Effect of chemical substitution on the surface charge of the photosynthetic Reaction Center from *Rhodobacter sphaeroides*: an *in-silico* investigation.

Jean-Pierre Dupont[1], Maria González Fernández[2], Eva Schmidt[3]

[1] Professor, Department of Sociology, Université Paris 1 Panthéon-Sorbonne, France

[2] Associate Professor, Department of Political Science, Universidad Complutense de Madrid, Spain

[3] Researcher, Institute for Social Research, Goethe University Frankfurt, Germany

The Reaction Centers (RCs) proteins are membrane proteins representing the key component so flight energy transduction in photosynthetic organisms. Upon photon absorption, these photoenzymes produce a long lasting intra protein hole electron couples whose charges are separated by 3 nanometers. The dipoles formed within the RCs can be effectively employed as transducing cores of several biological-organic hybrid devices whose design can accomplish photocurrents generation or act as phototransistor. To widen the application of the RCs to as many substrate as possible one valuable strategy is the bioconjugation of the protein with specific molecules ad-hoc selected to improve enzymatic performance and/or integration in proper scaffolding. In the present manuscript, we investigate the changes of the isoelectric point of the RC from the carotenoidless strain of the photosynthetic bacterium Rhodobacter sphaeroides R26 by inducing "in silico" mutations to predict on the role of the aminoacids involved in the bioconjugation.

INTRODUCTION

The integration of the stunning efficiency of photosynthesis with the rapidly developing field of organic electronics is opening a completely new scenario in the quest for biocompatible, environmental sustainable and efficient energy conversion. Photosynthetic microorganisms, active subcellular components, photosynthetic enzymes and pigments can be used as energy transducing units in properly designed bioelectrical hybrid devices[1]. For the purpose of the present manuscript, attention will be focused on the photochemical core, the so-called reaction center (RC), extracted from the green strain of the purple non-sulphur bacterium *Rhodobacter sphaeroides* R26. The RC is a specialized pigment-protein complex able to absorb light and transducing its energy in a intraprotein dipole formed by a positive and a negative charges separated by roughly

three nanometers. With the opportune precautions and chemical modifications[2-3], these photoenzymes can be integrate in biohybrid devices for several applicative purposes[4-6].The three-dimensional structure of the RC from *Rhodobacter sphaeroides* R26 and the role of the protein subunits and of the cofactors in the cascade of electron transfer reactions that eventually generates the long lasting intraprotein dipole are briefly here illustrated. TheRC is a membrane-spanning enzyme composed by three protein subunits named L, M, and H. Solubilization of the protein in aqueous buffer solution requires the presence of a toroid formed by a zwitterionic or non-ionic detergent surrounding the transmembrane alfa-helices. The protein scaffolding contains nine cofactors, namely four bacteriochlorophylls, two bacteriopheophytins, two ubiquinones, and one bivalent iron ion. The cofactors are arranged in two branches named A and B according with the crystallographic structure showed in Figure 1.[2,7-8].

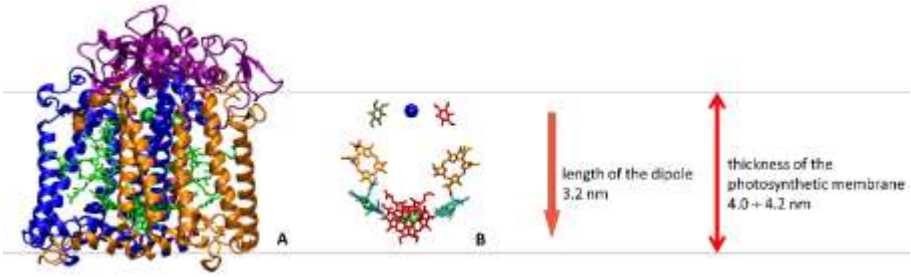

Figure 1. Representation of the three dimensional structures of the photosynthetic RC according to the crystallographic data from the protein databank, pdb code 2J8C [9]. (A) Protein scaffolding with the three subunits showed in orange (L), blue (M), and purple (H). The cofactors are shown in light green. (B) Space organization of the cofactors. The long aliphatic chains needed to anchor the cofactors to the protein subunits are omitted for clarity while the two branches are indicated. The bacteriochlorophyll dimer is represented in red; the two monomers of bacteriochlorophyll ($BChl_A$ and $BChl_B$) are represented in cyan. In all cases, the central magnesium is the van der Walls sphere in green. The bacteriopheophytins ($BPheo_A$ and $BPheo_B$) are represented in orange and the two ubiquinones are represented in dark green (Q_A) and red (Q_B) respectively. A non-heme low spin ferrous ion, indicated by a blue van der Waals sphere, sits between the two quinones. The two parallel horizontal lines represent the limits of the photosynthetic membrane. The dimer faces the periplasm while the quinone complex faces the cytoplasm. For comparison the estimated thickness of the bilayer and the length of the electric dipole generated after illumination are shown.

Upon illumination, the absorption of a photon promotes an electron of the bacteriochlorophyll dimer (D) that reaches the excited state D*. A consequent cascade of electron transfer reactions transfer the excited electrons from the dimer to the primary electron acceptor Q_A and finally the secondary acceptor Q_B, forming the charge separated-stated $D^+Q_B^-$. The intraprotein dipole has a length slightly longer than 3 nm and a lifetimes that depends upon the environmental conditions and ranges between one and three seconds[10-13].

In bioelectronics, the RC has been used as photoactive component in several devices architectures, such as silica/silicon electrodes[14-15] or in melanine-like systems[16] for photocurrent extraction, or in electrolyte-gated organic transistors for photovoltage generation[17-18].Specific chemical-physical properties of the RCs can be ameliorated *via* genetic manipulation [19-20] or chemical bioconjugation to ad-hoc designed molecules [2-3,21-23].

The chemical bioconjugation allows a selective protein functionalization on the chosen anchorage sites, most commonly the side chains of lysine (Lys), cysteine (Cys), and tyrosine (Tyr). Regardless on the target aminoacid, the bioconjugation alters some

physico-chemical properties[24] that may result of valuable importance in the interactions between the proteins and other proteins, or active surfaces, or biomimetic environments. From the bioelectronics point of view, the surface charge and isoelectric point (pI) of the RCs are both of great relevance. The pI, that represents the pH at which the net surface charge of a protein is zero, is assessed in-silico by substituting the target aminoacid with glycine (Gly), the simplest aminoacid that does not possess a lateral chain. The knowledge of pI is particularly valuable in the isolation protein procedures and in understanding the interaction between protein and electroactive surfaces.

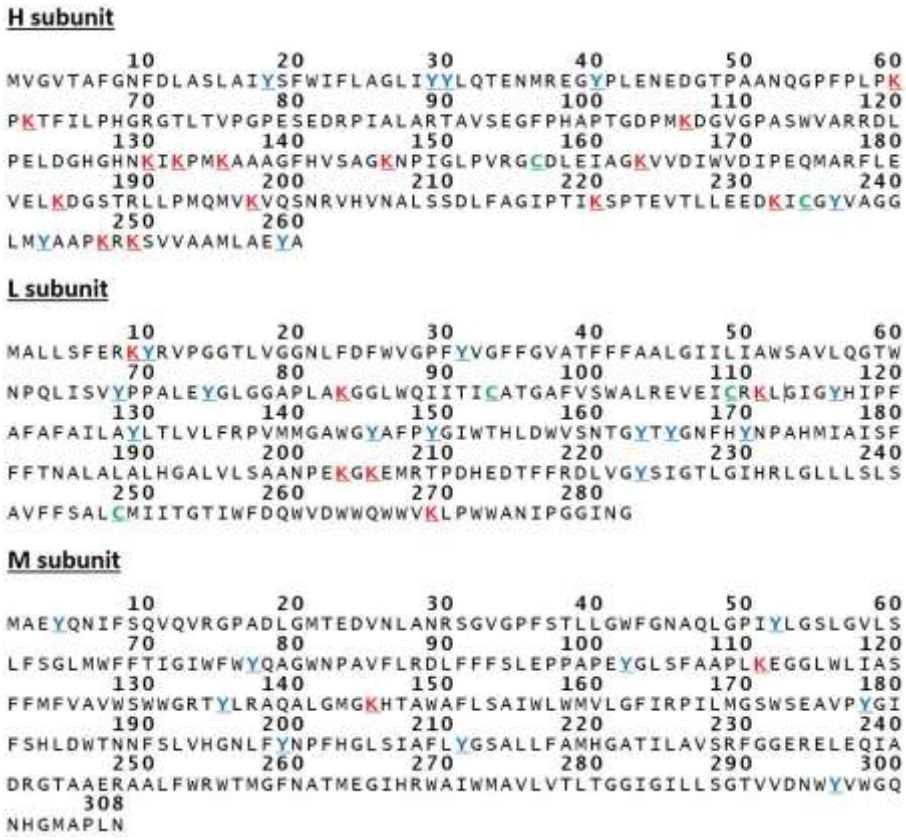

Figure 2. Primary sequence of the three subunits of the RC from *Rhodobacter sphaeroides* R26. The RC contains 22 Lysine residues (K) shown in red, 5 cysteine (C) residues shown in green, and 28 Tyrosine (Y) residues shown in bleu.

EXPERIMENTAL DETAILS

The pI values were obtain using the "Isoelectric Point Calculator" web service, recently developed by L.P. Kozlowski[25]. IPC is an independent program designed for the accurate estimation of protein pI using different sets of dissociation constant (pKa), and compare results with other calculator available online. Furthermore, the web service also uses several other algorithms as shown in Table 1. In each case, we evaluate the isoelectric point by progressively replacing the aminoacid of interest with glycine, one at the time, for each subunit of the model protein.

RESULTS AND DISCUSSION

Isoelectric point calculation

The photosynthetic RC from *Rhodobacter sphaeroides* R26 contains 848 aminoacids. The primary sequence of the three protein subunits[26-28] is shown in Figure 2 where each aminoacid is indicated with its one-letter code. Lysine (K) residues are in indicated in red, cysteine (C) residues in green, and tyrosine (Y) residues in blue. The native RC has a pI 6.1 as measured by isoelectric focusing (EIF)[29] . The pI of the whole RC was predicted by the Isoelectric Point Calculator[24] and compared to several other pI predicting software presented in the IPC webpage (http://isoelectric.org/) and indicated in Table 1 accordingly to the author[24]. The most accurate among the 11 algorithms is the IPC that, as shown in Table 1, predicts a value of 6.00, less than 2% divergent from the measured value. Furthermore, the dependence of the pI calculated value is very similar for all the tested algorithms, showing a strong internal consistency in the calculation. IPC will be hence used throughout this study.

Table 1: Calculation of the isoelectric point of non-mutated RC using different algorithms

	IPC	Toseland	Thurlkill	Nozaki_Tan	Dawson	EMBOSS	Grimsley	Rodwell	Solomon	Lehninger	ProMoST
RC	6.00	6.49	6.80	6.58	6.30	6.77	6.68	6.30	6.31	6.30	6.37

Figure 3. Spatial distribution within the three dimensional structure of the RC of (A) the lysine residues shown in blue, (B) cysteine residues shown in black and (C) tyrosine residues shown in green.

Lysine residues

The spatial distribution of lysine obtained from crystallographic data is shown in Figure 3. Residues L8, L110, and H130 are buried within the protein, not available to bioconjugation, and will be ignored. Residues H106, H132, H135, H146, H163, H184, H220, H232, H247, and H249 are located on the external hydrophilic portion of the H subunit and protrude in the cytoplasm. Residues M144, L202, L204, H60, H62 and H197 are in very close proximity to the detergent toroid surrounding L and M subunits. Residues M110, L83 and L268 lay in close proximity to the periplasm[2]. In Figure 4 is shown the predicted pI as function of the in-silico mutation of lysine to glycine. The effect of removing the protonable lysines on the isoelectric point of the three subunits in quite relevant. In particular, the effect is quite pronounced on the H subunits, where the pI diminishes by almost two units.

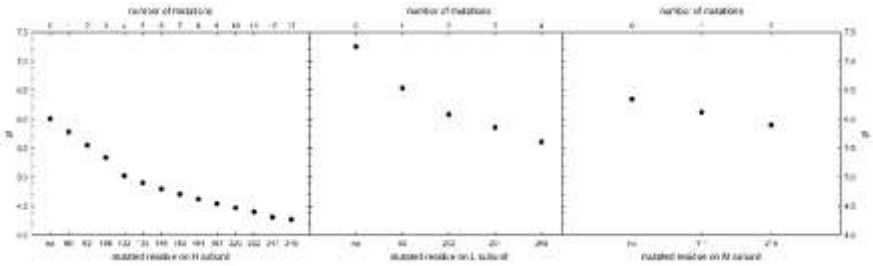

Figure 4. Predicted isoelectric point value of each subunit of the RC as function of increasing in-silico mutation of lysine to glycine residues.

Cysteine residues

The RC contains five cysteine residue located on the H and L subunits as shown in Figure 3B. Residues L93 and L247 are buried within the hydrophobic part and will be not taken into account since have no contribution to the surface charge of the protein. Residue L109 sits at the interface between the hydophobic and the hydrophylic regions, and residues H156 and H245 are located on the hydrophobic portion immerse in the cytoplasm. The effect of the changes in the pI due to mutations on the cysteins in glycines, shown in Figure 5, is much less pronounced than in the case of lysine reidues. In particular, the effect due to the cysteine in the H subunit is negligible.

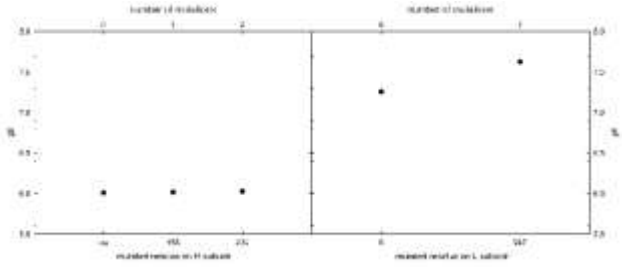

Figure 5. Predicted isoelectric point value of each subunit of the RC as function of increasing in-silico mutation of lysine to glycine residues.

Tyrosine residues

The RC contains 28 tyrosine residues distributed in the three subunits. Their spacial distribution is given in Figure 3C. As previously, buried Tyr (M52, M211, L10, L129). The dependence of the calculated pI values upon the Tyr mutation is negligible for all three proteins subunits. Data are presented in Figure 6. This finding is coherent with the phenolic group of Tyr remaining protonated up to pH 9.

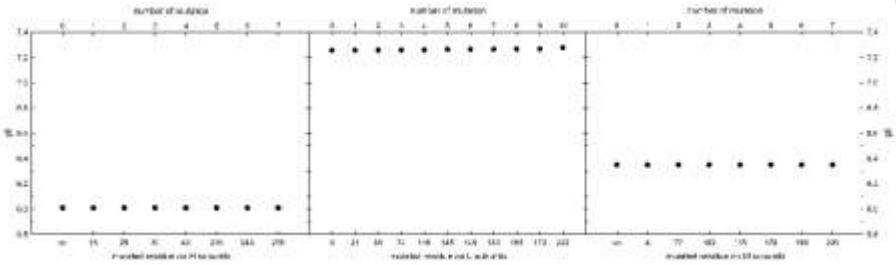

Figure 6. Predicted isoelectric point value of each subunit of the RC as function of increasing in-silico mutation of lysine to glycine residues

CONCLUSION

The correlation study of bioconjugation and protein isoelectric point can help to predict the best operating conditions for selective functionalization of amino acid when electrostatic interactions and surface charge are decisive factors for the implementation of proteins in electronic devices. Here we show how the pI behavior of the reaction center from *Rhodobacter sphaeroides* depends upon "in silico" mutations of the aminoacids typically involved in bioconjugation strategies. Isoelectric points of the RC have predicted by replacing the lysines, cysteines, and tyrosines with glycine. As expected, if the RC is deprived of surface lysine residues, a strong decrease of pI can be observed, associated with an increase of the negative charge of the overall subunits. This range corresponds to the pH decrease needed to reach the net overall surface charge of zero. On the contrary, when cysteine modification occurs, only minor variation of pI are reported. Finally, with tyrosine substitution there are no changes in isoelectric point and charge, despite the large amount of it in the RC sequence.

These results imply a significant difference in physico-chemical properties, depending on the type of bioconjugation reaction. Bioconjugation to lysines residues strongly influence the protein surface charge. These predictions can be helpful in planning molecular decoration of the photoenzyme and integration of RC in biohybrid devices.

ACKNOWLEDGEMENTS

Funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 800926 (HyPhOE, Hybrid Electronics Based on Photosynthetic Organisms).

References

[1] M. Rasmussen, S. D. Minteer, *J. Electrochem. Soc*. **2011,** *161*, H647-H655.
[2] F. Milano, R.R. Tangorra, O. Hassan Omar, R. Ragni, A. Operamolla, A. Agostiano, G.M. Farinola, and M. Trotta, *Angew. Chem. Int. Ed.* **2012**, *5*, 11019-11023.
[3] P. K. Dutta, S. Lin, A. Loskutov, S. Levenberg, D. Jun, R. Saer, J. T. Beatty, Y. Liu, H. Yan, N. W. Woodbury, *J. Am. Chem. Soc*. **2014**, *136*, 4599–4604.
[4] F. Milano, A. Punzi, R. Ragni, M. Trotta, G. M. Farinola, *Adv. Funct. Mater*. **2019**, *29*, 1805521.

[5] F. Milano, F. Ciriaco, M.Trotta, D. Chirizzi, V. De Leo, A. Agostiano, L.Valli, L.Giotta, M.R.Guascito, *Electrochim. Acta* **2019**, *293*, 105-115.
[6] Liu, J., Mantell, J., Jones, M.R., *ACS Nano* 2020, 14, 4536–4549.
[7] F. Milano, R.R. Tangorra, A. Agostiano, L. Giotta, V. De Leo, F. Ciriaco, M. Trotta, *MRS Adv.* **2018**, *3*, 1497-1507.
[8] J.P. Allen, G. Feher, T.O. Yeates, H. Komiya, D.C. Rees, *Proc. Natl. Acad. Sci*, **1988**, *85*, 8487-8491.
[9] J. Koepke, E.M. Krammer, A.R. Klingen, P. Sebban, G.M. Ullmann, G. Fritzsch, *J Mol Biol*. **2007**, 371, 396-409.
[10] R. R. Tangorra; Antonucci, A.; Milano, F.; la Gatta, S.; Farinola, G. M.; Agostiano, A.; Ragni, R.; Trotta, M. in **Handbook of Photosynthesis**, Third Edition CRC Press, **2016**, 201-220.
[11] R. Ragni, G. Leone, S. La Gatta, G. Rizzo, M. Lo Presti, V. De Leo, G.M. Farinola, *MRS Adv.* **2019**, *4*, 1143-1148.
[12] G. Feher, J.P. Allen, M.Y. Okamura, D.C. Rees, *Nature* **1989**, *339*, 111 – 116.
[13] M. Trotta, F. Milano, L. Nagy, A. Agostiano, *Mater. Science and Eng. C*, **2002**, *22*, 263-267.
[14] M. Lo Presti, D.Vona, G. Leone, G. Rizzo, R. Ragni, S. R. Cicco, F. Milano, F. Palumbo, M. Trotta, G. M. Farinola, *MRS Adv.* **2019**, *4*, 1741-1748.
[15] K. Hajdu, C. Gergely, M. Martin, T. Cloitre, L. Zimányi, K. Tenger, P. Khoroshyy, G. Palestino, V. Agarwal, K. Hernádi, Z. Németh, L. Nagy, *Langmuir*, **2012**, *28*, 11866-11873.
[16] M. Lo Presti, M.M. Giangregorio, R. Ragni, L. Giotta, M.R. Guascito, R. Comparelli, E. Fanizza, R.R. Tangorra, A. Agostiano, M. Losurdo, G.M. Farinola, F. Milano, M. Trotta, *Adv. Electron. Mater*. **2020**, 2000140.
[17] M. Di Lauro, S. La Gatta, C.A. Bortolotti, V. Beni, V. Parkula, S, Drakopoulou, M. Giordani, M. Berto, F. Milano, T. Cramer, M. Murgia, A. Agostiano, G.M. Farinola, M. Trotta, F. Biscarini, *Adv. Electron. Mater.* **2020**, ***6***, 1900888.
[18] M. Di Lauro, G. Buscemi, M. Bianchi, A. De Salvo, M. Berto, S. Carli, M. Trotta, *MRS Adv*. **2020**, *5*, 985-990.
[19] Y. Takahashi, K. Utsumi, Y. Yamamoto, A. Hatano, K. Satoh, *Plant and Cell Physiology*, **1996**, *37*, 161–168.
[20] E. Espiritu, K.D. Chamberlain, J.C. Williams, J.P. Allen. *Photosynthesis Research,* **2020**, *143,* 129-141.
[21] S. La Gatta, O. Hassan Omar, A. Agostiano, F. Milano, R.R. Tangorra, A. Operamolla, R. Ragni, *MRS Advances*, **2016**, *1*, 495-500.
[22] O. Heifler, C. Carmeli, I. Carmeli, *Langmuir*, **2020**, *36*, 4556-4562.
[23] E. Altamura, F. Milano, R.R. Tangorra, M. Trotta, O. Hassan Omar, P. Stano, F. Mavelli, *Proc. Natl. Acad. Sci. USA*, **2017**, *114*, 3837-3842.
[24] E. Audain,Y. Ramos, H. Hermjakob, D.R. Flower, Y. Perez-Riverol *Bioinformatics*, **2016**, *32,* 821–827.
[25] Kozlowski, L.P., *Biol. Direct* **2016**, *11*, 55.
[26] J.C. Williams, L.A. Steiner, R.C. Ogden, M.I. Simon, G. Feher, *Proc. Natl. Acad. Sci. USA*, **1983**, *80*, 6505-6509.
[27] J.C. Williams, L.A. Steiner, R.C. Ogden, M.I. Simon, G. Feher, *Proc. Natl. Acad. Sci. USA*, **1984**, *80*, 7303-7307.
[28] J.C. Williams, L.A. Steiner, G. Feher, *Proteins*, **1986**, *1*, 312-325.
[29] R.C. Prince, R.J. Cogdell, A.R. Crofts, *BBA - Bioenergetics*, **1974**, *347*, 1-13.