Modeling the Bacterial Photosynthetic Reaction Center. 4. The Structural, Electrochemical, and Hydrogen-Bonding Properties of 22 Mutants of *Rhodobacter sphaeroides*

Jason M. Hughes,[†] Michael C. Hutter,[†,§] Jeffrey R. Reimers,*[,†] and Noel S. Hush[‡]

Contribution from the School of Chemistry and Department of Biochemistry, University of Sydney, Sydney, NSW 2006, Australia

Received October 2, 2000. Revised Manuscript Received May 10, 2001

Abstract: Site-directed mutagenesis has been employed by a number of groups to produce mutants of bacterial photosynthetic reaction centers, with the aim of tuning their operation by modifying hydrogen-bond patterns in the close vicinity of the "special pair" of bacteriochlorophylls $P \equiv P_L P_M$. Direct X-ray structural measurements of the consequences of mutation are rare. Attention has mostly focused on effects on properties such as carbonyl stretching frequencies and midpoint potentials to infer indirectly the induced structural modifications. In this work, the structures of 22 mutants of *Rhodobacter sphaeroides* have been calculated using a mixed quantum-mechanical molecular-mechanical method by modifying the known structure of the wild type. We determine (i) the orientation of the 2a-acetyl groups in the wild type, FY(M197), and FH(M197) series mutants of the neutral and oxidized reaction center, (ii) the structure of the FY(M197) mutant and possible water penetration near the special pair, (iii) that significant protein chain distortions are required to assemble some M160 series mutants (LS(M160), LN(M160), LQ(M160), and LH(M160) are considered), (iv) that there is competition for hydrogen-bonding between the 9-keto and 10a-ester groups for the introduced histidine in LH(L131) mutants, (v) that the observed midpoint potential of P for HL(M202) heterodimer mutants, including one involving also LH(M160), can be correlated with the change of electrostatic potential experienced at P_L, (vi) that hydrogen-bond cleavage may sometimes be induced by oxidation of the special pair, (vii) that the OH group of tyrosine M210 points away from P_M, and (viii) that competitive hydrogen-bonding effects determine the change in properties of NL(L166) and NH(L166) mutants. A new technique is introduced for the determination of ionization energies at the Koopmans level from QM/MM calculations, and protein-induced Stark effects on vibrational frequencies are considered.

1. Introduction

Much interest in the photosynthetic reaction centers of purple bacteria has been generated in recent times owing to the availability of the three-dimensional X-ray crystal structure of reaction centers of *Rhodobacter* (*Rb.*) *sphaeroides* (1PCR)[1] and other purple bacteria including *Rhodopseudomonas* (*Rh.*) *viridis* (1PRC);[2] these structures have resulted from extensive development, see, for example, refs 1−3. Recently, the photophysics of photosynthesis and the structure and spectroscopy of the reaction centers have been reviewed.[3] In brief, the photooxidation process occurs in a membrane-bound protein which contains a number of cofactors, including four bacteriochlorophylls (BChl). Of these, the central two (P_L and P_M) overlap at ring I and are referred to as the special pair P, while the other two ($BChl_M$ and $BChl_L$) are referred to as "accessory" bacteriochlorophylls. There are also two bacteriopheophytins (BPh_M and BPh_L), two quinones (Q_A and Q_B), and a non-heme iron atom which together with the special pair, are organized in pseudo-C_2 symmetry forming two branches (L and M).

Light-initiated charge separation occurs between the special pair P (P_L and P_M) and the neighboring pigments, leading to a radical cation (P^+). Despite the quasi-symmetrical arrangement of the cofactors, the electrons are transported unidirectionally along the L-branch of the reaction center,[4−6] which suggests that the symmetry-breaking specific interactions with the protein are fundamentally important. The charge-transfer process is found to occur very rapidly from P toward the quinones with a quantum yield of near unity.[7] The driving force of the primary process of charge separation is related to the redox potential of P, as is the rereduction of the special pair radical cation by cytochrome c_2. This redox midpoint potential $E_m(P/P^+)$ is sensitive to the details of the protein environment and also to structural changes caused by mutagenesis; the primary aim of many studies of bacterial photosynthesis is to understand (and ultimately control) the variations of this property at the *molecular level*.

Much of the available molecular level information has been developed through examination of X-ray structures (see, e.g., refs 1, 2, and 8−12), through spin-resonance measurements of

[†] School of Chemistry.
[‡] Department of Biochemistry.
[§] Present address: Max-Planck Institute of Biophysics, D-60596 Frankfurt, Germany.

(1) Ermler, U.; Fritzsch, G.; Buchanan, S. K.; Michel, H. *Structure* **1994**, *2*, 925.
(2) Deisenhofer, J.; Michel, H. *Science* **1989**, *245*, 1463.
(3) Hoff, A. J.; Deisenhofer, J. *Phys. Rep.* **1997**, *287*, 1.
(4) Williams, L. C.; Taguchi, A. K. W. In *Anoxygenic Photosynthetic Bacteria*; Blankenship, R. E., Madigan, M. T., Bauer, C. E., Eds.; Kluwer: Dordrecht, 1995; p 1029.
(5) Kirmaier, C.; Holten, D.; Parson, W. W. *Biochim. Biophys. Acta* **1985**, *810*, 49.
(6) Lockhart, D. J.; Kirmaier, C.; Holten, D.; Boxer, S. G. *J. Phys. Chem.* **1994**, *94*, 6987.
(7) Boxer, S. G. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 267.

10.1021/ja0035710 CCC: $20.00 © 2001 American Chemical Society
Published on Web 08/09/2001

the spin distribution in the oxidized reaction center,[13−18] and through Fourier transform infrared (FTIR)[11,19−25] and Fourier transform resonance-Raman (FTRR)[26−30] spectroscopic studies. In particular, the X-ray, FTRR, and FTIR spectroscopies are powerful selective tools that enable one to probe the protein−cofactor interactions (such as hydrogen-bonding), leading to correlations between the molecular structure and the photosynthetic function.

Of equal importance to the above-mentioned techniques has been the development of a large number of mutant reaction centers (see, e.g., refs 4, 15, 17, and 31−36). These have been specifically designed either to (i) modify the hydrogen-bonding pattern between the special pair and the protein, (ii) change the electron-transport characteristics of the special pair directly (such as with the HL(M202) mutants where the P_M bacteriochlorophyll is replaced with a bacteriopheophytin), or (iii) alter the electron-transport capabilities of the protein between the special pair and the bacteriopheophytins (such as with the Y*x*(M210) mutants). To name them, we use the notation *yx*(*cn*) of Allen et al.,[27] where *cn* is the name of the mutated residue, chain *c* number *n*, *y* is the standard one-letter abbreviation for the amino acid present in the WT, and *x* is that as introduced in the mutant.

While much is now known of the structure around the special pair and the dependence of the photochemical process on this structure, many significant questions remain unanswered. Experimental X-ray structures are very difficult to obtain, and the FTIR and FTRR spectroscopic techniques often fail to resolve or detect important vibrational modes. As an example, only limited information is available concerning the orientations of the 2a-acetyl groups in the neutral and oxidized reaction centers, and the key P_M 9-keto frequency for P^+ is not observed[17,21,23,27,37,38] at all. The orientations of the acetyl groups in the WT, HF(L168), and F*x*(M197) mutants, as well as the nature of the hydrogen-bonding to P^+ for these and also the LH(L131) and L*x*(M160) mutants, remain unclear. In addition, for different mutants containing the LH(L131) modification, different specimens of the same mutant and different members of the family give varying spectroscopic results[17,21,23,27,32,38,39] for the effect of the mutation, and no simple interpretation of the observations has been proposed. A variation on this occurs, for example, with the FY(M197) and FH(M197) mutants for which differing techniques (vibrational spectroscopy[11,34,40] and midpoint potential[11,38−41]) suggest disparate results for the relative strengths of the hydrogen-bonding. In other cases, such as the NL(L166) and NH(L166) mutants, it is clear that the intended primary chemical effect is not systematically achieved, and the causes of this remain unknown.

A variety of approaches have historically been used to interpret the observed midpoint potentials of the mutant reaction centers,[22,29,30,32,39,41−44] and often these involve conflicting basic assumptions. Simpler analytical theories[41,43,44] are of great value as they lead to estimates of basic chemical properties such as intermolecular coupling strengths, but these models are necessarily simplistic and restricted in application. Sophisticated means of evaluating electrostatic energies for single configurations are available (see, e.g., refs 45 and 46); these are appropriate, provided that there is no structural rearrangement accompanying the oxidation of the special pair. More general are molecular dynamics approaches[45−51] which, in principle,

(8) Allen, J. P.; Feher, G.; Yeates, T. O.; Komiya, H.; Rees, D. C. *NATO ASI Ser., Ser. A* **1987**, *84*, 5730.
(9) Allen, J. P.; Feher, G.; Yeates, T. O.; Komiya, H.; Rees, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 5730.
(10) Yeates, T. O.; Komiya, H.; Chirino, A.; Rees, D. C.; Allen, J. P.; Feher, G. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7993.
(11) Kuglstatter, A.; Hellwig, P.; Fritzsch, G.; Wachtveitl, J.; Oesterhelt, D.; Mäntele, W.; Michel, H. *FEBS Lett.* **1999**, *463*, 169.
(12) McAuley, K. E.; Kyfe, P. K.; Cogdell, R. J.; Isaacs, N. W.; Jones, M. R. *FEBS Lett.* **2000**, *467*, 285.
(13) Plato, M.; Moebius, K.; Lubitz, W.; Allen, J. P.; Feher, G. *Jerusalem Symp. Quantum Chem. Biochem.* **1990**, *22*, 423.
(14) Lendzian, F.; Huber, M.; Isaacson, R. A.; Endeward, B.; Plato, M.; Bönigk, B.; Möbius, K.; Lubitz, W.; Feher, G. *Biochim. Biophys. Acta* **1993**, *1183*, 139.
(15) Rautter, J.; Gessner, C.; Lendzian, F.; Lubitz, W.; Williams, J. C.; Murchison, H. A.; Wang, S.; Woodbury, N. W.; Allen, J. P. *NATO ASI Ser., Ser. A* **1992**, *237*, 99.
(16) Rautter, J.; Lendzian, F.; Lubitz, W.; Wang, S.; Allen, J. P. *Biochemistry* **1994**, *33*, 12077.
(17) Rautter, J.; Lendzian, F.; Schulz, C.; Fetsch, A.; Kuhn, M.; Lin, X.; Williams, J. C.; Allen, J. P.; Lubitz, W. *Biochemistry* **1995**, *34*, 8134.
(18) Rautter, J.; Lendzian, F.; Lin, X.; Williams, J. C.; Allen, J. P.; Lubitz, W. *React. Cent. Photosynth. Bacter.: Struct. Dyn., Proc. Workshop* **1996**, 37.
(19) Mitchell, D. M.; Wang, Y.; Alben, J. O.; Shapleigh, J. P. *Biochim. Biophys. Acta* **1998**, *1409*, 99.
(20) Nabedryk, E.; Breton, J.; Wachtveitl, J.; Gray, K. A.; Oesterhelt, D. *NATO ASI Ser., Ser. A* **1992**, *237*, 147.
(21) Nabedryk, E.; Allen, J. P.; Taguchi, A. K. W.; Williams, J. C.; Woodbury, N. W.; Breton, J. *Biochemistry* **1993**, *32*, 13879.
(22) Nabedryk, E.; Breton, J.; Allen, J. P.; Williams, J. C. *Spectrosc. Biol. Mol.: Mod. Trends, [Eur. Conf.], 7th* **1997**, 107.
(23) Nabedryk, E.; Breton, J.; Williams, J. C.; Allen, J. P.; Kuhn, M.; Lubitz, W. *Spectrochim. Acta A* **1998**, *54*, 1219.
(24) Breton, J.; Nabedryk, E.; Clérici, A. *Vib.Spectrosc.* **1999**, *19*, 71.
(25) Leonhard, M.; Mäntele, W. *Biochemistry* **1993**, *32*, 4532.
(26) Albouy, D.; Kuhn, M.; Williams, J. C.; Allen, J. P.; Lubitz, W.; Mattioli, T. A. *Biochim. Biophys. Acta* **1997**, *1321*, 137.
(27) Allen, J. P.; Artz, K.; Lin, X.; Williams, J. C.; Ivancich, A.; Albouy, D.; Mattioli, T. A.; Fetsch, A.; Kuhn, M.; Lubitz, W. *Biochemistry* **1996**, *35*, 6612.
(28) Diers, J. R.; Bocian, D. F. *J. Phys. Chem.* **1994**, *98*, 12884.
(29) Mattioli, T. A.; Lin, X.; Allen, J. P.; Williams, J. C. *Spectrosc. Biol. Mol., Eur. Conf., 6th* **1995**, 199.
(30) Mattioli, T. A.; Allen, X. Lin, X.; Allen, J. P.; Williams, J. C. *Biochemistry* **1995**, *34*, 6142.
(31) Wang, S.; Williams, J. C.; Allen, J. P. *Res. Photosynth., Proc. Int. Congr. Photosynth., 9th* **1992**, 381.
(32) Williams, J. C.; Alden, R. G.; Murchison, H. A.; Peloquin, J. M.; Woodbury, N. W.; Allen, J. P. *Biochemistry* **1992**, *31*, 11029.
(33) Williams, J. C.; Woodbury, N. W.; Taguchi, A. K. W.; Peloquin, J. M.; Murchison, H. A.; Alden, R. G.; Allen, J. P. *NATO ASI Ser., Ser. A* **1992**, *237*, 25.
(34) Wachtveitl, J.; Rarchaus, J. W.; Das, R.; Lutz, M.; Robert, B.; Mattioli, T. A. *Biochemistry* **1993**, *32*, 12875.
(35) Rischel, C.; Spiedel, D.; Ridge, J. P.; Jones, M. R.; Breton, J.; Lambry, J.-C.; Martin, J.-L.; Vos, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12306.
(36) Reimers, J. R.; Craw, J. S.; Bacskay, G. B.; Hush, N. S. *BioSystems* **1995**, *35*, 107.
(37) Scherer, P. O. J. *J. Lumin.* **1992**, *53*, 133.
(38) Lin, X.; Murchisson, H. A.; Nagarajan, V.; Parson, W. W.; Allen, J. P.; Williams, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10265.
(39) Williams, J. C.; Alden, R. G.; Coryell, V. H.; Lin, X.; Murchison, H. A.; Peloquin, J. M.; Woodbury, N. W.; Allen, J. P. *Res. Photosynth., Proc. Int. Congr. Photosynth., 9th* **1992**, 377.
(40) Mattioli, T. A.; Williams, J. C.; Allen, J. P.; Robert, B. *Biochemistry* **1994**, *33*, 1636.
(41) Ivancich, A.; Artz, K.; Williams, J. C.; Allen, J. P.; Mattioli, T. A. *Biochemistry* **1998**, *37*, 11812.
(42) Zadorozhnyi, B. A.; Ishchenko, I. K. *Opt. Spectrosc.* **1965**, *19*, 306.
(43) Artz, K.; Williams, J. C.; Allen, J. P.; Lendzian, F.; Rautter, J.; Lubitz, W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13582.
(44) Reimers, J. R.; Hughes, J. M.; Hush, N. S. *Biochemistry* **2000**, 16185.
(45) Muegge, I.; Apostolakis, J.; Ermler, U.; Fritzsch, G.; Lubitz, W.; Knapp, E. W. *Biochemistry* **1996**, *35*, 8359.
(46) Alden, R. G.; Parson, W. W.; Chu, Z. T.; Warshel, A. *J. Phys. Chem.* **1996**, *100*, 16761.
(47) Apostolakis, J.; Muegge, I.; Ermler, U.; Fritzsch, G.; Knapp, E. W. *J. Am. Chem. Soc.* **1996**, *118*, 3743.
(48) Muegge, I.; Ermler, U.; Fritzsch, G.; Knapp, E. W. *J. Phys. Chem.* **1995**, *99*, 17917.
(49) Alden, R. G.; Parson, W. W.; Chu, Z. T.; Warshel, A. *J. Am. Chem. Soc.* **1995**, *117*, 12284.
(50) Treutlein, H.; Schulten, K.; Brünger, A. T.; Karplus, M.; Deisenhofer, J.; Michel, H. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 75.
(51) Marchi, M.; Gehlen, J. N.; Chandler, D.; Newton, M. *J. Am. Chem. Soc.* **1993**, *115*, 4178.

allow for structural relaxation, entropic effects, and thermal averaging. These have been successfully used[47] to study the rotation of the P_M 2a-acetyl group on oxidation and have also been able to interpret[46,47] anomalies in the midpoint potential of P for Y*x*(M210) and other mutants. However, the ultimate reliability of the molecular dynamics approach depends on the quality of the empirical force field used, and the downside of the full inclusion of thermal averaging effects is that it can be difficult to describe the results in simple terms using characteristic structures.

We seek to resolve the aforementioned issues using a computational strategy which focuses on the generation of one or at most two characteristic structures for a mutant. While this neglects possibly significant thermal averaging effects, it allows more direct comparisons between structure and function, possibly through subsequent quantum-chemical analysis of system substructures. In total some 22 mutants consisting of either single, double, triple, and heterodimer mutants of *Rb. sphaeroides* are considered, namely: FH(M197),[40] FY(M197),[11,34] FH(M197)+LH(M160)+LH(L131),[38] HF(L168),[40] HF(L168)+FH(M197),[38] HF(L168)+LH(L131),[38] HL(M202),[17] HL(M202)+FH(M197),[17] HL(M202)+HF(L168),[17] HL(M202)+LH(L131),[17] HL(M202)+LH(M160),[17] LH(L131),[40] LH(M160),[40] LH(M160)+FH(M197),[40] LH(M160)+LH(L131),[32,39] LN(M160),[43] LQ(M160),[43] LS(M160),[43] NL(L166),[52] NH(L166),[52] YF(M210),[53,54] and YW(M210).[54,55] A total of 30 representative structures are presented for these mutants, including those for the WT and also for the wild type of *Rh. viridis*. The relative locations of the sites of the considered mutations with respect to the special pair are sketched in Figure 1.

In overview, our computational procedure starts with the X-ray structure[1] of the wild-type reaction center protein, introduces a model[56] for the aspects of the actual structure which are not properly represented in the X-ray structure, refines the microscopic structure,[56] introduces the required mutation(s), searches for likely low-energy configurations of the mutant, and finally refines its microscopic structure. Such a procedure is capable of producing reasonable results only if the overall protein structure of the mutant is very similar to that of the wild type; in fact, the results of all previous mutation studies indicate that this is indeed the case. Clearly, however, the larger the perturbation introduced by the mutation, the less reliable will be the results of the calculations.

Refinement of the microscopic structure found in the X-ray crystallographic studies is an essential feature if the nuclear coordinates are subsequently to be used in quantum-chemical calculations of molecular properties. It is necessary as, from X-ray structures of proteins (unlike those for simple organic or inorganic compounds), the desired nuclear coordinates are insufficiently resolved to indicate the (crucial) subtle changes in bond lengths which accompany changes in chemical interactions (this effect is visually evident from, e.g., the unusual pyrrole ring geometries found in the crystal structures shown in the Figure 1 of Supporting Information). Further, use of X-ray coordinates has led to incorrect deductions of the primary structural forces holding together the special pair.[56] However,

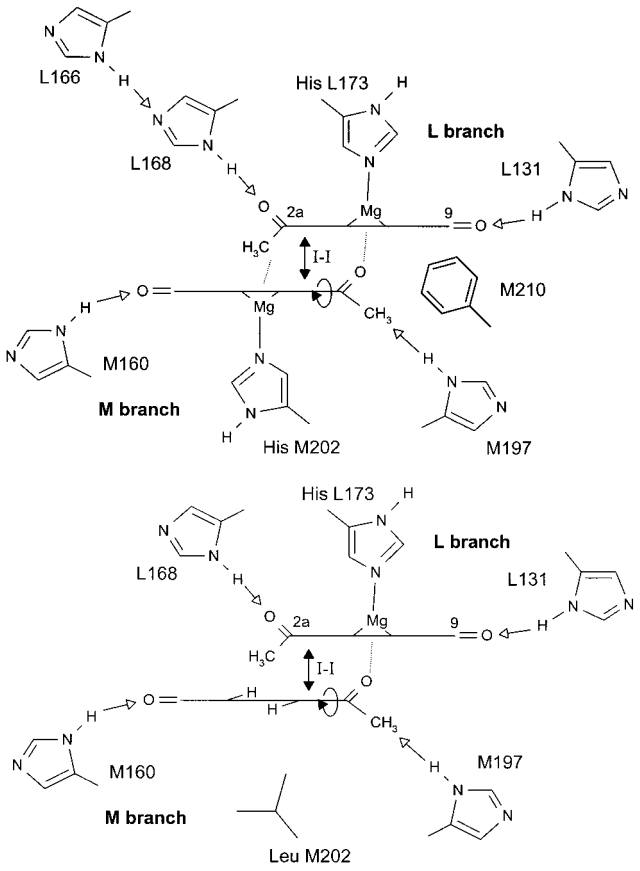

Figure 1. Sketch of the photosynthetic reaction center special pair, indicating the sites of the various individual mutations considered in this study. The upper frame shows mutations of the wild-type special pair, while the lower frame shows mutations of the heterodimer mutant special pair (HL(M202)) in which the bacteriochlorophyll at the P_M site is changed to a bacteriopheophytin.

very few calculations have been reported that optimize or refine the molecular structure of the cofactors in situ in the protein, and only a few studies (see, e.g., refs 46 and 47) have performed a systematic study for a series of different mutants. Our study[56] of the wild types of *Rb. sphaeroides* and *Rh. viridis* is the only one which has employed quantum-chemical methods to optimize the structures of the cofactors in situ. There have been a variety of molecular dynamic calculations using standard force fields which include both molecular relaxation and entropy of the entire protein.[45,47,49−51] Also a number of ab initio,[57−60] density-functional,[61−63] and semiempirical[64−68] calculations have been

(52) Ivancich, A.; Mattioli, T. A.; Artz, K.; Wang, S.; Allen, J. P.; Williams, J. C. *Biochemistry* **1997**, *36*, 3027.
(53) Gray, K. A.; Farchaus, J. W.; Wachtveitl, J.; Breton, J. *Oesterhelt. EMBO J.* **1990**, *9*, 2061.
(54) Nagarajan, V.; Parson, W. W.; Davis, D.; Schenck, C. C. *Biochemistry* **1993**, *32*, 12324.
(55) van Brederode, M. E.; van Mourik, F.; van Stokkum, I. H. M.; Jones, M. R.; van Grondelle, R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2054.
(56) Hutter, M. C.; Hughes, J. M.; Reimers, J. R.; Hush, N. S. *J. Phys. Chem. B* **1999**, *103*, 4906.
(57) Zhang, L. Y.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13603.
(58) Nakatsuji, H.; Hasegawa, J.; Ohkawa, K. *Chem. Phys. Lett.* **1998**, *296*, 499.
(59) Hasegawa, J.; Nakatsuji, H. *J. Phys. Chem. B* **1998**, *102*, 10420.
(60) Hasegawa, J.; Ohkawa, K.; Nakatsuji, H. *J. Phys. Chem. B* **1998**, *102*, 10410.
(61) Siegbahn, P. E. M.; Blomberg, M. R. A.; Pavlov, M. *Chem. Phys. Lett.* **1998**, *292*, 421.
(62) Siegbahn, P. E. M.; Blomberg, M. R. A. *Annu. Rev. Phys. Chem.* **1999**, *50*, 221.
(63) Blomberg, M. R. A.; Siegbahn, P. E. M.; Babcock, G. T. *J. Am. Chem. Soc.* **1998**, *120*, 8812.
(64) Thompson, M. A.; Zerner, M. C. *J. Am. Chem. Soc.* **1991**, *113*, 8210.
(65) Thompson, M. A.; Zerner, M. C.; Fajer, J. *J. Phys. Chem.* **1991**, *95*, 5693.
(66) Scherer, P. O. J.; Scharnagl, C.; Fischer, S. F. *Chem. Phys.* **1995**, *197*, 333.
(67) Reimers, J. R.; Hush, N. S. *J. Am. Chem. Soc.* **1995**, *117*, 1302.
(68) Ivashin, N.; Källebring, B.; Larsson, S.; Hansson, Ö. *J. Phys. Chem. B* **1998**, *102*, 5017.

Table 1: Average, Maximum Difference, and the RMS Deviation Found for the AM1 Cofactor Strain Energies (the Energy Required, in the Gas Phase, To Distort a Cofactor from Its Equilibrium Geometry to its in Situ Geometry), in kcal mol^{-1}

			difference	
cofactor	total no.	av	max	RMS
P_L (BChl)	29	6.8	5.6	1.4
P_M	23	8.3	2.9	1.0
$BChl_L$	29	10.0	2.4	0.4
$BChl_M$	29	13.8	1.9	0.4
all BChl	110	9.8	9.9	2.8
P_L (BPheo)	6	4.0	1.9	0.7
$BPheo_L$	29	3.4	0.5	0.1
$BPheo_M$	29	4.5	0.4	0.1
all BPh	64	4.0	2.3	0.7
Q_A	29	5.1	0.6	0.1
Q_B	29	2.3	0.7	0.1
all Q	58	3.7	3.7	1.4
SPO	29	0.8	0.9	0.3

performed at either X-ray or gas-phase geometries. Some of these have, however, included treatments of the protein−cofactor interaction either using small clusters, atomic point charges, or polarizable dielectrics such as self-consistent reaction fields. We use a mixed quantum-mechanical (QM)/molecular-mechanical (MM) model to optimize the structures of the mutant reaction centers. The MM part is used to describe the protein while the QM part is used to describe the cofactors; the influence of the MM part on the QM part is exerted[56] self-consistently through a combination of electrostatic point-charge and Lennard-Jones forces.

2. Computational Methods

All of the reaction center mutants' systems considered are based on the wild-type model of *Rhodobacter sphaeroides* as published earlier[56] which is derived from the X-ray structure[1] given in the Protein Data Bank.[69] The procedure for establishing this initial protein environment is given in detail elsewhere;[56] general procedures used in constructing structures for the mutant reaction centers are described in Supporting Information, as are some specific procedures used for situations in which the introduced residue is significantly different in volume from the original residue. All molecular mechanics calculations were performed using the HYPERCHEM[70] program while the QM/MM calculations were carried out using a modified version of the VAMP[71] program package. In addition, a new technique is introduced for the evaluation of ionization energies in the QM/MM approach, and this is described in Appendix 1. Finally, methods used to model Stokes shifts of carbonyl vibration frequencies are described in Appendix 2.

3. Results and Discussion

The QM/MM optimized coordinates for the various mutants are provided in full in Supporting Information; the variations found in the AM1 calculated energies of the cofactors are summarized in Table 1, while key intermolecular bond lengths are presented in Tables 2−6 for the L*x*(M160), F*x*(M197), HL(M202), LH(L131), and HF(L168) and L*x*(L166) and Y*x*(M210) mutant families, respectively. We analyze these structures in three ways: through quantitative improvements in geometrical properties, by the provision of semiquantitative analyses of observed spectroscopic and electrochemical data, and most importantly, by deciphering key qualitative features of the chemistry of the mutant reaction centers.

Table 1 shows the AM1 calculated strain energies for the cofactors; the strain energy of a cofactor is the energy required to distort it, in the gas phase, from its gas-phase equilibrium structure to that in the protein environment. These changes range up to 15 kcal mol^{-1}, a value which is quite small given the large size of the molecules. Per degree of freedom, the maximum strain energy is only 5% of the available thermal energy at 300 K; for any particular cofactor, the calculated strain energies vary by much less than this, however, at most 6 kcal mol^{-1} with root-mean-square deviations less than 2.8 kcal mol^{-1}, as shown in Table 1. In the protein, this energy is provided by the intermolecular solvation interactions. Typical electronic structure calculations on the reaction center protein and other biological species are performed using X-ray coordinates for the heavy atoms, with hydrogens added as appropriate. However, inter-heavy-atom distances are often very poorly represented in such X-ray structures, and hence calculated strain energies can be[56,72] very large, for example, 300 kcal mol^{-1}, and the resulting electronic structure is thus rather poor. Hence, we see that QM/MM (and similar) approaches can be employed to enhance significantly the high-resolution details of X-ray structures.

Some systematic trends are found in the calculated strain energies. Those for the special pair are 3−5 kcal mol^{-1} higher than those for the accessory BChls, while those for the BPheos and the remaining cofactors are smaller by that much again. Also, the strain energies for the active L-side cofactors are significantly *less* than those of their M-side counterparts by 2, 4, 1, and 3 kcal mol^{-1} for P, BChl, BPheo, the quinones, respectively.

Previously,[56] we optimized the structure for the WT; this calculation commenced at the X-ray coordinates (1PCR)[1] and conserved all key qualitative features of the structure. Here, for completeness, we revise these optimized coordinates using the same procedures that are applied to the mutants, and again only seemingly minor (but chemically significant) changes to the structure result. This procedure is also applied to the X-ray coordinates (1PRC)[2] of *Rh. viridis*. As a quantitative indication of the magnitude of the induced structural changes, we have evaluated the root-mean-square (RMS) change in the Cartesian coordinates of the heavy atoms of all residues which penetrate within 9 Å of any atom of the special pair (ca. 190 residues). For the wild types of *Rb. Sphaeroides* and *Rh. viridis*, this is 0.63 and 0.68 Å, respectively, changes which are well inside the experimental resolution of ca. 2.5 Å (although, in practice, the uncertainties in X-ray coordinates are typically much less than the resolution). Of the mutants which we consider, X-ray structures are available for two, FY(M197) and HL(M202) (also, a structure for YW(M210) should shortly become available[12]). We have compared our structures, estimated by modification of the WT, with these experimental structures and obtain RMS displacements of 0.79 Å for FY(M197) and 0.90 Å for HL(M202). These changes are again much less than the experimental resolution and support the hypothesis that useful structures for the mutants can be obtained simply by modifying the structure of the WT.

In Table 7 and Figure 2 we correlate the shortest calculated intermolecular bond length between the 2a-acetyl and 9-keto carbonyl oxygen atoms of P_L and P_M with the observed carbonyl vibrational frequencies. The shortest bond length is selected from the available possibilities shown in Tables 2−6. In overview,

(69) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucl. Acids Res.* **2000**, 235.

(70) *HYPERCHEM*, Release 5.0; Hypercube Inc.: Waterloo, Ontario, 1996.

(71) Rauhut, G.; Alex, A.; Chandrasekhar, J.; Steinke, T.; Sauer, W.; Beck, B.; Hutter, M.; Gedeck, P.; Clark, T. *VAMP*, Version 6.5; Universität Erlangen: Nürnberg, 1997.

(72) Sakuma, T.; Kashiwagi, H.; Takada, T.; Nakamura, H. *Int. J. Quantum Chem.* **1997**, *61*, 137.

Table 2: Key Intermolecular Bond Distances (Å) Involving the Cofactors of *Rb. sphaeroides* Wild Type (WT) and Its L*x*(M160) Single Mutants[a]

cofactor site	other site	WT	LS(M160)	LN(M160)	LQ(M160)	LH(M160)
P_M 2a-acetyl-O	TyrM210-HO	4.837	4.865	4.835	4.811	4.806
P_M 2a-acetyl-O	P_L Mg	2.426	2.427	2.436	2.430	2.424
P_M 2a-acetyl-H?	P_L Mg	4.394	4.389	4.389	4.386	4.391
P_M 9-keto-O	IleM284-H?	2.400	2.341	2.445	2.398	2.761
P_M 9-keto-O	xM160-H?	2.769	2.020	1.984	1.980	1.780
P_M Mg	HisM202-NE2	2.083	2.122	2.083	2.085	2.123
P_M Mg	P_L Mg	7.869	7.867	7.880	7.871	7.853
P_L 2a-acetyl-O	HisL168-HE2	1.894	1.839	1.932	1.921	1.934
P_L 2a-acetyl-H?	P_M Mg	2.735	2.667	2.658	2.667	2.612
P_L 9-keto-O	MetL248−2HE	2.666	2.655	2.665	2.663	2.675
P_L 10a-ester-O	CysL247-HS	3.744	3.713	3.656	3.629	3.003
P_L 10a-ester-O	SerL244-H?	2.287	2.309	2.291	2.295	2.311
P_L 10a-ester-O	LeuL131-H?	4.420	4.487	4.500	4.493	4.471
P_L Mg	HisL173-NE2	2.139	2.212	2.092	2.093	2.099
$BChl_M$ Mg	HisM182-NE2	2.086	2.101	2.083	2.083	2.113
$BChl_M$ 9-keto-O	water62-H?	2.295	2.242	2.234	2.109	1.908
$BChl_L$ Mg	HisL153-NE2	2.115	2.109	2.115	2.113	2.113
$BChl_L$ 9-keto-O	water59-H?	3.428	3.329	4.218	3.355	5.422
$BPhl_M$ 10a-ester	TrpM129-HE1	2.375	2.332	2.366	2.397	2.361
$BPhl_L$ 10a-ester	TrpL100-HE1	2.479	2.503	2.508	2.501	2.507
$BPhl_L$ 9-keto-O	TrpM252-HH2	2.417	2.406	2.408	2.408	2.408
$BPhl_L$ 9-keto-O	GluL104-HE2	1.806	2.063	1.834	1.804	1.808

[a] H? indicates the closest hydrogen.

Table 3: Key Intermolecular Bond Distances (Å) Involving the Cofactors of *Rb. sphaeroides* FY(M197) and FH(M197) Single, Double, and Triple Mutants[a]

cofactor site	other site	FY(M197)	FH(M197)	FH(M197)+ HF(L168)	FH(M197)+ LH(M160)	FH(M197)+ LH(M160)+ LH(L131)	FH(M197)+ LH(M160)+ $LH(L131)_{10a}$
P_M 2a-acetyl-O	P_L Mg	4.903	5.008	4.978	5.013	4.939	4.939
P_M 2a-acetyl-O	TyrM210-HO	4.773	5.104	5.074	5.124	4.851	4.884
P_M 2a-acetyl-O	xM197-H?	1.761	1.898	1.867	1.876	1.765	1.754
P_M 2a-acetyl-H?	P_L Mg	2.450	2.614	2.647	2.619	2.530	2.530
P_M 9-keto-O	xM160-H?	2.482	2.604	2.604	1.800	1.842	1.841
P_M 9-keto-O	IleM284-H?	2.325	2.333	2.374	2.362	2.221	2.222
P_M Mg	HisM202-NE2	2.086	2.063	2.063	2.064	2.060	2.129
P_M Mg	P_L Mg	8.177	8.380	8.380	8.350	8.464	8.464
P_L 2a-acetyl-O	P_M Mg	4.794	4.921	4.853	4.887	5.072	5.072
P_L 2a-acetyl-O	xL168-H?	1.929	1.847	2.188	1.888	1.929	1.934
P_L 2a-acetyl-H?	P_M Mg	2.525	2.562	2.547	2.517	2.697	2.697
P_L 9-keto-O	MetL248−2HE	2.785	2.613	2.610	2.610	2.791	2.854
P_L 9-keto-O	xL131-H?	2.629	2.673	2.686	2.699	1.879	2.896
P_L 10a-ester-O	CysL247-HS	3.763	3.782	3.704	3.633	3.909	5.356
P_L 10a-ester-O	SerL244-H?	2.264	2.267	2.287	2.307	2.132	4.024
P_L 10a-ester-O	LeuL131-H?	4.459	4.343	4.391	4.348	3.878	1.898
P_L Mg	HisL173-NE2	2.076	2.066	2.071	2.060	2.025	2.119
$BChl_M$ Mg	HisM182-NE2	2.081	2.086	2.090	2.090	2.072	2.108
$BChl_M$ 9-keto-O	water62-H?	2.026	2.258	2.239	1.937	2.216	2.164
$BChl_L$ Mg	HisL153-NE2	2.101	2.105	2.104	2.106	2.099	2.125
$BChl_L$ 9-keto-O	water59-H?	3.386	5.099	5.067	5.085	4.212	4.212
$BPhl_M$ 10a-ester	TrpM129-HE1	2.361	2.368	2.428	2.354	2.446	2.446
$BPhl_L$ 10a-ester	TrpL100-HE1	2.514	2.516	2.513	2.519	2.501	2.501
$BPhl_L$ 9-keto-O	TrpM252-HH2	2.398	2.387	2.394	2.400	2.454	2.454
$BPhl_L$ 9-keto-O	GluL104-HE2	1.802	2.061	1.817	1.804	1.839	1.836

[a] H? indicates the closest hydrogen.

Figure 2 shows that hydrogen-bonded interactions result in bond lengths *R* of typically 1.8−2.0 Å and result in significant red-shifts in the vibration frequency compared to non-hydrogen-bonding interactions, for which the bond lengths range over 2.2−2.7 Å. Quantitatively, the nature of the correlation depends on both the identity of the carbonyl group and the intrinsic nature of the surrounding protein, but qualitatively the clear pattern apparent in this figure is only obtained after a significant number of structural issues relating to the WT and its mutants are resolved, and these specific chemical issues are addressed in turn in the following subsections.

A. Orientation of the 2a-Acetyl Groups in the Wild Type, F*x*(M197), and HF(L168) Mutants. The 2a-acetyl groups of the special pair may each be oriented in one of two general directions, either with their oxygen atoms facing inward, forming a sixth ligand to the magnesium of the other half, or oriented outward forming a hydrogen-bond to the surrounding protein. In the optimized structure[56] of the wild-type reaction center sketched in Figure 1, the 2a-acetyl group of P_L is oriented outward, hydrogen-bonded to the histidine at L168, while as no site for hydrogen-bonding is available, the 2a-acetyl of P_M is oriented inward.

On a fine scale, both the inward-pointing and outward-pointing structures could be either planar, with the carbonyl conjugated to the porphryrinic ring, or twisted, optimizing the intermolecular interactions and minimizing internal methyl steric repulsions. We have found[73] through extensive ab initio calculation on model compounds in the gas phase that the energetic

Table 4: Key Intermolecular Bond Distances (Å) Involving the Cofactors of *Rb. sphaeroides* Heterodimer HL(M202) Single and Double Mutants, with P_M Now a BPhl Rather than BChl[a]

cofactor site	other site	HL(M202)	HL(M202)+ LH(L131)	HL(M202)+ $LH(L131)_{10a}$	HL(M202)+ HF(L168)	HL(M202)+ LH(M160)	HL(M202)+ FH(M197)
P_M 2a-acetyl-O	P_L Mg	2.347	2.349	2.359	2.394	2.327	5.046
P_M 2a-acetyl-O	TyrM210-HO	4.786	4.688	4.659	4.943	4.751	5.232
P_M 2a-acetyl-O	xM197-H?	4.984	5.109	5.026	4.905	4.974	1.932
P_M 2a-acetyl-H?	P_L Mg	4.322	4.456	4.437	4.502	4.413	2.632
P_M 9-keto-O	IleM284-H?	2.490	2.377	2.378	2.479	2.363	2.462
P_M 9-keto-O	xM160-H?	3.043	3.187	3.155	3.188	1.843	2.830
P_M Ring Centre	LeuM202-H?	2.523	2.585	2.570	2.601	2.283	2.721
P_M Ring Centre	P_L Mg	7.606	7.532	7.564	7.546	7.421	8.167
P_L 2a-acetyl-O	P_M Ring Centre	4.562	4.498	4.502	4.342	4.404	4.646
P_L 2a-acetyl-O	xL168-H?	1.929	1.953	1.966	2.243	1.976	1.886
P_L 2a-acetyl-H?	P_M Ring Centre	3.433	3.940	3.933	3.484	3.979	3.174
P_L 9-keto-O	MetL248−2HE	2.689	2.667	2.836	2.645	2.655	2.526
P_L 9-keto-O	xL131-H?	2.644	2.450	2.658	2.801	2.773	2.673
P_L 10a-ester-O	CysL247-HS	3.622	3.750	5.327	3.643	3.540	3.652
P_L 10a-ester-O	SerL244-H?	2.357	2.324	3.831	2.280	2.336	2.332
P_L 10a-ester-O	LeuL131-H?	4.443	3.480	2.062	4.518	4.599	4.388
P_L Mg	HisL173-NE2	2.116	2.177	2.190	2.104	2.163	2.062
$BChl_M$ Mg	HisM182-NE2	2.086	2.100	2.100	2.107	2.085	2.087
$BChl_M$ 9-keto-O	water62-H?	2.272	2.247	2.356	1.932	2.231	2.135
$BChl_L$ Mg	HisL153-NE2	2.114	2.109	2.109	2.113	2.116	2.105
$BChl_L$ 9-keto-O	water59-H?	5.882	5.839	5.839	5.816	5.704	5.334
$BPhl_M$ 10a-ester	TrpM129-HE1	2.346	2.349	2.349	2.348	2.276	2.329
$BPhl_L$ 10a-ester	TrpL100-HE1	2.507	2.504	2.504	2.503	2.505	2.521
$BPhl_L$ 9-keto-O	TrpM252-HH2	2.407	2.408	2.408	2.416	2.407	2.400
$BPhl_L$ 9-keto-O	GluL104-HE2	1.803	1.800	1.805	1.823	1.813	1.810

[a] H? indicates the closest hydrogen.

Table 5: Key Intermolecular Bond Distances (Å) Involving the Cofactors of *Rb. sphaeroides* LH(L131), HF(L168) and the Double Mutants HF(L168)+LH(L131) and LH(L131)+LH(M160)[a]

cofactor site	other site	LH(L131)	$LH(L131)_{10a}$	HF(L168)	$HF(L168)_{rot}$	HF(L168)+ LH(L131)	LH(L131)+ LH(M160)
P_M 2a-acetyl-O	TyrM210-HO	4.606	4.698	4.603	4.959	4.674	4.848
P_M 2a-acetyl-O	P_L Mg	2.379	2.392	2.358	2.351	2.361	2.426
P_M 9-keto-O	IleM284-H?	2.293	2.311	2.332	2.420	2.327	2.354
P_M 9-keto-O	xM160-H?	2.643	2.672	2.626	3.345	2.633	1.807
P_M Mg	HisM202-NE2	2.122	2.124	2.121	2.167	2.090	2.083
P_M Mg	P_L Mg	7.818	7.813	7.692	7.464	7.716	7.860
P_L 2a-acetyl-O	P_M Mg	4.863	4.868	4.678	2.301	4.695	4.868
P_L 2a-acetyl-O	xL168-H?	1.931	1.938	2.323	5.062	2.301	1.921
P_L 2a-acetyl-H?	P_M Mg	2.749	2.709	2.511	4.260	2.559	2.597
P_L 9-keto-O	MetL248−2HE	3.158	3.222	2.802	2.681	2.852	2.752
P_L 9-keto-O	xL131-H?	1.897	3.330	2.592	3.073	1.837	1.835
P_L 10a-ester-O	CysL247-HS	2.569	7.046	3.728	3.706	4.036	3.782
P_L 10a-ester-O	SerL244-H?	2.206	3.177	2.246	2.307	2.171	2.227
P_L 10a-ester-O	LeuL131-H?	4.280	2.030	4.530	4.706	5.419	4.797
P_L Mg	HisL173-NE2	2.180	2.208	2.186	2.191	2.147	2.102
$BChl_M$ Mg	HisM182-NE2	2.103	2.103	2.104	2.103	2.078	2.086
$BChl_M$ 9-keto-O	water62-H?	2.210	2.159	2.367	2.128	2.280	2.072
$BChl_L$ Mg	HisL153-NE2	2.110	2.110	2.116	2.106	2.102	2.112
$BChl_L$ 9-keto-O	water59-H?	3.370	3.370	3.414	3.070	3.366	4.091
$BPhl_M$ 10a-ester	TrpM129-HE1	2.308	2.308	2.339	2.273	2.319	2.379
$BPhl_L$ 10a-ester	TrpL100-HE1	2.655	2.629	2.520	2.408	2.496	2.504
$BPhl_L$ 9-keto-O	TrpM252-HH2	2.382	2.450	2.400	2.385	2.410	2.397
$BPhl_L$ 9-keto-O	GluL104-HE2	1.798	1.779	1.833	2.052	2.072	1.801

[a] H? indicates the closest hydrogen.

balance is finely tuned, with coupled-cluster theory supporting the AM1 result, evident in all of our calculated structures, that the carbonyl groups always twist to ca. 40° out of planarity; this occurs for *both* the inward-pointing and outward-pointing structures.

The F*x*(M197) mutants were designed to introduce a hydrogen bond to the 2a-acetyl group of P_M. The intermolecular-bond-length/vibration frequency correlation shown in Figure 2 for neutral reaction centers reveals two types of structures, the hydrogen-bonded mutants (minimum intermolecular distance $R = 1.8-1.9$ Å, $\nu = 1628-1639$ cm^{-1}) and the unmutated magnesium-coordinated ones ($R_{Mg-O} = 2.3-2.6$ Å, $\nu = 1653-1663$ cm^{-1}). Our calculated bond lengths and the observed vibration frequencies for the F*x*(M197) mutants fall into the hydrogen-bonded block and hence it is clear that rotation of the 2a-acetyl group is in fact induced by the mutation to the neutral reaction center.

For the oxidized reaction center P^+ in the F*x*(M197) mutants, it is not obvious that this rotation should persist, as oxidized P_L magnesium should prefer to be coordinated to the electron-rich

(73) Reimers, J. R.; Hutter, M. C.; Hughes, J. M.; Hush, N. S. *Int. J. Quantum Chem.* **2000**, *80*, 1224.

Table 6: Key Intermolecular Bond Distances (Å) Involving *Rb. sphaeroides* WT with the OH from M210 Rotated by ca. 180° and Also the N*x*(L166) and Y*x*(M210) Mutants[a]

site 1	site 2	NL(L166)	NH(L166)	WT_{210}	YF(M210)	YW(M210)
P_M 2a-acetyl-O	xM210-H?	4.845	4.830	3.841	4.620	4.224
P_M 2a-acetyl-O	P_L Mg	2.430	2.434	2.426	2.428	2.370
P_M 2a-acetyl-O	PheM197-HZ	4.928	4.947	4.947	4.928	4.846
P_M 9-keto-O	LeuM160-H?	2.737	2.731	2.769	2.793	2.776
P_M 9-keto-O	IleM284-H?	2.399	2.388	2.400	2.394	2.519
P_M Mg	HisM202-NE2	2.082	2.083	2.083	2.083	2.116
P_M Mg	P_L Mg	7.879	7.886	7.869	7.864	7.674
P_L 2a-acetyl-O	P_M Mg	4.921	4.899	4.920	4.932	4.772
P_L 2a-acetyl-O	HisL168-HE2	1.852	1.829	1.894	1.912	1.859
P_L 9-keto-O	MetL248−2HE	2.636	2.638	2.666	2.668	2.611
P_L 9-keto-O	LeuL131-H?	2.738	2.720	2.735	2.723	2.560
P_L 10a-ester-O	CysL247-HS	3.673	3.670	3.744	3.704	3.599
P_L 10a-ester-O	SerL244-H?	2.298	2.291	2.287	2.294	2.368
P_L 10a-ester-O	LeuL131-H?	4.429	4.425	4.420	4.389	4.726
P_L Mg	HisL173-NE2	2.099	2.106	2.139	2.093	2.146
HisL168-ND1	xL166-H?	2.745	2.523	2.009	1.970	2.056
$BChl_M$ Mg	HisM182-NE2	2.085	2.088	2.086	2.087	2.100
$BChl_M$ 9-keto-O	water62-H?	2.135	2.157	2.295	2.120	2.211
$BChl_L$ Mg	HisL153-NE2	2.115	2.115	2.115	2.115	2.133
$BChl_L$ 9-keto-O	water59-H?	3.393	3.400	3.428	3.403	3.416
$BPhl_M$ 10a-ester	TrpM129-HE1	2.388	2.379	2.375	2.397	2.384
$BPhl_L$ 10a-ester	TrpL100-HE1	2.495	2.484	2.479	2.487	2.311
$BPhl_L$ 9-keto-O	TrpM252-HH2	2.426	2.417	2.417	2.423	2.506
$BPhl_L$ 9-keto-O	GluL104-HE2	1.808	1.819	1.806	1.805	1.819

[a] H? indicates the closest hydrogen.

Table 7: Observed Carbonyl Vibration Frequencies ν, in cm^{-1}, Compared to the Calculated Minimum Carbonyl Oxygen-to-Protein Interatomic Separation Extracted from Tables 1−5, in Å

		P_L 9-keto		P_L 2a-acetyl		P_M 9-keto		P_M 2a-acetyl	
mutant	obs. ref	ν	*R*	ν	*R*	ν	*R*	ν	*R*
WT	30	1691	2.476	1620	1.894	1679	2.400	1653	2.426
HF(L168)	30	1688	2.420	1653	2.297	1679	2.332	1653	2.358
LH(M160)	30	1694	2.481	1618	1.934	1657	1.780	1657	2.424
LS(M160)	41	1693	2.487		1.839	1664	2.020		2.427
LN(M160)	41	1693	2.495		1.932	1665	1.984		2.436
LQ(M160)	41	1694	2.485		1.921	1656	1.980		2.430
FH(M197)	30	1693	2.317	1622	1.847	1679	2.333	1630	1.898
FH(M197)+HF(L168)	30	1689	2.328	1657	2.188	1682	2.374	1628	1.867
FH(M197)+LH(M160)	30	1696	2.338	1618	1.888	1661	1.800	1639	1.876
FH(M197)+LH(M160)+$LH(L131)_{10a}$	30	1674	2.250	1628	1.934	1663	1.841	1636	1.754
HL(M202)	23	1697	2.552	1619	1.929		3.156	1663	2.626
HL(M202)+$LH(L131)_{10a}$	23	1697	2.371		1.966		3.075		2.657
HL(M202)+HF(L168)	23	1695	2.441		2.243		2.985		2.579
HL(M202)+LH(M160)	23	1700	2.478	1619	1.976		3.090		2.585
HL(M202)+FH(M197)	23	1700	2.344	1619	1.886		3.324		2.695
LH(L131)	21	1657	1.897	1627	1.931	1684	2.293	1657	2.379
$LH(L131)_{10a}$	30	1673	2.457	1627	1.938	1684	2.311	1657	2.392
HF(L168)+LH(L131)	30	1664	1.837	1653	2.289		2.327	1653	2.361
NL(L166)	52		2.447	1618	1.852		2.399		2.430
NH(L166)	52		2.438	1616	1.829		2.388		2.434

carbonyl group rather than to the 2a methyl group. Molecular dynamics simulations[45,47] do indeed suggest that the 2a-acetyl rotates on oxidation, but without detailed verification, the potential surfaces used are not of sufficient accuracy to provide an authoritative result. We have considered the gas-phase energetics of this process using coupled-cluster calculations[73] for model compounds. These results show that a large driving force for hydrogen-bond breakage of ca. 7 kcal mol^{-1} is in fact provided by the preference of the magnesium of the charged BChl to be coordinated to a carbonyl group rather than to a methyl group. This will be sufficient to overcome the hydrogen-bonding with the protein, consistent with the molecular dynamics result.

The HF(L168) mutants, on the other hand, were designed to *remove* the hydrogen bond between the 2a-acetyl of P_L and histidine L168. For the wild type, our structures with 2a-acetyl-oxygen to protein−hydrogen-bond lengths of $R = 1.8-2.0$ Å (see Figure 2) confirm the spectroscopic results[74] that a hydrogen bond is present therein. This conclusion is actually contrary to that drawn from the interaction distance of 2.6 Å obtained[56] using the observed X-ray structure of the wild type and indicates a significant improvement in the geometry. In the HF(L168) mutants, hydrogen-bonding could only involve unlikely CO-to-CH interactions, and it is possible that the 2a-acetyl group actually rotates so that its oxygen forms a sixth ligand to the magnesium of P_M. To test for this possibility, two optimizations for the HF(L168) single mutant were performed, one starting in each of the two possible orientations. For the structure in which the 2a-acetyl is oriented with its CO toward the protein (akin to the WT structure), we find that the introduced

(74) Mattioli, T. A.; Lin, X.; Allen, J. P.; Williams, J. C. *Biochemistry* **1995**, *34*, 6142.

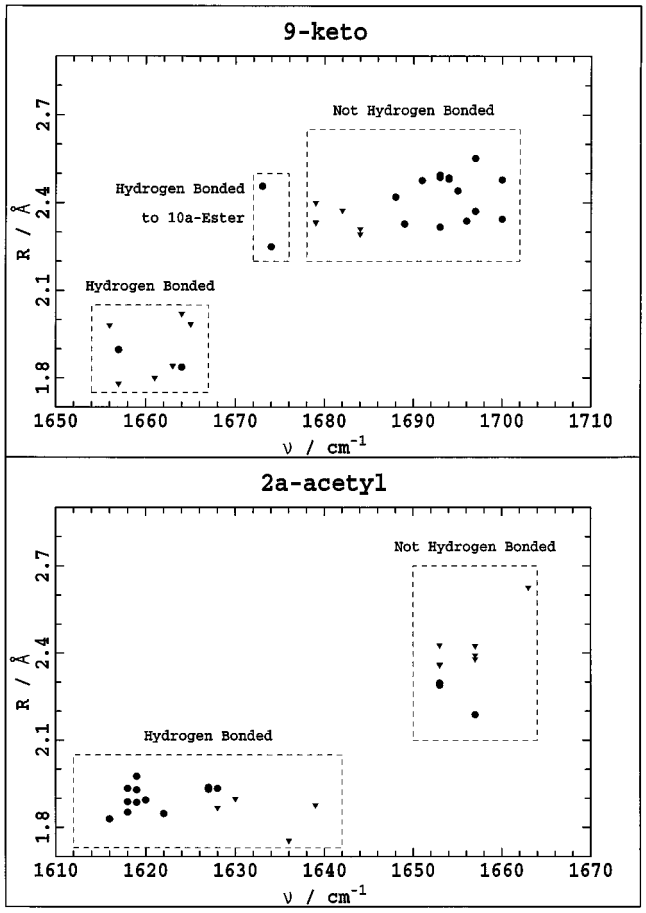

Figure 2. The calculated nearest intermolecular distance from the 2a-acetyl and 9-keto carbonyl oxygens of P_L (●) or P_M (▼) to an amino acid hydrogen atom, shown as a function of the observed carbonyl vibration frequencies. The data are from Table 7.

phenylalanine is quite close to the special pair, with intermolecular bond lengths of order $R = 2.2-2.3$ Å. While these are somewhat shorter than is typical for CO to CH interactions (ca. 2.4−2.7 Å), they nevertheless indicate that no hydrogen bond is present in the mutant reaction centers. Results from the alternate structure in which the group was rotated such that the carbonyl coordinates with the magnesium are labeled $HF(L168)_{rot}$ in Table 5 and the Supporting Information. Within the time scale of 5 ps, we find that both structures are stable in molecular dynamics simulations. Although these structures differ significantly in the relative orientations within the special pair, the AM1-calculated monomer strain energies are similar in each case. Note that simply rotating the acetyl from the HF(L168) unrotated structure to form the $HF(L168)_{rot}$ structure is not possible due to steric repulsions between the 2a-methyl group, the bulky phenylalanine, and $BChl_L$, but this is alleviated through the flexing of P_L when its 2a-acetyl ligates to the Mg of P_M. The rotation thus involves significant rearrangement of the surrounding structure.

Authoritative theoretical analysis of the energetics of the 2a-acetyl rotation of neutral reaction centers requires the evaluation of the associated free-energy change. Our ab initio calculations on gas-phase model dimer interactions[73] indicate that the rotated conformer is more stable by just 0.2 kcal mol^{-1}, a value which is smaller than the likely errors in the calculation as well as intermolecular interaction energies and possible entropic contributions. Hence, such free-energy calculations must be made using a very high quality potential-energy surface, and are not currently feasible. The observed CO vibration frequency, in principle, provides a means of resolving this issue, but no model for the interpretation of the available data has yet been proposed.

The situation is simpler for the (oxidized) special-pair radical cation of HF(L168) mutants, however. For these, a large fraction[15,17] of the charge is localized on P_M, and hence the electrostatic interaction[73] between P_M and the electron-rich carbonyl should be sufficient to guarantee the magnesium-coordinated configuration. By contrast, in the wild type and most other mutants, much less charge is delocalized onto P_M, reducing the attractiveness of the coordinated orientation. Also, for these, 2a-acetyl rotation on oxidation would additionally require that the hydrogen bond with L168 be broken, and the combination of these two features should ensure that in P^+ WT the 2a-acetyl of P_L remains oriented toward the protein. Hence, in summary, we conclude that the orientation of the acetyl groups of P_L and P_M in oxidized reaction centers in which hydrogen-bonding to the protein is possible is controlled by the degree of charge delocalization: for the 2a-acetyl of P_L (and for P_M of HF(L168) mutants), the cofactor on which the bulk of the charge localizes, hydrogen-bonding persists, while for that of P_M WT, carbonyl coordination to magnesium results.

B. The Structure of the FY(M197) Mutant. The FY(M197) mutant has attracted particular attention because its P_M 2a-acetyl frequency is[11] 1636 cm^{-1}, 6 cm^{-1} greater[30] than that of the related FH(M197) mutant. With the aid of X-ray structural information,[11] this has been interpreted in terms of reduced hydrogen-bonding caused by an alternate configuration for the protein and hence a modified hydrogen-bonding arrangement. Our results do indeed *predict* the observed alternate configuration and *further* indicate that there is considerable structural flexibility associated with the introduced tyrosine. These results are described in detail in Supporting Information.

C. Water Penetration in the Oxidized FY(M197) Mutant. The M197 residue is located adjacent to the special pair but forms part of the hydrophobic−hydrophilic border within the protein. In the X-ray structure of the wild type, two water molecules (HOH77 and HOH25 in our optimized structure) are located adjacent to this residue but distant from the special pair. The introduction of tyrosine at M197 offers the possibility of an additional water molecule being drawn to the other side of this reside, near the special pair; such a molecule could form three hydrogen bonds, interacting with asparagine M195, serine M158, and HOH28. The X-ray structure of the neutral reaction center does not indicate that a water is located there; a large void capable of accommodating a water molecule is apparent in this region, however, and there is some chance that one is indeed present but has been missed due to the difficulty of detection of water in protein X-ray structures. Regardless, one indeed may be present in the oxidized reaction center P^+ as it would act to solvate the charge, thus significantly lowering the midpoint potential. Further, as the hydrogen bond from tyrosine M197 to the special pair is expected to be broken in P^+, the tyrosine itself is free to reorient and hence form a *fourth* hydrogen bond with the introduced water. This very strongly suggests that water penetration near the oxidized special pair does indeed occur.

The observed midpoint potential of P in FY(M197) is anomalously low,[11] 0.146 V lower than that in FH(M197); this shift is actually very large, as the observed maximum difference in midpoint potential for the entire, much studied, L*x*(M160) series is just 0.057 eV. We have calculated a structure for the neutral species, named $FY(M197)_{H2O}$, containing an additional water molecule, and its coordinates are given in Supporting Information. Using the method developed in Appendix 1 for

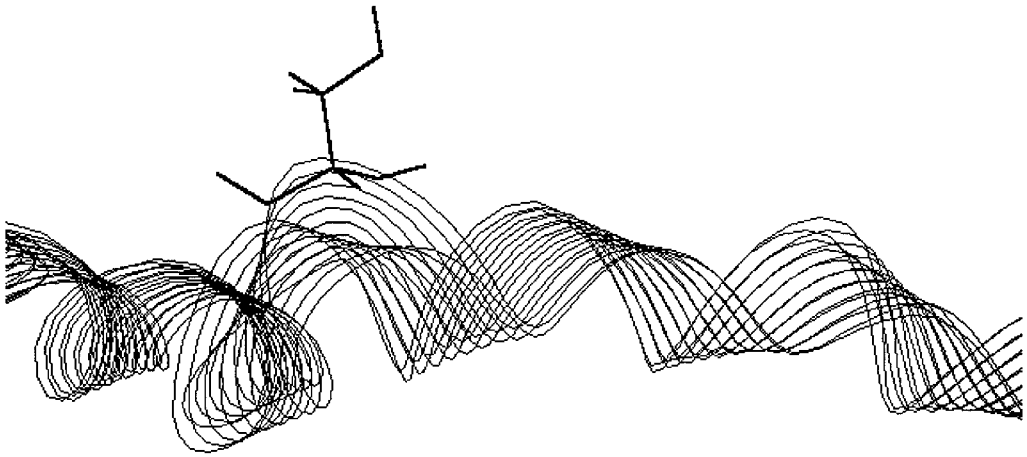

Figure 3. Overlay of the optimized structures for the wild-type and LS(M160) mutant, indicating the change in the loop structure required to hydrogen-bond the small serine residue (shown) to the 9-keto group of P_M.

determining Koopmans-level ionization energies from QM/MM wave functions, we calculate that the effect of the water penetration on the midpoint potential of P is to lower it by 0.10 eV. In addition, reorientation of the OH group of tyrosine M197 toward the introduced water and other structural relaxation effects would be expected to provide a substantial additional lowering of the midpoint potential, toward the observed value. While speculative, this calculation provides a plausible explanation for the observed highly anomalous value.

D. Chain Distortion in L*x*(M160) Mutants. The L*x*(M160) series of mutants have been developed with the intent of introducing hydrogen bonds to the 9-keto groups of P_M, and it is clear from spectroscopic evidence[38,41,43] that the desired results have indeed been obtained. These mutants have formed an important part of the overall reaction-center research through correlations which have been obtained linking their special-pair midpoint potentials to fundamental intermolecular interaction properties.[41,43,44]

Structurally, these mutants raise interesting questions as many (e.g., LS(M160) and LN(M160)) involve the replacement of a large residue with a small one. Clearly, if the skeletal backbone remains largely invariant, then for the small residues, no hydrogen bond could form. Our results, summarized in Figure 2, have bond lengths $R = 2.3-2.4$ Å to the unmutated leucine ($\nu = 1679-1684$ cm^{-1}) but bond lengths of 1.8−2.0 Å ($\nu = 1656-1664$ cm^{-1}) to the mutated residues. However, significant changes to the protein backbone were required to obtain separations as short as these for LS(M160) and LN(M160). An overlay of the backbone of the wild-type and LS(M160) mutant is shown in Figure 3, along with the serine of LS(M160). Large displacements of up to 1.7 Å are evident for the protein loop involving site M160, tightening the curvature on the neighboring loops. Smaller changes at much larger distances are also evident; while these do not appear significant when viewed using the scale of Figure 3, if relaxation of the entire helical protein strand and its neighboring strands is not included in the calculations, then the introduced hydrogen bond was always found to break as the protein returned to its original configuration.

Molecular dynamics calculations such as ours on proteins do not fully scan the available configuration space, and hence are not expected to have fully relaxed the structure. For the short-chain M160 mutants we observed quite large dynamical fluctuations during the early stages of the optimization, suggesting that our final structures may indeed be significantly different from equilibrium ones. The final structures, obtained after many relaxations of order 100 ps each, are stable, however. What is clear is that qualitatively these structures do indicate the (minimum) extent of protein modifications required to form the hydrogen bonds at M160, and significant changes are required. It is thus remarkable that the observed midpoint potential of P for these mutants can[41,43,44] be correlated with properties of the special pair using simplistic models.

E. Competition for Hydrogen-Bonding to the Introduced Histidine between the 9-Keto and 10a-Ester Groups in L*x*(M160) and LH(L131) Mutants (and the Stark Effect of the Protein on the 9-Keto Frequency). The LH(L131) mutants have been developed with the intent of introducing hydrogen bonds to the 9-keto groups of P_L and are hence analogous to the LH(M160) mutant previously discussed. However, the available spectroscopic data[17,21,32,38,39] indicates that sometimes the desired hydrogen bond is introduced while sometimes it is not, and no interpretation of this variance has as yet been proposed. Our results shown in Figure 2 indicate that the observed data may be interpreted qualitatively in terms of *three* bonding patterns. Reaction centres without the LH(L131) mutation appear in the region of this figure with $R = 2.3-2.6$ Å, $\nu = 1688-1700$ cm^{-1}, while mutants in which the hydrogen-bonding to the 9-keto group is properly achieved have $R = 1.8-1.9$ Å, $\nu = 1657-1664$ cm^{-1}. The third group has $R \approx 2.2-2.5$ Å, $\nu \approx 1673$ cm^{-1} and in these the introduced histidine forms a hydrogen bond to the neighboring 10a-ester carbonyl of P_L rather than the 9-keto group. It is possible to interpret the available experimental data by classifying the mutants LH(L131) at 15 K,[40] HL(M202)+LH(L131),[23] and FH(M197)+LH(M160)+LH(L131)[38] as ones containing this alternate hydrogen-bonding, while HF(L168)+LH(L131)[38] and LH(L131) at 100 K[21] are classified as mutants in which the expected hydrogen bond is found.

For three mutants, LH(L131), HL(M202)+LH(L131), and FH(M197)+LH(M160)+LH(L131), we perform two sets of simulations, one with the hydrogen-bonding established to the 9-keto group and the other, labeled LH(L131)$_{10a}$, HL(M202)+LH(L131)$_{10a}$, and FH(M197)+LH(M160)+LH(L131)$_{10a}$, respectively, with hydrogen-bonding established to the 10a-acetyl group. All of the simulations resulted in structures stable on the 5 ps time scale except for LH(L131)$_{10a}$ for which the hydrogen bond reverts to the 9-keto group. The optimized structures for LH(L131) and LH(L131)$_{10a}$ are shown in Figure 4. These differ by a reorientation of the introduced histidine, accompanied by a ca. 180° rotation of the 10a-ester group about its link to ring V. We find that, for each mutant, hydrogen-bonding to the 9-keto group breaks following manual reorienta-

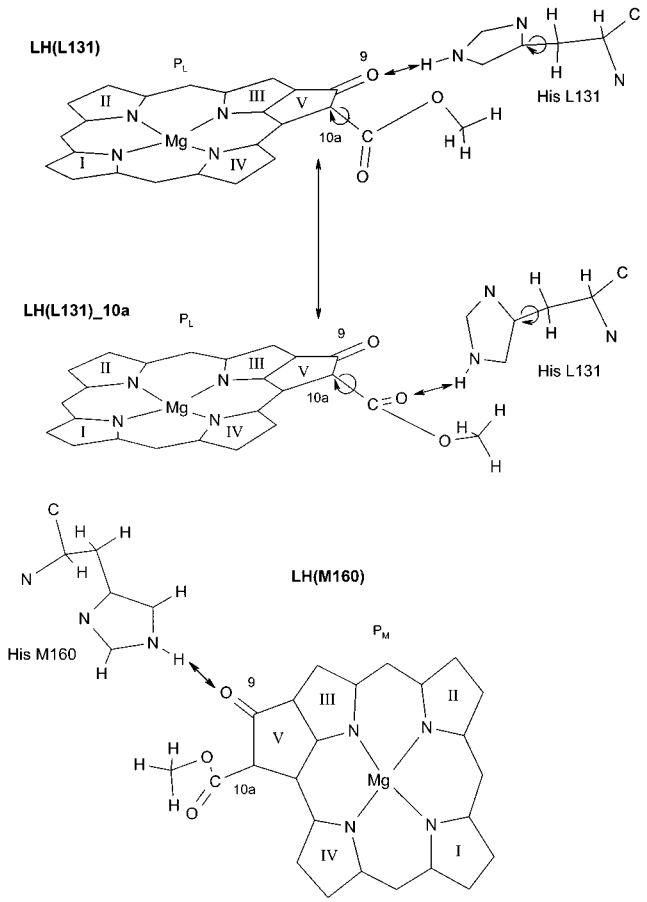

Figure 4. Optimized structures for the introduced histidine in the LH(L131) mutant, showing hydrogen-bonding to the 9-keto of P_L, its $LH(L131)_{10a}$ variant in which the hydrogen bond forms to the neighboring 10a-ester instead, and the analogous bonding in LH(M160). For LH(M160), the introduced histidine comes from the side, and only the 9-keto group is accessible, whereas both the 9-keto and 10a-ester groups are accessible for LH(L131).

tion of the 10a-ester in the appropriate direction. This result can be rationalized by noting that the 9-keto bond dipole is reduced significantly from that of a typical carbonyl through conjugation with the porphyrinic ring, allowing for stronger hydrogen-bonding with the equally accessible 10a-ester. When the 10a-ester is oriented away from L131, it falls into a stable pocket in which it is located between serine L244, cystine L247 and methionine L248. The net interaction with these groups is fairly weak, the closest hydrogen atoms being those of CH from the serine (2.41 Å) and the methionine (2.61 Å); the cystine SH is distant at 2.71 Å, while the serine OH is actually oriented in the opposite direction, forming a hydrogen bond to the 7c-keto oxygen of P_L. While the possibility of this OH group hydrogen-bonding to the 10a-ester in the wild type has been discussed,[75] it is currently believed that no such bond actually forms.[34] Combining all of these features, we conclude that the energetics of the 10a-ester rotation appear finely balanced, accounting for the variability observed in the experimental results for LH(L131) mutants.

A key qualitative feature to our interpretation of the experimental data is the assumption that the P_L 9-keto vibration frequency when L131 to 10a-ester hydrogen-bonding occurs should be intermediate between the cases of no hydrogen-bonding at all and hydrogen-bonding to the 9-keto group. Stark-effect calculations described in Appendix 2 support this. Hence, by combining structural and vibrational arguments, we present a plausible explanation of variability of the observed 9-keto vibration frequencies for LH(L131) mutants.

The analogous geometrical structure obtained for the LH(M160) mutant is also shown in Figure 4. For this mutant, hydrogen-bonding of the histidine to the 10a-ester group is structurally inhibited as the histidine is tied into the protein backbone at a location that permits access only to the 9-keto region. Another possible explanation for the observation of hydrogen-bonding only to the 9-keto group of P_M is that the neighboring 10a-ester is inhibited from rotating into the required conformation by steric or hydrogen-bonding interactions with the protein. However, our calculated structures indicate that the 10a-ester orientation is only 30° different from that which leads to hydrogen-bonding with the histidine at L131, and thus the required conformation may well be accessible. Hydrogen-bonding of the 10a-ester with the OH from serine M205 could occur, and hence we started the MD simulations in favorable configurations, but in no case did a hydrogen bond actually form. Rather, we found that conformational changes occurred which allowed serine M205 to strongly hydrogen bond to the carbonyl skeletal oxygen of valine M276. It is thus concluded that the protein attachment site variation shown in Figure 4 is the cause of the asymmetry between the results found for the LH(L131) and L*x*(M160) mutants.

In the oxidized reaction center P^+, charge is withdrawn from the 9-keto groups owing to their conjugation with the porphyrinic rings. The extent to which this effect relates to the two halves P_L and P_M is dependent on the degree of charge localization within the special pair. We have analyzed[44] the midpoint potentials of P for the L*x*(M160) mutants using a model which has as a parameter the oxidation potential of P_M. The deduced values for this quantity were found to correlate roughly with the ground-state hydrogen-bond energy between the 9-keto group and the introduced residue, estimated from the observed change in carbonyl frequency. On the basis of this correlation, for the hypothetical case in which all of the charge in P^+ is localized on P_M, hydrogen-bond energies to its 9-keto group would be expected to decrease by over 3 kcal mol^{-1}. Such a change would be sufficient to break the hydrogen bond in a competitive environment. In the available L*x*(M160) mutants,[43,44] most of the charge is in fact localized on P_L rather than P_M, and therefore rupture of the hydrogen bonds is unlikely. Alternatively, for the LH(L131) mutants in which the neutral species does form a hydrogen bond to the 9-keto group of P_L, the significant charge buildup on P_L in P^+, combined with the availability of hydrogen-bonding options other than to the 10a-ester, should be sufficient to guarantee the rupture of the hydrogen bond.

F. The Midpoint Potential of P for the Heterodimer (HL(M202)) Mutants. The study of the midpoint potentials for the oxidation of P to P^+ has produced valuable information[41,43,44] about the interaction between the two BChl molecules in the special pair and the extent of charge localization in the cation. For the heterodimer mutants, the cation is essentially fully charge-localized[17,27] on P_L, and these thus form a chemically simple subset. We have evaluated their ionization energies at the Koopmans level on the basis of the QM/MM wave functions using a new technique which is described in Appendix 1. These calculations assume that the effect of solvent relaxation is the same for all mutants, allowing the ionization energy to be estimated from only the electronic wave function of the neutral special pair. Variation of the ionization potential then

(75) El-Kabbani, O.; Chang, C. H.; Tiede, D.; Norris, J.; Schiffer, M. *Biochemistry* **1991**, *30*, 5361.

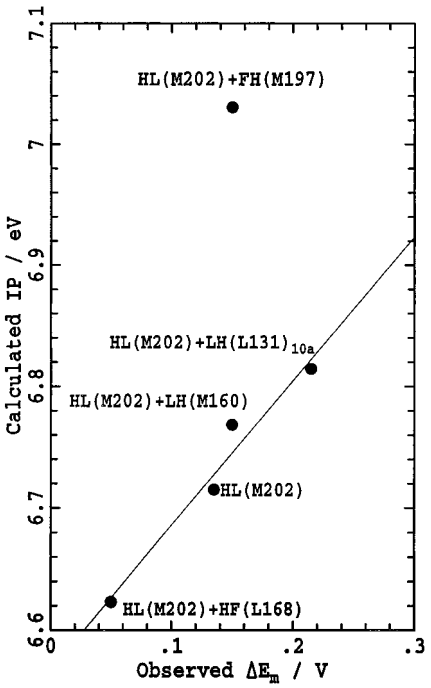

Figure 5. Plot of the QM/MM ionization potential (IP), evaluated at the Koopmans level, as a function of the observed[17,27] (see also ref 76) change ΔE_m in midpoint potential of P from the wild type, for various heterodimer mutants of *Rh. sphaeroides*.

arises purely from the changes in equilibrium geometry, changes in the external field due to the protein, and cofactor polarization effects induced by the changes in the external field.

The QM/MM calculated ionization potentials are shown in Figure 5 as a function of the observed[17,27] (see also ref 76) change in midpoint potential from the wild type, ΔE_m. Four of the five points on this graph fall on a straight line of slope 1.187; the RMS error in this fit is 14 mV. Naively, unit slope is expected, and the (small) deviation of the fitted slope from this value suggests that the most appropriate value for the high-frequency dielectric constant to be used in the QM/MM calculations is 2.4 rather than the value actually used, 2.00.

The midpoint potential for the fifth mutant, HL(M202)+FH(M197) lies quite dramatically some 0.28 eV above the expected line. For the FH(M197) mutants, we have already argued that in the oxidized form the 2a-acetyl group of P_M rotates away from the introduced histidine back to coordinate to the magnesium of P_L. This is a relaxation effect which is not anticipated in the other mutants; hence, a key assumption underlying the expectation of a linear relationship is not valid for this mutant. Actually, the value of the offset, 6.5 kcal mol^{-1}, can be interpreted simply as the relaxation energy associated with the 2a-acetyl rotation.[45,47,48] Elsewhere,[73] on the basis of high-level computations for the properties of the gas-phase BChl molecules combined with empirical estimates of the histidine to 2a-acetyl hydrogen-bond energy, we estimated this relaxation energy to be somewhat less, 3−4 kcal mol^{-1}. These two independent estimates both qualitatively indicate that there is a significant driving force for 2a-acetyl rotation on oxidation.

By implication, HL(M202) and its double mutants with HF(L168), LH(M160), and LH(L131) are interpreted as having in common all relaxation processes accompanying oxidation. It is not trivial that this should be so, and we discuss each case in turn. First, for HL(M202)+LH(M160), the additional hydrogen bond is added to the 9-keto group of P_M, but as little charge is delocalized onto P_M in P^+, hydrogen-bond rupture is not expected. Second, if the introduced hydrogen bond in HL(M202)+LH(L131) forms to the 9-keto group of P_L, then as significant charge is removed from this carbonyl in P^+, hydrogen-bond rupture is expected[44] and the associated reorganization energy would need be taken into account. However, we have already argued that the hydrogen bond in this mutant actually forms to the 10a-ester group (see structure HL(M202)+LH(L131)$_{10a}$); as charge is not removed from the 10a-ester group in P^+, this hydrogen bond is therefore expected to remain intact on oxidization. Last, we consider the double mutant with HF(L168) for which the primary issue is the nature of the orientation of the 2a-acetyl group of P_L in the neutral and oxidized forms of the wild-type and the mutants. We have argued that in the WT this group is directed outward toward the protein in both the neutral and oxidized reaction centers and that, for most HF(L168) mutants, the energetics of rotation for the neutral species are finely balanced, while for the oxidized species the large charge delocalized on P_M should guarantee the inward-pointing orientation. However, the heterodimer HL(M202)+HF(L168) mutant is exceptional in that the magnesium to which the 2a-acetyl can coordinate is absent. Hence, it is most likely from energetic considerations[44] that the 2a-acetyl still points outward to the protein in the neutral species, and as now *no* charge is delocalized onto P_M in the cation, no rotation is expected, and thus the same relaxation processes are expected for HL(M202)+HF(L168) as for the other mutants.

Overall, we see that the results indicate that direct effects of mutations, even those including the LH(M160) modifications to P_M, do influence the oxidation potential of P_L. This is in marked contrast to a key assumption commonly used in analyses[41,43,44] of the midpoint potentials P for LH(M160) mutants.

G. Orientation of the Tyrosine M210 Hydroxyl Group Deduced from Properties of Y*x*(M210) Mutants. In the Y*x*(M210) mutants, a tyrosine group, which lies in the hydrophobic region of the protein and contains a non-hydrogen-bonded phenolic OH group, is replaced with aprotic residues, see Figure 1. These mutants have been well studied both experimentally and theoretically and hence permit our method to be tested in terms of its ability to make both quantitative and qualitative predictions. We solve unresolved issues concerning the orientation of the OH group in the WT raised by the works of Gunner et al.,[77] Alden et al.,[46] and Apostolakis et al.,[47] advancing a plausible scenario to explain the available results and hence provide a consistent physical model. For this purpose, modified structures named WT$_{210}$ in the results tables, etc., are obtained in which the orientation alternate to that believed to be correct is investigated. Results are described in detail in Supporting Information.

H. Competitive versus Cooperative Hydrogen-Bonding Involving Histidine L168 in N*x*(L166) Mutants. To modify the effect of the hydrogen-bonding of histidine L168 to the 2a-acetyl group of P_L, the NH(L166) and NL(L166) mutants have been introduced.[52] As pictured in Figure 1, suitable residues at site L166 may form an additional hydrogen bond to histidine L168. In the wild type, a weak hydrogen-bonding ligand, asparagine, is present at L166, while in the mutants either a non-hydrogen-bonding residue, leucine (in NL(L166)), or a strong hydrogen-bonding one, histidine (in NH(L166)), is introduced. If cooperative effects are important, then the 2a-

(76) Laporte, L. L.; Palaniappan, V.; Davis, D. G.; Kirmaier, C.; Schenck, C. C.; Holten, D.; Bocian, D. F. *J. Phys. Chem.* **1996**, *100*, 17696.

(77) Gunner, M. R.; Nicholls, A.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 4277.

acetyl carbonyl frequency should decrease as the L166 to L168 hydrogen-bond strength increases: a hydrogen-bonding residue at L166 would polarize L168 and hence cooperatively increase the strength of the L168 to 2a-acetyl hydrogen bond. However, it was observed that small decreases of the frequency occurred for *both* mutants.

Our simulations indicate that the naive sketch of the structure of the NH(L166) mutant shown in Figure 1, which leads to the expectation that cooperative effects should control the frequency shifts, is highly misleading. The two residues, L166 and L168, are close together, and their mutual configurations are thus restricted, preventing the development of a linear hydrogen bond such as that sketched in the figure. Indeed, for NH(L166) no hydrogen bond between the introduced histidine at L166 and histidine L168 is produced. The hydrogen gets within 2.5 Å of its target nitrogen (see Table 6), but the angle between the NH vector of L166 and the vector which bisects the CNC angle is 51°, not the expected value of 180° for a linear hydrogen bond. Various attempts were made to find an alternate configuration in which a hydrogen bond is present; these led initially to the partial breakage of the L168 to P_L 2a-acetyl hydrogen bond before finally reverting to the original structure. The NH(L166) and NL(L166) mutants are thus found to be similar in that neither mutation leads to formation of a hydrogen bond between L166 and L168. Alternatively, the WT clearly does have such a hydrogen bond, with calculated bond lengths (see Tables 2−6) of ca. 1.9 Å and hydrogen-bond angles in excess of 140°.

The experimental observation that both mutants have slightly *lower* frequencies than the wild type shows that the cooperative model for frequency modulation is incomplete. Our simulations suggest a reason for this. For the 20 structures produced with a histidine in position L168, the average L168 to 2a-acetyl hydrogen-bond length from the results given in Tables 2−6 is 1.916 ± 0.034 Å, while for the NH(L166) and NL(L166) mutants this is 1.829 and 1.852 Å, respectively, values which, on average, are over two standard deviations shorter. The hydrogen-bond between L166 and L168 present in the wild type is thus seen to act *in competition* to the L168 to 2a-acetyl hydrogen bond, this again being necessitated by the close proximity of the L166 and L168 residues. Such competitive hydrogen-bonding would lead to the hydrogen-bonded species having higher vibration frequencies, as observed.

I. The Configuration of the Special Pair. Many of the functional properties of the reaction center are controlled by the intermolecular coupling between the two BChl molecules in the special pair, and our structures provide a systematic investigation of the effects of mutation on the dimer configuration. For all of the mutants studied, as well as the WT and that for *Rh. viridis*, key dimer properties obtained from both the optimized and X-ray structures are given in Table 8. These are the distances, R_{N-N}, between the centers of the four pyrrolic nitrogens, and R_{I-I}, between the centers of ring-I fragments, as well as the angles, α_N and α_I, between the planes of the pyrrolic nitrogens and the ring-Is, respectively. We deduce these angles by forming planes of best fit either to the four pyrrole nitrogens or to the five atoms of ring-I, reporting the angle between the normals of the corresponding planes of P_L and P_M.

Comparison of the optimized R_{N-N} distances with the X-ray ones show some variability arising from the flat nature of the intermolecular potential-energy surface with respect to slippage of the porphyrin planes, the differences being −0.06 Å for the WT of *Rb. sphaeroides*, −0.07 Å for HL(M202), 0.23 Å for FY(M197), and 0.72 Å larger for the WT of *Rh. viridis*. The X-ray inter ring-I separation is smaller for *Rh. viridis* than for *Rb. sphaeroides*, consistent with the perceived[78] larger coupling in this strain; this effect is reproduced in the optimized structures, but the inter-ring separations are considerably expanded. Further, the planes containing the pyrrole nitrogens, and the ring-I planes, are essentially parallel (α_N and $\alpha_I < 6°$ in most cases) in the X-ray structures but misalign somewhat in the optimized ones (typically $9° < \alpha_N, \alpha_I < 18°$). The misalignment and inter ring-I expansion are accentuated for the heterodimer mutants.

Table 8: Intermolecular Structure of the Special Pair in a Gas-Phase Dimer, *Rb. sphaeroides* WT and Its Mutants, and *Rh. viridis*[a]

mutant	R_{N-N}	R_{I-I}	α_N	α_I
BChl-*a* dimer B3LYP/3-21G	7.72	3.72	5	5
WT	7.66	3.97	10	13
WT X-ray	7.72	3.45	4	6
WT_{210}	7.66	3.97	10	13
LS(M160)	7.66	3.96	10	12
LN(M160)	7.67	3.96	10	12
LQ(M160)	7.66	3.96	10	12
LH(M160)	7.65	3.95	10	12
FY(M197)	7.98	4.05	13	10
FY(M197) X-ray	7.75	3.73	6	3
FH(M197)	8.10	4.15	11	7
FH(M197)+HF(L168)	8.11	4.15	11	8
FH(M197)+LH(M160)	8.08	4.14	10	7
FH(M197)+LH(M160)+LH(L131)	8.14	4.31	8	4
FH(M197)+LH(M160)+$LH(L131)_{10a}$	8.14	4.31	8	4
HL(M202)	7.55	4.08	13	16
HL(M202) X-ray	7.62	3.73	4	2
HL(M202)+LH(L131)	7.48	3.98	9	18
HL(M202)+$LH(L131)_{10a}$	7.49	4.00	10	18
HL(M202)+HF(L168)	7.46	3.94	10	17
HL(M202)+LH(M160)	7.37	3.91	9	18
HL(M202)+FH(M197)	8.01	4.33	15	9
LH(L131)	7.65	3.91	9	12
$LH(L131)_{10a}$	7.64	3.88	8	12
HF(L168)	7.59	3.84	12	11
$HF(L168)_{rot}$	7.59	3.64	9	6
HF(L168)+LH(L131)	7.60	3.87	11	11
LH(M160)+LH(L131)	7.65	3.94	10	11
NL(L166)	7.67	3.97	10	12
NH(L166)	7.68	3.97	10	12
YF(M210)	7.65	3.96	10	13
YW(M210)	7.53	3.88	11	13
Rh. viridis	8.09	3.90	14	10
Rh. viridis X-ray	7.37	3.34	11	5

[a] R_{N-N} is the separation of the centers of the pyrrolic nitrogens, in Å; R_{I-I} is the separation between the centers of the two ring-I's, in Å; α_N is the angle between the best-fit planes containing the four nitrogen atoms, in degrees; and α_I is the angle between the best-fit planes containing the five ring-I atoms, in degrees.

To investigate the causes of these changes we have performed AM1 and extensive ab initio and density functional calculations on the structure of the gas-phase BChl-*a* dimer and the ring-I interaction. Preliminary results[56,73] indicate that the dimer is held together by significant attractions between the magnesiums and the 2a-acetyl groups coupled with a weak ring-I-to-ring-I attraction. AM1, however, treats the ring-I-to-ring-I interaction as purely repulsive in nature. Our QM/MM calculations reflect the results of this internal tensioning, producing large values for R_{I-I} and nonparallel molecular planes, the effect being magnified for the heterodimer mutants in which only one magnesium to 2a-acetyl interaction is present. The net result is that the coupling between P_L and P_M is not well described at this level. However, this has marginal influence on the many protein−cofactor interactions described earlier. It is, in principle, possible to obtain more realistic intradimer geometries using QM/MM or related methods in which AM1 is replaced with a

(78) Reimers, J. R.; Hush, N. S. *Chem. Phys.* **1995**, *197*, 323.

higher-level method such as density functional theory. Results obtained for the BChl-*a* dimer, optimized using the B3LYP density functional[79] with the 3-21G basis are also shown in Table 8. These are much more consistent with the X-ray structures, although this method (and indeed all density functionals) have difficulties with the 2a-acetyl rotational potential[73] and hence are less suited to other aspects of the computation.

4. Conclusions

We have constructed representative structures for some 22 reaction centers of mutant strains of *Rh. sphaeroides* and used these to interpret a large range of experimental observations, drawing conclusions about the nature of the wild type and the effects of the introduced mutations. For the two mutants studied for which X-ray structural data is available, good overall agreement between the key qualitative features of the calculated and observed structures is found. All structures correlate well with observed carbonyl frequency changes caused by the mutations, and we find physical causes for unexpected observed frequencies. Also, we demonstrate that, for simple systems in which the location of the hole in P^+ is known a priori, these structures alone can be used to interpret changes in observed midpoint potentials, thus isolating important chemical effects. So as to facilitate other such investigations of the properties of the cofactors using electronic structure techniques, all optimized coordinates are provided in the Supporting Information. However, while our QM/MM technique is shown to provide highly self-consistent, realistic cofactor *intramolecular* geometries with low strain energies (unlike those obtained from X-ray structures[56]) and while here we show that it also provides qualitatively descriptive hydrogen-bonding and other structures, we find that the intermolecular special-pair dimer configurations produced for the special pair are subtly deformed, rendering them inappropriate for detailed intradimer coupling analyses.

For the wild type, we find that a specific solvation effect, the rotation of the phenolic OH of tyrosine M210 following oxidation of P to P^+, is important in determining the experimental midpoint potential. Also, despite it being known[47] that the 2a-acetyl group of P_M prefers to coordinate to the magnesium of P_L in P^+ irrespective of whether hydrogen-bonding to the protein is available, we conclude that the orientation of the 2a-acetyl of P_L in P^+ is variable: it is directed outward in the wild type, inward in most HF(L168) mutants, but outward in the heterodimer HL(M202)+HF(L168) mutant even though hydrogen-bonding with the protein is not available.

There has been much discussion about the structural asymmetry of the protein and its effects on controlling the photosynthetic process. Here, we show that small asymmetries in the location of the leucine residues adjacent to the 9-keto groups (M160 and L131) lead to qualitatively different chemical behavior on mutation to histidine. On the L side, the group is located between the 9-keto and 10a-acetyl groups, allowing for competitive hydrogen-bonding which is effectively controlled by rotation of the 10a-acetyl group. Alternatively, on the M side, the leucine is located only near the 9-keto group only, and hence L*x*(M160) mutants can form only one type of hydrogen bond with the BChl. This explains the irregular pattern of 9-keto vibration frequencies which has been observed for LH(L131) mutants.

The L*x*(M160) mutants present their own interesting structural anomalies, however, in that, as some of the introduced residues are much smaller than the original leucine, significant rearrangements in the protein backbone are required to form the observed hydrogen bond to the 9-keto group of P_M. While these changes may be larger than could reliably be modeled using our approach, they do indicate an intrinsic complexity for these mutants which is not considered in simple interpretations[41,43,44] of their midpoint potentials. Also, we are able to model the midpoint potentials for the HL(M202) heterodimer mutants satisfactorily on the assumptions that the charge in the cation radical is fully localized on P_L and that the main effect of mutations is to directly modify the local field experienced at P_L. This last assumption is also contrary to one which is used in modeling the L*x*(M160) special-pair midpoint potentials, and clearly a more sophisticated approach is required to provide a comprehensive description of both series of mutants.

Our results also provide explanations for anomalies seen for other mutations. The NH(L166) mutation was intended to form a strong hydrogen bond to histidine L168 but this is shown not to be realized due to the geometrical constraints of the nearby residues, while in the WT the effect of the weak hydrogen bond present is to pull L168 away from the 2a-acetyl of P_L and hence competitively reduce the strength of this hydrogen bond. Also, it is suggested that the anomalous midpoint potential of P for FY(M197) compared to FH(M197) arises both from the changed hydrogen-bond strength and from the introduction of an additional water near the special pair in P^+ facilitated by the extension of the hydrophilic region of the protein which this mutation provides.

Acknowledgment. We thank the Australian Research Council for funding this research, and Dr. Günter Fritzsch and Dr. Andreas Kuglstatter from the Max Planck Institute for Biophysics (Frankfurt) for kindly providing us with coordinates from the X-ray structure[11] of FY(M197).

Appendix 1

Evaluating QM/MM Ionization Potentials at the Koopmans Level. In a method based on Hartree−Fock self-consistent field (SCF) theory such as AM1, ionization energies may be estimated using Koopmans' theorem.[80] This states that the ionization energy, obtained from the energy of the ion obtained with the geometry and molecular orbitals of the neutral species, is given simply as the negative of the energy of the orbital from which oxidation occurs. Improved energies can be obtained by both relaxing the electronic wave function, allowing for a fully self-consistent description of the ion, and further by relaxation of the nuclear coordinates. In QM/MM methods such as the one employed in VAMP,[71] Koopmans' theorem does not hold, and the orbital eigenvalues have no simple interpretation. To evaluate ionization energies, it is thus customary to perform the full energy calculation for the radical cation. Such a calculation may proceed using either spin-unrestricted Hartree−Fock (UHF) or spin-restricted open-shell (ROHF) methodologies, but only the UHF procedure is available in VAMP. For porphyrinic systems such as BChl, UHF is inappropriate as it incorrectly depicts the ground state of the porphyrin as a ^3Q state, with the ground-state of the cation radical being ^4Q.

To calculate ionization energies at the QM/MM level, we thus developed a scheme which calculates the ionization energy at the Koopmans' level but which does not make use of orbital eigenvalues. To do this we evaluated the electronic density for the cation radical from the molecular orbitals of the neutral, and entered this into VAMP. The source code was modified to prevent update of this density, so that, as a result, the SCF energy

(79) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(80) Koopmans, T. A. *Physica* **1939**, *1*, 104.

after the second cycle is just the original SCF energy of the neutral molecule less the desired Koopmans-level ionization energy.

Appendix 2

The Stark Effect on the P_L 9-Keto Frequency in LH(L131) Mutants. Our structures indicate the existence of three distinct possible environments for the 9-keto group of P_L: the structures labeled LH(L131) in which the histidine introduced at residue L131 forms a hydrogen bond to the 9-keto group, the structures labeled $LH(L131)_{10a}$ in which this histidine preferentially hydrogen-bonds to the neighboring 10a-ester group, and structures for which the WT is characteristic that have no hydrogen-bonding involving the L131 residue. From Table 7 and Figure 2, the two structures assigned to the second category have observed P_L 9-keto vibration frequencies of 1673 and 1674 cm^{-1}, frequencies which fall between those with hydrogen-bonding to the 9-keto group (1657−1664 cm^{-1}) and those with no hydrogen-bonding possible (1688−1700 cm^{-1}). While the intermediate band appears at first inspection to be located closer to the lower (hydrogen-bonded) one, it should be noted that the two observed frequencies for the same mutant (LH(L131)) under different conditions are 1657 and 1673 cm^{-1}, and thus the frequency increase caused by the breaking of the direct hydrogen bond to the 9-keto group in an otherwise similar environment is significant, 16 cm^{-1}. From the alternate viewpoint, the members of the intermediate group have frequencies which are 14−27 cm^{-1} lower than the other non-hydrogen-bonded 9-keto frequencies.

This scenario can be rationalized by considering the Stark effect of the protein on the CO frequency. Using a high-frequency dielectric constant of $\epsilon = 2$ (as is used in the QM/MM calculations), the changes in the electric field in the direction of the CO bond from that of the WT caused by the introduction of the $LH(L131)_{10a}$ mutation and that caused by the subsequent formation of the hydrogen bond to the 9-keto group in the LH(L131) structure are 23 mV cm^{-1} (0.0044 au) and 34 mV cm^{-1} (0.0066 au), respectively. Qualitatively, this indicates that the $LH(L131)_{10a}$ structure should have a 9-keto vibration frequency intermediate between those of the WT and LH(L131) structures.

To obtain quantitative estimates of the Stark effects, we have evaluated[81] the change in the dipole moment on vibrational excitation, $\Delta\mu$, for acetone using the coupled-cluster[82] (CCSD and CCSD(T)) and coupled electron-pair approximation[83] (CEPA) techniques. The CEPA result obtained using the aug-cc-pVDZ basis set[84] is −0.024 D (−0.0096 au) compared to the progressively more accurate CCSD and CCSD(T) results of −0.024 D (−0.0094 au) and −0.022 D (−0.0086 au), respectively. Using the larger aug-cc-pVTZ basis set,[84] the CEPA value increases to −0.036 D (−0.014 au), however, and thus a best-possible estimate would be −0.033 D (−0.013 au). This value is similar to that found in related systems[85−87] and should be indicative of that expected for the BChl 9-keto stretch. Using it, progressive frequency shifts of −13 cm^{-1} and −19 cm^{-1} are predicted for the WT-to-$LH(L131)_{10a}$-to-LH(L131) modifications. These values are of the correct order, and hence the observed vibration frequencies are consistent with the structural model presented for the LH(L131) family of mutants.

Supporting Information Available: (a) The QM/MM optimization strategy, (b) comparison of X-ray and deduced structures for FY(M197), and (c) comparison of deduced structures for Y*x*(M210) mutants with those from previous simulations, deducing the hydroxyl orientation of tyrosine M210 (PDF). Optimized coordinates for a total of 30 structures depicting all 22 mutants studied as well as the analogously optimized structures for the wild type (*Rb. sphaeroides*) and also for the wild type of another strain, *Rh. viridis* (ASCII "hin" format).This material is available free of charge via the Internet at http://pubs.acs.org.

JA0035710

(81) Hush, N. S.; Reimers, J. R. *J. Phys. Chem.* **1995**, *99*, 15798.
(82) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
(83) Meyer, W. *Int. J. Quantum Chem. Symp.* **1971**, *5*, 341.
(84) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
(85) Reimers, J. R.; Zeng, J.; Hush, N. S. *J. Phys. Chem.* **1996**, *100*, 1498.
(86) Reimers, J. R.; Hush, N. S. *J. Phys. Chem. A* **1999**, *103*, 10580.
(87) Andrews, S. S.; Boxer, S. G. *J. Phys. Chem. A* **2000**, *104*, 11853.