Electrostatic Potential Derived Charges for Enzyme Cofactors: Methods, Correlations, and Scaling for Organic Cofactors

John E. Wampler

Department of Biochemistry, University of Georgia, Athens, Georgia 30602

The correlations between electrostatic potential (ESP) derived atomic charges for a wide range of different quantum mechanical approaches and basis sets have been investigated for both small and large organic structures including several enzyme cofactors. The previously observed linear correlation between ESP charges calculated by different approaches has been verified to extend to many different basis sets and procedures including effective core potential (ECP) basis sets, density functional theory (DFT) approaches, a hybrid Hartree−Fock (HF)/DFT approach, and inclusion of electron correlation corrections. Above a threshold level of complexity, most procedures and basis sets give results that correlate very well (linear correlation coefficients >0.99), including several procedures that have reasonable computational costs for large molecules. These procedures have been used to calculate ESP charges for five different types and forms of enzyme cofactors: biotin, pyridoxal-5′-phosphate, pyridoxamine-5′-phosphate, and the flavin mononucleotide in two different oxidation states.

INTRODUCTION

Electrostatic potentials calculated from atom centered point charges are increasingly important computations in their own right[1−6] and as one component of the nonbonded interactions for many molecular mechanics based calculations (for reviews see refs 7−9). With proteins and other biomolecules the electrostatic energy component in most molecular mechanics force fields is a major component of the total potential. Point charge parameters for these various calculations can be assigned from quantum mechanical calculation,[10−14] by analysis of electron density from X-ray crystal diffraction pattern analysis,[15] or by an empirical approach based on fitting observed fluid characteristics.[16] Limited recent data suggest that ESP charges from quantum mechanical calculations correlate well with those from the empirical approaches.[15,17]

The usual approach to evaluating atom-centered charges for molecular mechanics of organic molecules is to calculate a set of potentials in the surrounding space using a quantum mechanical approach and then to fit the atomic charges to give that set of potentials.[14,18] Thus, the charges are defined by the potentials that they must reproduce and are variously referred to as electrostatic potential (ESP) charges or molecular electrostatic potential (MEP) charges. It must be pointed out that the resulting "charges" are intended for one purpose, to predict electrostatic interaction potentials. They do not seem to be as good as other types of charges in evaluating chemical and bonding effects.[19] However, ESP charges do predict dipole moments and energetics that correlate well with experimental values.[13,17,19−24] Another advantage of ESP charges is that charge sets calculated with widely different quantum mechanical approaches have been found to be linearly correlated.[20,22,24] As a result, a fairly well accepted standard for ESP charges has been based on self-consistent-field (SCF) *ab initio* calculations with a 6-31G* basis set at the Hartree−Fock (HF) level. In cases where the size of the problem and available resources

preclude such an approach, charges from less rigorous calculations are then scaled to this level.

A number of more analytical approaches to atom-centered charges are based on defining each atom's component of the electronic charge distribution. The problem in defining point charges in this way lies in partitioning the electron density between bonded atoms. Even with an experimental approach, this partitioning is ambiguous.[15,25] With semiempirical and *ab initio* calculations several methods have been used including partitioning based on the occupancy of the atomic orbitals of the basis set,[10] on the occupancy of natural atomic orbitals as defined by Reed *et al.*,[26] and on integration of charge density within physical boundaries.[11,27,28] Finally, Cioslowski[29] (see also refs 19 and 30) has described a set of effective charges based on atomic polar tensors. In their recent analysis of many of these charge types, Wiberg and Rablen[19] again support the use of ESP charges for the purpose of representing electrostatic potentials indicating that no one type of charge is best for all purposes.

For molecular mechanics and electrostatic potential calculations of bioorganic molecules a number of charge sets have already been calculated and scaled to the HF/6-31G* standard.[14,21,24,31−34] This includes all 20 amino acid residues, the common nucleic acids, and many other biomolecules. However, similar charge sets for many of the common organic prosthetic groups are either not available or have been calculated using approaches where the correlation with and scaling to this standard level have not been investigated. The focus of this paper is first to investigate methods to define the level of complexity required for accurate evaluation of ESP charges for large organic enzyme cofactors. Since most of the literature evaluating methods and defining the *de facto* standard (HF/6-31G* level calculations) has used small inorganic and organic models with only a few heavy atoms, preliminary work using small and medium size organic structures has been used to test a wide range of procedures including both standard *ab initio* and density functional theory (DFT)[35,36] approaches and of basis sets including effective core potential (ECP) basis sets. Based

---
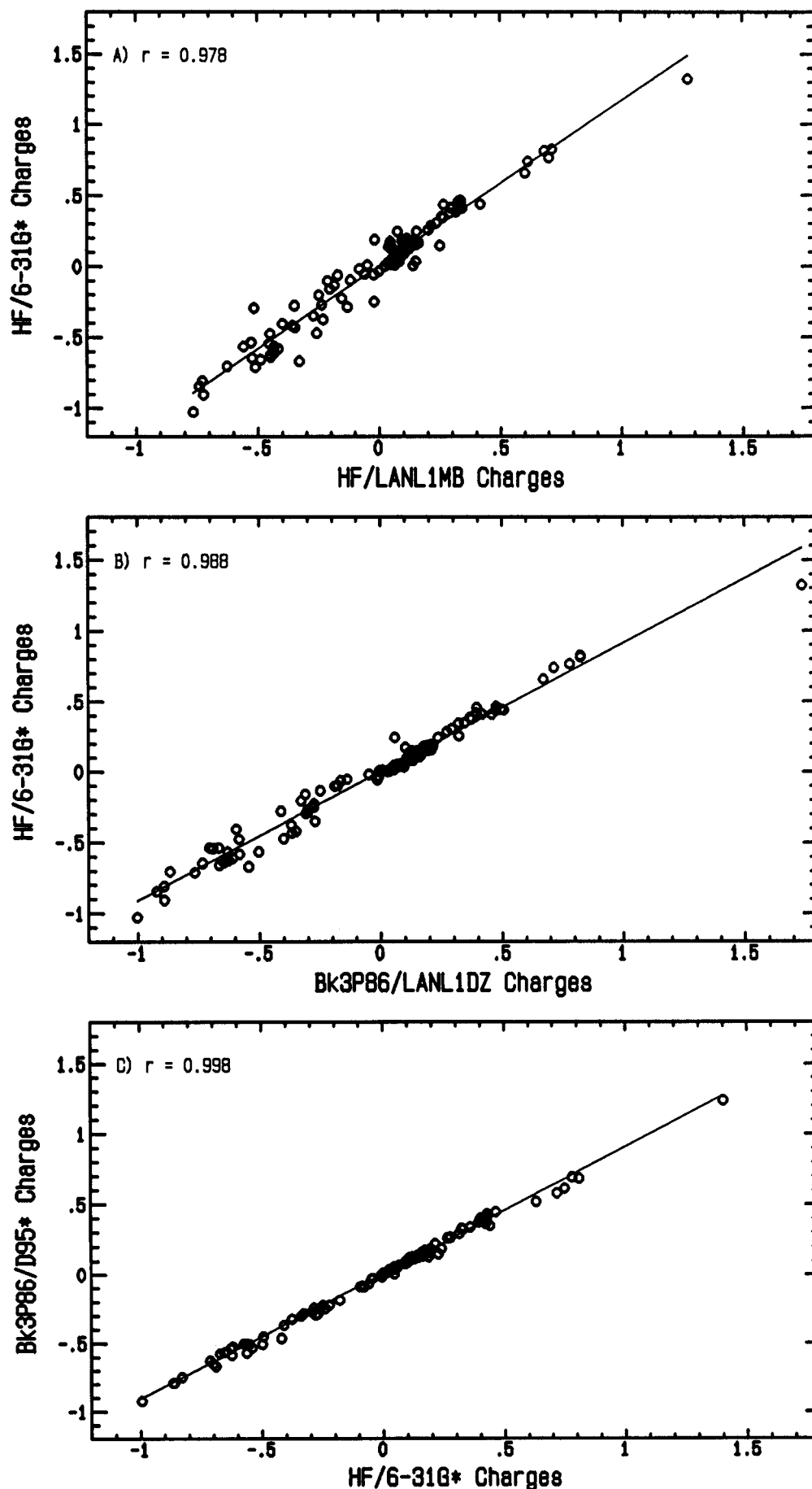
Figure 1. ESP charge comparisons for 123 atoms in nine test structures. Panel A, charges from HF/6-31G* calculations (y axis) were fit to charges from HF/LANL1MB calculations (x axis). Panel B, HF/6-31G* charges versus Bk3P86/LANL1DZ charges. Panel C, Bk3P86/ D95* charges versus HF/6-31G* charges. In all panels the calculated values are given by the circle symbols, the linear least squares fit given by the straight line, and the linear correlation coefficient of the fit is given in the upper left corner (r value).

Table 1. Grouped Data Comparison for Various Approaches To Calculate ESP Charges in Order of Increasing Precision[a]

method	total, energies	total, times	dipole moment[b] differences		linear fit[c] versus HF/6-31G*			linear fit[c] versus Bk3P86/D95*		
			⟨Δμ⟩	RMS	SD of fit	r-corl.	slope ± SD	SD of fit	r-corl.	slope ± SD
HF/STO-3G	−3299	229	0.823	0.484	0.0862	0.9753	1.124 ± 0.023	0.0781	0.9753	1.017 ± 0.021
HF/LANL1MB	−2585[d]	246	0.914	0.529	0.0815	0.9780	1.169 ± 0.022	0.0780	0.9754	1.054 ± 0.022
HF/CED-121G	−495[d]	3168	0.810	0.535	0.0643	0.9864	0.824 ± 0.012	0.0728	0.9786	0.739 ± 0.014
HF/LANL1DZ	−2618[d]	3607	0.860	0.604	0.0631	0.9870	0.812 ± 0.012	0.0690	0.9808	0.730 ± 0.013
MP2/STO-3G	−3302	837	0.997	0.596	0.0904	0.9729	1.236 ± 0.027	0.0768	0.9761	1.122 ± 0.023
MP2/LANL1DZ	−2182[d]	19833	0.487	0.401	0.0571	0.9893	0.899 ± 0.012	0.0578	0.9866	0.810 ± 0.012
Xα/LANL1DZ	−2600[d]	40842	0.422	0.133	0.0865	0.9752	0.929 ± 0.019	0.0563	0.9873	0.851 ± 0.012
HF/3-21G	−3322	914	0.516	0.459	0.0494	0.9920	0.882 ± 0.010	0.0531	0.9886	0.795 ± 0.011
HF/D95*	−3341	17431	0.405	0.334	0.0351	0.9960	0.984 ± 0.008	0.0504	0.9898	0.885 ± 0.012
BLYP/LANL1DZ	−2634[d]	66992	0.314	0.119	0.0690	0.9843	0.957 ± 0.016	0.0463	0.9914	0.872 ± 0.010
Bk3P86/3-21G	−3348	11123	0.209	0.119	0.0645	0.9863	1.014 ± 0.015	0.0440	0.9922	0.923 ± 0.010
MP2/3-21G	−3327	7272	0.312	0.210	0.0476	0.9926	1.024 ± 0.011	0.0438	0.9923	0.926 ± 0.010
Bk3P86/LANZ1DZ	−2642[d]	21482	0.493	0.264	0.0603	0.9880	0.912 ± 0.013	0.0416	0.9931	0.829 ± 0.009
MP2/D95*	−3346	115003	0.112	0.085	0.0322	0.9966	1.082 ± 0.008	0.0367	0.9946	0.977 ± 0.009
Xα/D95*			0.090	0.035	0.0718	0.9795	1.108 ± 0.022	0.0323	0.9948	1.004 ± 0.010
HF/6-31G*	−3341	9249	0.398	0.329				0.0339	0.9954	0.901 ± 0.007
HF/6-311G**	−3342	15956	0.417	0.344	0.0272	0.9976	0.978 ± 0.006	0.0332	0.9956	0.883 ± 0.007
HF/6-31G**	−3341	7334	0.397	0.322	0.0624	0.9999	1.002 ± 0.001	0.0313	0.9961	0.903 ± 0.007
HF/6-311+G**	−3342	82186	0.397	0.244	0.0252	0.9979	1.002 ± 0.006	0.0248	0.9975	0.906 ± 0.005
Bk3LYP/6-31G**	−3358	27931	0.121	0.048	0.0302	0.9970	1.164 ± 0.008	0.0233	0.9978	1.054 ± 0.006
MP2/6-311G**	−3351	135036	0.150	0.114	0.0353	0.9959	1.086 ± 0.009	0.0190	0.9986	0.985 ± 0.005
BLYP/D95* [e]			0.128	0.040	0.0433	0.9926	1.183 ± 0.014	0.0100	0.9995	1.063 ± 0.003
Bk3P86/D95*	−3367	182847			0.0375	0.9954	1.100 ± 0.009			

[a] Order based on the linear correlation coefficient for the fit to the ESP Charges from the Becke3P86/D95* calculation, the lowest collective energy approach. [b] Statistics presented are the mean vector difference, ⟨Δμ⟩, and RMS vector difference relative to the Bk3P86/D95* values. [c] Linear fit of charges of the 123 atoms of the nine test models, i.e., $y = ax$ where y is either the Becke3P86/D95* charges or the 6-31G* charges, x is the set being compared, and a is the slope of the fit line. The statistics reported are the standard deviation of the data from the fit line, the linear correlation coefficient, and the standard deviation of a. [d] ECP energies to not correlate with those calculated by other methods. [e] Calculation of ESP Charges using this method could not be completed for one of the structures (phospho-2-aminoethanol). Fit statistics are given for the other eight structures (107 atoms) for comparison.

on these tests, selected methods have been applied to evaluate ESP charges for the vitamin B_6 cofactors, pyridoxal-5′-phosphate, and pyridoxamine-5′-phosphate; the vitamin biotin, precursor to the CO_2 binding biotinyl cofactor; and the widely studied redox cofactor, flavin mononucleotide (FMN) in its oxidized and half-reduced forms. An accompanying paper (manuscript in preparation) extends this work to consider transition metal cofactors.

METHODS

Ab initio quantum mechanical calculations were performed using the widely available Gaussian-92/DFT[37] program (Gaussian, Inc., Pittsburgh, PA). For general reference to the methods see Hehre et al.[38] and the program user manuals available from Gaussian Inc.[39,40] Except where noted, program defaults were used.

Basis Sets. The basis sets employed included seven full, all electron basis sets:[38] the minimal STO-3G set, the double-split valence 3-21G, the larger double-split valence 6-31G* with polarization orbitals for the heavy atoms, 6-31G** with polarization for all atoms, the full double-ζ Dunning/Huzinaga set with polarization for heavy atoms[41−43] (referred to as D95*), the triple-split valence 6-311G** with all atom polarization, and the 6-311+G** with added diffuse functions. Note that the 6-311G** basis set[44] has been optimized for use with second-order Møller−Plesset electron correlation (MP2).[45]

In addition to these full basis sets, three effective core potential (ECP) sets were also used. These were the Hay and Wadt[46−48] LANL1MB set combining STO-3G for first row atoms and an ECP form for heavier atoms up to Bi, the LANL1DZ set combining a Dunning/Huzinaga double-ζ valence model (D95V) for first row atoms with the ECP

model for the heavier atoms, and the CEP-121G triple-split basis set of Stevens et al.[49]

Computational Approaches. Calculations were performed using both Hartree−Fock (HF) methods and density functional theory (DFT) methods. For DFT, the Slater (HFS),[50] Xα,[50] and Becke (B)[51] DFT exchange functionals were used.

Electron correlation corrections for the Hartree−Fock calculations were carried out using the Møller−Plesset perturbation theory (MP2).[52,53] In addition, the DFT correlation functionals LYP[54] and P86[55] were used. Becke's hybrid HF/DFT method[56] with his three-parameters functional combined with LYP and Perdew corrections was also employed (referred to as Bk3P86 and Bk3LYP in this paper). After preliminary tests with a wide range of combinations of functionals, the procedures and the basis sets that were employed in the full study were as follows: MP2 correlation with STO-3G, 3-21G, D95*, 6-311G**, and LANL1DZ basis sets; Xα, BLYP, and Bk3P86 DFT all with LANL1DZ and D95* basis sets. The Bk3P86 functional was also used with the 3-21G basis set, and the Bk3LYP functional was used with the 6-31G** basis set.

Single point SCF direct calculations were carried out on a cluster of IBM RISC workstations and a single Silicon Graphics Indy workstation. In some cases, quadratic convergence SCF[57] was used or the virtual orbital energies were shifted by 0.200 hartrees (VSHIFT = 200) in order to obtain convergence. Timing data were scaled to a common machine type (IBM RS/6000, Model 560) without regard for subtle differences in the calculations.

Models. To investigate the issues outlined in the introduction above, nine small organic molecules (eight derived
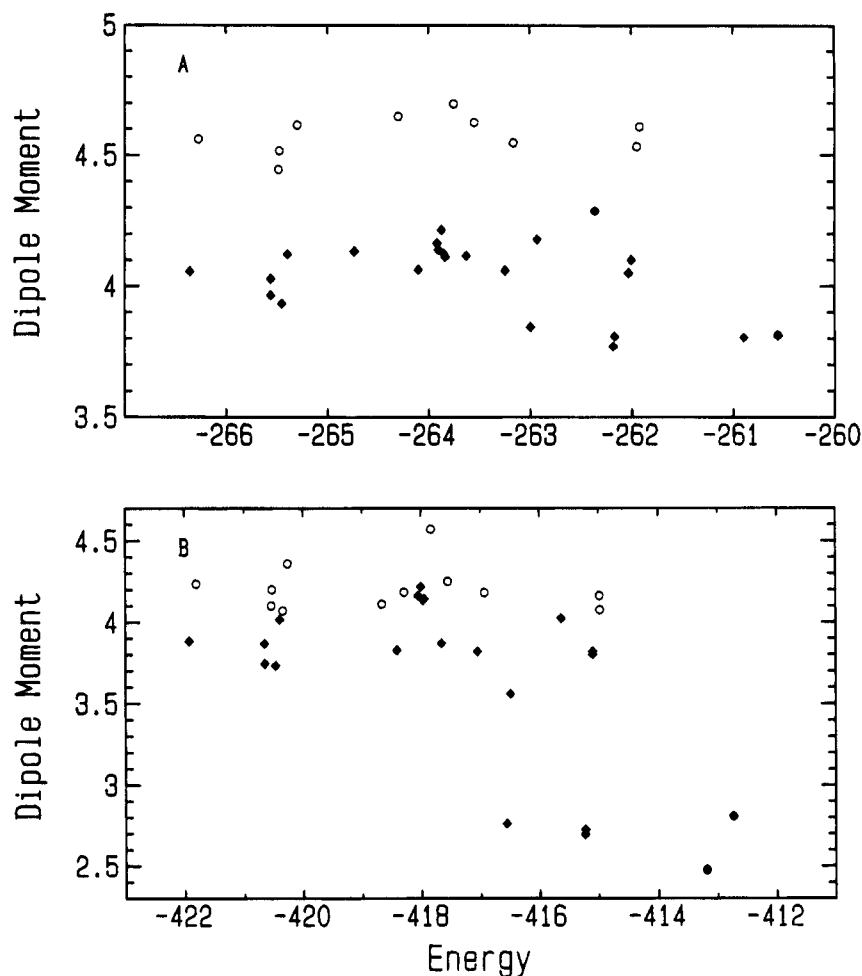
Figure 2. Dipole moments calculated from the ESP charges versus SCF energy. Panel A, dipole moments and energies for HF and DFT calculations with and without electron correlation corrections for 4-methylimidazole. Panel B, dipole moments and energies for the various procedures for L-propylamide. In both panels the filled diamond symbols are for full basis set calculations, and the open circles are for ECP basis sets.

from amino acids and amino acid side chain structures and one from the Cambridge structure database) were used which contained a variety of functional groups. They also included aromaticity (represented by p-methylphenol) and both saturated and aromatic heterocyclic compounds (represented by prolylacetamide and 4-methylimidazole). This group of structures was chosen to be of comparable size and complexity to those used in other studies but more specifically homologous to the atom types and functional groups common to organic enzyme cofactors. The complete list is as follows: 2-aminophosphoethanol, ethanol, formamide, 4-methylimidazole, p-methylphenol, methyl ethyl thioether, n-acetylglycineamide, L-prolylacetamide, and propionic acid amide. The elements in these structures, C, N, H, O, S, and P, include all of those common to the organic enzyme prosthetic groups and cofactors used by biochemical systems. The geometries and conformations for these test compounds were taken from the sybyl 6.0 amino acid database (Tripos Associates, Inc., St. Louis, MO) with unwanted atoms deleted and hydrogens added using the Sybyl "add-atom" function. Care was taken to assure that the geometry of the heavy atoms was not changed by any of the modeling operations. The structure of 2-aminophosphoethanol was obtained from the Cambridge Structure database[58] based on the structure reported by Weber et al.[59] No geometry optimization was performed for any of these structures.

Structures for the organic cofactors were obtained from the Brookhaven Protein Data Bank (PDB) files.[60,61] As with the test molecules, these structures were not modified (see exception below) or optimized except for the addition of hydrogen atoms using the SYBYL6.0 molecular graphics program. The structures used, PDB file names, and the primary references to each are as follows: the oxidized and semiquinone forms of flavin mononucleotide (FMN) from *Clostridium* MP flavodoxin,[62] files "3fxn" and "4fxn"; a second oxidized FMN structure was also taken from the *Chondrus crispus* flavodoxin structure,[63] file "2fcr"; pyridoxal 5-phosphate (PLP) structures were taken from rabbit muscle glycogen phosphorylase[64,65] (files "1gpa" & "1gpb"), *Salmonella typhimurium* tryptophan synthase[66] (file "1wsy"), and *Escherichia coli* aspartate aminotransferase[67] (file "3aat"); pyridoxamine-5'-phosphate (PMP) from *Eschericia coli* aspartate aminotransferase[68] (file "2aat"); and biotin from the streptavidin/biotin complex structure,[69] file "1stp". In the case of PLP, the covalently attached cofactor (pyridoxyl-5'-phosphate aldimine) was removed from the protein preserving the bond direction and angle of the connection but replacing it with a carbonyl oxygen using the standard SYBYL C=O bond length.

Charges. Both Mulliken[10] and ESP charges[14,21] were determined using functions built into Gaussian 92/DFT. Dipole moments calculated from both types of charges were also taken from the program's output.

Table 2. Statistics for Enzyme Cofactor Calculations

compd[a]	PDB file[b]	method	basic func	primitives	energy	dipole moment	time (s)
biotin	1stp	HF/STO-3G	99	297	−1104.438 0	21.898 8	387
biotin	1stp	HF/3-21G	178	297	−1112.126 4	21.235 5	966
biotin	1stp	MP2/3-21G	178	297	−1113.669 2	21.038 1	22608
biotin	1stp	HF/6-31G*	274	532	−1118.119 1	21.171 8	5935
biotin	1stp	HF/6-31G**	319	577	−1118.151 4	21.127 1	8352
biotin	1stp	HF/D95*	294	552	−1118.257 2	21.230 6	12477
biotin	1stp	HF/6-311G**	386	630	−1118.331 7	21.018 5	9522
biotin	1stp	Bk3P86/3-21G	178	297	−1119.636 4	20.636 0	63344
FMN	3fxn	HF/STO-3G	178	534	−1861.967 9	58.745 2	3025
FMN	3fxn	HF/3-21	321	534	−1876.290 0	48.398 0	6885
FMN	3fxn	HF/6-31G*	507	968	−1886.749 7	47.806 3	40798
FMN	3fxn	HF/6-31G**	564	1025	−1886.798 1	47.731 1	40152
FMN	3fxn	HF/6-311G**	680	1127	−1887.173 9	47.698 7	114563
FMN	2fcr	HF/6-31G**	564	1025	−1886.919 4	51.043 4	43253
FMNH	4fxn	HF/STO-3G	179	537	−1863.346 1	17.400 8	4131
FMNH	4fxn	HF/3-21G	323	537	−1877.092 3	57.400 5	7361
FMNH	4fxn	HF/6-31G*	509	972	−1887.551 7	56.024 1	34360
FMNH	4fxn	HF/6-31G**	569	1032	−1887.600 5	55.823 3	54715
PLP	1gpb	HF/STO-3G	92	276	−1136.524 0	25.435 6	502
PLP	1gpb	HF/3-21G	164	276	−1145.270 0	26.551 7	1091
PLP	1gpb	MP2/3-21G	164	276	−1146.892 9	26.541 8	7960
PLP	1gpb	HF/6-31G*	260	504	−1151.636 3	25.992 9	6121
PLP	1gpb	HF/6-31G**	284	528	−1151.654 4	25.959 9	7638
PLP	1gpb	HF/D95*	280	524	−1151.792 2	25.863 4	13570
PLP	1gpb	HF/6-311G**	344	574	−1151.872 3	25.853 7	18170
PLP	3aat	HF/6-31G**	284	528	−1151.816 5	28.790 7	7799
PLP	1gpa	HF/6-31G**	284	528	−1151.770 8	28.902 7	8306
PLP	1wsy	HF/6-31G**	284	528	−1151.840 8	27.148 4	8369
PMP	2aat	HF/STO-3G	96	288	−1119.028 6	31.318 9	780
PMP	2aat	HF/3-21G	172	288	−1127.502 8	33.571 9	1372
PMP	2aat	MP2/3-21G	172	288	−1129.080 5	33.015 1	17591
PMP	2aat	HF/6-31G*	268	520	−1133.729 1	33.277 2	8081
PMP	2aat	HF/6-31G**	304	556	−1133.757 8	33.244 5	10425
PMP	2aat	HF/D95*	288	540	−1133.874 8	33.211 4	17431
PMP	2aat	HF/6-311G**	368	606	−1133.965 6	33.304 7	21736

[a] Structures for biotin, FMN (flavin mononucleotide), FMNH (flavin mononucleotide semiquinone), PLP (pyridoxal-5′-phosphate), and PMP (pyridoxamine-5′-phosphate) were taken from the enzyme/cofactor X-ray structures as described in the text. [b] The file name are for X-ray crystal structure files from the Brookhaven Protein Data Bank (see references in text).

The final ESP charge sets for the cofactor structure models were calculated using ESP charges from HF/6-31G** calculations; The charge sets were rounded to three decimal places and charges were adjusted proportionally (largest charges getting largest adjustments) if necessary to obtain the appropriate formal charge. The charges on equivalent atoms of symmetrical groups (e.g., hydrogens on methyl groups) were averaged and any small adjustments (0.001 charge) needed to maintain formal charge were compensated on the atom to which they were attached. Thus, group charges were kept constant.
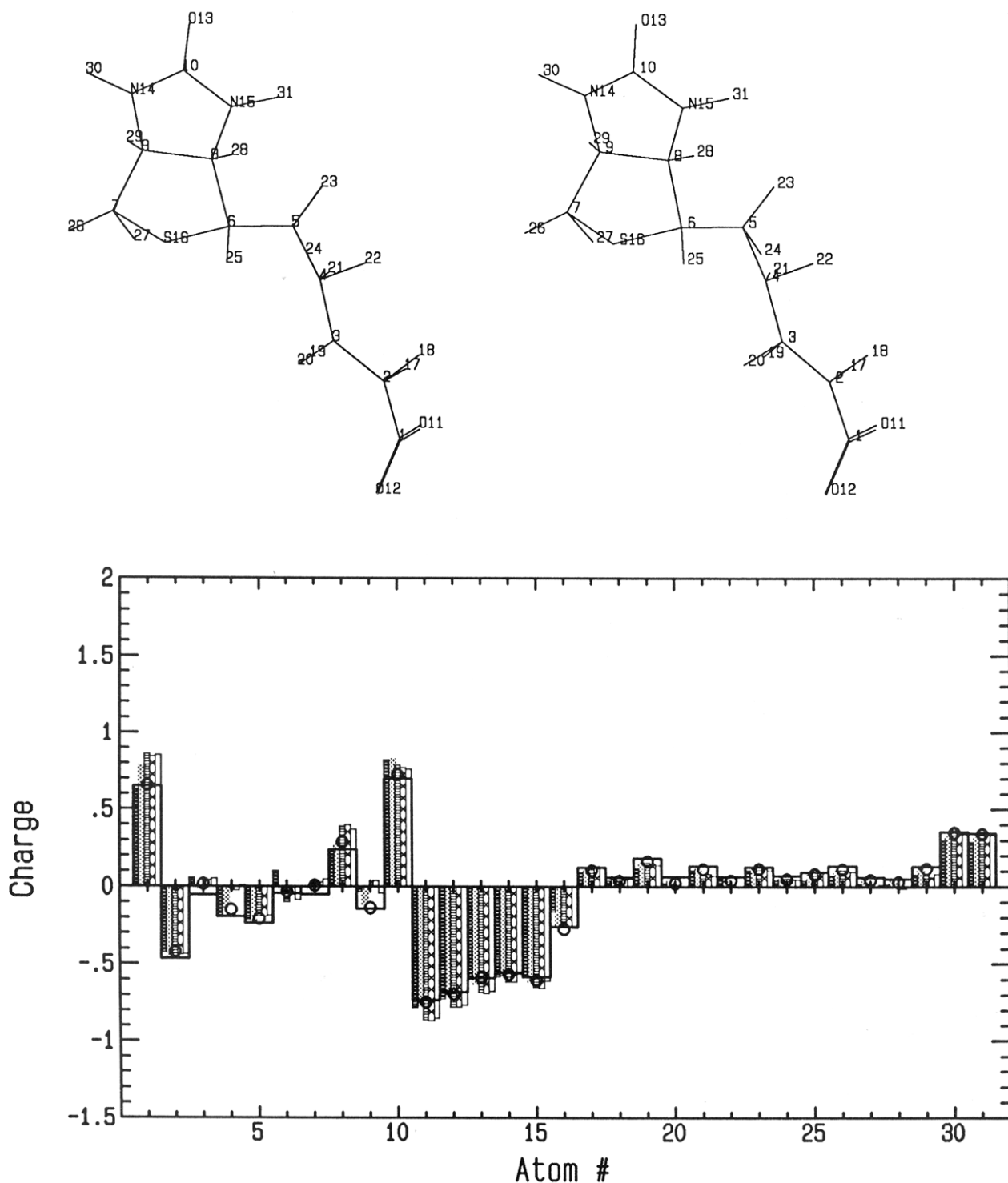
RESULTS

Correlation and Scaling. With small organic molecules, ESP charges derived with and without electron correlation are linearly correlated (Table 1) with all of the basis sets used. However, the quality of the correlation varies as indicated by the correlation coefficients and standard deviations (see Table 1). The two correlations reported are for the procedure with the lowest net energy (Bk3P86/D95*) and the HF/6-31G* approach which has become a *de facto* standard. A better understanding of the quality of these correlations comes from examining the plots of the fit lines and data points (Figure 1). These graphs show the atom-by-atom comparisons of ESP charges for all 123 atoms of the nine test structures. The three fits shown in the figure cover a range of values of linear correlation coefficient and are shown to put the numbers in Table 1 (and Tables 3 and

7) in perspective. Note: Mulliken charges for these molecules were more poorly correlated in all cases, and the scaling factors varied over a wider range (data not shown) in complete agreement with previous studies.[13,17,20,22]

In this comparison (Table 1), the total of energies by the various methods can be used as a rough gauge of their quality (except in comparisons between ECP and full basis sets), and the total computer time can be used to gauge the cost (and/or slow convergence) of the computations. The dipole moment statistics for the nine structures are a coarse indicator of variations in charge distribution between the various methods and the values from the Bk3P86/D95* calculation. Graphics examination (example in Figure 2) of the dipole moment data for each structure reveals that their magnitudes tend to converge with the higher level calculations as has been previously observed.[20,22,23] With the ECP basis sets, convergence was also found but at a different value of the dipole moment. Note, in both cases the "higher-level" calculations are to the left in Figure 2.
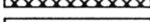
Electron correlation using the second order Møller−Plesset perturbation and DFT approaches improved the energy and made small differences in the charge distribution. The magnitude of the dipole moment is typically slightly lower compared to HF level calculations with the same basis set as previously observed by others.[22,23] Indeed with the full basis sets the individual charges for the atoms of the test molecules correlated very well with scaling factors near unity for all of the lowest energy methods with or without
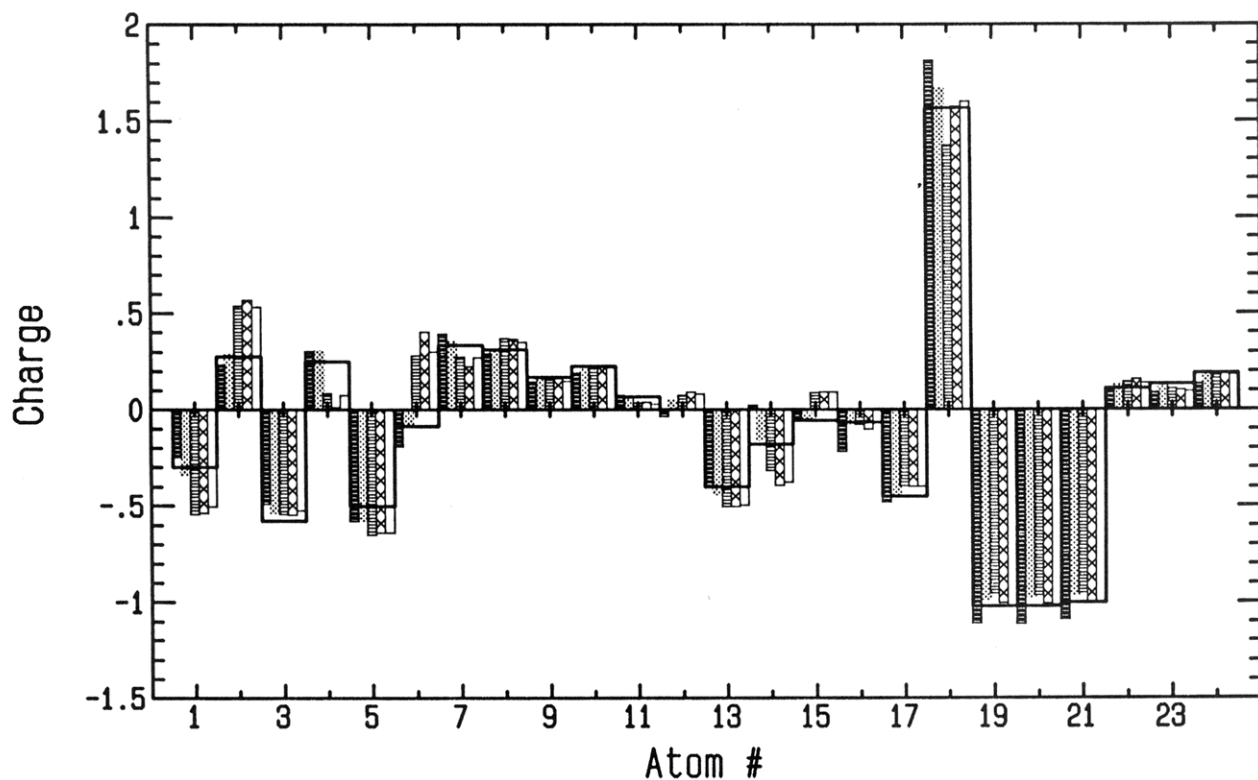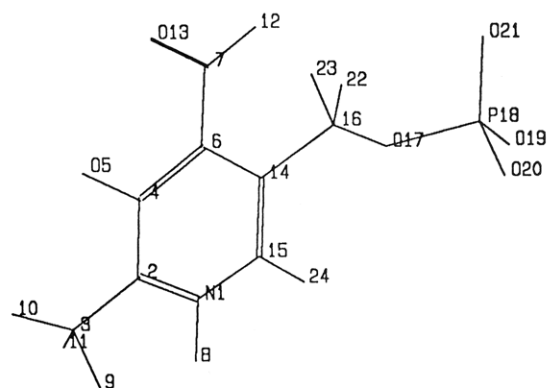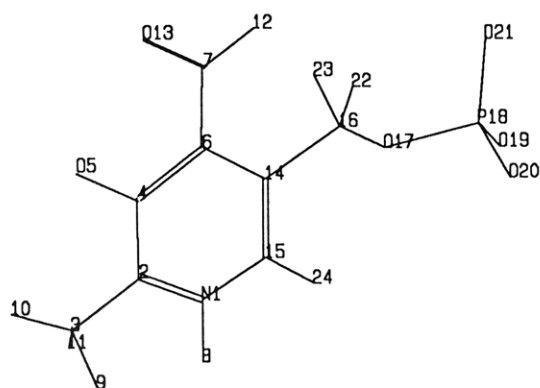
Figure 3. ESP charges for the atoms of biotin (stereodiagram) from different *ab initio* procedures after scaling. The HF calculation results are shown using narrow width bars on the graph (solid bar = HF/STO-3G; shaded = HF/3-21G; horizontal hatching = HF-6-31G**; cross hatching = HF/D95*; unfilled = HF/6-311G**). The DFT/HF hybrid results (Bk3P86/3-21G) are shown with the full width, unfilled bar and the MP2/3-21G results by a filled circle symbol. In all cases the charges were scaled to be comparable to HF/6-31G* charges using the factors of Table 1.

correlation corrections (see Table 1). With the double-ζ ECP basis set, the linear correlation to higher level HF and DFT

charge sets was also good when electron correlation corrections were included, but the scaling factors were uniformly
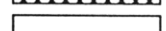
Figure 4. ESP charges for the atoms of pyridoxal-5′-phosphate (stereodiagram) from different *ab initio* procedures after scaling. The model used was abstracted from the X-ray crystal structure of aspartate amino transferase (PDB file 3aat). The HF calculation results are shown using narrow width bars on the graph (solid bar = HF/STO-3G; shaded = HF/3-21G; horizontal hatching = HF/6-31G**; cross hatching = HF/D95*; unfilled = HF/6-311G**). The MP2/3-21G result are shown with the full width bars. In all cases the charges were scaled to be comparable to HF/6-31G* charges using the factors of Table 1.

lower, consistent with the other comparisons between ECP and full basis sets.

In Table 1, the division between the higher quality procedures and the rest is indicated by the dividing blank
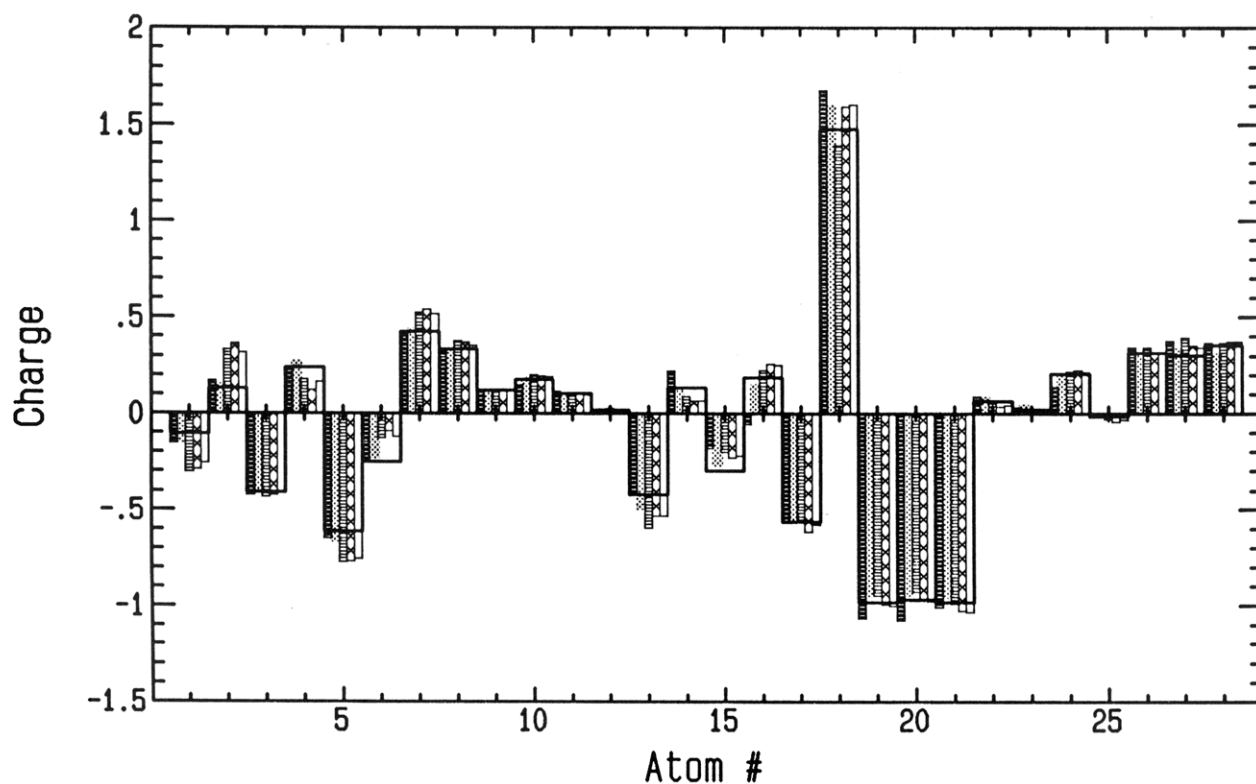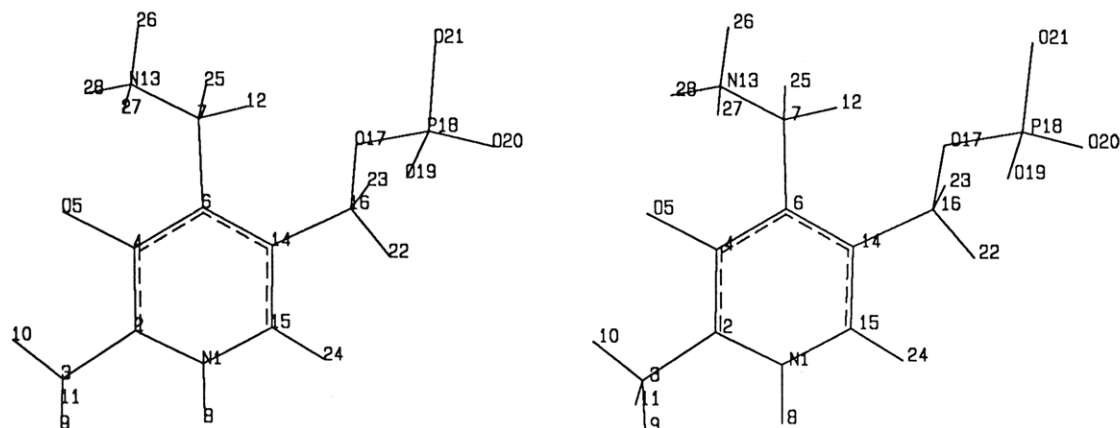
Figure 5. ESP charges for the atoms of pyridoxamine-5′-phosphate (stereodiagram) from different *ab initio* procedures after scaling. The model used was abstracted from the X-ray crystal structure of aspartate amino transferase (PDB file 2aat). The HF calculation results are shown using narrow width bars on the graph (solid bar = HF/STO-3G; shaded = HF/3-21G; horizontal hatching = HF-6-31G**; cross hatching = HF/D95*; unfilled = HF/6-311G**). The MP2/3-21G results are shown with the full width bars. In all cases the charges were scaled to be comparable to HF/6-31G* charges using the factors of Table 1.

line which separates procedures having a correlation coefficient greater than 0.99 and SD values less than 0.04 (lower half of table). This break-point is roughly comparable to the obvious devision between the results of the dipole

Table 3. Comparison of Linear Correlations[a] between ESP Charges for the Enzyme Cofactors and the Test Set of Small Organic Molecules

method	model	atoms	SD	r_{ab}	f_{ab}
HF/STO-3G	biotin	31	0.0771	0.9818	1.238 ± 0.044
	FMN	50	0.1455	0.9642	0.992 ± 0.039
	FMNH	51	0.1396	0.9648	1.078 ± 0.042
	PLP	24	0.1818	0.9446	0.955 ± 0.071
	PMP	28	0.1102	0.9784	1.047 ± 0.043
	test set	123	0.0863	0.9753	1.124 ± 0.023
HF/3-21G	biotin	31	0.0514	0.9920	0.938 ± 0.022
	FMN	50	0.0544	0.9951	0.873 ± 0.012
	FMNH	51	0.0671	0.9920	0.860 ± 0.016
	PLP	24	0.1362	0.9693	0.830 ± 0.045
	PMP	28	0.0823	0.9880	0.876 ± 0.027
	test set	123	0.0494	0.9920	0.882 ± 0.010
HF/D95*	biotin	31	0.0162	0.9992	0.990 ± 0.007
	PLP	24	0.0442	0.9968	0.915 ± 0.016
	PMP	28	0.0368	0.9976	0.933 ± 0.013
	test set	123	0.0351	0.9960	0.984 ± 0.008
HF/6-31G**	biotin	31	0.0051	0.999	1.002 ± 0.002
	FMN	50	0.0043	1.0000	1.001 ± 0.001
	FMNH	51	0.0116	0.9998	1.006 ± 0.003
	PLP	24	0.0022	1.0000	1.001 ± 0.001
	PMP	28	0.0077	0.9999	1.001 ± 0.003
	test set	123	0.0062	0.9999	1.002 ± 0.001
HF/6-311G**	biotin	31	0.0196	0.9988	1.008 ± 0.009
	FMN	50	0.0407	0.9972	0.945 ± 0.010
	PLP	24	0.0413	0.9972	0.914 ± 0.014
	PMP	28	0.0404	0.9971	0.932 ± 0.014
	test set	123	0.0272	0.9976	0.978 ± 0.006
MP2/3-21G	biotin	31	0.0570	0.9901	1.142 ± 0.030
	PLP	24	0.1363	0.9693	0.978 ± 0.053
	PMP	28	0.0861	0.9869	1.036 ± 0.033
	test set	123	0.0475	0.9926	1.024 ± 0.011
Bk3P86/3-21G	biotin	31	0.0748	0.9829	1.129 ± 0.039
	test set	123	0.0645	0.9863	1.014 ± 0.015

[a] Linear fits of charges, atom-by-atom, i.e., $y = f_{ab}x$ where y is the HF/6-31G* charge set, x is the set being compared, and f_{ab} is the slope of the fit line. The statistics reported are the standard deviation of the data from the fit line, SD, the linear correlation coefficient, r_{ab}, and the standard deviation of the slope.

Table 4. Biotin ESP Charge Set[a]

atom[b]	charge	atom	charge	atom[b]	charge
1	0.854	11	-0.823	21	0.039
2	-0.451	12	-0.823	22	0.039
3	-0.002	13	-0.684	23	0.078
4	-0.008	14	-0.619	24	0.078
5	-0.193	15	-0.653	25	0.097
6	-0.100	16	-0.260	26	0.102
7	0.022	17	0.091	27	0.010
8	0.387	18	0.091	28	-0.004
9	-0.007	19	0.080	29	0.074
10	0.785	20	0.080	30	0.360
				31	0.350

total charge = -1.000

[a] HF/6-31G** based ESP charges were derived by rounding values to three decimal places, adjusting the total charge to -1.000 by modifying the largest magnitude values (adjustments of -0.001) and averaging charges on identical atoms of symmetrical groups, but maintaining constant group charge. [b] Atoms identified by number in structure from Figure 3.

Table 5. ESP Charge Sets[a] for Pyridoxal-5'-phosphate and Pyridoxamine-5'-phosphate[b] from HF/6-31G** Calculations

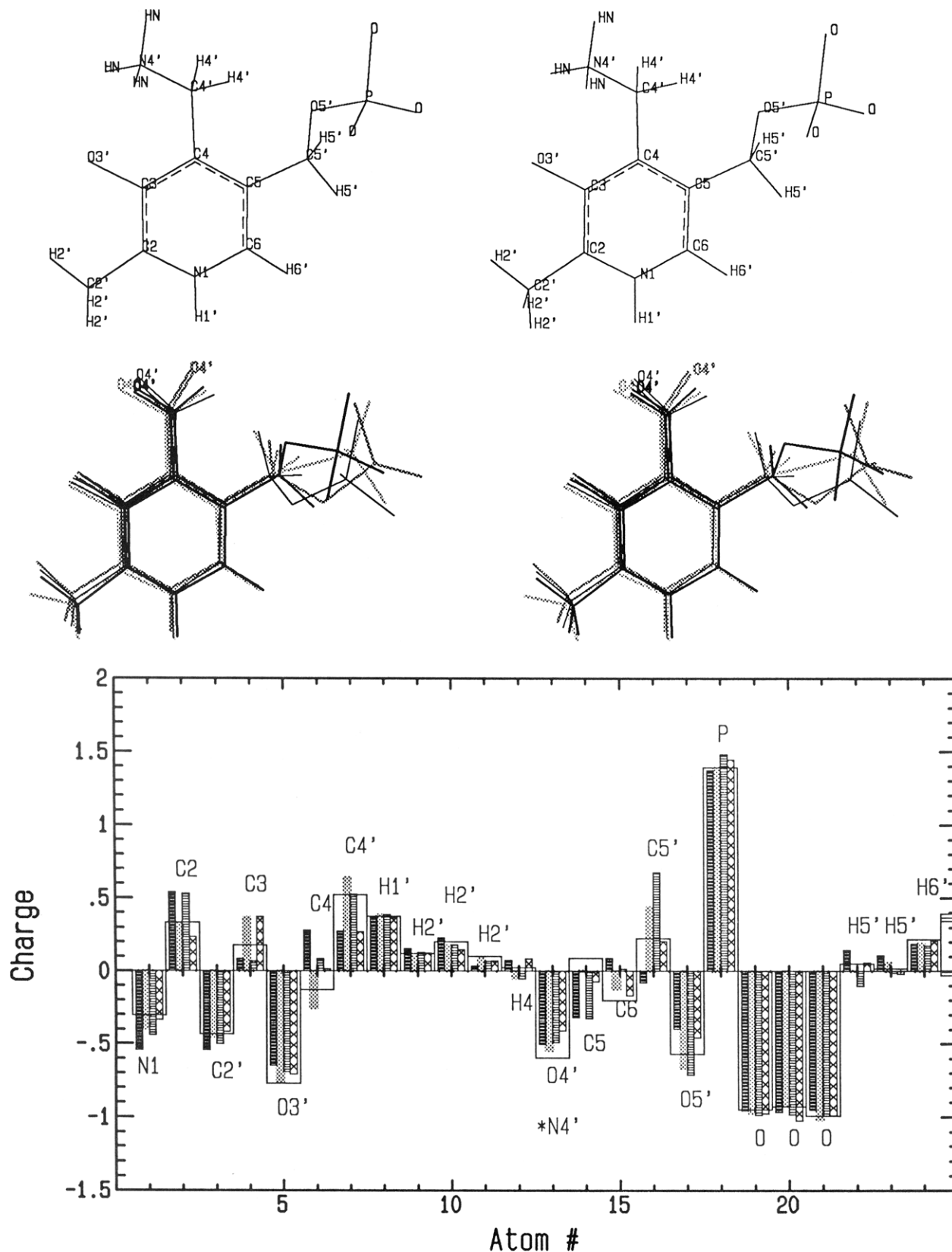| no.[c]	ID[c]	charge	no.[c]	ID[c]	charge	no.[c]	ID[c]	charge
1	N1	-0.430	11	H2'	0.123	21	O	-0.978
2	C2	0.403	12	H4' [b]	0.015	22	H5'	0.029
3	C2'	-0.479	13	O4' [b]	-0.488	23	H5'	0.029
4	C3	0.222	14	C5	-0.178	24	H6	0.189
5	O3'	-0.704	15	C6	-0.048			
6	C4	0.030	16	C5'	0.310	for PMP	C4'	0.504
7	C4' [b]	0.427	17	O5'	-0.558		H4'	-0.004
8	H1	0.374	18	P	1.422		N4'	-0.604
9	H2'	0.123	19	O	-0.978		HN	0.354
10	H2'	0.123	20	O	-0.978			

total for PLP = -2.000; for PMP = -1.000

[a] HF/6-31G** ESP charges for four pyridoxal-5'-phosphate models (see text) were averaged and rounded to three decimal places. Charges on identical atoms of symmetrical groups were then averaged with group charge held constant. [b] For pyridoxamine-5'-phosphate the 4'-amine group ($-CH_2NH_3^+$) replaces the 4'-carbonyl ($-CHO$) of pyridoxal-5'-phosphate. The 4' group charge (+1) was distributed according to the ESP charges for this group in a separate HF/6-31G** calculation of a PMP model (see text) combined with the 4' group charge (-0.046) from pyridoxal-5'-phosphate. Adjustments to the calculated charges were distributed proportionally. [c] For atom numbers and IDs see Figures 4 and 5.

appropriate of the ones tested for general use in developing charge sets for organic cofactors.

ESP Charges for Some Enzyme Cofactors. Table 2 gives the summary statistics for the various calculations for the organic enzyme cofactors. These are all very large *ab initio* calculations as indicated by the number of basis functions and primitives. For the open shell system FMNH, the highest level calculation that was feasible within our computational limitations was UHF/6-31G**. The charges for the smaller cofactors (all scaled to the 6-31G* level) are shown in Figures 3–5. The comparison of Table 3 shows the correlation between charge sets for each individual cofactor with the different methods compared to their HF/6-31G* set. The table (Table 3) also lists the corresponding statistics for the correlations for the test models of Table 1. It is clear from these figures (Figures 3–5) and Table 3 that the conclusions drawn from the calculations of the nine small organic structures discussed above also pertain to these much larger structures. These conclusions are (1) that scaling can be used to obtain general agreement between charges calculated with different methods, (2) that the more rigorous procedures give very similar charge sets, and (3) that electron correlation decreases the charge separation. This latter point can be seen in comparing the HF/3-21G results with the MP2/3-21G and Bk3P86/3-21G results.

For the single biotin model, the higher level methods converge toward a very similar set of charges when scaled using the factors of Table 1. The HF/6-31G** ESP charge set of Table 4 should serve the needs of electrostatic and molecular modeling calculations for this cofactor. This charge set was derived with rounding to three decimal places, with adjustment in the third decimal position of the largest magnitude values so that the total charge sums to -1 and with averaging between identical atoms of symmetrical groups.

Figures 4 and 5 show differences in charge distribution for the two forms of the vitamin B_6 cofactor. The chemical difference is, of course, in the 4' position where pyridoxal-5'-phosphate has an sp^2 carbon as a carbonyl group and pyridoxamine-5'-phosphate has an sp^3 carbon with a primary

moments plots (e.g., Figure 2). Thus, while all procedures tested give reasonable linear correlations of the ESP charges, those in the less accurate group yield more variation in charge distribution. Those in the high accuracy group all give good correlations and small standard deviations and should give charges that can be accurately scaled to each other. It is suggested, therefore, that these methods are the most

Figure 6. Comparison of HF/6-31G** ESP charges for one pyridoxamine-5′-phosphate mode (PMP from PDB file 2aat; upper stereodiagram and full width, unfilled bars) and four different pyridoxal-5′-phosphate (PLP) models (lower stereodiagram). The charges for the atoms of the PLP models (with pdb file names) are indicated by narrow bars: filled bars = PLP from 1gpb; shaded bar = PLP from 3aat; horizontal hatched bars = PLP from 1gpa; cross hatched bars = PLP from 1wsy. Atom numbers are given on the diagrams of Figures 3 and 4. Atom names are indicated on the diagrams of this figure.

amine substituent. However, in addition to this chemical difference there are conformational differences in the two structures, particularly in the 5′ side chains. Figure 6 shows the charge distributions calculated for four different PLP models (lower set of stereoviews) taken from different X-ray

crystal structures compared with the ESP charge distribution for the PMP model (upper stereoview) over the common atom positions. It is clear from this comparison that the differences in distribution due to conformation are much greater than the difference between any one of the PLP
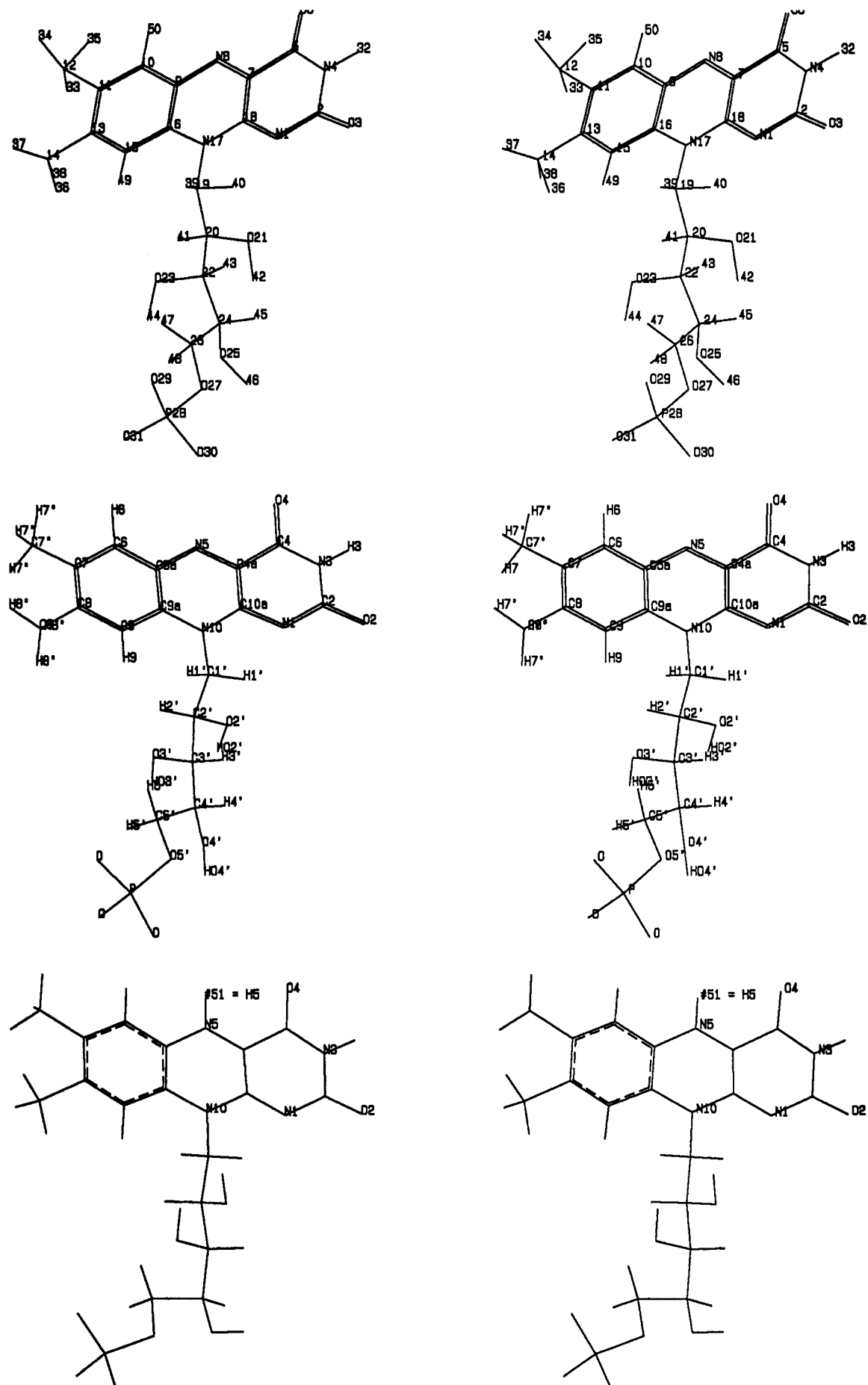
Figure 7. Flavin mononucleotide (FMN) structures used for the *ab initio* calculations. Top stereo pair, FMN abstracted from *Clostridial* flavodoxin (PDB file 3fxn). Middle stereo pair, FMN abstracted from *Chondrus crispus* flavodoxin (PDB file 2fcr). Bottom stereo pair, the flavin mononucleotide semiquinone (FMNH) from *Clostridial* flavodoxin (PDB file 4fxn). Atom numbers are indicated in the top diagram for both FMN and FMNH except for the H5 proton (see bottom diagram). Atom names are indicated in the middle diagram.
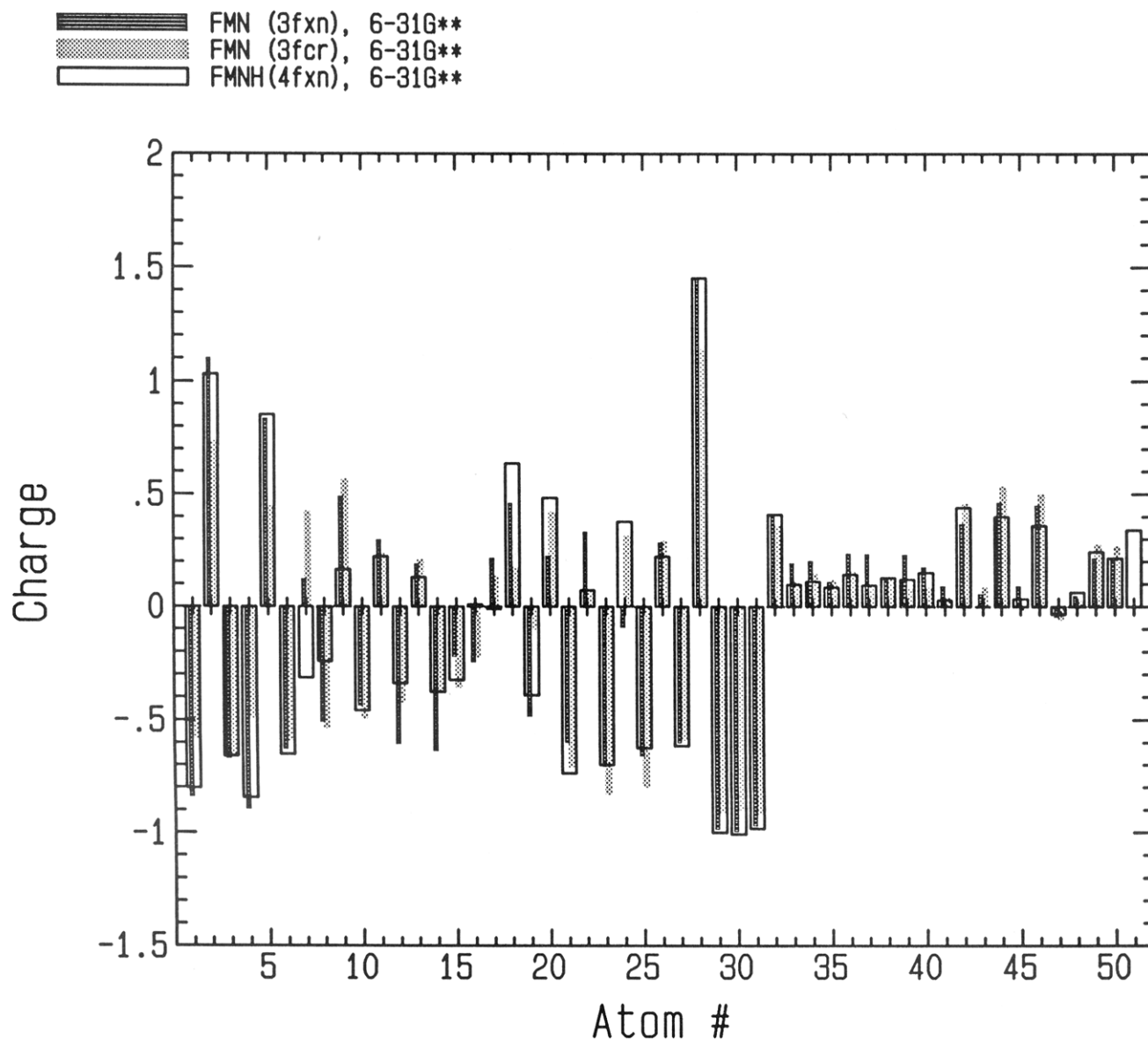
Figure 8. Comparison of the HF/6-31G** ESP charges for the atoms of two oxidized flavin mononucleotide models (narrow, filled bar = FMN from *Clostridial* flavodoxin, PDB file 3fxn; narrow, shaded bar = FMN from *Chondrus crispus* flavodoxin, PDB file 2fcr) and one flavin mononucleotide semiquinone model (wide, unfilled bar = FMNH from *Clostridial* flavodoxin).

models and the PMP model. Thus, a charge set for molecular modeling taken over the conformational average should be equally valid for both PLP and PMP cofactors except for the 4′ groups. The suggested charge sets are then given in Table 5. The charges for the 4′ groups were scaled to a common group total charge (−0.046) plus the additional +1 formal charge in the case of PMP. The distribution over the 4′-methylamine group from the PMP calculation (group charge = 1.013) was scaled down to 0.954 by adjusting each charge proportionally.

Of the several models of proteins containing flavin mononucleotide cofactors in the Brookhaven Protein Data Bank files, two that are identified as containing fully oxidized FMN and one as containing the semiquinone, partially reduced form, FMNH, have been selected. The structures of these three abstracted FMN cofactors are shown in stereo in Figure 7. The ESP charge distributions for these models using HF/6-31G** calculations are given in Figure 8. The two models, one oxidized and one semiquinone, taken from the same protein, *Clostridial* flavodoxin, give similar charge distributions in spite of the difference in oxidation state (wide bars versus narrow filled bars in Figure 8). Again, as with

the larger sample of structures in the PLP study (Figure 6), the conformational differences between the two FMN models give greater differences in charge distribution than those between the two chemical forms. Only at the C7 position is the charge in the FMNH model dramatically different (opposite sign) than that for both FMN models. Indeed, the added electron charge seems to be distributed primarily over the atoms of the central ring (C7, N8, C9, C16, N17, and C18) and the additional hydrogen atom (H8′) by the criterion of a consistent difference between the FMNH atoms and both FMN models. Without a greater sample of experimental conformations for both forms, it is difficult to be more specific or very confident about the differences between them in the biological setting. Scaled to net zero charge, the approximate difference charges and the total H8 charge for these seven atoms are $\Delta C7 = -0.446$, $\Delta N8 = 0.258$, $H8 = 0.331$, $\Delta C9 = -0.331$, $\Delta C16 = 0.245$, $\Delta N17 = -0.232$, $\Delta C18 = 0.175$. Thus the FMN charge set of Table 6 derived from the average of the charges for the two conformations can be converted for use with FMNH by adding these factors to the appropriate atomic charges.

Table 6. Flavin Mononucleotide (FMN) and Lumiflavin Charge Comparison

no.	ID	6-31G** ESP Charges			Del Re + Huckel[b]	MINDO/3 lumiflavin[c]
		3fxn	2fcr	av[a]		
1	N1	-0.841	-0.578	-0.710	-0.760	-0.413
2	C2	1.099	0.730	0.915	0.458	0.734
3	O2	-0.671	-0.660	-0.666	-0.522	-0.548
4	N3	-0.895	-0.488	-0.692	-0.083	-0.328
32	H3	0.413	0.357	0.385	0.209	0.112
5	C4	0.832	0.443	0.637	0.335	0.641
6	O4	-0.632	-0.592	-0.612	-0.479	-0.545
7	C4a	0.120	0.419	0.270	0.178	-0.106
8	N5	-0.508	-0.530	-0.519	-0.344	-0.013
9	C5a	0.487	0.561	0.524	0.132	0.001
10	C6	-0.437	-0.489	-0.463	-0.052	0.028
50	H6	0.204	0.262	0.233	0.053	0.001
11	C7	0.293	0.230	0.262	-0.002	-0.025
12	C7''	-0.607	-0.419	-0.512		0.088
33	H7''	0.190	0.106	0.143		-0.022
34	H7''	0.201	0.140	0.143		-0.024
35	H7''	0.108	0.114	0.143		-0.024
C7α	CH3	-0.108	-0.059	-0.083	0.019	
13	C8	0.187	0.205	0.196	0.049	0.059
14	C8''	-0.639	-0.365	-0.502		0.070
36	H8''	0.234	0.152	0.156		-0.016
37	H8''	0.229	0.091	0.156		-0.017
38	H8''	0.115	0.114	0.156		-0.020
C8α	CH3	-0.061	-0.008	-0.034	0.025	
15	C9	-0.222	-0.353	-0.288	-0.115	-0.068
49	H9	0.212	0.271	0.242	0.053	0.001
16	C9a	-0.244	-0.221	-0.232	0.133	0.098
17	N10	0.214	0.132	0.173	0.267	-0.045
18	C10a	0.455	0.170	0.312	0.353	0.292
19	C1'	-0.485	-0.097	-0.291	-0.013	0.140
39	H1'	-0.228	0.078	0.140	0.046	0.008
40	H1'	-0.173	0.081	0.140	0.046	-0.031
20	C2'	0.223	0.417	0.320	0.099	
41	H2'	0.090	0.039	0.065	0.051	
21	O2'	-0.602	-0.705	-0.653		
42	HO2'	0.365	0.452	0.408		
OH2'		-0.237	-0.253	-0.245	-0.158	
22	C3'	0.330	0.067	0.198	0.105	
43	H3'	0.054	0.081	0.068	0.052	
23	O3'	-0.709	-0.826	-0.767		
44	HO3'	0.460	0.528	0.494		
OH3'		-0.249	-0.298	-0.273	-0.157	
24	C4'	-0.094	0.310	0.108	0.110	
45	H4'	0.088	-0.014	0.037	0.052	
25	O4'	-0.661	-0.793	-0.727		
46	HO4'	0.448	0.495	0.471		
OH4'		-0.213	-0.298	-0.256	-0.156	
26	C5'	0.283	0.286	0.285	0.037	
47	H5'	-0.045	-0.054	-0.008	0.058	
48	H5'	0.047	0.021	-0.008	0.058	
27	O5'	-0.600	-0.584	-0.592	-0.214	
28	P	1.454	1.135	1.294	0.659	
29	O	-0.984	-0.910	-0.944	-0.868	
30	O	-0.993	-0.898	-0.944	-0.868	
31	O	-0.969	-0.912	-0.944	-0.868	
total				-2.000	-2.022	0.028
linear correlation vs ESP set					0.898	0.687
std dev of fit vs ESP set					0.219	0.318
slope of fit vs ESP set					1.299	1.132

[a] ESP average charges with equivalent atoms of symmetrical groups averaged. [b] From Vinayaka and Rao, ref 77. [c] From Hall *et al.*, ref 82.

DISCUSSION

In terms of computer time, the least expensive computations for the small organic molecules in the more stable grouping (bottom half of Table 1) were the MP2/3-21G and HF/6-31G* (also HF/6-31G**) approaches. Note the good performance of the small 3-21G basis set which has been previously pointed out[20] for smaller molecules (<4 heavy atoms). The added polarization on hydrogens of the HF/6-

31G** makes no significant difference; both HF/6-31G* and HF/6-31G** give essentially identical ESP charges (see also ref 20). Thus, use of HF/6-31G* (or HF/6-31G**) calculations as a standard approach for these small organic molecules is again supported. Another well established point from smaller molecule calculations that is consistent with this study is the reduction in charge distribution and dipole moment when electron correlation is included.[22,23,70] Indeed, these data compare favorably with previous studies of ESP charges and their correlations (Table 7) with HF/6-31G* based charges. ECP basis sets also were found to give scalable charges but with greater charge separation unscaled.

The DFT and hybrid DFT approaches gave results that correlate well with those from the traditional *ab initio* methods. The hybrid Becke functional[56] with correlation corrections seemed to perform better than either Hartree–Fock or MP2 level calculations using the same basis set. In a few cases, they also showed a clear advantage in speed.

The data from Table 1 suggests that for the larger cofactor molecules the most efficient calculations in the more precise group ($r_{ab} > 0.99$) would be Bk3P86/3-21G, MP2/3-21G, HF/6-31G* (or HF/6-31G**), and HF/6-311G**. The statistics of Tables 2 and 3 generally support these choices. However, even though the test set of small molecules contained representative functional groups and structures found in these cofactor molecules, the linear correlations for the atoms of the large molecules compared individually are slightly but significantly different in some cases (Table 3). The midrange Hartree–Fock approaches give the best correlations.

ESP charges from HF/6-31G* level calculations give dipole moments that are too large when compared with experimental values.[13,20] Therefore, the charges used in the AMBER molecular mechanics program[73-75] have been reduced by a factor of 0.91. In cases where formal charges are involved other adjustments are necessary.[32,34]

The author was not able to find previous *ab initio* or semiempirical calculations for biotin. As with the other models tested, the ESP charges for this cofactor in the experimental geometry were consistently calculated with different levels of theory and electron correlation tended to decrease charge separation.

Del Vado *et al.*[76] recently used semiempirical methods to investigate the conformations of pyridoxal Shiff's bases but did not report charge calculations from their studies. The geometries used here sample the variation in conformations of enzyme bound pyridoxal cofactors (Figure 6) and indicate, as would be expected, that the calculated charges are conformationally dependent. Indeed, the parts of the molecule demonstrating the most conformational variation, the C4' and C5' groups, also show the most variation in ESP charges, i.e., the charge variation is accommodated locally. Similarly, the chemical difference between PMP and PLP models only appears to cause local changes in the ESP charge distribution of the C4' group, however, without a greater sample of PMP conformations this conclusion is only preliminary.

The ESP charges for oxidized FMN (Figure 8) show considerable differences from the previous charge set of Vinayaka and Rao[77] (Table 6). The earlier charge set[77] was the combination of Del Re σ charges[78] and Huckel π charges.[79] They were used as molecular mechanics parameters for a study of FMN conformations. They have also been used for AMBER molecular mechanics, dynamics, and

Table 7. Correlations and Scaling Factors for Atom Centered Point Charges for Organic and Inorganic Molecules from Different Computational Methods

method	atoms and atom types[a]	F_{ab}[b]	r_{ab}[b]	ref
	Miscellaneous			
X-ray	C, H, N, O, Aro	1.01[c]		Coppens[15]
EN	C, H, N, O, Hcy	1.17[d]		Mullay[31]
	Semiempirical			
AM1	C, H, N, O, Aro	1.39	0.89	Besler et al.[21]
AM1	C, H, N, O, F, Aro, Hcy	1.33	0.88	Aleman et al.[24]
AM1/NDDO	C, H, N, O, F.Aro, HCy	1.13	0.98	Alhambra et al.[71]
AM1	C, H, N, O, Aro, Hcy	1.22	0.96	Wang and Ford[72]
MNDO	C, H, N, O, S, Aro	1.42	0.97	Besler et al.[21]
MNDO	C, H, N, O, F.Aro, Hcy	1.31	0.98	Aleman et al.[24]
MNDO/NDDO	C, H, N, O, F.Aro, Hcy	0.99	0.94	Alhambra et al.[71]
PM3	C, H, N, O, F.Aro, Hcy	1.41	0.88	Aleman et al.[24]
PM3/NDDO	C, H, N, O, F.Aro, Hcy	1.10	0.95	Alhambra et al.[71]
	Ab Initio: Hartree−Fock SCF: Full Basis Sets			
STO-3G	C, H, N, O, B	1.13[e]		Cox and Williams[13]
STO-3G	C, H, N, O, Cl, F, P, S		0.985	Chirlian and Francl[20]
STO-3G	C, H, N, O, S, Aro	1.10	0.96	Besler et al.[21]
STO-3G	C, H, N, O, P, S, Aro, Hcy	1.12	0.975	this work
3-21G	C, H, N, O, Cl, F, P, S	0.86[e]	0.997	Chirlian and Francl[20]
3-21G	C, H, N, O, P, S, Aro, Hcy	0.88	0.992	this work
6-31G	C, H, N, O, B	0.82[e]		Cox and Williams[13]
6-31G	C, H, N, O, Cl, F, P, S		0.995	Chirlian and Francl[20]
6-31G**	C, H, N, O, P. S, Aro, Hcy	1.00	1.000	this work
6-311G**	C, H, N, O, P, S, Aro, Hcy	0.98	0.998	this work
6-311+G**	C, H, N, O, P, S, Aro, Hcy	1.00	1.000	this work
D95*	C, H, N, O, P, S, Aro, Hcy	0.98	0.996	this work
	ECP Basis Sets			
LANL1MB	C, H, N, O, P, S, Aro, Hcy	1.17	0.978	this work
CED-121G	C, H, N, O, P, S, Aro, Hcy	0.82	0.986	this work
LANL1DZ	C, H, N, O, P, S, Aro, Hcy	0.81	0.987	this work
	DFT SCF			
Xα/ECP-DZ	C, H, N, O, P, S, Aro, Hcy	0.94	0.971	this work
Xα/D95*	C, H, N, O, S, Aro, Hcy	1.11	0.980	this work
	HF Methods with Electron Correlation			
	Møller−Plesset			
MP2/STO-3G	C, H, N, O, P, S, Aro, Hcy	1.24	0.973	this work
MP2/3-21G	C, H, N, O, P, S, Aro, Hcy	1.02	0.993	this work
MP2/ECP-DZ	C, H, N, O, P, S, Aro, Hcy	0.90	0.989	this work
MP2/D95*	C, H, N, O, P, S, Aro, Hcl	1.08	0.997	this work
MP2/6-311G**	C, H, N, O, P, S, Aro, Hcl	1.09	0.996	this work
	Configuration−Interaction			
CIPSI/G/6-31G*	C, F, H, N, O	1.02	0.995	Luque et al.[22]
CIPSI/G+M/6-31G*	C, F, H, N, O	1.03	0.993	Luque et al.[22]
	DFT with Correlation			
BLYP/ECP-DZ	C, H, N, O, P, S, Aro, Hcy	0.96	0.984	this work
BLYP/D95*	C, H, N, O, S, Aro, Hcy	1.18	0.993	this work
	Hybrid HF/DFT with Correlation			
Bk3P86/3-21G	C, H, N, O, P, S, Aro, Hcy	1.01	0.986	this work
Bk3P86/ECP-DZ	C, H, N, O, P, S, Aro, Hcy	0.94	0.990	this work
Bk3LYP/6-31G**	C, H, N, O, P, S, Aro, Hcy	1.16	0.997	this work
Bk3P86/D95*	C, H, N, O, P, S, Aro, Hcy	1.11	0.994	this work

[a] Atoms and molecule types: Aro = aromatic, hcy = heterocycle, Cyc = simple cyclic. [b] Scale factor, F_{ab}, linear correlation coefficient, r_{ab}. F_{ab} is factor to convert charges, q, to the ESP charges from *ab initio* HF/6-31G* or HF/6-31G** calculations, i.e., $q_{6-31G*} \simeq F_{ab}q$. [c] Comparison from this paper and STO-3G results of Singh and Kollman[14] on 9-methylguanine (15 atoms total) gave a slope of 0.916 and a correlation coefficient of 0.87. Slope was then scaled to 6-31G* by a factor of 1.1. [d] Comparison in this paper to STO-3G scaled to 6-31G* by a factor of 1.1. [e] Scale factor, F_{ab}, is for scaling to 6-31G** level, i.e., $q_{6-31G**} = F_{ab}q$.

homology modeling of flavodoxin.[80,81] The comparison between the two charge sets (correlation coefficient, SD, and slope) shows that they do not correlate well. The second comparison in the table (Table 6) is with the MNDO/3 charges from calculations on lumiflavin[72] which show even less correlation with the FMN ESP charge set.

The semiquinone charge distribution from the 6-31G** calculation is shown in Figure 8 compared to similar calculations for the two oxidized FMN models. It seems that the consistent differences between the semiquinone results and both oxidized models are in the ESP charges distributed between the central ring atoms. As other structures for oxidized, semiquinone and reduced FMN containing proteins become available,[83] a better picture of the effect of oxidation state on the ESP charge set will emerge.

A number of *ab initio* and semiempirical calculations have been reported on flavin and flavin analogs[82,84–91] with their focus primarily on the structure and electronic states. The calculations of Song[84,85] and Platenkamp et al.[86] focused on changes between the ground state and singlet and triplet excited states. The HF/3-21G *ab initio* calculations and MINDO/3 semiempirical calculations of Vazques et al.[91] showed the difference in Mulliken populations between oxidized and semiquinone models of lumiflavin where the main changes found were also in the charge distribution in the central ring. However, the individual changes in charges for these ring atoms by the three methods do not correlate. The ESP charge changes from the HF/6-31G** calculations of this paper are uniformly larger in magnitude and more asymmetric than the Mulliken population changes for either of the methods of Vazques et al.[91] The work of Teitell and Fox[89] investigated the charge distribution in reduced states, while the MINDO/3 calculations of Hall et al.[82] and the *ab initio* calculations of Platenkamp et al.[87] gave charges for fully oxidized and fully reduced species. Thus, as with the PLP/PMP comparison, a definitive analysis of the contributions of conformation and oxidation state to the distribution of ESP charges in flavins will require further work with more conformations.

ACKNOWLEDGMENT

REFERENCES AND NOTES

(1) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the Electrostatic Potential of Molecules in Solution: Method and Error Assessment. *J. Comput. Chem.* **1987**, *9*, 327–335.
(2) Warshel, A.; Creighton, S. Microscopic Free Energy Calculations in Solvated Macromolecules as a Primary Structure-Function Correlator and the MOLARIS Program. In *Computer Simulation of Biomolecular Systems*; van Gunsteren, W. F., Weiner, P. K., Eds.; ESCOM: Leiden, 1989; pp 120–138.
(3) Warshel, A.; Åqvist, J. Electrostatic and Macromolecular Function. *Ann. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267–298.
(4) Xu, Y. W.; Wang, C. X.; Shi, Y. Y. Improvements on the Protein-Dipole Langevin-Dipole Model. *J. Comput. Chem.* **1992**, *13*, 1109–1113.
(5) Gilson, M. K.; Honig, B. H. Calculation of the Total Electrostatic Energy of a Macromolecular System: Solvation Energies, Binding Energies and Conformational Analysis. *Proteins* **1988**, *4*, 7–18.
(6) Lee, F. S.; Chu, Z. T.; Warshel, A. Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the POLARIS and ENZYMIX Programs. *J. Comput. Chem.* **1993**, *14*, 161–185.
(7) Harvey, S. C. Treatment of Electrostatic Effects in Macromolecular Modeling. *Proteins* **1989**, *5*, 79–92.
(8) Davis, M. E.; McCammon, J. A. Electrostatics in Biomolecular Structure and Dynamics. *Chem. Rev.* **1990**, *90*, 509–521.
(9) Wampler, J. E. Computational Chemistry and Molecular Modeling of Electron-Transfer Proteins. In *Methods in Enzymology*; Peck, H. D., Jr.; LeGall, J., Eds.; Academic Press: New York, 1994; pp 559–607.
(10) Mulliken, R. S. Electronic Population Analysis on LCAO-MO Molecular Wave Functions. *J. Chem. Phys.* **1955**, *23*, 1833–1840.
(11) Case, D.; Karplus, M. The calculation of one-electron properties from Xα multiple scattering wavefunctions. *Chem. Phys. Lett.* **1976**, *39*, 33–38.

(12) Momany, F. A. Determination of Partial Atomic Charges from Ab Initio Molecular Electrostatic Potentials. Application to Formamide, Methanol, and Formic Acid. *J. Phys. Chem.* **1978**, *82*, 592–601.
(13) Cox, S. R.; Williams, D. E. Representation of the Molecular Electrostatic Potential by a Net Atomic Charge Model. *J. Comput. Chem.* **1981**, *2*, 304–323.
(14) Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5*, 128–145.
(15) Coppens, P. Electron Density for X-Ray Diffraction. *An. Rev. Phys. Chem.* **1992**, *43*, 663–692.
(16) Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Functions for Proteins. Energy Minimization for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
(17) Carlson, H. A.; Nguyen, T. B.; Orozco, M.; Jorgensen, W. L. Accuracy of Free Energies of Hydration for Organic Molecules from 6-31G*-Derived Partial Charges. *J. Comput. Chem.* **1993**, *14*, 1240–1249.
(18) Breneman, C. M.; Wiberg, K. B. Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials. The Need for High sampling Density in Formamide Conformational Analysis. *J. Comput. Chem.* **1990**, *11*, 361–373.
(19) Wiberg, K. B.; Rablen, P. R. Comparison of Atomic Charges Derived via Different Procedures. *J. Comput. Chem.* **1993**, *13*, 1504–1518.
(20) Chirlian, L. E.; Francl, M. M. Atomic Charges Derived from Electrostatic Potentials: A Detailed Study. *J. Comput. Chem.* **1987**, *8*, 894–905.
(21) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431–439.
(22) Luque, F. J.; Orozco, M.; Illas, F.; Rubio, J. Effect of Electron Correlation on the Electrostatic Potential Distribution of Molecules. *J. Am. Chem. Soc.* **1991**, *113*, 5203–5211.
(23) Wiberg, K. B.; Hadad, C. M.; LePage, T. J.; Breneman, C. M.; Frisch, M. J. Analysis of the Effect of Electron Correlation on Charge Density Distributions. *J. Phys. Chem.* **1992**, *96*, 671–679.
(24) Aleman, C.; Luque, F. J.; Orozco, M. Suitability of the PM3-Derived Molecular Electrostatic Potentials. *J. Comput. Chem.* **1993**, *14*, 799–808.
(25) Su, Z.; Coppens, P. Electrostatic Properties of Molecules from the X-ray Charge Density. Application to Deuterated Benzene, 1-Alanine and d,l-Histidine. *Zeit. Natur. Sect. A* **1993**, *48*, 85–90.
(26) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural population analysis. *J. Chem. Phys.* **1985**, *83*, 735–746.
(27) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor. Chim. Acta* **1977**, *44*, 129–138.
(28) Bader, R. F. W. *Atoms in Molecules. A Quantum Theory*; Clarendon Press: Oxford, 1990.
(29) Cioslowski, J. A new population analysis based on atomic polar tensors. *J. Am. Chem. Soc.* **1989**, *111*, 8333–8336.
(30) Dinur, U.; Hagler, A. T. Direct evaluation of non-bonding interactions from *ab initio* calculations. *J. Am. Chem. Soc.* **1989**, *111*, 5149–5151.
(31) Mullay, J. A Method for Calculating Atomic Charges in Large Molecules. *J. Comput. Chem.* **1988**, *9*, 399–405.
(32) Merz, K. Analysis of a Large Data Base of Electrostatic Potential Derived Atomic Charges. *J. Comput. Chem.* **1992**, *13*, 749–767.
(33) Stouch, T. R.; Williams, D. E. Conformational Dependence of Electrostatic Potential-Derived Charges: Studies of the Fitting Procedure. *J. Comput. Chem.* **1993**, *14*, 858–866.
(34) Cannon, J. F. AMBER Force-Field Parameter for Guanosine-triphosphate and its imido and methylene analogs. *J. Comput. Chem.* **1993**, *14*, 995–1005.
(35) Labanowski, J. K.; Andzelm, J. W. *Density Functional Methods in Chemistry*; Springer-Verlag: New York, 1991.
(36) Ziegler, T. Approximate Density Functional Theory as a Practical Tool in Molecular Energetics and Dynamics. *Chem. Rev.* **1991**, *91*, 651–667.
(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Wong, M. W.; Foresman, J. B.; Robb, M. A.; Head-Gordin, M.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92/DFT, Revision F.2*; Gaussian, Inc.: Pittsburgh, PA, 1993.
(38) Hehre, W. J.; Radom, L.; Schleyer, R. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.
(39) Frisch, M.; Foresman, J.; Frisch, A. E. *Gaussian 92 User's Guide*; Gaussian Inc.: Pittsburgh, PA, 1992.
(40) Foresman, J. B.; Frisch, A. E. *Exploring Chemistry with Electronic Structure Methods: A Guide To Using Gaussian*; Gaussian Inc.: Pittsburgh, PA, 1993.
(41) Dunning, T. H. Gaussian Basis Functions for Use in Molecular Calculations. I. Contraction of (9s5p) Atomic Basis Sets for the First-Row Atoms. *J. Chem. Phys.* **1970**, *53*, 2823–2833.
(42) Dunning, T. H. Gaussian Basis Functions for Use in Molecular Calculations. IV. The Representation of Polarization Functions for the First Row Atoms and Hydrogen. *J. Chem. Phys.* **1971**, *55*, 3958–3966.

632 *J. Chem. Inf. Comput. Sci., Vol. 35, No. 3, 1995*

WAMPLER

(43) Huzinaga, S. Gaussian-Type Functions for Polyatomic Systems; *J. Chem. Phys.* **1965**, *42*, 1293–1302.

(44) Krishnan, R.; Frisch, M. J.; Pople, J. A. Contribution of triple substitutions to the electron correlation energy in fourth order perturbation theory. *J. Chem. Phys.* **1980**, *72*, 4244–4249.

(45) Hehre, W. J.; Pople, J. A. Molecular orbital theory of the electronic structure of organic compounds. XV. The protonation of benzene. *J. Am. Chem. Soc.* **1972**, *94*, 6901–6904.

(46) Hay, P. J.; Wadt, W. R. Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J. Chem. Phys.* **1985**, *82*, 270–283.

(47) Hay, P. J.; Wadt, W. R. Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.* **1985**, *82*, 299–310.

(48) Wadt, W. R.; Hay, P. J. Ab initio effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. *J. Chem. Phys.* **1985**, *82*, 284–298.

(49) Stevens, W.; Basch, H.; Krauss, J. Compact effective potentials and efficient shared-exponent basis sets for the first and second-row atoms. *J. Chem. Phys.* **1984**, *81*, 6026–6033.

(50) Slater, J. C. *The Self-Consistent Field for Molecules and Solids*; McGraw-Hill: New York, 1974.

(51) Becke, A. D. Density-functional exchange energy approximation with current asymptotic behavior. *Phys. Rev.* **1988**, *A38*, 3098–3100.

(52) Møller, C.; Plesset, M. S. Note on the approximation treatment for many electron systems. *Phys. Rev.* **1934**, *46*, 618–622.

(53) Binkley, J. S.; Pople, J. A. Møller–Plesset Theory for Atomic Ground State Energies. *Int. J. Quantum Chem.* **1975**, *9*, 229–236.

(54) Lee, C.; Wang, W.; Parr, R. G. Local softness and chemical reactivity in the molecules CO, SCN⁻ and H₂CO. *Theochem.* **1987**, *163*, 305–313.

(55) Perdew, J. P. Density functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev.* **1986**, *B33*, 8822–8824; Erratum, *Phys. Rev.* **1986**, *B34*, 7404.

(56) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(57) Bacskay, G. B. A quadratically convergent Hartree–Fock (QC-SCF) Method. Applications to closed shell systems. *Chem. Phys.* **1981**, *61*, 385–404.

(58) Allen, F. H.; Bellar, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummeling-Peters, B. G.; Kennard, O.; Motherwell, M. D. S.; Rodgers, J. R.; Watson, D. G. The Cambridge Crystallographic Data Centre: Computer-Based Search, Retrieval, Analysis and Display of Information. *Acta Crystallogr.* **1979**, *B35*, 2331–2339.

(59) Weber, H.-P.; McMullan, R. K.; Swaminathan, S.; Craven, B. M. The structure and thermal motion of phosphorylethanolamine at 122 K from neutron diffraction. *Acta Crystallogr., B (Struct. Sci.)* **1984**, *40*, 506–511.

(60) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tamusi, M. The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112*, 535–542.

(61) Abola, E. E.; Bernstein, F. C.; Bryant, S. H.; Koetzle, T. F.; Weng, J. Protein Data Bank. In *Crystallographic Databases-Information Content, Software Systems, Scientific Applications*; Allen, F. H., Bergerhoff, G., Sievers, R., Eds.; Data Commission of the International Union of Crystallography: Bonn, 1987; pp 107–132.

(62) Smith, W. W.; Burnett, R. M.; Darling, G. D.; Ludwig, M. L. Structure of the semiquinone form of flavodoxin from Clostridium MP. Extension of 1.8 Angstroms resolution and some comparisons with the oxidized state. *J. Mol. Biol.* **1977**, *117*, 195–225.

(63) PDB structure listed to Fukuyama et al., to be published; previous reference: Fukuyama, K.; Wakabayashi, S.; Matsubara, H.; Rogers, L. J. Tertiary structure of oxidized flavodoxin from an eukaryotic red alga *Chondrus crispus* at 2.35-Angstroms resolution. Localization of charged residues and implication for interaction with electron transfer partners. *J. Biol. Chem.* **1990**, *265*, 15804–15812.

(64) Acharya, K. R.; Stuart, D. I.; Varvill, K. M.; Johnson, L. N. Glycogen phosphorylase B: Description of the protein structure. In *Glycogen Phosphorylase B*; World Scientific Publishing Co.: Singapore, 1991.

(65) Barford, D.; Hu, S.-H.; Johnson, L. N. Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP. *J. Mol. Biol.* **1991**, *218*, 233.

(66) Hyde, C. C.; Miles, E. W. The tryptophan synthase multienzyme complex. Exploring structure-function relationships with x-ray crystallography and mutagenesis. *Biotechnology* **1990**, *8*, 27–31.

(67) Danishefsky, A. T.; Onnufer, J. J.; Petsko, G. A.; Ringe, D. Activity and structure of the active-site mutants R386Y and R386F of

(68) PDB structure listed to Smith et al., to be published; previous reference: Smith, D. L.; Ringe, D.; Finlayson, W. L.; Kirsch, J. F. Preliminary x-ray data for aspartate aminotransferase from *Escherichia coli. J. Mol. Biol.* **1986**, *191*, 301–302.

(69) Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. Structural origins of high-affinity biotin binding to Strepavidin. *Science* **1989**, *243*, 85–88.

(70) Carpenter, J. E.; McGrath, M. P.; Hehre, W. J. Effect of Electron Correlation on Atomic Electron Populations. *J. Am. Chem. Soc.* **1989**, *111*, 6154–6156.

(71) Alhambra, C.; Luque, F. J.; Orozco, M. Comparison of NDDO and Quasi-*ab-intio* Approaches to Compute Semi-empirical Molecular Electrostatic Potentials. *J. Comput. Chem.* **1994**, *15*, 12–22.

(72) Wang, B.; Ford, G. P. Atomic charges derived from a fast and accurate method for electrostatic potentials based on modified AM1 calculations. *J. Comput. Chem.* **1994**, *15*, 200–207.

(73) Weiner, P. K.; Kollman, P. A. AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. *J. Comput. Chem.* **1981**, *2*, 287–303.

(74) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. K. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.

(75) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230–252.

(76) del Vado, M. A. G.; Echevarria, G. R.; Blanco, F. G.; Blanco, J. G. S. On the stability of the Schiff-bases of pyridoxal 5′-phosphate with polypeptides containing 1-lysine. *Helv. Chim. Acta* **1991**, *74*, 1749–1756.

(77) Vinayaka, C. R.; Rao, V. S. R. A Conformational Approach to Study the Mode of Binding of Flavin Mononucleotide to Flavodoxin. *J. Biomol. Struct. Dynam.* **1984**, *2*, 663–674.

(78) Del Re, G. A simple MO-LCAO method for the calculation of charge distribution in saturated organic molecules. *J. Chem. Soc. Part IV* **1958**, 4031–4040.

(79) Pullman, B.; Pullman, A. *Quantum Biochemistry*; Interscience: New York, 1963; pp 104–109.

(80) Stewart, D. E. *The Structure, Interactions, and Dynamics of Electron Transport Proteins from Desulfovibrio: A Molecular Modeling and Computational Study.* Ph.D. Dissertation, University of Georgia, Athens, GA, 1989.

(81) Palma, P. N.; Moura, I.; LeGall, J.; van Beeumem, J.; Wampler, J. E.; Moura, J. J. G. Evidence for a Ternary Complex formed between Flavodoxin and Cytochrome c₃. *Biochemistry*, in press.

(82) Hall, L. H.; Orchard, B. J.; Tripathy, S. K. The Structure and Properties of Flavins: Molecular Orbital Study Based on Totally Optimized Geometries. II. Molecular Orbital Structure and Electron Distribution. *Int. J. Quantum Chem.* **1987**, *31*, 217–242.

(83) Watt, W.; Tulinsky, A.; Swenson, R. P.; Watenpaugh, K. D. Comparison of the crystal structures of a flavodoxin in its three oxidation states at cryogenic temperatures. *J. Mol. Biol.* **1991**, *218*, 195.

(84) Song, P.-S. Electronic Structure and Photochemistry of Flavins. IV. σ-Electronic Structure and the Lowest Triplet Configuration of a Flavin. *J. Phys. Chem.* **1968**, *72*, 536–542.

(85) Song, P.-S. On the Basicity of the Excited State of Flavins. *Photochem. Photobiol.* **1968**, *7*, 311–313.

(86) Platenkamp, R. J.; Palmer, M. H.; Visser, A. J. W. G. Molecular Orbital Studies of Flavin Radicals and the Lowest Triplet State of Isoalloxazine. *J. Mol. Struct.* **1980**, *67*, 45–64.

(87) Platenkamp, R. J.; Palmer, M. H.; Visser, A. J. W. G. Ab initio molecular orbital studies of closed shell flavins. *Eur. Biophys. J.* **1987**, *14*, 393–402.

(88) Palmer, M. H.; Simpson, I.; Platenkamp, R. J. The Electronic Structure of Flavin Derivatives. *J. Mol. Struct.* **1980**, *66*, 243–263.

(89) Teitell, M. F.; Fox, J. L. MO Study of Flavin Reduction. *Int. J. Quantum Chem.* **1982**, *22*, 583–594.

(90) Hall, L. H.; Orchard, B. J.; Tripathy, S. K. The Structure and Properties of Flavins: Molecular Orbital Study Based on Totally Optimized Geometries. I. Molecular Geometry Investigations. *Int. J. Quantum Chem.* **1987**, *31*, 195–216.

(91) Vasques, C. A.; Andrews, J. S.; Murray, C. W.; Amos, R. D.; Handy, N. C. An Investigation of the Three Oxidation Forms of Luminflavin. *J. Chem. Soc., Perkin Trans.* **1992**, *2*, 889–895.

Escherichia coli aspartate aminotransferase. *Biochemistry* **1991**, *30*, 1980–1985.

CI940261Q