Electromagnetic optimization of light-harvesting proteins

P. Etchegoin* and R. C. Maher
The Blackett Laboratory, Imperial College London, Prince Consort Road, SW7 2BW London, United Kingdom
(Received 22 October 2002; published 13 February 2003)

The problem of *electromagnetic optimization* of the structure of light-harvesting proteins is studied within a simplified diffraction scheme. It is suggested that evolutionary pressures for protein antennae has resulted in molecular strucures that optimize the absorption in the light-sensitive pigments. Bacteriorhodopsin is shown to be an exceptional case in terms of optimization, a fact that can be connected to its extremely high sensitivity as a light detector down to a single photon level.

DOI: 10.1103/PhysRevE.67.022902 PACS number(s): 87.14.Ee, 87.15.Aa

Nature has developed very sophisticated molecular systems, typically termed *light-harvesting proteins*, and mechanisms for the collection of light [1] either for vision purposes in animals or as energy resources in plants. In the latter, the conversion of photons into electrochemical energy by electron transfer in the plant photosystems I and II (PI/II) has been extensively studied [1]. In the former, bacteriorhodopsin (bR), a much more chemically stable variety (with respect to human rhodopsin) found in bacteria [2], has played a leading role in understanding the primary photochemical processes in vision. The activation of a signal pathway through conformational changes in the protein backbone by the all-trans/cis-isomerization of retinal has been a very active field of research over the last decades [3].

A remarkable feature of all light-harvesting systems is that, despite the enormous amount of work done to date, the fine details of the photo-electro-chemical processes at a molecular level are still a source of intense work and discussion. Key breakthroughs are in some cases relatively recent; the structure of the core and reaction center of PII, for example, was not known until 1997 [4], and the real three dimensional (3D) structure of several varieties of animal rhodopsin have not been determined yet. The lack of structural information is a generalized problem for membrane proteins like rhodopsin and in many cases only outlines are known from electron microscopy and secondary structure predictions based on the aminoacid sequence.

Both the plant photosystems PI/II and photoreceptors in vision have one feature in common: the astounding quantum efficiency for light collection. Rhodopsin can be considered as a detector with sensitivity down to the quantum limit. A single photon can be collected and detected by retinal. Plants have, in a sense, a different strategy for the collection of photons. The cross section for absorption of the photosystems themselves would be very small to collect reasonable amounts of energy from the sun and, accordingly, plants have developed several families of light-harvesting protein complexes, which act effectively as antennae. Light is absorbed by both antenna clrorophyls and accessory pigments (carotenoids, phycocyanin, phycoerythrin, etc.) and rapidly transfered to the special pair of the photosynthetic reaction centres by multiple resonance (Föster) energy tranfer [5].

Light-harvesting protein (LHP) complexes vary in size and geometry and in the use of different pigments. This paper concentrates on a very simple question: Is there something peculiar about the geometry and location of absorbing centres in light-harvesting systems? Or has nature chosen (through evolutionary pressures) structures of light harvesting complexes that maximize their absorption of light? Alternatively, we could ask: Is there any ''electromagnetic'' optimization at a molecular level in the structure of these complexes? The answer to this question is not straightforward and sets the basis for this discussion.

Moving away briefly from the biological considerations surrounding these problems and considering only the physical aspects of these systems the question can be formulated within the field of optical properties of macromolecules as a question of *local fields* in the absorbing centres. The protein backbone of LHP's does not absorb light in the visible range at the frequencies where the pigments are active, but they do affect the local distribution of electromagnetic energy inside the macromolecule through its polarizability. The role of the protein backbone in these complexes is one of the least understood aspects in LHP's. Many of the pigments (like retinal, for example) have a very different absorption spectrum when they are on their own or as part of a LHP, and key mutations in PI/II close to the special pair drastically change the electron transfer capabilities of the reaction centres.

It is natural that people have asked about the role of the ''*protein scaffolding*'' in all these molecules associated with the harvesting of light, but this remains one of the outstanding unresolved issues in this field. In addition, electromagnetic local field (LF) problems are among the most difficult in condensed matter in general. In solids, for example, the real solution to the LF problem requires the inversion of an infinite matrix [6] and this is only possible due to the presence of translational invariance. LF's have been calculated under certain truncation approximations only in a handful of cases (mainly basic semiconductors). For nonlinear optical susceptibilities the problem is even more difficult and only empirical anzats are used in general [7].

In macromolecules there is the additional complication of inhomogeneity. For the static (dc) susceptibility there have been attempts to develop *effective medium* models for the dielectric properties inside proteins [8,9]. These models create an effective medium from the average of the intrinsic polarizabilities and point dipoles inside the protein, which

*Electronic address: p.etchegoin@ic.ac.uk

acts as a boundary condition for a small region where a local dielectric function is defined. Extensions of the model to take into account the intrinsic anisotropy of the bonds have not been yet developed [8] and it is in general a very difficult undertaking. To the best of our knowledge, there is no effective medium theory available for proteins in the optical range.

From this discussion it is clear that in order to evaluate any local field contribution to the absorption in a light-harvesting complex (LHC) an approximation must be made. As a result, the overall calculated result will not be accurate. However, the major aim of this communication is to identify trends that may point to a preferential organization of the LHC-structure related to its electromagnetic function as ''light detector'' and absorber. Based on these considerations we have chosen the following scheme: The optical properties in the transparency region of any condensed matter system are always dominated by the first dipole-allowed gap. The first absorption edge is responsible for the dispersion in the optical constants in the transparency region, while all the gaps above that can be considered to contribute with a dispersionless constant (the so called Penn gap in the physics of semiconductors [10]). We consider the local fields inside a protein to be dominated by the interference effect of the reemission from the lowest dipole-allowed transitions along the protein backbone. In other words, a light wave in the visible range is diffracted by each individual dipole-allowed transition in the backbone. The coherent superposition of the near fields of these dipoles [13] and the external electromagnetic field is what produces the local field at any given position. This approach assumes intrinsically a valence-optical scheme (widely developed by Volkenstein for macromolecules [14]), whereby the total polarizability of a macromolecule is the result of the coherent sum of the appropriate characteristics of its valence bonds. It is well known in proteins that the first dipole-allowed transitions in the near uv are dominated by the amide chromophore with its $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ electronic transitions at ~ 190 and ~ 220 nm, respectively [11,12]. These two transitions show small changes in energy and intensity for different secondary conformations in proteins (like β sheets or α helices) and they are sometimes used for structural analysis with circular dichroism [12].

The first approximation is then to consider the combined effect of $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ as a single transition (as seen from the transparency region) with a dipole moment almost parallel to the (O═C)-(CH) bond of the peptide amide group. The calculation proceeds as follows: (i) A structure of a LHC is extracted from its protein data bank (PDB) file [15]. (ii) The absorption pigments (or prosthetic groups in general) of the protein are extracted from the structure to leave the bare protein backbone. (iii) The coordinates of each carbonyl bond (O═C)-(CH) of the peptide amide group are identified and assigned a polarizability α in the direction of the bond, i.e., there will be an induced dipole $\vec{p} = \alpha(\vec{u} \cdot \vec{E}_{loc})\vec{u}$ in each amide bond when a local electric field $\vec{E}_{loc}$ is present. The unit vector $\vec{u}$ is along the bond. (iv) A random direction for the incoming electric field $\vec{E}_0$ of the photon is chosen initially and the induced dipole $\vec{p}_0 = \alpha(\vec{u} \cdot \vec{E}_0)$ is calculated for each bond. (v) The field is now calculated self-consistently; i.e., the electric field of all bonds except a reference one is obtained by adding all the near fields of the dipoles, and $\vec{p}$ on the reference dipole is recalculated with the total field $\vec{E}_{tot} = (\vec{E}_{ext} + \vec{E}_{loc})$. (vi) Once self-consistency is achieved, the intensity over the coordinates of the absorbing pigments (I) is evaluated and compared to the bare intensity (I_0) it would have had without the protein scaffolding (obtained for $\alpha = 0$). (vii) The result is presented as a percentage change in intensity $[\Delta I = (I - I_0)/I_0]$ showing what the screening effect of the protein backbone is on the absorption centres of the LHC. When several chromophores are present, absorption in all of them is taken into account. Qualitatively similar results are obtained if instead of considering the 3D structure of the amide bonds we take the conventional backbone of the protein scaffolding (connecting α-carbon sites). This would suggests that the result is only sensitive to the overall 3D structure of the protein. The polarizability α is assumed dispersionless, for simplicity, and the same when different proteins are compared.

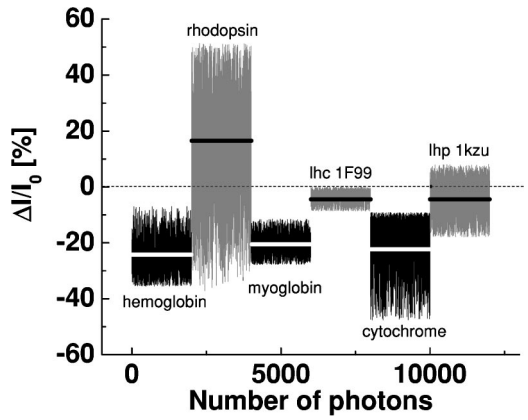

FIG. 1. Relative intensity changes $\Delta I/I_0$ at the active sites of several proteins for 2×10^3 photons with randomly chosen incident directions and polarizations. The calculations have been performed for $|E_{ext}|^2 = 1$ and $\alpha = 2.7$ in all cases. Changes are with respect of the bare active sites (hemes in the case of hemoglobin, myoglobin, and cytochrome; retinal in the case of rhodopsin; and different types of pigments in the LHC 1F99 or the light-harvesting protein 1KZU). The horizontal lines (either black or white) are the corresponding averages. A systematic increase in the contribution of the LF's to the intensity in the active centers has been obtained in these and similar proteins. This simple calculation shows that bacteriorhodopsin outperforms several light-sensitive proteins in the efficiency to couple light with its active center. See the text for further details.

Figure 1 shows six different calculations for different proteins. We have chosen three proteins where light absorption plays no role in the biological functioning and three LHC's where it does have an effect on the function. In the case of myoglobin and hemoglobin (Hb), in fact, light absorption in the hemes can be detrimental to its function through the initiation of a photolytic process, even though the proteins may not be (in a real situation) exposed to the levels of light intensity needed to achieve ligand unbinding. In any case,

myo or hemoglobin and the electron transfer protein cytochrome *C* do not have any advantage through optimization of the LF's around the prosthetic heme groups because their functioning is completely unrelated to this property. The calculations are presented for 2×10^3 photons chosen with a random polarizations. Since we are handling a long-wavelength limit of the electromagnetic field the wavevector of the light plays no role here. In principle, therefore, we can say that choosing a random polarization implies automatically a selection of a random incident direction (perpendicular to the polarization). Phase differences for the different contributions of bonds at a specific site are ignored; they are known to exist and constitute the basis of the Mie light-scattering theory for macromolecules [16] in the far-field, but are minor corrections for the internal LF's in this case. This approximation is equivalent to say that retardation effects from different parts of the protein backbone can be ignored for the local fields, a result which stems from the fact that the size of a typical protein is only 5–10 % of the wavelength of the light. Figure 1 suggests that optimization of the 3D structure around the active centres takes place; proteins which are related to the absorption or detection of light have much better coupling (on average) of their active centres with external electromagnetic radiation. Bacteriorhodopsin (PDB-file 1AP9) is particularly interesting, for it shows a substantial increase in the coupling of retinal to the external electromagnetic field through its LF; bR can be called a real molecular antenna in that respect. The fact that LF's enhance the coupling in bR much more than in the other LHC's is, perhaps, expected from their very different biological functions. Bacteriorhodopsin has to detect light at a specific point to reconstruct an image through the vision machinery (it is of no use to detect the photon in another region for the reconstruction of an image even if the total collected energy is the same) and, in addition, if the intensity is too high the eye has developed other strategies (iris) to control the situation. In LHC's for light collection in plants the situation is very different. The problem is here to gather energy, and the exact region where it has been collected is not an issue. A usual strategy of plants is the use of leaves: extended areas that maximize the cross section for light absorption while keeping mass to a minimum. Optimization at a molecular level in the receptors would obviously be advantageous, but it is probably less critical than in bR. Moreover, excessive optimization may be detrimental to the extent that plants have, in fact, molecular mechanisms to quench photosynthesis if the collection of energy is too high and exceeds the plant capacity for fixation of CO_2 [17,18] (playing the role of the iris in the eye but at a molecular level).

Further evidence that the overall 3D structures of these complexes may be related to their function as electromagnetic absorbers or detectors come from the analysis of the polarization of the LF's themselves or the incoming photons. In Fig. 2 we show an example of this for the LHC 1F99 (PDB). During the calculation in Fig. 1 we kept track of the polarization direction of the incoming photon that produces the maximum coupling to all the absorption centres simultaneously. The maximum coupling is obtained for the direction shown in the figure. This complex has a 3D structure that

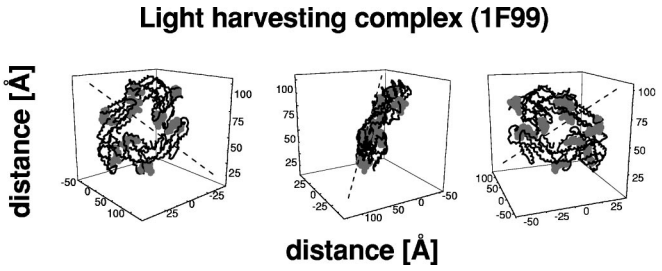

FIG. 2. Three side views of the light-harvesting complex 1F99. The dashed line is the polarization direction that maximizes the transfer of intensity from the external electromagnetic wave into the active pigments; it is almost parallel to the main plane of the complex and in agreement with the main direction for the incoming light in this quasiplanar complex. This complex may have evolved to optimize the collection of light impinging in the direction perpendicular to its main plane and can be called a molecular electromagnetic antenna in that sense.

maximizes its absorption in the active pigments for photons impinging normal to its main plane, as expected. One may argue that all polarizations parallel to the main plane of the complex should be favored, but this complex is expected to be distributed over very many different orientations in the plane in a real situation. This seems to be a general property of LF's in several LHC's, i.e., the maximum coupling of the external radiation to the active pigments is always in the direction where light is expected to impinge from, based on very simple geometrical considerations. Both the polarization characteristics of the LF's and the enhanced coupling in LHC's with respect to other proteins not related to light absorption seem to indicate an underlying electromagnetic optimization of the molecular geometries.

In no sense can the protein structure around the active pigments in LHC's *only* be associated with its electromagnetic function. On the contrary, the results in this paper show that there may be *in addition* a geometrical optimization of the structure in that respect added to the more conventional functions of the ''protein shield.'' One obvious function of the protein scaffolding is, of course, chemical in nature, i.e., protection of the active centres. We can take examples from other groups of proteins in this respect. In hemo or myoglobin, for example, it is well known that the $\alpha_{1,2}-\beta_{1,2}$ protein structures (globins) around the hemes prevent the irreversible autooxidation of the Fe ion. Fe(II) heme by itself (extracted from the protein) is incapable of binding O_2 reversibly. The protein scaffolding in LHC's around the active pigments provides, in the first instance, chemical and electronic protection from the environment to achieve fast and efficient electron transfer of the photoexcited carriers without excessive direct recombination (or nonradiative recombination through unwanted channels). However, for a given local chemical environment around the pigments, there are several possibilities for the tertiary or quaternary structure of large complexes. Our results here indicate that there may be an additional optimization of these structures in terms of their electromagnetic function. This is, in a sense, not surprising if we take into account that multiple optimizations of physicochemical

functions are known to exist in other proteins systems. Coming back to the case of Hb, the protein scaffolding has not only evolved to provide a chemical shield for autooxidation of the hemes but has also optimized a subtle ''mechanical'' function responsible for the allosteric cooperative activity of Hb during oxygen intake. A possible view in LHC's is that once the chemical problem has been solved, those complexes with a better 3D geometry for light absorption will have an evolutionary advantage. A plant could grow and reproduce faster if its molecular machinery for the collection of energy is better optimised in the chemical and electromagnetic sense, thus producing the fixation of random choices through the successive progenies that led to the highly sophisticated LHC's we see today. Our results strongly suggest that we can learn more about *the geometries that nature has chosen to use as molecular antennae* to optimize the collection or detection of light. Admittedly, the results have been obtained under drastic approximations in view of the complexity of the LF problem: (i) only the influence of one transition along the backbone is considered (amide cromophore), (ii) the polarizability is dispersionless, (iii) the absorption pigments (or prosthetic groups in general) are removed and not explicitly considered for the calculation of the LF, and (iv) phase differences are ignored. Even within these severe approximations, there are indications that there is an electromagnetic optimization in the 3D structure of light sensitive proteins. We hope our work here stimulates more sophisticated LF-calculations taking into account a more realistic picture of the electronic structure of the protein backbone and/or a more sophisticated electromagnetic formalisms.

We are indebted to L. F. Cohen (IC) for support and comments and Hilary Hartigan [National Physical Laboratory (NPL), Teddington, UK] for a critical reading of the manuscript. P.E. and R.C.M. acknowledge support from NPL under a Surface Enhanced Raman Scattering for Biomolecules project. P.E. acknowledges partial support by EPSRC under Grant No. GR/R28775.

[1] *The Photosynthetic Reaction Center*, edited by J. Deisenhofer and J.R. Norris (Academic, New York, 1993), Vols. I and II.
[2] bR is normally obtained from *Halobacterium salinarum*.
[3] An account on the early work on photosynthesis can be found in *Biological Physics*, edited by E.V. Mielczarek, E. Greenbaum, and R.S. Knox (AIP, New York, 1993).
[4] K.H. Rhee, E.P. Morris, D. Zheleva, B. Hankamer, W. Kühlbrandt, and J. Barber, Nature (London) **389**, 522 (1997).
[5] L.E. Andréasson and T. Vänngard, Annu. Rev. Plant Physiol. Plant Mol. Biol. **39**, 379 (1988).
[6] S.L. Adler, Phys. Rev. **126**, 413 (1962); N. Wiser, *ibid.* **129**, 62 (1963).
[7] Y.R. Shen, *Principles of Nonlinear Optics* (Wiley, New York, 1984).
[8] D. Voges and A. Karshikoff, J. Chem. Phys. **108**, 2219 (1998).
[9] X. Song, J. Chem. Phys. **116**, 9359 (2002).
[10] P. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin, 1996).
[11] H. Basch, M.B. Robin, and N.A. Kuebler, J. Chem. Phys. **47**, 1201 (1967).
[12] M.G. Mulkerrin, in *Spectroscopic Methods for Determining Protein Structure in Solution*, edited by H. A. Havel (VHC Publishers Inc., New York, 1995).
[13] A. López Dávalos and D. Zanette, *Fundamentals of Electromagnetism (Vacuum Electrodynamics, Media, and Relativity)* (Springer Verlag, Berlin, 1999).
[14] M.V. Volkenstein, *Molecular Biophysics* (Academic, New York, 1977).
[15] Structures are downloaded from http://www.rcsb.org/pdb/.
[16] *Laser Light Scattering in Biochemistry*, edited by S.E. Harding, D.B. Sattelle, and V.A. Bloomfield (Royal Soc. of Chemistry, Cambridge, 1992).
[17] B. Demming-Adams and W.W. Adams III, Nature (London) **403**, 371 (2000).
[18] X.P. Li, O. Björkman, C. Shih, A.R. Grossman, M. Rosenquist, S. Jansson, and K.K. Niyogi, Nature (London) **403**, 391 (2000).