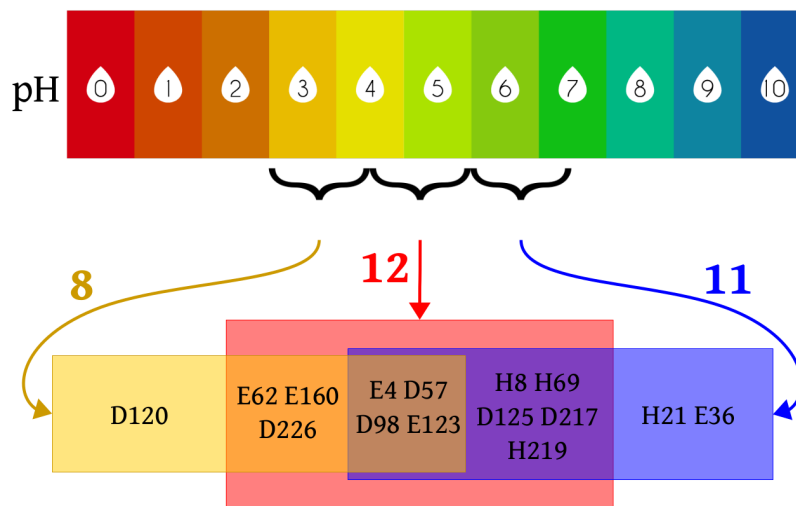


Figure 6: Selection of the titrated residues in the three narrow pH windows 3.0-4.5, 4.5-6.0 and 6.0-7.5. In a given window, the protonation state of the other titratable residues is fixed and decided according to their deprotonated fractions.

Each pH window is treated independently of the other ones. As a consequence of their

different numbers of titrating residues, and the resultant fixed protonation state of the non-titrated amino acids, the 3 different potential energy surfaces yield two  $\lambda_{\max}$  discontinuities at pH=4.5 and pH=6.0 (Figure 7).

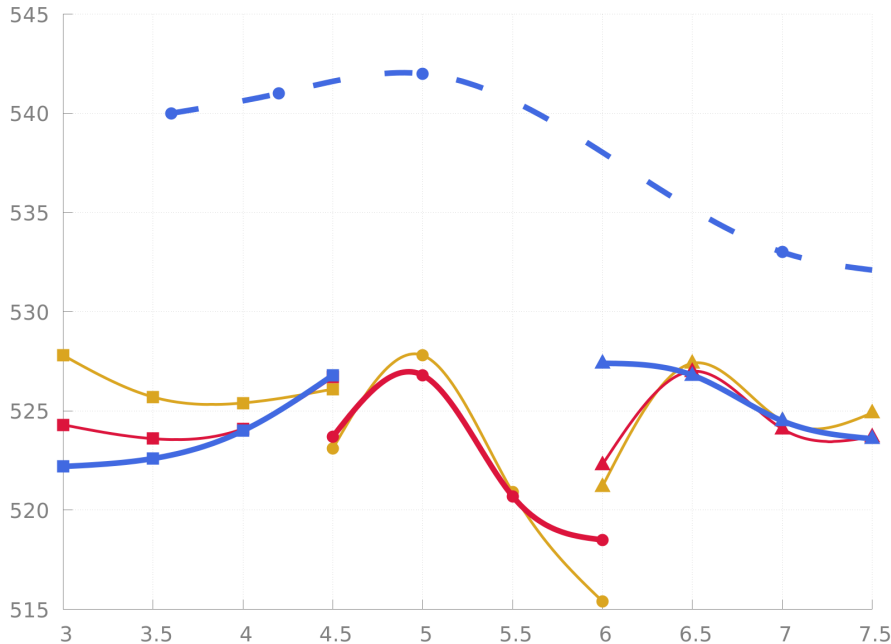


Figure 7: Experimental (dashed line) and computed (full lines) ASR 13C  $\lambda_{\max}$  (in nm) variation with the pH, increasing the CpHMD trajectory length: 10 ns in yellow, 20 ns in red, 30 ns in blue.

At first glance, our CphMD-then-QM/MM protocol is able to catch the qualitative features of  $\lambda_{\max}$  variations induced by the considered pH variations, even using a reduced number of titrating residues in each window. As a matter of fact, our results show a red-shift between pH=3 and pH=5, and then a blue-shift between pH=5 and pH=7.5. As shown in Figure 7 and confirmed by a bootstrapping analysis (see SuppInfo for the details), the  $\lambda_{\max}$  convergence in each window is improved by increasing the CpHMD sampling, while keeping reasonable the consumed computational resources.

**Global electrostatic analysis of the ASR  $\lambda_{\max}$  pH-dependency.** The retinal chromophore and its counterion (D75) keeping the same protonation states, the overall ASR structure being conserved at all pH values, the pH-dependency of the ASR absorption spec-

trum is probably due to pH-dependent electrostatic interactions between titrating residues and retinal. In order to test this hypothesis, we have calculated the average external potential felt by the retinal atoms for each considered pH value, as reported in Figure 8a. Two

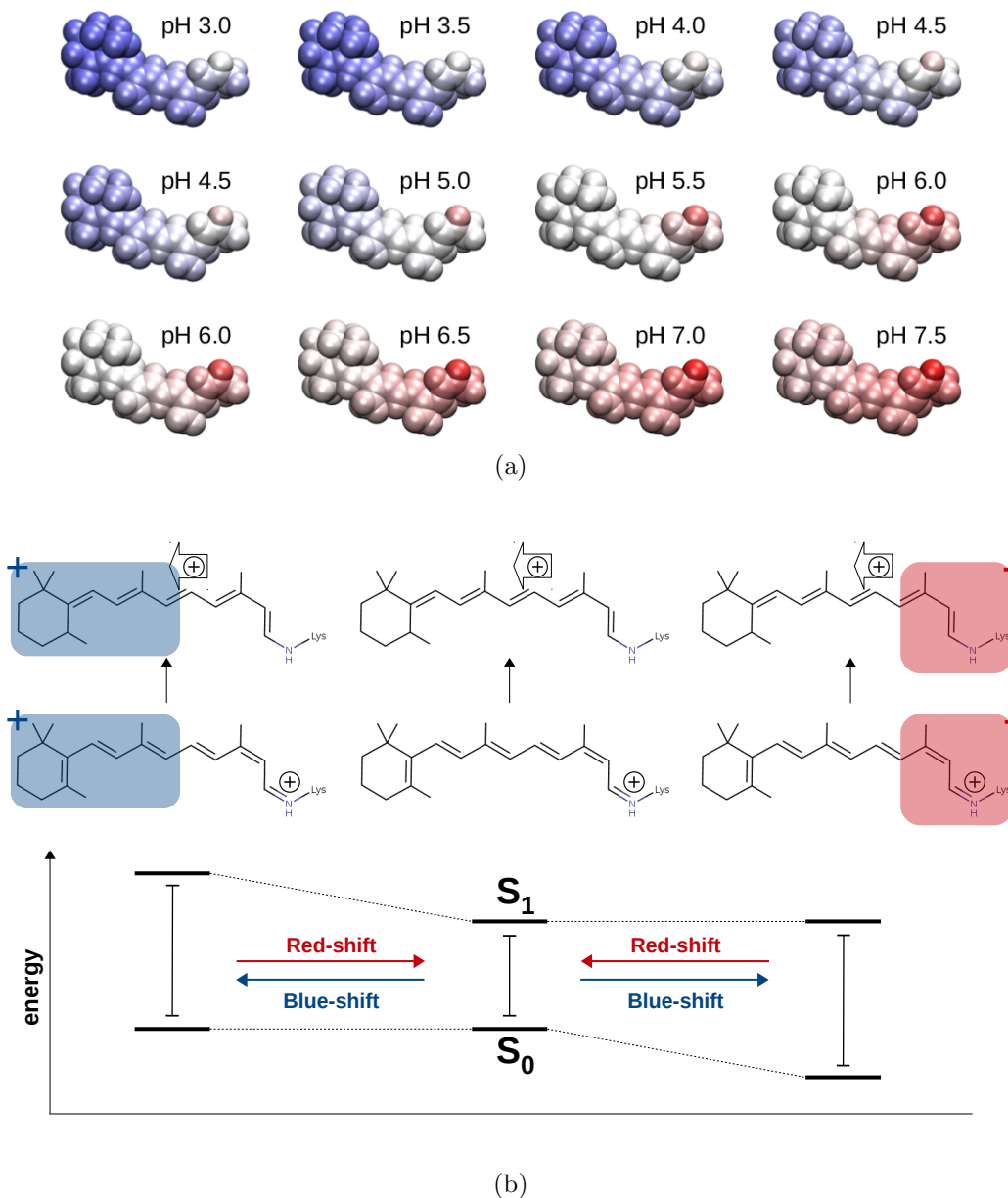


Figure 8: (a): Average external electrostatic potential felt by the retinal atoms, from pH=3.0 to pH=7.5. From blue (positive potential), to red (negative potential). (b): Stabilization/destabilization of the retinal ground ( $S_0$ ) and excited ( $S_1$ ) states triggered by either a positive external potential (left, in blue) in the  $\beta$ -ionone ring region and/or a negative external potential (right, in red) in the protonated Schiff base region, with respect to the reference situation in which no external potential is applied (middle).

main trends can be highlighted. The external potential felt by the Schiff base is very small at acidic pH and becomes more negative when the pH increases. Conversely, the external potential felt by the  $\beta$ -ionone ring is very positive at acidic pH and decreases to reach small negative intensity at pH=7.5. Remembering that the retinal positive charge, mainly located in the Schiff base region in the electronic ground state ( $S_0$ ), is translocated towards the  $\beta$ -ionone ring when retinal is excited to  $S_1$ ,<sup>58</sup> we can immediately infer that both  $S_0$  and  $S_1$  are more and more stabilized when the pH increases from acidic to neutral (Figure 8b). Of course, the pH-induced stabilizations of each electronic state have no reason to be equal. Accordingly, we can rationalize the acidic  $\lambda_{\max}$  red-shift by a slightly larger stabilization of  $S_1$  than  $S_0$ . Conversely, the  $\lambda_{\max}$  blue-shift between pH=5 and 7.5 can be attributed to a larger stabilization of  $S_0$  than  $S_1$ .

As a final check of the electrostatic origin of the pH-dependence of ASR 13C absorption spectrum, the  $\lambda_{\max}$  values at pH=3, 5 and 7 have been computed in the absence of the external potential in the PM7 calculations, keeping frozen the retinal structures extracted from the corresponding CpHMD trajectories. Whatever the pH, retinal  $\lambda_{\max}$  is 632 nm ( $\pm 0.4$ ), ie red-shifted by almost 100 nm, in agreement with experimental maximum absorption wavelength of retinal in vacuo.<sup>59</sup>

**Deciphering the molecular origin of the ASR 13C  $\lambda_{\max}$  pH-dependence.** Correlations between computed QM/MM vertical excitation energies and the individual protonation states of all the titrating residues have been determined for each pH window, using MLR analysis. First, it must be remembered that using three different pH windows with different sets of titrating residues results in exploring three different potential energy surfaces. As a consequence, some MLR coefficients are significantly different in different windows. For instance, D57 deprotonation induces a red-shift in both the acidic and intermediate windows and a tiny blue-shift in the neutral one. However, in the latter window, its deprotonated fraction is larger than 90%, casting doubts on the statistical relevance of the blue-shift. In

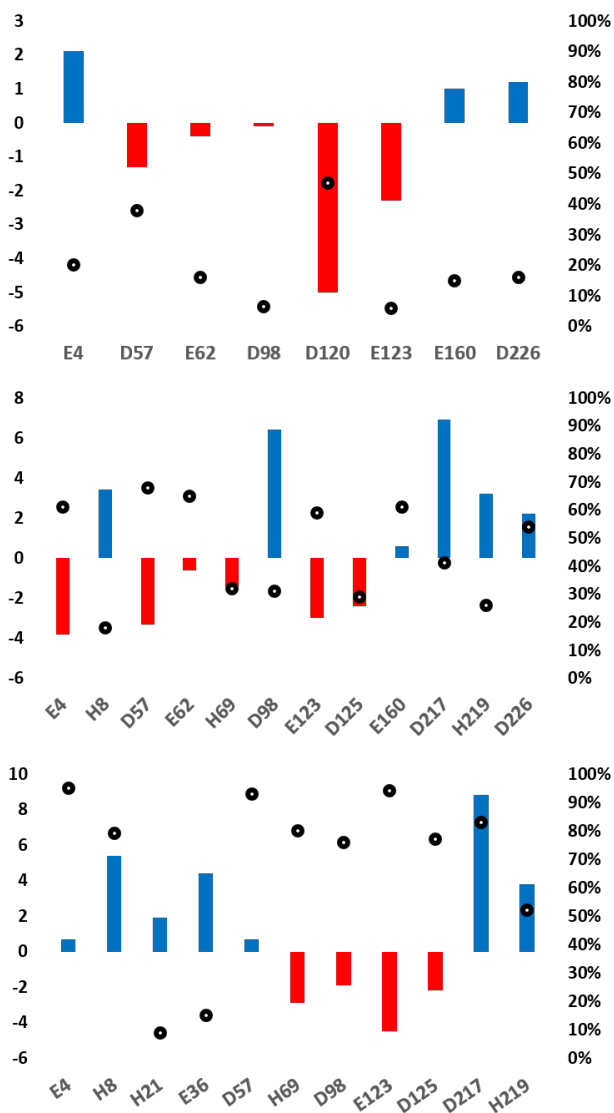


Figure 9: Left axis: individual contributions to  $\lambda_{\max}$  variations (bars, in nm). Right axis: deprotonated fraction of each titrating residue (black dots). Top: acidic window, middle: intermediate window, bottom: neutral window.

the case of E4, its deprotonation induces a blue-shift in the acidic window, a red-shift in the intermediate one, and a tiny blue-shift in the neutral one. Again, we don't consider the latter one, because of the corresponding large deprotonated fraction. The discrepancy between the acidic and intermediate windows can be related to the non-titration of H8 and H69 in the acidic one, despite the spatial proximity of these three amino-acids.

In the acidic window (pH=3.0–4.5), deprotonated fractions of D120 and D57 are 47%

and 38% respectively. Both titrating residues contribute to  $\lambda_{\max}$  variations with a red-shift. Three blue-shifting contributions are due to E4, D226 and E160, all with a deprotonated fraction comprised between 15 and 20%. Accordingly, we propose D120 as the main molecular origin of  $\lambda_{\max}$  red-shift. This residue is located at about 8Å from the retinal  $\beta$ -ionone ring. Accordingly, upon photo-excitation, the negative charge of the D120 deprotonated form will stabilize the retinal excited state, whose positive charge moves away from the Schiff base in  $S_1$  with respect to  $S_0$ , ultimately inducing the reported  $\lambda_{\max}$  red-shift.

In the middle window (pH=4.5–6.0), many important contributions are found, either red-shifting (D57, H69, E123, D125) or blue-shifting (D217, D226) ASR 13C retinal  $\lambda_{\max}$ . This complicated situation reflects the large number of titrating amino-acids in this window, i.e. many carboxylic acids and histidines showing deprotonated fractions comprised between 20% and 70%. Accordingly, the presence of the  $\lambda_{\max}$  peak can be roughly rationalized using the different stabilizations of both the retinal ground and excited state energies, as already presented above.

In the neutral window (pH=6.0–7.5), D217 gives the largest blue-shifting contribution, while being mainly deprotonated. Other residues significantly contribute to the blue-shift: H8, E36, H219. Residues H69, D98, E123, D125 induce a red-shift, however smaller. With the exception of H8, all the blue-shifting residues are spatially close and are possibly interacting, also with H21. Accordingly, we propose that the titrating residues H21, E36, D217 and H219 behave as a single acid-base cluster, being characterized by the experimental  $pK_a = 6.4$  value.<sup>14</sup>

**Confirmation of the electrostatic origin using the Minimal Electrostatic Model.**

Recalling that the Minimal Electrostatic Model is based on the assumption that pH-effects are mainly electrostatic (Equation 4), we have used different structures extracted from the CpHMD trajectories in the acidic and neutral pH windows and fitted the calculated  $\lambda_{\max}$  to the deprotonated fractions of each titrated residue. As expected from the low structural fluc-

tuations highlighted above, the results of the present analysis are qualitatively independent of the selected molecular model (see SuppInfo). In Figure 10, we present the energetic contribution of each titrating residue as a function of pH. In the acidic window, D120 is clearly the most  $\lambda_{\max}$  red-shifting residue. In the neutral window, D217 contributes the most to the  $\lambda_{\max}$  blue-shift. However, H8, H21, E36 and H219 also contribute to the blue-shift while H69, E123 and D125 are found to red-shift  $\lambda_{\max}$ . All these results are in line with the ones presented in the previous section and confirm the electrostatic origin of the pH effects on ASR  $\lambda_{\max}$ .

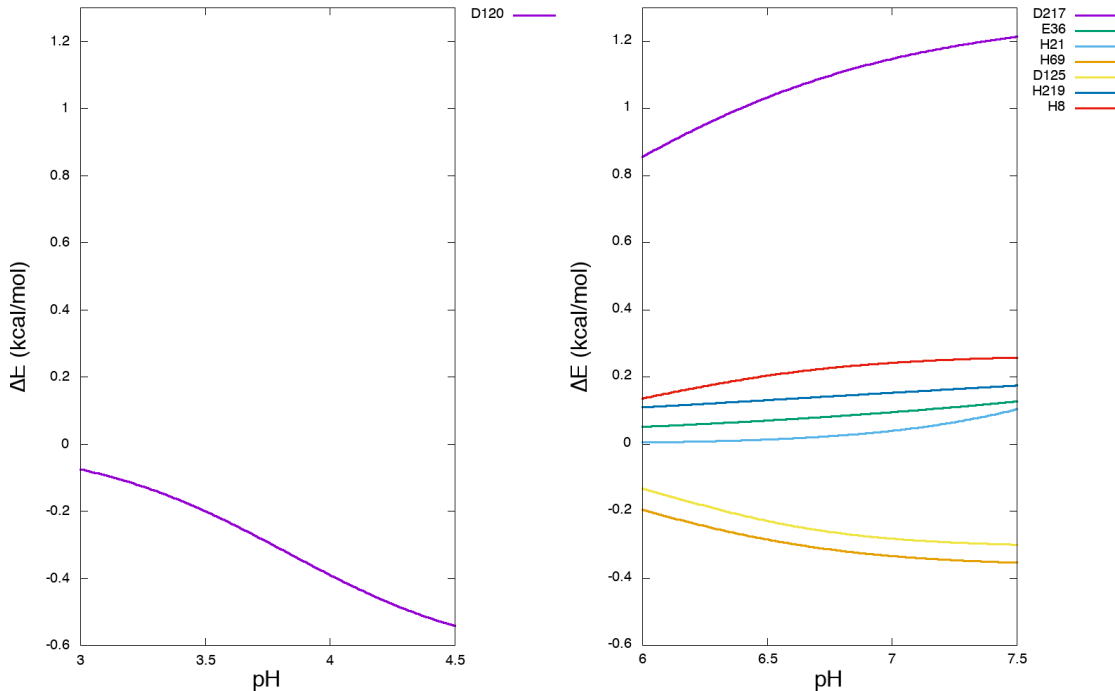


Figure 10: Minimal Electrostatic Model energy contributions of each important residue in the acidic and neutral pH windows. Negative (resp. positive) contribution translates to  $\lambda_{\max}$  red-shift (resp. blue-shift).

**The AT case.** The same CpHMD-then-QM/MM protocol has been applied to the AT retinal case, with the further hypothesis:  $^{13}\text{C}$   $pK_a$  values and protonated/deprotonated fractions for individual titrating residues are transferable to AT. As a consequence, the definition of the 3 pH windows, including the number and list of titrating residues per



window, is kept the same. The final  $\lambda_{\max}$  evolution as a function of the pH is reported in Figure 11. Qualitatively, the same trends are observed in AT and 13C: red-shift, a peak at

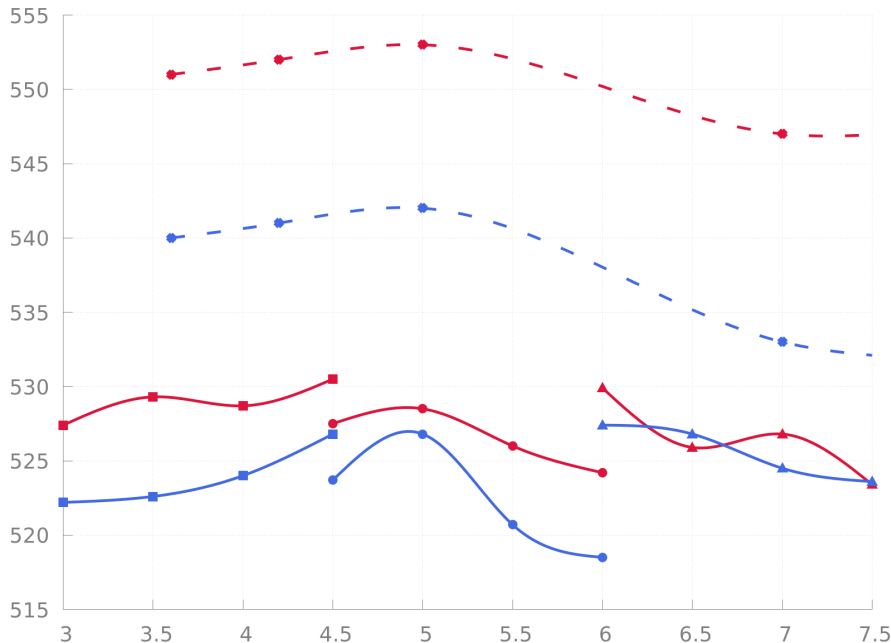


Figure 11: Experimental (dashed lines) and computed (full lines) ASR 13C (in blue) and AT (in red)  $\lambda_{\max}$  (in nm) variations with the pH.

pH=5, then blue-shift. Moreover, the ordering of 13C and AT  $\lambda_{\max}$  is correctly reproduced, even if the energy gap between the two retinal conformations is too small with respect to the experimental data, probably due to the qualitative QM level of theory. Nevertheless, the same convergence issues are met in both AT and 13C isomers.

## 4 Conclusions

In this article, we have reported the successful application of our CpHMD-then-QM/MM protocol to the study of the pH-dependent ASR absorption spectrum between pH=3 and pH=7.5.

The main result is the identification of the titrating residue responsible for the  $\lambda_{\max}$  red-shift in acidic conditions, namely D120 which was never suggested in the available literature.

We have also confirmed the implication of D217 in the  $\lambda_{\max}$  blue-shift between pH=5 and pH=7.5. However, we have found out that other residues, close to D217, are also involved. Accordingly, we suggest that the experimental  $pK_a$  (about 6.5) has to be attributed to a cluster of titrating residues located on the ASR cytoplasmic side, instead of D217 alone. Overall, it is demonstrated that pH-induced  $\lambda_{\max}$  tuning in ASR can be reproduced by means of simple opsin – retinal electrostatic interactions and achieved by residues located far from its retinal chromophore, as seen in figure 12. More involved and challenging studies would be needed in order to understand the very nature of these long-range interactions.<sup>34,36,37</sup> This finding suggests that the ASR structure is actually designed for reducing as much as possible the effect of pH on its absorption spectrum, and maybe on other photochemical properties.

In principle, the methodology used in this work could be applied to understand the origin of the  $pK_a = 8.5$  value experimentally reported for the 13C retinal isomer only.<sup>14</sup> However, major complications arise. First, at this pH value, other titratable residues can enter into the game: lysine and maybe some cysteine or tyrosine amino-acids. Second, it would be necessary to use a different starting structure (the one used in the current work, 1XIO, was obtained at pH=5.6), for instance the NMR-based 2M3G structure<sup>60</sup> obtained at pH=9.

On the methodological side, we have shown evidence that atomistic models based on a single protonation microstate are not relevant when pH-dependent properties are investigated. As a matter of fact, the predominant microstate represents less than 5% of the populated microstates, even at pH=7.5 which is the most favorable case in the case of ASR.

The CpHMD-then-QMMM protocol remains a very expensive tool, requiring a lot of computational resources. Accordingly, it is important to reduce the number of titrating residues, using a screening procedure based on the determination of correlations between the targeted property and the individual (de)protonation events. The minimal electrostatic model<sup>55</sup> endorses the electrostatic origin and the identification of the microscopic sources of the ASR pH-dependent absorption spectrum.

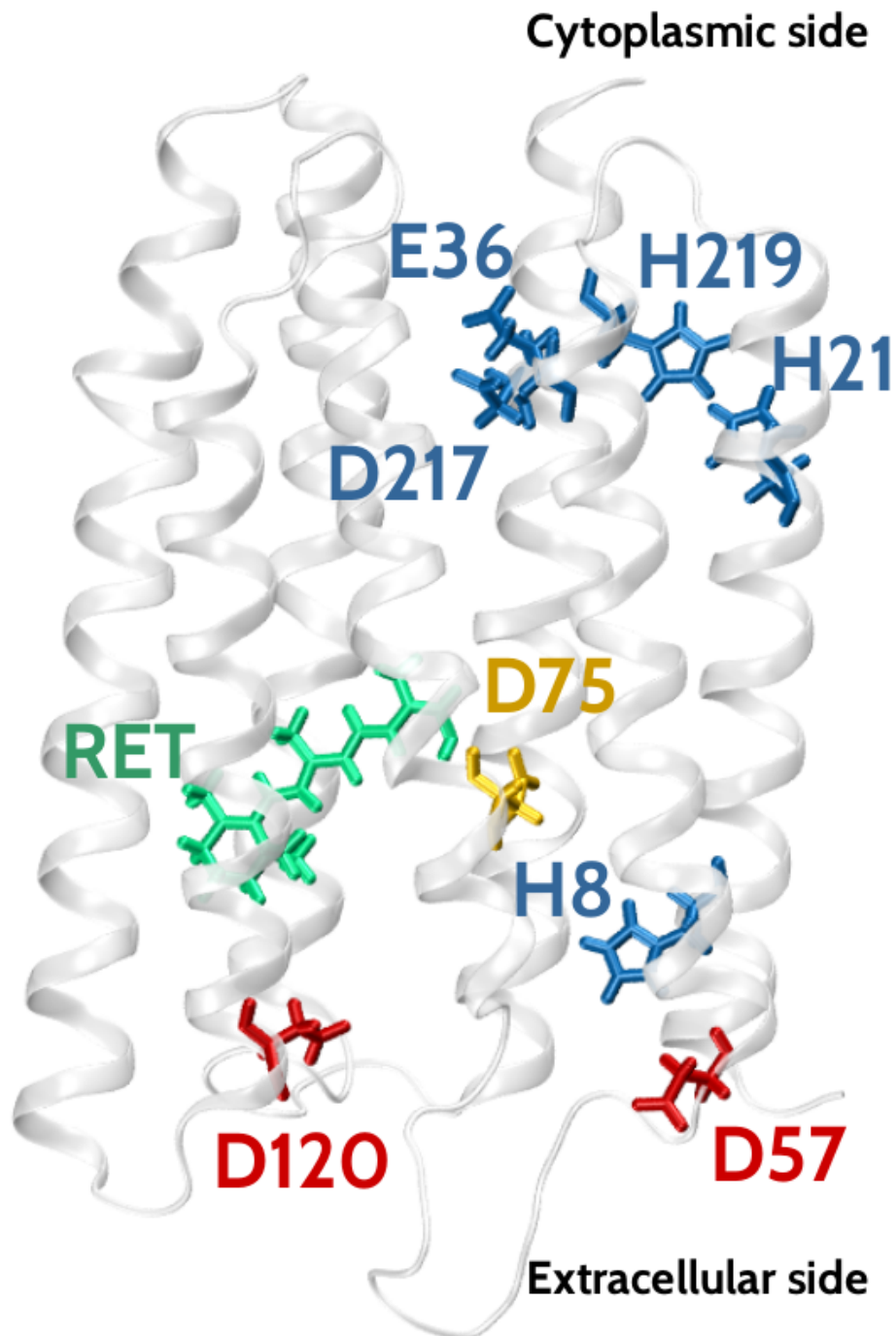


Figure 12: ASR representation featuring its retinal chromophore (RET), its counterion (D75) and its most important titrating residues. D120 (mainly) and D57 are responsible for the  $\lambda_{\max}$  red-shift between pH=3 and pH=5. H8, H21, E36, D217 and H219 are collectively responsible for the  $\lambda_{\max}$  blue-shift between pH=5 and pH=7.

## Acknowledgement

EP, VL, MS and NF thank the French Agence Nationale de la Recherche for funding (grant

ANR-14-CE35-0015-02, project FEMTO-ASR). This work was granted access to the HPC resources of Aix-Marseille Université financed by the project Equip@Meso (ANR-10-EQPX-29-01) of the program "Investissements d'Avenir" supervised by the Agence Nationale de la Recherche. FD is grateful to Contrat Plan Etat Région (CPER) IT2MP and Fonds Européen de Développement en Région (FEDER).

## Supporting Information Available

Computational details, CpHMD convergence analysis (titration curves, structural modifications), details regarding the 3 pH windows (selection, distance distributions, bootstrapping analysis, multilinear regression analysis, minimal electrostatic model analysis). This material is available free of charge via the Internet at <http://pubs.acs.org/>.

## References

- (1) Di Russo, N. V.; Estrin, D. A.; Martí, M. A.; Roitberg, A. E. pH-Dependent Conformational Changes in Proteins and Their Effect on Experimental pKas: The Case of Nitrophorin 4. *PLoS Comput. Biol.* **2012**, *8*, e1002761.
- (2) Harris, T. K.; Turner, G. J. Structural Basis of Perturbed pKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life* **2002**, *53*, 85–98.
- (3) Franco, L. F. M.; Pessôa Filho, P. d. A. On the solubility of proteins as a function of pH: Mathematical development and application. *Fluid Phase Equilibr.* **2011**, *306*, 242–250.
- (4) Jordan, I. K.; Kondrashov, F. A.; Adzhubei, I. A.; Wolf, Y. I.; Koonin, E. V.; Kondrashov, A. S.; Sunyaev, S. A universal trend of amino acid gain and loss in protein evolution. *Nature* **2005**, *433*, 633–638.

- (5) Ullmann, G. M.; Bombarda, E. *Protein Modelling*; Springer International Publishing: Cham, 2014; pp 135–163.
- (6) Banci, L.; Bertini, I.; Luchinat, C.; Viezzoli, M. S. pH-dependent properties of SOD studied through mutants on Lys-136. *Inorg. Chem.* **1993**, *32*, 1403–1406.
- (7) O’Keefe, D. O.; Cabiaux, V.; Choe, S.; Eisenberg, D.; Collier, R. J. pH-dependent insertion of proteins into membranes: B-chain mutation of diphtheria toxin that inhibits membrane translocation, Glu-349—Lys. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6202–6206.
- (8) DiGiammarino, E. L.; Lee, A. S.; Cadwell, C.; Zhang, W.; Bothner, B.; Ribeiro, R. C.; Zambetti, G.; Kriwacki, R. W. A novel mechanism of tumorigenesis involving pH-dependent destabilization of a mutant p53 tetramer. *Nat. Struct. Biol.* **2001**, *9*, 12–16.
- (9) Horng, J.-C.; Cho, J.-H.; Raleigh, D. P. Analysis of the pH-dependent Folding and Stability of Histidine Point Mutants Allows Characterization of the Denatured State and Transition State for Protein Folding. *J. Mol. Biol.* **2005**, *345*, 163–173.
- (10) Ormazabal, V.; Zuiga, F. A.; Escobar, E.; Aylwin, C.; Salas-Burgos, A.; Godoy, A.; Reyes, A. M.; Vera, J. C.; Rivas, C. I. Histidine Residues in the Na<sup>+</sup>-coupled Ascorbic Acid Transporter-2 (SVCT2) Are Central Regulators of SVCT2 Function, Modulating pH Sensitivity, Transporter Kinetics, Na<sup>+</sup> Cooperativity, Conformational Stability, and Subcellular Localization. *J. Biol. Chem.* **2010**, *285*, 36471–36485.
- (11) del Carmen Marín, M.; Agathangelou, D.; Orozco-Gonzalez, Y.; Valentini, A.; Kato, Y.; Abe-Yoshizumi, R.; Kandori, H.; Choi, A.; Jung, K.-H.; Haacke, S.; Olivucci, M. Fluorescence Enhancement of a Microbial Rhodopsin via Electronic Reprogramming. *J. Am. Chem. Soc.* **2018**, *141*, 262–271.
- (12) Sineshchekov, O. A.; Trivedi, V. D.; Sasaki, J.; Spudich, J. L. Photochromicity of

- Anabaena Sensory Rhodopsin, an Atypical Microbial Receptor with a cis-Retinal Light-adapted Form. *J. Biol. Chem.* **2005**, *280*, 14663–14668.
- (13) Tahara, S.; Kato, Y.; Kandori, H.; Ohtani, H. PH-Dependent Photoreaction Pathway of the All-Trans Form of Anabaena Sensory Rhodopsin. *J. Phys. Chem. B* **2013**, *117*, 2053–2060.
- (14) Rozin, R.; Wand, A.; Jung, K.-H.; Ruhman, S.; Sheves, M. pH Dependence of Anabaena Sensory Rhodopsin: Retinal Isomer Composition, Rate of Dark Adaptation, and Photochemistry. *J. Phys. Chem. B* **2014**, *118*, 8995–9006.
- (15) Shi, L.; Yoon, S. R.; Bezerra, A. G.; Jung, K.-H.; Brown, L. S. Cytoplasmic Shuttling of Protons in Anabaena Sensory Rhodopsin: Implications for Signaling Mechanism. *J. Mol. Biol.* **2006**, *358*, 686–700.
- (16) Hayashi, S.; Tajkhorshid, E.; Pebay-Peyroula, E.; Royant, A.; Landau, E. M.; Navarro, J.; Schulten, K. Structural Determinants of Spectral Tuning in Retinal Proteins Bacteriorhodopsin vs Sensory Rhodopsin II. *J. Phys. Chem. B* **2001**, *105*, 10124–10131.
- (17) Bucher, D.; Guidoni, L.; Rothlisberger, U. The Protonation State of the Glu-71/Asp-80 Residues in the KcsA Potassium Channel: A First-Principles QM/MM Molecular Dynamics Study. *Biophys. J.* **2007**, *93*, 2315–2324.
- (18) Rangarajan, R.; Galan, J. F.; Whited, G.; Birge, R. R. Mechanism of Spectral Tuning in Green-Absorbing Proteorhodopsin. *Biochemistry* **2007**, *46*, 12679–12686.
- (19) Frähmcke, J. S.; Wanko, M.; Phatak, P.; Mroginiski, M. A.; Elstner, M. The Protonation State of Glu181 in Rhodopsin Revisited: Interpretation of Experimental Data on the Basis of QM/MM Calculations. *J. Phys. Chem. B* **2010**, *114*, 11338–11352.

- (20) Vanni, S.; Neri, M.; Tavernelli, I.; Rothlisberger, U. A Conserved Protonation-Induced Switch can Trigger Ionic-Lock Formation in Adrenergic Receptors. *J. Mol. Biol.* **2010**, *397*, 1339 – 1349.
- (21) Grante, I.; Actins, A.; Orola, L. Protonation effects on the UV/Vis absorption spectra of imatinib: A theoretical and experimental study. *Spectrochim. Acta A* **2014**, *129*, 326 – 332.
- (22) Driant, T.; Nachon, F.; Ollivier, C.; Renard, P.-Y.; Derat, E. On the Influence of the Protonation States of Active Site Residues on AChE Reactivation: A QM/MM Approach. *ChemBioChem* **2017**, *18*, 666–675.
- (23) Molakarimi, M.; Mohseni, A.; Taghdir, M.; Pashandi, Z.; Gorman, M. A.; Parker, M. W.; Naderi-Manesh, H.; Sajedi, R. H. QM/MM simulations provide insight into the mechanism of bioluminescence triggering in ctenophore photoproteins. *PLOS ONE* **2017**, *12*, 1–19.
- (24) Gozem, S.; Luk, H. L.; Schapiro, I.; Olivucci, M. Theory and Simulation of the Ultrafast Double-Bond Isomerization of Biological Chromophores. *Chem. Rev.* **2017**, *117*, 13502–13565.
- (25) Iijima, E.; Gleeson, M. P.; Unno, M.; Mori, S. QM/MM Investigation for Protonation States in a Bilin Reductase PcyA-Biliverdin IX $\alpha$  Complex. *ChemPhysChem* **2018**, *19*, 1809–1813.
- (26) Ling, B.; Wang, X.; Su, H.; Liu, R.; Liu, Y. Protonation state and fine structure of the active site determine the reactivity of dehydratase: hydration and isomerization of  $\beta$ -myrcene catalyzed by linalool dehydratase/isomerase from *Castellaniella defragrans*. *Phys. Chem. Chem. Phys.* **2018**, *20*, 17342–17352.
- (27) Swails, J. M.; Roitberg, A. E. Enhancing Conformation and Protonation State Sampling

- of Hen Egg White Lysozyme Using pH Replica Exchange Molecular Dynamics. *J. Chem. Theory Comput.* **2012**, *8*, 4393–4404.
- (28) Huang, Y.; Chen, W.; Wallace, J. A.; Shen, J. All-Atom Continuous Constant pH Molecular Dynamics With Particle Mesh Ewald and Titratable Water. *J. Chem. Theory Comput.* **2016**, *12*, 5411–5421.
- (29) Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. Calculating Absorption Shifts for Retinal Proteins: Computational Challenges. *J. Phys. Chem. B* **2005**, *109*, 3606–3615.
- (30) Fujimoto, K.; Hasegawa, J.; Hayashi, S.; Kato, S.; Nakatsuji, H. Mechanism of color tuning in retinal protein: SAC-CI and QM/MM study. *Chem. Phys. Lett.* **2005**, *414*, 239–242.
- (31) Altun, A.; Yokoyama, S.; Morokuma, K. Spectral Tuning in Visual Pigments: An ONIOM(QM:MM) Study on Bovine Rhodopsin and its Mutants. *J. Phys. Chem. B* **2008**, *112*, 6814–6827, PMID: 18473437.
- (32) Tomasello, G.; Olaso-González, G.; Altoè, P.; Stenta, M.; Serrano-Andrés, L.; Merchán, M.; Orlandi, G.; Bottoni, A.; Garavelli, M. Electrostatic Control of the Photoisomerization Efficiency and Optical Properties in Visual Pigments: On the Role of Counterion Quenching. *J. Am. Chem. Soc.* **2009**, *131*, 5172–5186.
- (33) Rajamani, R.; Lin, Y.-L.; Gao, J. The opsin shift and mechanism of spectral tuning in rhodopsin. *J. Comput. Chem.* **2011**, *32*, 854–865.
- (34) y. Hasegawa, J.; Fujimoto, K. J.; Kawatsu, T. A Configuration Interaction Picture for a Molecular Environment Using Localized Molecular Orbitals: The Excited States of Retinal Proteins. *J. Chem. Theory Comput.* **2012**, *8*, 4452 – 4461.



- (35) Campomanes, P.; Neri, M.; Horta, B. A. C.; Röhrig, U. F.; Vanni, S.; Tavernelli, I.; Rothlisberger, U. Origin of the Spectral Shifts among the Early Intermediates of the Rhodopsin Photocycle. *J. Am. Chem. Soc.* **2014**, *136*, 3842–3851.
- (36) Guareschi, R.; Valsson, O.; Curutchet, C.; Mennucci, B.; Filippi, C. Electrostatic versus Resonance Interactions in Photoreceptor Proteins: The Case of Rhodopsin. *J. Phys. Chem. Lett.* **2016**, *7*, 4547–4553.
- (37) Yanai, K.; Ishimura, K.; Nakayama, A.; ya Hasegawa, J. First-Order Interacting Space Approach to Excited-State Molecular Interaction: Solvatochromic Shift of p-Coumaric Acid and Retinal Schiff Base. *J. Chem. Theory Comput.* **2018**, *14*, 3643–3655.
- (38) Pedraza-González, L.; Vico, L. D.; del Carmen Mariín, M.; Fanelli, F.; Olivucci, M. a-ARM: Automatic Rhodopsin Modeling with Chromophore Cavity Generation, Ionization State Selection, and External Counterion Placement. *J. Chem. Theory Comput.* **2019**, *15*, 3134–3152.
- (39) Udvarhelyi, A.; Olivucci, M.; Domratcheva, T. Role of the Molecular Environment in Flavoprotein Color and Redox Tuning: QM Cluster versus QM/MM Modeling. *J. Chem. Theory Comput.* **2015**, *11*, 3878–3894.
- (40) Bombarda, E.; Ullmann, G. M. pH-Dependent pKa Values in Proteins – A Theoretical Analysis of Protonation Energies with Practical Consequences for Enzymatic Reactions. *J. Phys. Chem. B* **2010**, *114*, 1994–2003.
- (41) Ullmann, G. M.; Bombarda, E. pKa values and redox potentials of proteins. What do they mean? *Biol. Chem.* **2013**, *394*.
- (42) Pieri, E.; Ledentu, V.; Huix-Rotllant, M.; Ferré, N. Sampling the protonation states: the pH-dependent UV absorption spectrum of a polypeptide dyad. *Phys. Chem. Chem. Phys.* **2018**, *20*, 23252–23261.

- (43) Vogele, L.; Sineshchekov, O. A.; Trivedi, V. D.; Sasaki, J.; Spudich, J. L.; Luecke, H. Anabaena Sensory Rhodopsin: A Photochromic Color Sensor at 2.0 Å. *Science* **2004**, *306*, 1390–1393.
- (44) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R.; Wu, X.; Xiao, L.; Kollman, P. *AMBER 2016*; University of California: San Francisco, 2016.
- (45) Swails, J. M.; York, D. M.; Roitberg, A. E. Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation. *J. Chem. Theory Comput.* **2014**, *10*, 1341–1352.
- (46) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (47) Hayashi, S.; Tajkhorshid, E.; Schulten, K. Structural Changes during the Formation of Early Intermediates in the Bacteriorhodopsin Photocycle. *Biophys. J.* **2002**, *83*, 1281–1297.
- (48) Stewart, J. J. P. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *J. Mol. Mod.* **2013**, *19*, 1–32.
- (49) Armstrong, D. R.; Fortune, R.; Perkins, P. G.; Stewart, J. J. P. Molecular orbital theory for the excited states of transition metal complexes. *J. Chem. Soc., Farad. Trans. 2* **1972**, *68*, 1839.

- (50) Stewart, J. Mopac2016. Stewart Computational Chemistry, Colorado Springs, CO, USA, <http://openmopac.net>.
- (51) Barbatti, M.; Ruckebauer, M.; Plasser, F.; Pittner, J.; Granucci, G.; Persico, M.; Lischka, H. Newton-X: a surface-hopping program for nonadiabatic molecular dynamics. *WIREs Computat. Mol. Sci.* **2014**, *4*, 26–33.
- (52) Barbatti, M.; Granucci, G.; Ruckebauer, M.; Plasser, F.; Crespo-Otero, R.; Pittner, J.; Persico, M.; Lischka, H. NEWTON-X: a package for Newtonian dynamics close to the crossing seam. 2015; [www.newtonx.org](http://www.newtonx.org).
- (53) Melaccio, F.; del Carmen Marín, M.; Valentini, A.; Montisci, F.; Rinaldi, S.; Cherubini, M.; Yang, X.; Kato, Y.; Stenrup, M.; Orozco-Gonzalez, Y.; Ferré, N.; Luk, H. L.; Kandori, H.; Olivucci, M. Toward Automatic Rhodopsin Modeling as a Tool for High-Throughput Computational Photobiology. *J. Chem. Theory Comput.* **2016**, *12*, 6020–6034.
- (54) Finley, J.; Malmqvist, P.-A.; Roos, B. O.; Serrano-Andrés, L. The multi-state CASPT2 method. *Chem. Phys. Lett.* **1998**, *288*, 299–306.
- (55) Stenrup, M.; Pieri, E.; Ledentu, V.; Ferré, N. pH-Dependent absorption spectrum of a protein: a minimal electrostatic model of Anabaena sensory rhodopsin. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14073–14084.
- (56) Pfaff, B.; Darrington, J.; Stover, J. GNU PSPP. 2017; [www.gnu.org/software/pspp/](http://www.gnu.org/software/pspp/).
- (57) Weiss, J. N. The Hill equation revisited: uses and misuses. *The FASEB Journal* **1997**, *11*, 835–841.
- (58) Demoulin, B.; El-Tahawy, M. M. T.; Nenov, A.; Garavelli, M.; Le Bahers, T. Intramolecular photo-induced charge transfer in visual retinal chromophore mimics: elec-

tron density-based indices at the TD-DFT and post-HF levels. *Theor. Chem. Acc.* **2016**, *135*, 96.

- (59) Knudsen, J. L.; Kluge, A.; Bochenkova, A. V.; Kiefer, H. V.; Andersen, L. H. The UV-visible action-absorption spectrum of all-trans and 11-cis protonated Schiff base retinal in the gas Phase. *Phys. Chem. Chem. Phys.* **2018**, *20*, 7190–7194.
- (60) Wang, S.; Munro, R. A.; Shi, L.; Kawamura, I.; Okitsu, T.; Wada, A.; Kim, S.-Y.; Jung, K.-H.; Brown, L. S.; Ladizhansky, V. Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nat. Methods* **2013**, *10*, 10071012.