# Absence of Selection for Quantum Coherence in the Fenna–Matthews–Olson Complex: A Combined Evolutionary and Excitonic Study

ACS central science

Research Article

Cite This: ACS Cent. Sci. 2017, 3, 1086-1095

http://pubs.acs.org/journal/acscii

# Absence of Selection for Quantum Coherence in the Fenna–Matthews–Olson Complex: A Combined Evolutionary and Excitonic Study

Stéphanie Valleau,*[,†] Romain A. Studer,[‡] Florian Häse,[†] Christoph Kreisbeck,[†] Rafael G. Saer,[§] Robert E. Blankenship,[§] Eugene I. Shakhnovich,[†] and Alán Aspuru-Guzik*[,†,∥]

[†]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States

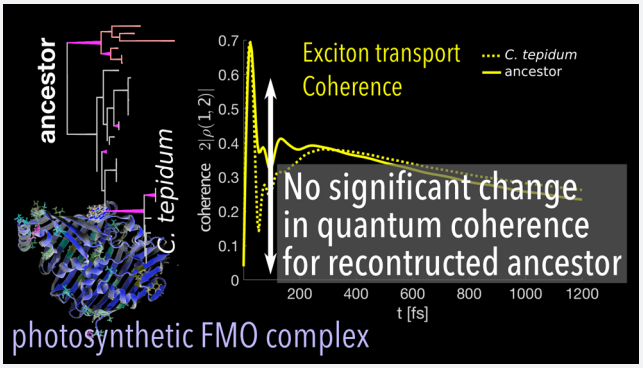
[‡]European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SD, U.K.

[§]Departments of Biology and Chemistry, Washington University in Saint Louis, One Brookings Drive, St. Louis, Missouri 63130, United States

[∥]Bio-inspired Solar Energy Program, Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada

Supporting Information

**ABSTRACT:** We present a study on the evolution of the Fenna–Matthews–Olson bacterial photosynthetic pigment–protein complex. This protein complex functions as an antenna. It transports absorbed photons—excitons—to a reaction center where photosynthetic reactions initiate. The efficiency of exciton transport is therefore fundamental for the photosynthetic bacterium's survival. We have reconstructed an ancestor of the complex to establish whether coherence in the exciton transport was selected for or optimized over time. We have also investigated the role of optimizing free energy variation upon folding in evolution. We studied whether mutations which connect the ancestor to current day species were stabilizing or destabilizing from a thermodynamic viewpoint. From this study, we established that most of these mutations were thermodynamically neutral. Furthermore, we did not see a large change in exciton transport efficiency or coherence, and thus our results predict that exciton coherence was not specifically selected for.




## INTRODUCTION

The Fenna–Matthews–Olson complex is a light-harvesting protein complex found in green sulfur bacteria. These bacteria first appeared about 1.6 billion years ago, in the Proterozoic era.[1] Green sulfur bacteria survive through anoxygenic photosynthesis; they use sulfide and other reduced sulfur compounds or hydrogen as photosynthetic electron donors.[2]

In the past decade, much research has been dedicated to understanding the excitation energy transfer in the Fenna–Matthews–Olson (FMO) protein complex.[3−13] Once absorbed, photons become molecular excitations, or excitons. These excitons are then transported through the complex due to their interactions with neighboring excitations, the protein scaffold, and the environment. A renewed interest in this complex arose when low temperature 2D-spectroscopy experiments suggested the presence of quantum coherence effects[14,15] in the exciton dynamics. Quantum coherent effects can be thought of as concurrent beats between electronic excitations which do not occur classically. Following these experiments, much theoretical work was carried out[6,16−19] with the goal of understanding whether quantum effects were present and, if so, how they contributed to the exciton transport.

Most of the studies have relied on the existence of an X-ray crystal structure of the FMO complex of either *Chlorobaculum tepidum* or *Prosthecochloris aestuarii*.

FMO is found in all of the anaerobic Chlorobi phyla, and recently it has also been found in aerobic *Chloracidobacterium thermophilum* of the Acidobacteria phyla.[20] In these organisms, the FMO complex forms a homo trimer (see Figure 1). Each monomer consists of 7/8 bacteriochlorophyll *a* (BChl-a) pigment molecules enclosed in a protein scaffold. The BChl pigment molecules interact with the protein scaffold through coordination and hydrogen bonds. The complex is thought to act as an excitonic wire, funneling the excitation from the baseplate to the reaction center, where charge separation occurs. This charge separation enables reactions which are fundamental to the organisms' survival.

**Received:** June 22, 2017
**Published:** August 30, 2017




DOI: 10.1021/acscentsci.7b00269
ACS Cent. Sci. 2017, 3, 1086−1095

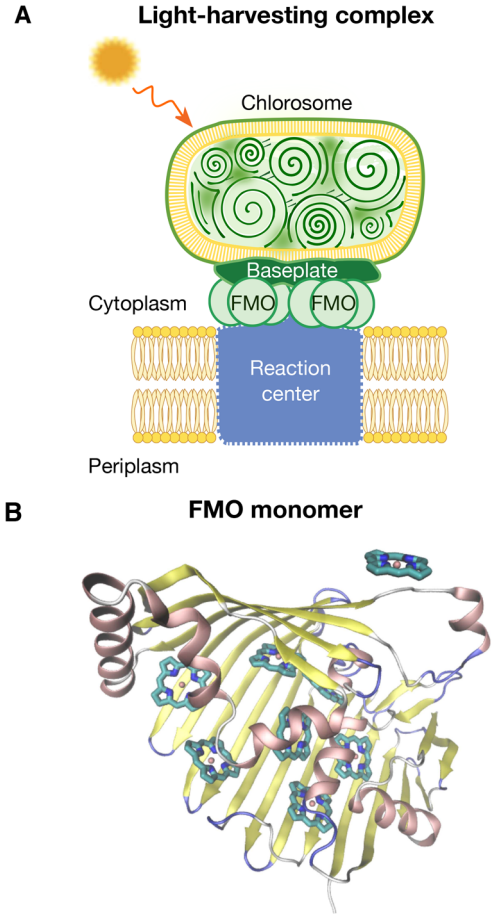



**Figure 1.** In panel A, an image describes the location of the FMO complex in the light-harvesting complex of green sulfur bacteria. The photons are absorbed by the chlorosome and transported as excitations to the FMO complex and ultimately to the reaction center. In panel B one can see the crystal structure of the FMO complex monomer for *Chlorobaculum tepidum* (PDB: 3ENI). The protein scaffold is colored in pink for alpha helices, in yellow for beta sheets, and in blue for loops. The 8 bacteriochlorophyll *a* (BChl-a) pigments are colored in cyan (their side chains are not shown for simplicity).

The biological origin of the complex remains a mystery. Olson et al.[21] speculated that the FMO complex might have come from an ancient reaction center. They looked for homology between the FMO protein and PscA, the reaction center protein of green sulfur bacteria. They found a signature sequence common to both; however, the sequences only had a 13% identity score. More recently, the role of some specific site mutations[22] and of cysteines in the FMO protein has been investigated as well.[23] It was found that cysteines appear to be fundamental for the photoprotection mechanism of the protein complex.

Regarding exciton transport in this system, there have only been few efforts to understand whether and how protein evolution might have influenced it. For instance, Müh et al.[24] computed the BChl transition energies and looked at the effect of including the point charges coming from the protein. They also looked at how changes in the polar groups of amino acids, such as the amino or hydroxyl groups, influence BChl transition energies. They found that the charges coming from the alpha helices seem to influence the energies strongly. On the other hand, no large change in the energies was observed when they modified the charge distribution of single amino acids as a means to emulate single point mutations. Experimentally, the Blankenship group[25] looked at comparing the optical properties of FMO coming from three different species to understand the effect of the protein scaffold. They observed spectral differences, and these were assigned to the way the BChl molecules bind to the protein scaffold in each species. This suggests that one may probe exciton dynamics and the biological role of coherence through point mutations. In this sense, we are most interested in mutations occurring in the vicinity of BChl binding sites as the effect of their charge is expected to have a stronger influence on the BChl energies.

For the above reasons, we have carried out a theoretical study of the protein complex from an evolutionary perspective. We aimed to identify whether quantum coherence in exciton transport and thermodynamic stability (free energy change upon folding) changed during the evolution of the protein complex. Hence, we carried out calculations which trace evolution of exciton transport as well as evolution of thermodynamic stability. For the thermodynamics, we studied the stability of the current day protein complex to mutations by computing and comparing the folding free energy change of the mutant to that of the nonmutated protein complex. Subsequently, we constructed an ancestor for the FMO complex and computed the folding free energies due to point mutations along the phylogenetic tree. Thus, with these two results, we were able to compare the stability of *C. tepidum* to the stability of its ancestors and see how it changed along the tree. To investigate the role of exciton transport, we computed and compared exciton transport in the ancestor with that of *C. tepidum*. Because two-dimensional photon echo (2D-echo) spectroscopy is the experimental tool employed to reveal coherent phenomena, we computed this spectrum to determine the role of coherence in the evolution from ancestor to *C. tepidum*. Finally, from the reconstructed phylogenetic tree, we identified sites under positive selection and determined whether positively selected amino acids were close to BChl molecules and could influence exciton transport.

## RESULTS

**Robustness of Current Day FMO Proteins to Mutations.** We computed the free energy change upon folding, $\Delta\Delta G = \Delta G^{\text{fold}}_{\text{mutant}} - \Delta G^{\text{fold}}_{\text{wild}}$, as a measure of robustness of the FMO complex to mutations. Calculations were carried out using FoldX[26] for the trimers of *C. tepidum* and *P. aestuarii* for all possible single point mutations (see Methods and Computational Details for details). In our model, each single point mutation occurs at the same time in all monomers of the trimer. We found that most single point mutations would be destabilizing ∼60% (see Figure 2). The overall landscape of mutations is the same for *C. tepidum* and *P. aestuarii* (see Figure S4). This is qualitatively similar to the general statistics of the effects of mutations on stability.[27,28]

We also looked at how the free energy change varies depending on the location of the mutation in the protein complex monomers of *P. aestuarii* and *C. tepidum* and found that mutations at the interface of monomers are among the most destabilizing (see Figure S5). These observations rely on the assumption that each single point mutation (in each monomer) does not impede the assembling of a trimeric structure.

**Ancestor Reconstruction, Selection of Mutations, and Variation of the Free Energy Landscape.** A phylogenetic tree was generated using the multiple sequence alignment built


DOI: 10.1021/acscentsci.7b00269
ACS Cent. Sci. 2017, 3, 1086−1095

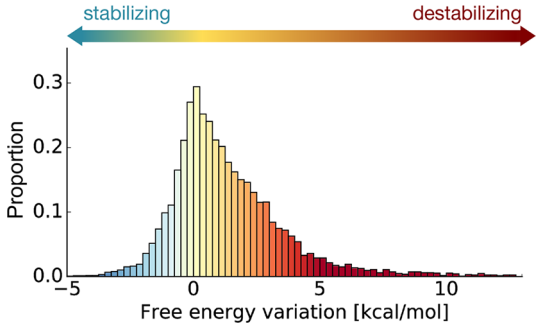



**Figure 2.** Histogram of free energy variation upon folding, $\Delta\Delta G = \Delta G^{fold}_{mutant} - \Delta G^{fold}_{wild}$, for all possible single point mutations in the FMO complex trimer of *P. aestuarii* (PDB: 3EOJ). The coloring is shaded to be more red for destabilizing mutations (positive values) and blue for stabilizing mutations (negative values). The bins in the histogram are normalized by the area under the histogram.

from the *fmoA* amino acid sequences (see details in Methods and Computational Details). The amino acid sequences were obtained by translating the *fmoA* monomer gene sequences. In Figure 3 we show the cladogram tree (phylogenetic tree is also shown in Figure S2).

The tree divides in two main clades: saltwater (e.g., *P. aestuarii*) and freshwater (e.g., *C. tepidum*) bacteria.

At each node of the tree, the sequence and structure of the protein scaffold were reconstructed. The ancestral amino acid sequences were obtained using Bayesian inference, and the scaffold structures were computed using homology modeling (see Methods and Computational Details). The pie charts above each branch indicate, based on the diameter, how many mutations occurred between the two branch nodes, and based on the color, whether they were stabilizing (blue), neutral (gray), or destabilizing (red). Most mutations were neutral in terms of modifying the free energy change upon folding.

Given the inferred ancestor amino acid sequence at the root, we constructed a guess for the full ancestral structure (protein scaffold and BChl molecules) through a combination of

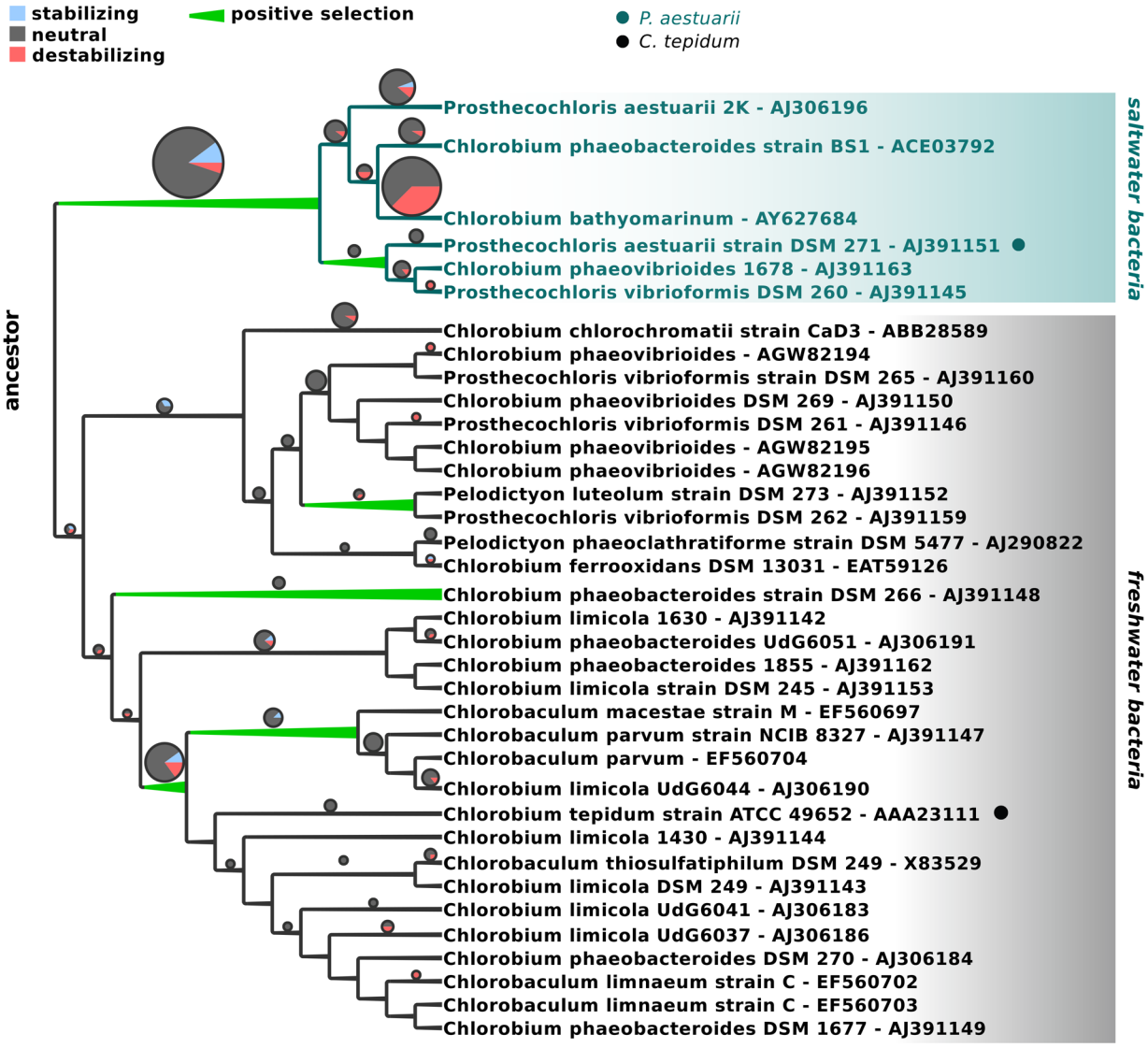



**Figure 3.** Rooted cladogram for *fmoA* computed using MAFFT and TranslatorX[29] for the alignment and PHYML.[30] Above each branch we show a pie-chart label with a diameter proportional to the number of single point mutations and color based on the associated free energy change $\Delta\Delta G$. Red represents destabilizing mutations, blue represents stabilizing mutations, and gray represents neutral mutations.



DOI: 10.1021/acscentsci.7b00269
*ACS Cent. Sci.* 2017, 3, 1086−1095

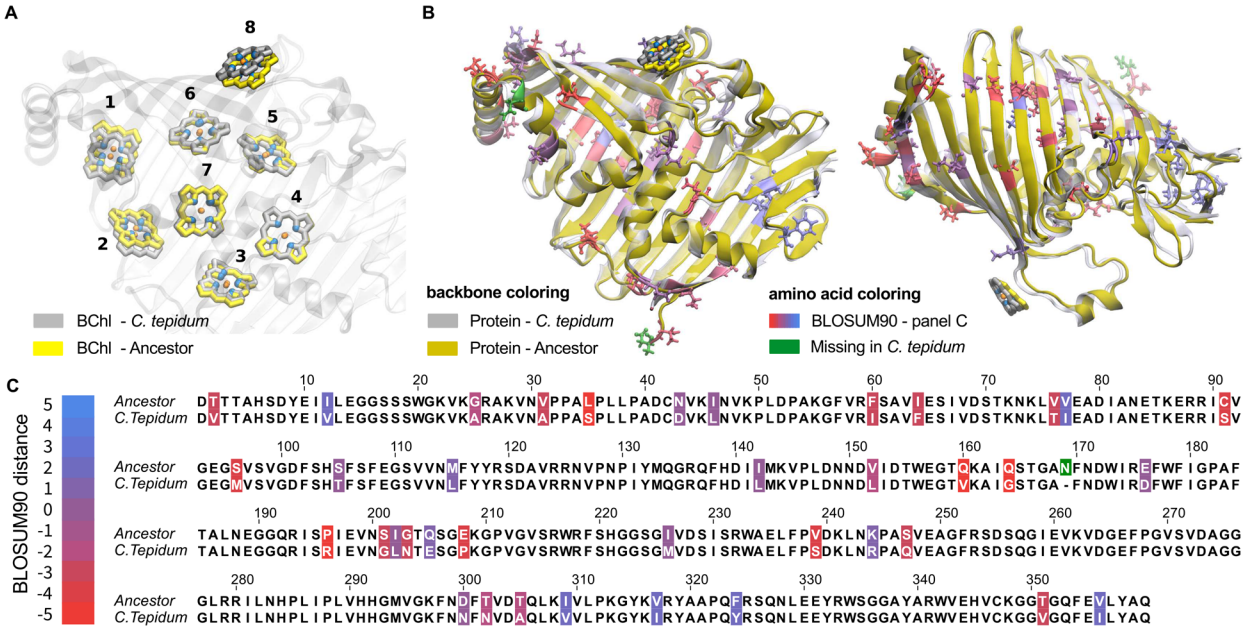



**Figure 4.** In panel A, we show the structure of the BChl molecules in a monomer of the ancestral structure (yellow) overlaid with the BChl molecules in the current day structure of *C. tepidum* (light gray). The images were generated from snapshots of the molecular dynamics simulations at 300 K. In panel B, for the same snapshots we compare the backbone of the ancestral FMO structure (yellow) to that of the current day *C. tepidum* structure (light gray backbone). Some amino acids are colored based on the BLOSUM90 matrix, which quantifies how different they are (positive numbers indicate easy substitution while negative numbers indicate difficult substitution). These amino acids are also highlighted in the legend in panel C. The structures are snapshots from the molecular dynamics simulations at 300 K. The amino acids which differ in the two structures are colored based on the BLOSUM90 matrix which quantifies how likely it is to substitute one amino acid with the other. The amino acid mutants in the ancestral structure are also shown using the bond and atom representation.

homology modeling and molecular dynamics simulations (see Methods and Computational Details). An image of the structure of the putative ancestor compared to that of *C. tepidum* for one monomer is shown in Figure 4. The images in each panel are constructed from a snapshot of the molecular dynamics simulations at 300 K after equilibration. In panel B, we show two orientations of one monomer of the protein complex. No large difference in the backbone of the protein is observed. This is due to the fact that the protein scaffold was constructed with homology modeling. The differences come from the location of side chains and the change in charge of the mutated amino acids. In panel B, mutations are highlighted using a color scheme based on the BLOSUM90 distance. We see a few red mutations (large negative BLOSUM90 distance) on the $\beta$ sheet which could be a problem for the formation of the trimer. Another effect of the mutations is the change in orientation of BChl molecules (see panel A). This change in orientation of the BChls influences the coupling among their first excited states (see Table S1). In particular, the coupling between the first excited states of BChl 3 and 4 is weaker in the ancestor than in *C. tepidum* and, on the other hand, the coupling between the first excited states of BChl 4 and 5 is stronger in the ancestor than in *C. tepidum*. Finally, as can be seen in Table S2, the average distance of the amino acids from the BChl molecules changes between the ancestor and *C. tepidum*, and the differences are on the order of 0.1−0.4 Å.

Given the phylogenetic tree, we also identified branches under positive selection (bright green) by using the branch−site model (details in Methods and Computational Details). In Figure 5, we show mutations which were positively selected for along these branches in the structure of *C. tepidum*. Most positively selected mutations are distant from the BChl

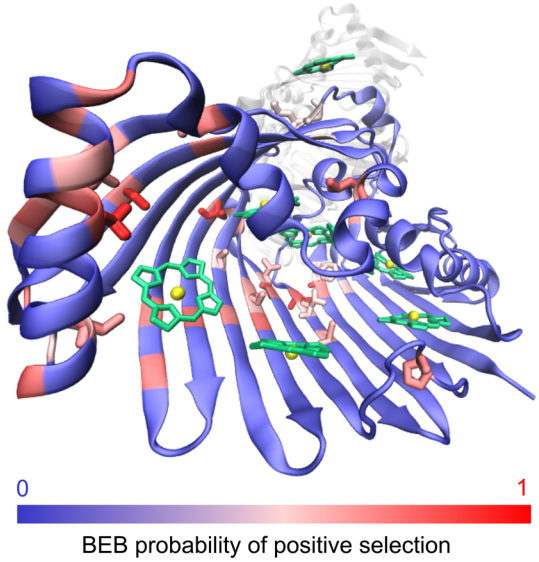



**Figure 5.** Positively selected sites shown in the structure of *C. tepidum* (PDB: 3ENI). The color gradient corresponds to the strength of positive selection (darkest red, strongest selection; blue, no positive selection). We only plot the side chains of positively selected residues which are within 6 Å of the magnesium in the BChl molecules.

residues, and therefore we currently do not believe that they had a strong role in modifying exciton transport. Given the high computational cost associated with constructing the full structure (with BChl molecules) at each node, we did not investigate the role of each positively selected mutation using atomistic *ab initio* methods. However, we computed the free-energy changes associated with the mutations.



DOI: 10.1021/acscentsci.7b00269
*ACS Cent. Sci.* 2017, 3, 1086−1095

In Figure 6, panel A, we show the histogram of free energy changes which connects the ancestor to current day species in red and, in cyan, the histogram of folding free energy variation for all possible single point mutations in the FMO complex trimer of *C. tepidum*. We found that 65% of mutations are destabilizing for *C. tepidum*. On the other hand only 32% of mutations of the ancestral FMO complex along the phylogenetic tree were destabilizing while 21% were stabilizing and 47% were neutral. This could indicate that mutations which favored the folding free energy change variation during evolution led to more stable current day protein structures. However, we cannot claim this in general as the trend changes for different current day species. For instance, as can be seen by comparing the two figures in panel B of Figure 6, the distribution of folding free energy changes for mutations which connect the ancestor to *P. aestuarii* is significantly different from that for mutations which connect the ancestor to *C. tepidum*. This difference was confirmed by the Kolmogorov−Smirnov test. Furthermore, we computed that, for the mutations which connect the ancestor to *P. aestuarii*, 25% are stabilizing and 19% are destabilizing. However, in the case of mutations which lead to *C. tepidum*, 47% are destabilizing while 22% are stabilizing. This could indicate that *C. tepidum* is less stable than the ancestor while *P. aestuarii* is more stable than the ancestor. For a more accurate comparison, however, one would need to take into account two environments for the ancestor (in salt water versus fresh water) as well as account for the coevolution of the baseplate, chlorosome, and reaction center which are located above and below the FMO complexes in the light-harvesting complex.

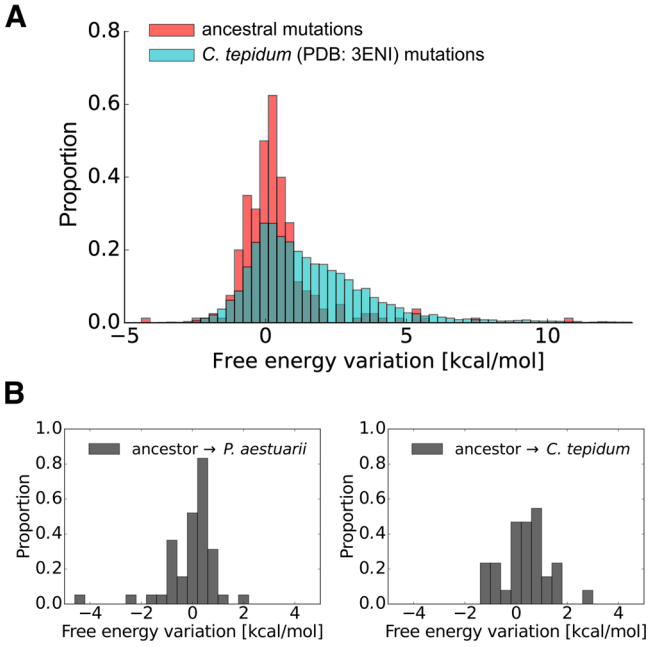



**Figure 6.** Panel A: In blue we show a histogram of the folding free energy changes for all possible single point mutations of the current day *C. tepidum* FMO complex structure. In red, we show a histogram of the folding free energy changes for all the single point mutations which are necessary to connect the ancestral FMO structure to the current day structures (i.e., all single point mutations along all branches). Panel B, right-hand side: histogram of the folding free energy changes for the mutations which connect *P. aestuarii* to the ancestor. Left-hand side: histogram of the folding free energy changes for the mutations which connect *C. tepidum* to the ancestor. Note that, for all histograms, each bin contains the number of points with energy in that interval normalized by the area under the entire histogram.

**Exciton Transport and Coherence.** *Absorption, Circular Dichroism, and Linear Dichroism.* In Figure 7 we compare the absorption spectrum of the ancestor to that of *C. tepidum*. The ancestor spectrum shows three main absorption peaks at 807, 819, and 830 nm while *C. tepidum* shows strong absorption at 804, 814, and 826 nm. These trends and the corresponding circular dichroism spectra (see Supporting Information) of the ancestor and of *C. tepidum* are very similar, respectively, to the spectra of FMO in *Chlorobium thiosulfatophilum* and *Chlorobium limicola*.[31] However, we do not see strong evidence of a similarity between the ancestor spectrum and the reaction center absorption spectra reported in that work. The small bump at about 845 nm (see Figure 7) is most likely due to noise, as it does not correspond to any of the frequencies of the equivalent Fermi golden rule absorption spectrum. Therefore, our current results do not confirm that the ancestor is related to current day PscA, the reaction center protein of green sulfur bacteria; however, we cannot exclude that it may have been related to ancestral reaction centers.

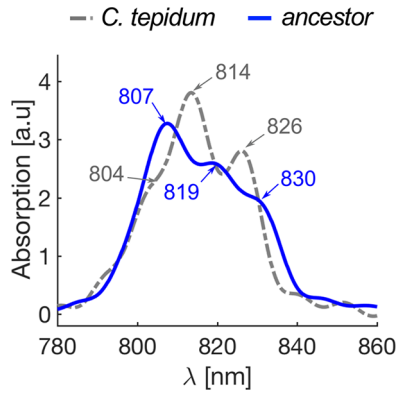



**Figure 7.** Comparison of simulated absorption spectra of the ancestral FMO complex (blue line) to that of current day *C. tepidum* at 300 K (gray dashed line) in arbitrary units. Spectra were computed using the Qy transitions obtained by using QM/MM. See Methods and Computational Details for more details.

In the Supporting Information we also show the comparison of our *C. tepidum* spectra to the experimental absorption, linear dichroism, and circular dichroism spectra.

*Coherence and 2D Spectra.* Using the hierarchy equation of motion (HEOM) non-Markovian master equation method,[32−34] we computed the 2D-echo spectra of the reconstructed ancestral FMO complex and of *C. tepidum* at 150 K. We computed the spectra at this temperature given that coherent effects are usually larger at lower temperatures. Therefore, this will allow us to make a clearer comparison between the ancestral FMO complex and that of *C. tepidum*. We also computed the population dynamics and coherences at 300 K, and these are shown in the Supporting Information. To this end, we employed the *QMaster* software package[35] (for more details see Methods and Computational Details). The spectral densities which were employed are described in Methods and Computational Details. These were obtained by fitting the original QM/MM spectral densities to Drude−Lorentz spectral densities with three peaks. The spectral


DOI: 10.1021/acscentsci.7b00269
ACS Cent. Sci. 2017, 3, 1086−1095

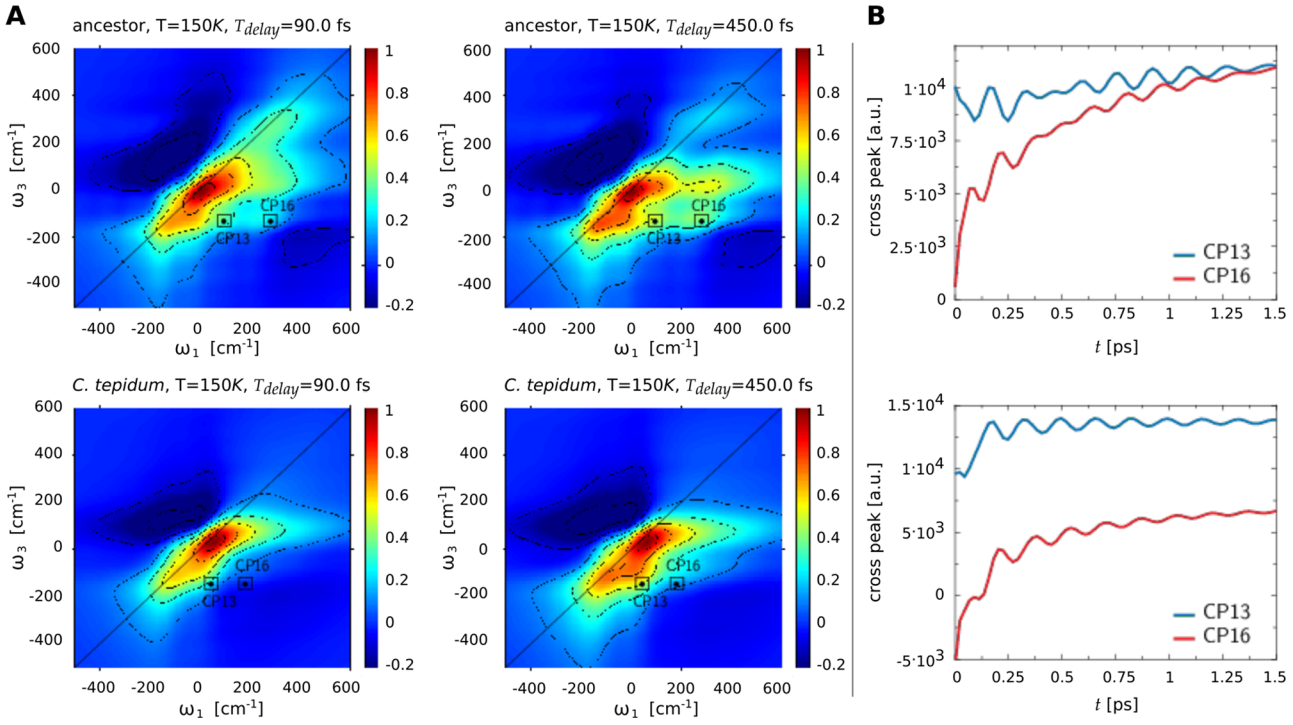



**Figure 8.** Panel A: Contour plots of the total computed signal for the 2D-echo spectra at different delay times at 150 K. The figures in the top row show results for the ancestor at a delay time of 90 and 400 fs, and the bottom figures show the analogous plots for *C. tepidum*. Panel B: The amplitude of oscillation of two specific cross peaks. The top image is for the ancestor, and the bottom image is for *C. tepidum*.

densities indicated a similar coupling strength of the system to the bath for the ancestral and *C. tepidum* FMO complexes for almost all modes. The total signal of the 2D-echo spectra, i.e., sum of the stimulated emission (SE), ground state bleaching (GB), and excited state absorption (ESA) pathways, is shown in Figure 8. From the oscillations of the cross peaks in panel B we cannot deduce an improvement in quantum coherence or a significant difference between the ancestor and *C. tepidum*. For all species, cross-peak beatings in the 2D-echo spectra are dominated by ground state vibrations[36] superposed by minor vibronic contributions. This suggests that optimizing exciton transport through quantum coherence was not part of the evolution of the FMO complex.

## CONCLUSION

We have reconstructed a structure for the ancestral Fenna–Matthews–Olson complex and found that mutations of its amino acids over time have influenced the free energy upon folding of the protein complex. In particular, the current day species of *P. aestuarii* was obtained through stabilizing mutations while mutations which lead to the current FMO complex in *C. tepidum* were less stabilizing. Regarding exciton transport, we do not observe a significant change or improvement in the efficiency of exciton transport or quantum coherent transport. These results suggest that the complex did not evolve to optimize for quantum coherent transport. It is also possible that the ancestor was already sufficiently optimized for exciton transport and therefore further evolution or optimization was not necessary. One of the results we find is that mutations most likely led to better binding of the FMO complex to the baseplate and reaction center. This is confirmed by the larger instability associated with mutating residues in those locations. This leaves an open question: perhaps the overall exciton transport, from the chlorosome to the reaction center, was optimized by favoring mutations in FMO which led to a stronger binding of FMO to its neighboring protein systems.

For future work, we are interested in looking at the exciton transport at each node of the phylogenetic tree and further in expressing the ancestral sequence in current day *C. tepidum* to obtain an experimental structure and 2D-echo spectra for the ancestor. Finally it would be interesting to look at the role of photoprotection to see whether the mutations could have influenced it.

## METHODS AND COMPUTATIONAL DETAILS

**Stability of the *C. tepidum* and *P. aestuarii* Fenna–Matthews–Olson Complex to Mutations.** Each amino acid of the trimer and monomer of the FMO complex of *P. aestuarii* and *C. tepidum* was mutated to each of the other possible 19 amino acids using FoldX.[26] The corresponding free energy variation upon folding $\Delta\Delta G = \Delta G_{mut} - \Delta G_{wt}$ was obtained for each mutation. In order to account for the effect of mutations at the interface between monomers within the trimer, the mutations were carried out simultaneously in each monomer of the trimer, and the resulting free energy variation was normalized by dividing by three.

In FoldX, mutations are modeled as follows. The initial crystal structure is optimized to remove any eventual steric clashes. Then, the residue of interest is mutated and its nearest neighbors are mutated to themselves and conformationally relaxed to remove any local clash. The nearest neighbors are mutated to themselves so that their geometry may be optimized together with that of the central residue. During this procedure the backbone of the protein is kept fixed, and all other residues that are far from the one of interest are also kept fixed. The stability, $\Delta G_{wt}$, of this relaxed structure is obtained by using an effective energy function. Subsequently, the residue of interest


DOI: 10.1021/acscentsci.7b00269
ACS Cent. Sci. 2017, 3, 1086–1095

is mutated to each of the other 19 amino acids and its neighboring amino acids to themselves, and the various $\Delta G_{mut}$ are computed.

The procedure described above was repeated 5 times for each mutation to ensure that the minimum energy conformations of large residues that have many rotamers were identified. The effective energy function ("effective energy" here refers to the Helmoltz free energy of a system (protein + solvent) for a fixed protein conformation) in FoldX has been optimized for amino acid sequences; thus all BChl-*a* molecules could not be included directly in this calculation. To account for their presence, the structures given in input to FoldX were obtained by homology modeling with Modeller.[37] In the Modeller simulations, the BChl molecules were inserted and kept fixed as hard spheres. In the FoldX optimization the residues that are known to bind to BChl-*a* molecules from the crystal structure analysis of FMO complexes from current day species were also kept fixed. A similar procedure was recently employed successfully for the case of RubisCO where Mg atoms are present.[38]

**Determination of Positively Selected Sites: Site and Branch−Site Models.** *Phylogenetic Tree.* All *fmoA* gene sequences from the EMBL database were gathered and employed to generate an alignment using translatorX with MAFFT.[29] Phylogenetic trees were constructed using PhyML[30] with 5024 bootstraps. The parameters and settings used in PhyML were as follows. The LG substitution model was chosen. We selected to have four substitution rate categories. The alpha parameter for the gamma distribution of sites was set to be estimated by the code. Both NNIs and SPR methods were used to search for the optimal tree topology, and finally, tree topology, branch lengths, and rate parameters were chosen to be optimized by the code. A set of three trees were constructed. In the first tree, all sequences were included excluding FJ210646, as it was missing about 100 residues. Therefore, including it would have introduced more error in the tree reconstruction phase. For the second tree, we removed *Chloroacidobacterium thermophilum* (ABV27353) as it had the lowest sequence identity percent and thus also introduced more uncertainty. In the third tree, *Chloroherpeton thalassium* (ACF13179) was removed as it was the second most distant sequence from the rest. We computed ancestral sequences and searched for positively selected sites on all three trees, but we present results for the third tree as we believe it has the smallest error.

*Site and Branch−Site models.* The branches under positive selection were identified using the branch−site model[39] as implemented in CodeML. This model allows for variation of $\omega = \frac{dN}{dS}$ (the ratio of nonsynonymous to synonymous mutations) among branches and sites. For each simulation, one branch is selected as a foreground branch and, using CodeML, we computed the likelihood that the branch was under positive selection. The input DNA sequence alignment was the same as the one generated to deduce the phylogenetic tree. Given a branch under positive selection, we determined which amino acids were under positive selection by using the Bayes empirical Bayes (BEB) model.[40] We corrected the probability values *p*-values for false discovery rate (FDR) by using the *q*-value.[41,42]

**Reconstruction of the Ancestral Protein Structure of FMO.** In the Supporting Information, we have included Figure S1, which reports the procedure employed to reconstruct the ancestral sequence and structure computationally.

*Sequence Reconstruction.* An ancestral sequence was reconstructed using maximum likelihood as implemented in FastML[43] and the LG+G model for amino acid substitution. The LG+G model was established to be the best model under the Bayesian Information Criterion (BIC). With FastML, we obtained the most probable sequences, together with the posterior probabilities for each character and indel at each sequence position for each internal node of the tree. As inputs, we used the phylogenetic trees described in the previous section and rooted them based on the most distant sequence. The alignments used in the input were the same as those employed to generate the trees. For the last tree that excludes *C. thermophilum* and *C. thalassium*; we used the midpoint root, as most sequences are similar, (see Figure 3) and the branching is almost identical to that obtained when *C. thalassium* or *C. thermophilum* are included. In Figure S3, we report a plot of the marginal probability for each amino acid reconstructed in the ancestral sequence.

*Structure Reconstruction and Molecular Dynamics.* The ancestral structure was built by using homology modeling with satisfaction of spatial restraints as implemented in Modeller.[37] The method of modeling by satisfaction of spatial restraints works as follows. Constraints are generated on the structure of the target sequence using its alignment to the related protein structures as a guide. The restraints are obtained based on the assumption that the structures will be similar in the aligned regions. In Modeller, the form of the restraints was obtained from a statistical analysis of the relationships between similar protein structures in a database of 105 alignments that included 416 proteins of known 3D structure.[37] These restraints are supplemented by stereochemical restraints such as bond lengths, angles, etc., which are obtained from a force field. Once all these restraints are established, the model structure is obtained in Cartesian space by minimizing the violations of all restraints.

We used the structure of *P. aestuarii* (PDB: 3EOJ) as the homologue and generated 100 possible structures with slow refinement. The ancestral structure was chosen to be the one with the best molpdb factor. The initially reconstructed structure only included the protein scaffold, as there is no current parametrization for BChl's in Modeller. Thus, BChl molecules were included subsequently using minimization as implemented in NAMD.[44] The chromophores were initially positioned as in the homologous structure, and the structure was optimized by minimizing the energy in two steps. In the first step, the backbone was kept fixed, and for the second minimization, the backbone was free to move. The structure was subsequently equilibrated without constraints using the Amber14 software package,[45] with the Amber ff99SB force field.[46] The BChl-a parameters employed are taken from previous work.[11,47] The protonation states of all amino acids were determined with the H++ 3.0 software,[48] under neutral pH conditions. All complexes were solvated using TIP3P periodic water boxes,[49] with a minimum distance of 15 Å between the complex and the box boundaries. Charges were neutralized by adding sodium ions for the ancestor and *P. aestuarii* and chloride for *C. tepidum*. Shake constraints were applied to all bonds containing hydrogen. Minimizations were carried out for 2000 steps for *P. aestuarii* and *C. tepidum* and for 10000 steps for the ancestor. Minimizations were followed by 200 ps adaptation runs to impose a temperature of 300 K and a


DOI: 10.1021/acscentsci.7b00269
ACS Cent. Sci. 2017, 3, 1086−1095

pressure of 1 atm on the systems. All three complexes were equilibrated for 50 ns in the same environmental conditions. Long range electrostatic interactions were calculated using the particle−mesh Ewald method.[50] For each complex we then carried out 40 ps production runs with a 1 fs integration step.

**Correlating Biological Stability and Evolution to Exciton Dynamics.** *QM/MM Hamiltonian.* For the exciton dynamics, the FMO complex is simulated as a system coupled to a bath (the protein environment). The system Hamiltonian is defined as

$$H_{ex} = \sum_{i=1}^{8} \varepsilon_i |i\rangle\langle i| + \sum_{i\neq j}^{8} V_{ij} |i\rangle\langle j| \quad (1)$$

where $\varepsilon_i$ is the first excited state energy of the $i$th BChl. The $V_{ij}$ terms correspond to the excitonic coupling between the excited states for the $i$th and $j$th molecules.

Site energies were computed using TDDFT in QChem[51] with the PBE0 functional[52] and 3-21G basis set for 10000 frames (40 ps) of the MD production run. Couplings were calculated using two approaches: the point-dipole approximation (PDA) and transition charges from the electrostatic potential (TrEsp).[53] TrEsp has been shown to be the most accurate, although it is computationally more demanding than using the PDA.[53] The computed couplings are given in the Supporting Information.

The average first excited state energy of each BChl is shown in Figure 9. We notice that there is no large change in the trend of the excited state energies for the BChl's.

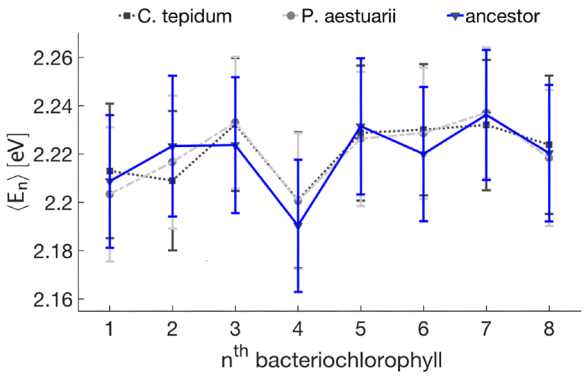



**Figure 9.** Average first excited state energies $\langle E_n \rangle$ for each BChl molecule in the FMO complexes of *C. tepidum* (black) and *P. aestuarii* (gray) and for the ancestral structure (blue). Values were obtained by averaging over the production QM/MM trajectory at 300 K.

*Exciton Coherence and 2D Spectrum.* The 2D-echo spectra in Figure 8 were calculated as was done by Hein et al.[54] In two-dimensional electronic spectroscopy the sample is probed by a sequence of three laser pulses. Adjusting the delay time $T_{delay}$ between the second and the third pulse yields a time-resolved picture of the exciton dynamics. Within the impulsive limit (assuming $\delta$-pulses), the 2D-echo spectrum is related to the third order response function $S^{(3)}(t_1, T_{delay}, t_3)$. Phase matching ensures that the components of the detected signal go along distinct directions, and thus can be experimentally separated. We considered the rephasing (RP) component of the signal

$$I_{RP}(\omega_1, T_{delay}, \omega_3) = \iint_0^\infty dt_1\, dt_3\, e^{i\omega_3 t_3 - i\omega_1 t_1}\, S^{(3)}_{RP}(t_1, T_{delay}, t_3) \quad (2)$$

for which coherent phenomena show as a unique oscillatory pattern in the cross-peak dynamics as a function of delay time. Using the notation of double sided Feynman diagrams,[55,56] the total signal is given by three separate pathways: stimulated emission (SE), ground state bleaching (GB), and excited state absorption (ESA). The latter involves the double exciton manifold. We numerically evaluated $I_{RP}(\omega_1, T_{delay}, \omega_3)$ by propagating the reduced density matrix, describing the exciton degrees of freedom, along the double sided Feynman diagrams. Hereby, each interaction with the laser pulses requires an interruption of the dynamics and a multiplication of the reduced density matrix with the dipole operator, either from the right or from the left. The propagation was done using *QMaster*,[35] a high-performance implementation of HEOM.[57,58] The rotational average over random orientations of the probed sample was included by sampling 20 laser polarization vectors aligned with the vertices of a dodecahedron.[54]

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscentsci.7b00269.

Procedures, phylogenetic tree, marginal probabilities, thermodynamic results of free energy changes, Hamiltonian and couplings, spectral densities and fitting parameters, absorption, CD, and LD spectra, population dynamics, and coherences (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Authors

*E-mail: svalleau@alumni.harvard.edu.
*E-mail: alan@aspuru.com.

### ORCID

Stéphanie Valleau: 0000-0003-0499-2054
Alán Aspuru-Guzik: 0000-0002-8277-4434


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

S.V. would like to thank Pouria Dasmeh and Adrian Serohijos for their interesting discussions on methods to compute the free energy changes for mutations and phylogenetic tree reconstruction. The authors thank Nicolas Sawaya for sharing his code for the calculation of atomic partial charges from the transition densities. The computations in this paper were completed on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University. S.V., F.H., and A.A.-G. acknowledge support from the Center for Excitonics, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science and Office of Basic Energy Sciences under Award Number DE-SC 0001088. A.A.-G. acknowledges the generous support from the Canadian Institute for Advanced Research. F.H. acknowledges support from the Herchel Smith Graduate Fellowship. C.K. acknowledges funding from the National Science Foundation under Award Number CHE-146486. This work was supported as part of the Photosynthetic Antenna Research Center (PARC), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC 0001035 to R.E.B. R.G.S. was supported by the PARC grant.



DOI: 10.1021/acscentsci.7b00269
ACS Cent. Sci. 2017, 3, 1086−1095

## REFERENCES


(1) Brocks, J. J.; Love, G. D.; Summons, R. E.; Knoll, A. H.; Logan, G. A.; Bowden, S. A. Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea. *Nature* **2005**, *437*, 866−870.

(2) Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*, 2nd ed.; Wiley-Blackwell: Oxford, U.K., 2014.

(3) Adolphs, J.; Renger, T. How proteins trigger excitation energy transfer in the FMO complex of green sulfur bacteria. *Biophys. J.* **2006**, *91*, 2778−2797.

(4) Mohseni, M.; Rebentrost, P.; Lloyd, S.; Aspuru-Guzik, A. Environment-assisted quantum walks in photosynthetic energy transfer. *J. Chem. Phys.* **2008**, *129*, 174106-1−174106-9.

(5) Rebentrost, P.; Mohseni, M.; Kassal, I.; Lloyd, S.; Aspuru-Guzik, A. Environment-assisted quantum transport. *New J. Phys.* **2009**, *11*, 033003-1−033003-12.

(6) Ishizaki, A.; Fleming, G. R. Theoretical examination of quantum coherence in a photosynthetic system at physiological temperature. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 17255−17260.

(7) Rebentrost, P.; Mohseni, M.; Aspuru-Guzik, A. Role of Quantum Coherence and Environmental Fluctuations in Chromophoric Energy Transport. *J. Phys. Chem. B* **2009**, *113*, 9942−9947.

(8) Panitchayangkoon, G.; Hayes, D.; Fransted, K. A.; Caram, J. R.; Harel, E.; Wen, J.; Blankenship, R. E.; Engel, G. S. Long-lived quantum coherence in photosynthetic complexes at physiological temperature. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 12766−12770.

(9) Rebentrost, P.; Shim, S.; Yuen-Zhou, J.; Aspuru-Guzik, A. Characterization and quantification of the role of coherence in ultrafast quantum biological experiments using quantum master equations, atomistic simulations, and quantum process tomography. *Procedia Chem.* **2011**, *3*, 332−346.

(10) Ritschel, G.; Roden, J.; Strunz, W. T.; Aspuru-Guzik, A.; Eisfeld, A. Absence of Quantum Oscillations and Dependence on Site Energies in Electronic Excitation Transfer in the Fenna−Matthews−Olson Trimer. *J. Phys. Chem. Lett.* **2011**, *2*, 2912−2917.

(11) Shim, S.; Rebentrost, P.; Valleau, S.; Aspuru-Guzik, A. Atomistic study of the long-lived quantum coherences in the Fenna−Matthews−Olson complex. *Biophys. J.* **2012**, *102*, 649−660.

(12) Valleau, S.; Eisfeld, A.; Aspuru-Guzik, A. On the alternatives for bath correlators and spectral densities from mixed quantum-classical simulations. *J. Chem. Phys.* **2012**, *137*, 224103-1−224103-13.

(13) Dostál, J.; Pšenčík, J.; Zigmantas, D. In situ mapping of the energy flow through the entire photosynthetic apparatus. *Nat. Chem.* **2016**, *8*, 705−710.

(14) Engel, G. S.; Calhoun, T. R.; Read, E. L.; Ahn, T.-K.; Mancal, T.; Cheng, Y.-C.; Blankenship, R. E.; Fleming, G. R. Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature* **2007**, *446*, 782−786.

(15) Scholes, G. D.; et al. Utilizing Coherence to Enhance Function in Chemical and Biophysical Systems. *Nature* **2017**, *543*, 647−656.

(16) Fassioli, F.; Olaya-Castro, A. Distribution of entanglement in light-harvesting complexes and their quantum efficiency. *New J. Phys.* **2010**, *12*, 085006−1−085006−15.

(17) Kreisbeck, C.; Kramer, T. Long-Lived Electronic Coherence in Dissipative Exciton Dynamics of Light-Harvesting Complexes. *J. Phys. Chem. Lett.* **2012**, *3*, 2828−2833.

(18) Olaya-Castro, A.; Lee, C.; Olsen, F.; Johnson, N. Efficiency of energy transfer in a light-harvesting system under quantum coherence. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *78*, 085115-1−085115-7.

(19) Olbrich, C.; Jansen, T. L. C.; Liebers, J.; Aghtar, M.; Strümpfer, J.; Schulten, K.; Knoester, J.; Kleinekathöfer, U. From atomistic modeling to excitation transfer and two-dimensional spectra of the FMO light-harvesting complex. *J. Phys. Chem. B* **2011**, *115*, 8609−8621.

(20) Wen, J.; Tsukatani, Y.; Cui, W.; Zhang, H.; Gross, M. L.; Bryant, D. A.; Blankenship, R. E. Structural model and spectroscopic characteristics of the FMO antenna protein from the aerobic chlorophototroph, Candidatus Chloracidobacterium thermophilum. *Biochim. Biophys. Acta, Bioenerg.* **2011**, *1807*, 157−164.

(21) Olson, J. M. The FMO Protein. *Photosynth. Res.* **2004**, *80*, 181−187.

(22) Saer, R. G.; Stadnytskyi, V.; Magdaong, N. C.; Goodson, C.; Savikhin, S.; Blankenship, R. E. Probing the excitonic landscape of the Chlorobaculum tepidum Fenna−Matthews−Olson (FMO) complex: a mutagenesis approach. *Biochim. Biophys. Acta, Bioenerg.* **2017**, *1858*, 288−296.

(23) Saer, R.; Orf, G. S.; Lu, X.; Zhang, H.; Cuneo, M. J.; Myles, D. A.; Blankenship, R. E. Perturbation of bacteriochlorophyll molecules in Fenna−Matthews−Olson protein complexes through mutagenesis of cysteine residues. *Biochim. Biophys. Acta, Bioenerg.* **2016**, *1857*, 1455−1463.

(24) Müh, F.; Madjet, M. E.-A.; Adolphs, J.; Abdurahman, A.; Rabenstein, B.; Ishikita, H.; Knapp, E.-W.; Renger, T. Alpha-helices direct excitation energy flow in the Fenna Matthews Olson protein. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16862−16867.

(25) Tronrud, D. E.; Wen, J.; Gay, L.; Blankenship, R. E. The structural basis for the difference in absorbance spectra for the FMO antenna protein from various green sulfur bacteria. *Photosynth. Res.* **2009**, *100*, 79−87.

(26) Guerois, R.; Nielsen, J. E.; Serrano, L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.* **2002**, *320*, 369−387.

(27) Zeldovich, K. B.; Chen, P.; Shakhnovich, E. I. Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16152−16157.

(28) Tokuriki, N.; Stricher, F.; Schymkowitz, J.; Serrano, L.; Tawfik, D. S. The Stability Effects of Protein Mutations Appear to be Universally Distributed. *J. Mol. Biol.* **2007**, *369*, 1318−1332.

(29) Abascal, F.; Zardoya, R.; Telford, M. J. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* **2010**, *38*, W7−13.

(30) Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307−321.

(31) Olson, J. M.; Ke, B.; Thompson, K. H. Exciton interaction among chlorophyll molecules in bacteriochlorophyll a proteins and bacteriochlorophyll a reaction center complexes from green bacteria. *Biochim. Biophys. Acta, Bioenerg.* **1976**, *430*, 524−537.

(32) Tanimura, Y.; Kubo, R. Time Evolution of a Quantum System in Contact with a Nearly Gaussian-Markoffian Noise Bath. *J. Phys. Soc. Jpn.* **1989**, *58*, 101−114.

(33) Ishizaki, A.; Tanimura, Y. Quantum dynamics of system strongly coupled to low-temperature colored noise bath: reduced hierarch equations approach. *J. Phys. Soc. Jpn.* **2005**, *74*, 3131−3134.

(34) Ishizaki, A.; Fleming, G. R. Unified treatment of quantum coherent and incoherent hopping dynamics in electronic energy transfer: Reduced hierarchy equation approach. *J. Chem. Phys.* **2009**, *130*, 234111-1−234111-10.

(35) Kreisbeck, C.; Kramer, T.; Aspuru-Guzik, A. *J. Chem. Theory Comput.* **2014**, *10*, 4045−4054.

(36) Kreisbeck, C.; Kramer, T.; Aspuru-Guzik, A. Disentangling Electronic and Vibronic Coherences in Two-Dimensional Echo Spectra. *J. Phys. Chem. B* **2013**, *117*, 9380−9385.

(37) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.

(38) Studer, R. A.; Christin, P. A.; Williams, M. A.; Orengo, C. A. Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 2223−2228.

(39) Zhang, J.; Nielsen, R.; Yang, Z. Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level. *Mol. Biol. Evol.* **2005**, *22*, 2472−2479.

(40) Yang, Z.; Wong, W. S.; Nielsen, R. Bayes Empirical Bayes Inference of Amino Acid Sites Under Positive Selection. *Mol. Biol. Evol.* **2005**, *22*, 1107−1118.




DOI: 10.1021/acscentsci.7b00269
*ACS Cent. Sci.* 2017, 3, 1086−1095

(41) Storey, J. D.; Bass, A. J.; Dabney, A.; Robinson, D. *qvalue: Q-value estimation for false discovery rate control*; R package version 2.6.0, 2015.
(42) Anisimova, M.; Yang, Z. Multiple Hypothesis Testing to Detect Lineages under Positive Selection that Affects Only a Few Sites. *Mol. Biol. Evol.* **2007**, *24*, 1219−1228.
(43) Ashkenazy, H.; Penn, O.; Doron-Faigenboim, A.; Cohen, O.; Cannarozzi, G.; Zomer, O.; Pupko, T. FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res.* **2012**, *40*, W580−W584.
(44) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(45) Case, D. A.; et al. *AMBER 14*; 2014.
(46) Cheatham, T. E.; Cieplak, R.; Kollman, R. A. A Modified Verion of the Cornell et. al. Force Field with Improved Sugar Pucker Phases and Helical Repeat. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.
(47) Chandrasekaran, S.; Aghtar, M.; Valleau, S.; Aspuru-Guzik, A.; Kleinekathöfer, U. Influence of Force Fields and Quantum Chemistry Approach on Spectral Densities of BChl a in Solution and in FMO Proteins. *J. Phys. Chem. B* **2015**, *119*, 9995−10004.
(48) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **2012**, *40*, W537−W541.
(49) Jorgensen, W. L.; Madura, J. D. Solvation and Conformation of Methanol in Water. *J. Am. Chem. Soc.* **1983**, *105*, 1407−1413.
(50) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1993**, *10*, 10089-1−10089-18.
(51) Shao, Y.; et al. Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172−3191.
(52) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **1996**, *105*, 9982−9985.
(53) Madjet, M. E.; Abdurahman, A.; Renger, T. Intermolecular Coulomb Couplings from Ab Initio Electrostatic Potentials: Application to Optical Transitions of Strongly Coupled Pigments in Photosynthetic Antennae and Reaction Centers. *J. Phys. Chem. B* **2006**, *110*, 17268−17281.
(54) Hein, B.; Kreisbeck, C.; Kramer, T.; Rodriguez, M. Modelling of Oscillations in Two-Dimensional Echo-Spectra of the Fenna−Matthews−Olson Complex. *New J. Phys.* **2012**, *14*, 023018-1−023018-21.
(55) Mukamel, S. *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: Oxford, 1999.
(56) Cho, M. *Two-Dimensional Optical Spectroscopy*; CRC Press: Boca Raton, FL, 2009.
(57) Kreisbeck, C.; Kramer, T.; Rodríguez, M.; Hein, B. *J. Chem. Theory Comput.* **2011**, *7*, 2166−2174.
(58) Kreisbeck, C.; Kramer, T. *Exciton Dynamics Lab for Light-Harvesting Complexes (GPU-HEOM)*; https://nanohub.org/resources/gpuheompop, 2013; DOI: 10.4231/D3RB6W248.



DOI: 10.1021/acscentsci.7b00269
*ACS Cent. Sci.* 2017, 3, 1086−1095