Theor Chem Acc (2003) 109:213–219
DOI 10.1007/s00214-002-0400-8

Theoretical
Chemistry Accounts

Regular article

Accuracy assessment of semiempirical molecular electrostatic potential of proteins

Victor M. Anisimov[1], **Nikolay Anikin**[2], **Vladislav Bugaenko**[2], **Vladimir Bobrikov**[2], **Alexey Andreyev**[2]

[1]FQS Poland, Starowislna 13–15, 31-038 Krakow, Poland
[2]Quantum Biochemistry Group, Konstantina Fedina-3/24, 105215 Moscow, Russian Federation

Received: 26 April 2002 / Accepted: 19 September 2002 / Published online: 14 February 2003

Abstract. Accurate electrostatic maps of proteins are of great importance in research of protein interaction with ligands, solvent media, drugs, and other biomolecules. The large size of real-life proteins imposes severe limitations on computational methods one can use for obtaining the electrostatic map. Well-known accurate second-order Møller–Plesset and density functional theory methods are not routinely applicable to systems larger than several hundred atoms. Conventional semiempirical tools, as less resource demanding ones, could be an attractive solution but they do not yield sufficiently accurate calculation results with reference to protein systems, as our analysis demonstrates. The present work performs a thorough analysis of the accuracy issues of the modified neglect of differential overlap type semiempirical Hamiltonians AM1 and PM3 on example of the calculation of the molecular electrostatic potential and the dipole moment of natural amino acids. Real capabilities and limitations of these methods with application to protein modeling are discussed.

Keywords: Semiempirical method – Electrostatic maps – Dipole moment – Coulomb potential – Biomolecules

Introduction

Protein systems are characterized by the great diversity of various weak energy contributing terms. These are long-range Coulomb interactions, hydrogen bonds, van der Waals interactions, polarization induced by ionized fragments, and the effects of the solvent environment. Each of them locally contributes very little to the total energy of the system, but because of the large size of proteins a subtle balance of numerous weak interactions makes proteins very complex systems for molecular modeling.

The nonadditive nature of the interaction forces introduces unpredictable errors when these interactions are modeled in a simplistic fashion. This usually results in a limited parameter transferability and loss of accuracy. In order to be able to challenge protein modeling the modeling functional must be physically correct.

Up to this moment quantum mechanical theory remains the only proven tool for dealing with nonadditive phenomena. Undoubtedly, high-level ab initio quantum-chemical calculations, like second-order Møller–Plesset (MP2) and coupled clusters, provide a reliable technique for the prediction of the electronic structure and geometry of proteins; however, these methods place enormous requirements on computer resources and cannot be applied in a routine fashion to systems larger than 100 atoms.

While we see significant progress in the development of nonempirical methods [1] and fast progress in the improvement of computer hardware, resource requirement still remains a bottleneck in the path of routine industrial application.

Alternatively, semiempirical methods also employing quantum mechanical methodology could be a practical replacement for expensive nonempirical tools in solving the electronic structures of proteins. Unfortunately, modern semiempirical methods provide less accurate results for the prediction of protein structure in comparison with existing classical mechanics force-field methods. The situation might be improved by development of more physically correct semiempirical methods, which we believe this work may inspire.

There is growing interest in the semiempirical community to improve the situation with the insufficient accuracy. The publication of results of PM5 parame-

Contribution to the 8th Electronic Computational Chemistry Conference, 2002

Electronic Supplementary Material to this paper (full text of the lecture in html as given at the ECCC8 conference) can be obtained by using the SpringerLink server located at http://dx.doi.org/10.1007/s00214-002-0400-8.

Correspondence to: V. M. Anisimov
e-mail: victor@quantumbiochem.org

terization [2] shows progress toward better accuracy. The PM5 parameterization results in a fourfold improvement in the prediction of the absolute heat of formation; while other molecular properties, for example, the dipole moment, require further development.

Arguably, the progress achieved in the description of the properties of small molecules does not necessarily guarantee that the method will be as successful when applied to large protein systems. Plenty of protein-specific interactions are not reflected in the common semiempirical training sets, for example, long-range Coulomb interactions. Thus there is a great need in thorough verification of real capabilities of the semiempirical methods before their application to protein modeling.

Computational part

Accuracy comparison critically depends on the use of reliable reference data. Traditionally experiment was the sole source of the trustworthy prediction of molecular properties. The situation has changed recently with significant progress taking place in computer hardware and *ab initio* methods. Additionally high-level theoretical calculation can provide much more detailed information about the molecular electronic structure which far exceeds the capabilities of the best experimental techniques.

We took MP2 for our computational tests as a reference method among the other lower resource demanding post-Hartree–Fock methods as a reasonably reliable one.

In order to test capabilities of AM1 and PM3 in electrostatic property prediction we prepared a very large set of points around each amino acid where the molecular electrostatic potential (MESP) was evaluated by semiempirical and nonempirical methods. After generation of the cube grid it was modified in such way to ensure that the distance between the given grid point and the closest atom center was between 3 and 15 Å. Three distinct grid steps were set to 0.5, 1.0, and 2.0 Å for distance zones 3–4.5, 5–8, and 9–15 Å, respectively. The total number of grid points generated was in the range 12,000–15,000 points per molecule depending on geometric size of the amino acid. The large size and high density of the mesh guarantee accurate validation of the MESP predicting capabilities of the computational methods tested.

An error in the semiempirical MESP in comparison to the reference potential contributes to the error function, R:

$$R = \sqrt{\sum_{i}^{N} \frac{\left(U_i^{\mathrm{Ref}} - U_i\right)^2}{N}},$$

where N is number of MESP points around of the given amino acid and $U_{\mathrm{i}}^{\mathrm{Ref}}$ and U_{i} represent the MESP value at a point i calculated by reference and test methods, respectively.

In this work we investigate AM1 and PM3 semiempirical Hamiltonians and HF/6-31G* level of theory in their ability to describe quantitatively the electrostatic potential and the electric dipole moment for short peptide systems.

The semiempirical calculations were performed using AM1 [3] and PM3 [4] Hamiltonians as implemented in MOPAC6 program [5]. Nonempirical calculations were performed using the GAMESS program [6]. The molecules investigated in the current research are shown in Table 1. These are 20 natural amino acids terminated by acetyl and *N*-methyl groups in the β-sheet and α-helix conformation of the peptide molecules representing the main building blocks of proteins. The structure of these molecules was optimized at the HF/6-31G* level of theory [7, 8]. Single-point calculations were performed at the MP2 level of theory [9] with the 6-31 + Gx basis set augmented by diffuse orbitals and *d*-orbital polarization functions placed on heavy atoms. The *d*-orbital exponent was set to 0.2, which is denoted by the small "x" by analogy with the popular "*" symbol. As demonstrated elsewhere [10], the reduced *d*-orbital exponent is more suitable for description of long-range interactions in contrast to the conventional Pople exponent. Whenever an MP2 calculation with the 6-31 + Gx basis set was problematic we decreased the basis to a 6-31Gx one. For single-point calcuations of several relatively large molecules we used the B3LYP method [11, 12, 13] with the 6-31Gx basis set. All the calculations were performed on an IBM-compatible desktop computer.

Discussion

The long-range Coulomb potential plays an important role in protein structure and function; therefore, accurate description of the MESP is a critical component of the protein modeling method. Before involving ourselves in time-consuming calculations of proteins a lot can be learned from MESP research of single amino acids. If the computational method is unable to reproduce a MESP of these molecules with sufficient accuracy, then the calculation of large proteins will definitely result in unsatisfactory predictions. The larger the protein the bigger error is expected to be, because of the long-range nature of the Coulomb potential. Therefore, the aim of this work is a thorough investigation of electrostatic property prediction of AM1 and PM3 for single amino acids first.

The total root-mean-square (rms) errors of the MESP presented in Table 2 demonstrate that both AM1 and PM3 exhibit comparable accuracy, although PM3 shows slightly better results on average. Computationally expensive *ab initio* HF in the popular 6-31G* basis set gives only a moderate improvement of 1.4 times.

The computational results demonstrate that for all amino acids except positively charged arginine and histidine PM3 is more accurate than AM1. The accuracy of PM3 in this case is closer to HF in contrast to the common opinion about the better accuracy of AM1.

The test results taken from Table 2 are depicted in Fig. 1. The molecules with numbers 30–38 represent ionized amino acids. The improvement in accuracy of PM3 versus AM1 is better seen on neutral amino acids than for ionized ones. These errors of PM3 for ionized

Table 1. Gas-phase HF/6-31G* optimized structures of natural amino acids; *1* and *2* denote β-sheet and α-helix conformations, respectively

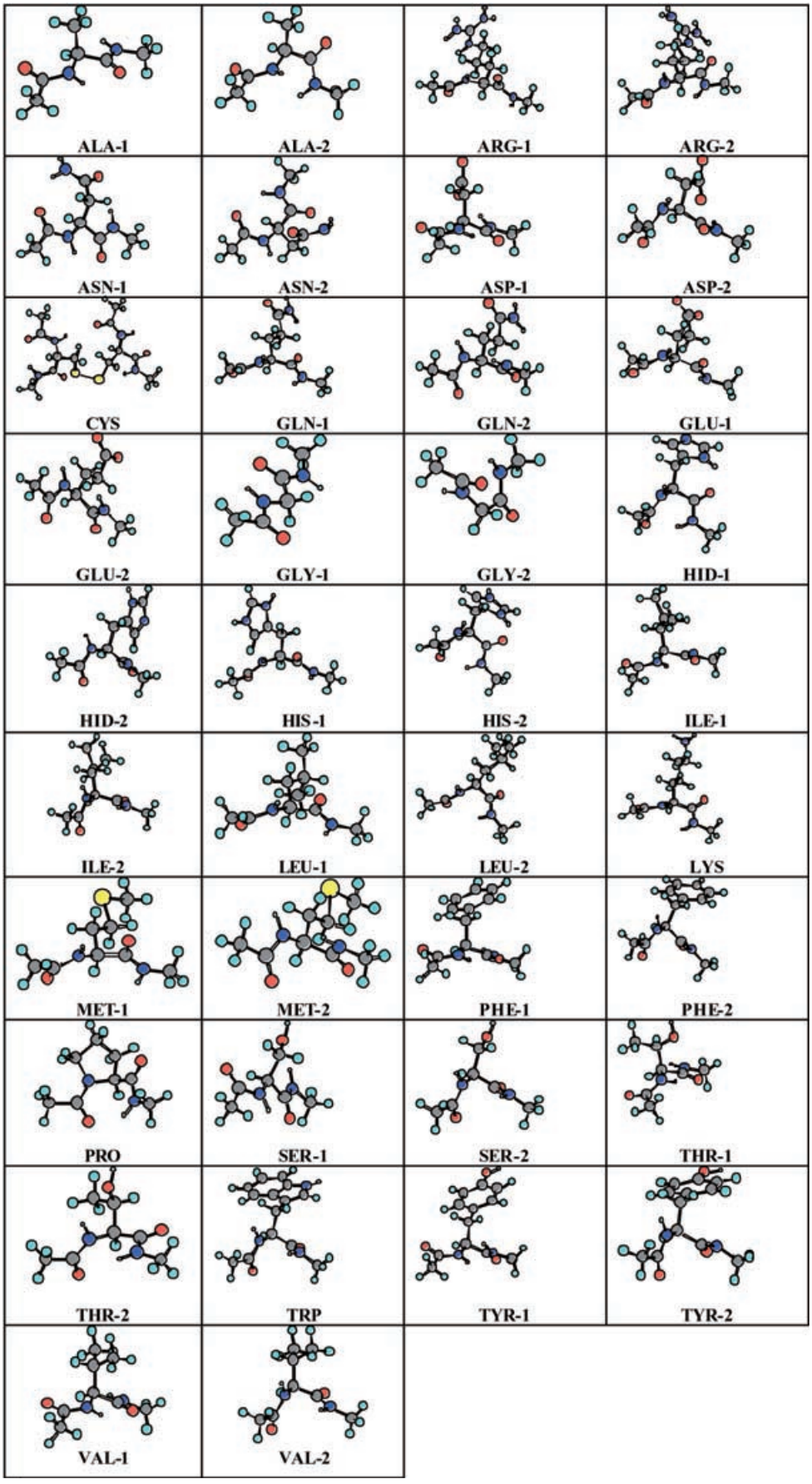

Table 2. Comparison of root mean square (*rms*) of the molecular electrostatic potential in atomic units calculated by AM1, PM3, and HF/6-31G* with MP2/6-31 + Gx level of theory

Number	Molecule	AM1	PM3	HF
Neutral amino acids				
1	Ala-1	0.00149	0.00117	0.00076
2	Ala-2	0.00123	0.00083	0.00081
3	Asn-1	0.00086	0.00101	0.00100
4	Asn-2	0.00096	0.00098	0.00099
5	Cys[a]	0.00115	0.00100	0.00116
6	Gln-1	0.00167	0.00134	0.00069
7	Gln-2	0.00104	0.00095	0.00074
8	Gly-1	0.00166	0.00124	0.00071
9	Gly-2	0.00133	0.00092	0.00084
10	Hid-1	0.00130	0.00119	0.00085
11	Hid-2	0.00158	0.00134	0.00101
12	Ile-1	0.00135	0.00109	0.00069
13	Ile-2	0.00103	0.00076	0.00086
14	Leu-1	0.00150	0.00117	0.00077
15	Leu-2	0.00114	0.00079	0.00097
16	Met-1[b]	0.00128	0.00095	0.00072
17	Met-2[b]	0.00093	0.00063	0.00115
18	Phe-1[a]	0.00111	0.00087	0.00077
19	Phe-2[b]	0.00086	0.00068	0.00131
20	Pro[a]	0.00072	0.00058	0.00097
21	Ser-1	0.00137	0.00090	0.00074
22	Ser-2	0.00132	0.00109	0.00110
23	Thr-1	0.00150	0.00109	0.00065
24	Thr-2	0.00120	0.00085	0.00082
25	Trp[a]	0.00110	0.00099	0.00105
26	Tyr-1[b]	0.00131	0.00103	0.00077
27	Tyr-2[b]	0.00098	0.00084	0.00132
28	Val-1	0.00131	0.00105	0.00069
29	Val-2	0.00104	0.00076	0.00087
Partial rms error		0.00124	0.00099	0.00091
Ionized amino acids				
30	Asp-1	0.00148	0.00126	0.00100
31	Asp-2	0.00143	0.00121	0.00093
32	Glu-1[c]	0.00166	0.00167	0.00069
33	Glu-2[c]	0.00160	0.00148	0.00081
34	Arg-1[b]	0.00163	0.00228	0.00082
35	Arg-2[b]	0.00178	0.00222	0.00080
36	His-1	0.00184	0.00231	0.00116
37	His-2	0.00148	0.00189	0.00086
38	Lys[b]	0.00173	0.00161	0.00077
Partial rms error		0.00163	0.00182	0.00088
Total rms error		0.00134	0.00123	0.00090

[a] B3LYP/6-31Gx method was used as a reference
[b] MP2/6-31Gx method was employed as a reference
[c] MP2/6-31Gx method with diffuse orbitals placed on oxygen atoms was employed as a reference

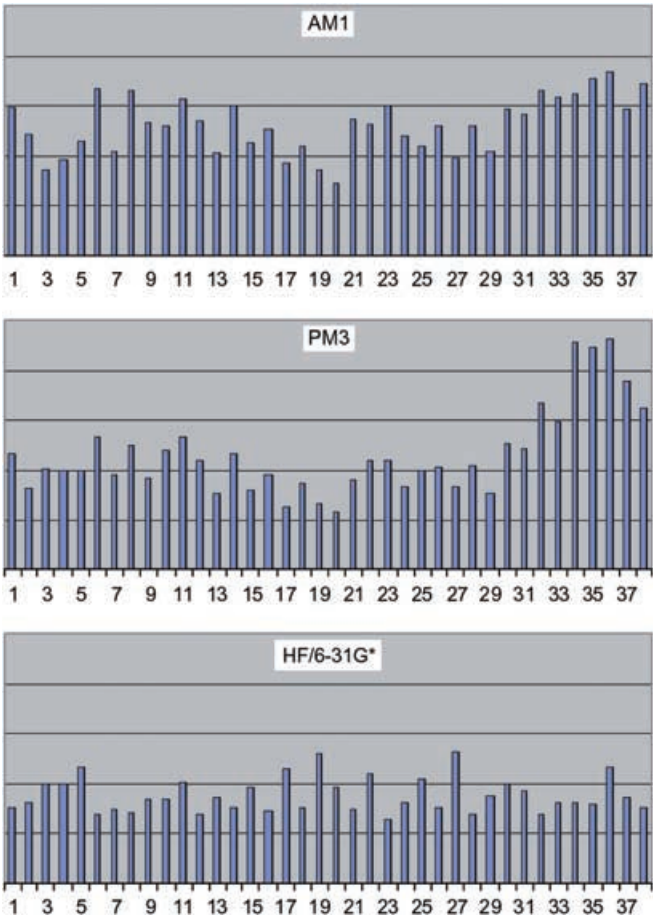

Fig. 1. *AM1*, *PM3*, and *HF/6-31G** molecular electrostatic potential (*MESP*) errors for neutral (1–29) and ionized (30–38) amino acids with reference to MP2/6-31 + Gx MESP data as presented in Table 2

amino acids could perhaps be explained by the majority of neutral molecules in the PM3 training set although ionized amino acids are a hard test for both semiempirical Hamiltonians. As expected, HF has no such problems with differentiation of neutral and ionized amino acids.

The observed better accuracy performance of the PM3 parameterization over the AM1 one for the neutral amino acids and the very small presence of the ionized ones in the real protein structure could lead to an arguable conclusion that PM3 will also be more accurate for real proteins. On the other hand, amino acid ionization and loss of the PM3 accuracy is expected for the protein outer shell contacting with solvent molecules and representing the most chemically active part of the protein. Currently there is not enough computational information to verify whether the errors of AM1 or PM3 will prevail and computational tests on real-size proteins are necessary.

The AM1 and PM3 MESP errors are relatively close to the HF/6-31G* ones but the significance of the close proximity should not be overestimated. As we could see, the average MESP error of the HF method measured for any of the 15 000 points is 0.00090 au, or 0.56 kcal/mol. The best semiempirical result shows an error of 0.00123 au, or 0.77 kcal/mol, for any point of the 3D space. These are quite big errors and accumulated they may result in a very big total error and incorrect prediction of the enzymatic reaction profile. In other words, the MESP accuracy shown by HF/6-31G* is not sufficient for protein modeling. This observation also questions the use of HF/6-31G* in the development of Coulomb interaction models for classical force-field methods [14, 15].

Further details regarding the electrostatic property prediction by the AM1, PM3, and HF methods can be learned from the analysis of their capabilities of calculation of the dipole moment of the amino acids. The results of these calculations are collected in Table 3.

The reference data are obtained as before from MP2/6-31 + Gx calculations.

The total rms error in the prediction of the dipole moment by AM1, PM3, and HF are 1.02, 0.83, and 0.71 D, respectively. Again the PM3 results are closer to the HF predictions than the AM1 ones. In spite of the average improvement achieved by PM3 in the dipole moment prediction it loses to AM1 for most of the ionized amino acids.

The close proximity of the AM1 and especially the PM3 dipole moments to the HF ones cannot be interpreted as a success of the modified neglect of differential overlap type Hamiltonians. More detailed analysis of semiempirical dipole moments reveals some hidden problems. The nature of these problems can be learned from comparison of the components of the dipole moment calculated by AM1 and PM3 and the reference method, MP2. The data for two selected amino acids,

Table 3. Comparison of dipole moment calculated by AM1, PM3, and HF/6-31G* with MP2/6-31 + Gx level of theory

Molecule	AM1	PM3	HF	MP2	AM1 Std. Error[d]	Real error[e]	Angle[f]	PM3 Std. Error[d]	Real error[e]	Angle[f]	HF Std. error[d]	Real error[e]	Angle[f]
Neutral amino acids													
Ala-1	2.02	2.44	2.58	3.17	1.15	1.15	2	0.73	0.74	2	0.59	0.59	1
Ala-2	4.15	4.24	4.90	4.51	0.36	0.75	9	0.27	0.50	6	−0.39	0.63	6
Asn-1	1.29	0.97	1.49	0.68	−0.60	0.61	5	−0.28	0.31	10	−0.81	0.82	8
Asn-2	3.88	3.86	4.54	4.08	0.19	0.61	8	0.21	0.42	5	−0.47	0.81	9
Cys [a]	3.67	3.58	4.01	3.14	−0.53	0.96	14	−0.45	0.68	9	−0.87	1.21	13
Gln-1	4.30	4.73	5.26	5.51	1.21	1.44	9	0.78	0.91	5	0.25	0.48	4
Gln-2	2.28	2.32	2.75	2.55	0.27	0.62	13	0.23	0.37	7	−0.20	0.52	10
Gly-1	2.16	2.60	2.87	3.42	1.26	1.26	2	0.82	0.82	2	0.54	0.54	1
Gly-2	4.41	4.48	5.25	4.86	0.45	0.76	8	0.38	0.53	5	−0.38	0.64	6
Hid-1	4.95	5.16	5.60	5.65	0.70	0.76	3	0.48	0.52	2	0.05	0.58	6
Hid-2	8.77	9.09	10.03	9.55	0.78	1.07	5	0.46	0.81	4	−0.48	0.75	3
Ile-1	1.92	2.31	2.45	2.99	1.08	1.10	5	0.69	0.72	4	0.54	0.55	3
Ile-2	5.14	5.23	5.92	5.34	0.20	0.64	7	0.11	0.41	4	−0.58	0.73	4
Leu-1	1.50	1.90	2.11	2.74	1.24	1.25	4	0.84	0.86	5	0.62	0.63	3
Leu-2	5.34	5.42	6.18	5.53	0.19	0.73	7	0.12	0.47	5	−0.65	0.81	5
Met-1[b]	3.06	3.62	3.59	4.01	0.95	1.05	7	0.39	0.51	5	0.42	0.50	4
Met-2[b]	3.73	3.64	4.62	3.82	0.10	0.58	9	0.19	0.35	5	−0.80	0.94	7
Phe1[b]	1.96	2.33	2.46	2.88	0.92	0.93	2	0.56	0.58	4	0.42	0.43	1
Phe-2[b]	5.33	5.42	6.28	5.40	0.06	0.52	5	−0.02	0.32	3	−0.88	0.98	4
Pro[b]	2.99	2.86	3.34	2.60	−0.39	0.40	2	−0.26	0.34	5	−0.74	0.76	4
Ser-1	2.11	2.55	2.73	3.05	0.94	1.04	10	0.50	0.50	1	0.33	0.57	9
Ser-2	6.76	6.74	8.05	7.32	0.57	0.86	5	0.58	0.67	3	−0.73	0.84	3
Thr-1	2.46	2.88	3.27	3.56	1.10	1.18	9	0.67	0.69	3	0.28	0.48	7
Thr-2	3.70	3.86	4.46	3.95	0.25	0.79	11	0.09	0.47	7	−0.50	0.67	6
Trp[a]	4.90	5.07	5.92	5.54	0.64	0.77	5	0.47	0.55	3	−0.38	0.64	5
Tyr-1[b]	2.55	2.95	3.25	3.67	1.12	1.21	9	0.71	0.74	3	0.42	0.53	5
Tyr-2[b]	5.56	5.57	6.53	5.64	0.08	0.74	7	0.08	0.49	5	−0.89	1.04	5
Val-1	1.91	2.29	2.40	2.92	1.01	1.03	5	0.63	0.66	4	0.52	0.54	3
Val-2	5.21	5.30	6.07	5.51	0.31	0.65	6	0.21	0.41	4	−0.56	0.70	4
Partial rms error					0.75	0.91	7	0.49	0.59	5	0.57	0.71	6
Ionized amino acids													
Asp-1	10.03	9.99	9.80	9.39	−0.64	1.14	6	−0.60	0.96	4	−0.41	0.81	4
Asp-2	9.48	9.29	8.96	8.53	−0.95	1.03	3	−0.76	0.91	3	−0.43	0.74	4
Glu-1[c]	15.15	15.58	14.80	14.58	−0.57	1.19	4	−1.00	1.24	3	−0.22	0.57	2
Glu-2[c]	17.10	17.45	16.98	16.64	−0.46	0.91	3	−0.81	1.08	2	−0.33	0.67	2
Arg-1[b]	18.04	18.39	17.56	17.05	−0.99	1.37	3	−1.34	1.50	2	−0.51	0.73	2
Arg-2[b]	16.17	16.25	14.94	14.79	−1.38	1.54	3	−1.46	1.55	2	−0.15	0.55	2
His-1	12.29	12.50	11.83	10.96	−1.33	1.64	5	−1.54	1.71	4	−0.87	1.06	3
His-2	10.85	10.82	10.24	9.82	−1.03	1.06	1	−1.00	1.09	2	−0.43	0.51	2
Lys[b]	19.62	19.71	18.11	18.08	−1.53	1.67	2	−1.63	1.70	2	−0.02	0.52	2
Partial rms error					1.05	1.31	4	1.18	1.34	3	0.44	0.70	3
Total rms error					0.83	1.02	7	0.71	0.83	4	0.54	0.71	5

[a] B3LYP/6-31Gx method was used as a reference
[b] MP2/6-31Gx method was employed as a reference
[c] MP2/6-31Gx method with diffuse orbitals placed on oxygen atoms was employed as a reference
[d] The error in the total dipole moment, an old standard for assessment of the quality of the dipole moment prediction
[e] The error in the components (orientation) of the dipole moment. Calculated as the square root of $(x_r - x)^2 + (y_r - y)^2 + (z_r - z)^2$, where x_r,y_r,z_r are components of the dipole from the reference method (MP2) and x,y,z are components of the dipole from the semiempirical or *ab initio* Hartree–Fock calculation
[f] The angle in degrees between the test and the reference dipole moment vectors

Table 4. Components of dipole moment of Phe-2 calculated by the AM1, PM3, and MP2 methods

Method Error	*x*	y	*z*	Total
MP2	−1.2857	−1.5360	−5.0107	5.3963
AM1	−1.4064	−1.9669	−4.7529	5.3326
AM1 error	0.1207	0.4309	−0.2578	0.0637
PM3	−1.3420	−1.8338	−4.9169	5.4166
PM3 error	0.0563	0.2978	−0.0938	−0.0203

Table 5. Components of dipole moment of Tyr-2 calculated by AM1, PM3, and MP2 methods

Method Error	*x*	*y*	*z*	Total
MP2	−1.4778	−0.4337	−5.4278	5.6421
AM1	−1.8265	−1.0215	−5.1565	5.5649
AM1 error	0.3487	0.5878	−0.2713	0.0772
PM3	−1.5728	−0.8895	−5.2641	5.5655
PM3 error	0.0950	0.4558	−0.1637	0.0766

phenylalanine and tyrosine, are shown in Tables 4 and 5, respectively.

As can be seen from Tables 4 and 5 the total dipole errors are very small. This kind of result was traditionally considered as a good achievement; however, such analysis based on the comparison of the total dipole only is incomplete, because the dipole is a vector value and hence has orientation, which is ignored here.

The errors in the dipole moment components are substantially larger than the errors in the total dipole. This indicates problems in the prediction of the dipole orientation of the Phe-2 and Tyr-2 samples, while the test and reference vectors have fortuitously the same length. And the error is quite significant, almost 1 order of magnitude difference between the total dipole error and the error in the dipole moment components. In other words, the semiempirical and the reference dipole vectors do not have coinciding orientations. Other examples where AM1 and PM3 show considerable errors in the components of the dipole moment are the Asn-2, Gln-2, Ile-2, Leu-2, Met-2, Thr-2, and Val-2 molecules.

This observation seems to have its natural explanation in the traditional technique of the semiempirical parameter optimization. Normally the dipole moment value for parameterization is taken from experiment, which understandably does not provide vector orientation. Therefore, the error not visible for small molecules included into the training set becomes a cause of errors for larger molecules. The bigger the protein, the bigger the error in the dipole moment to be expected.

A better criterion of the assessment of the error in the dipole moment is the length of the error vector between the reference and the semiempirical dipole vectors. These error values are given in Table 3 as real errors along with traditionally evaluated errors in the total dipole moment. The computational data obtained demonstrate that real errors are systematically larger than the traditionally expected errors. The column in Table 3 after the real error is the angle in degrees between the semiempirical and the reference dipole vectors. As the calculations demonstrate, both AM1 and PM3 show nonsystematic deviations from the accurate dipole vector orientation, although PM3 performs better on average. However, there are examples where AM1 dipole moments are more accurate.

Surprisingly big errors are observed in the dipole moment orientation for the HF/6-31G* level of theory. Here the total rms error in the angle is 5°. The maximum error of 13° is observed for the cysteine dimer. Augmentation of the 6-31G* basis set with additional diffuse and polarization functions did not improve the situation; therefore, we connect the accuracy problem with the lack of effects of electron correlation. This finding may have some impact on the development of Coulomb interaction parameters for molecular mechanics methods where HF/6-31G* predicted electrostatic properties are employed as the *ab initio* standard [14,15].

Conclusions

The results obtained in this work lead to the following conclusions:

1. AM1 and PM3 semiempirical Hamiltonians were tested to evaluate their applicability to describe quantitatively electronic properties of protein systems. The PM3 method was found to provide better quality electrostatic maps and dipole moments than AM1 for neutral amino acids, although the PM3 parameterization shows larger dipole moment errors for ionized molecules. The latter observation still leaves open a question about the relative accuracy of AM1 and PM3 for real complex proteins containing both neutral and ionized amino acids.
2. While reproducing the total dipole length quite well, both AM1 and PM3 have certain problems in the prediction of the dipole orientation of amino acids. Similar errors are observed for the HF/6-31G* level of theory. In order to avoid common mistakes in the assessment of the quality of the dipole moment, the analysis of errors should be based on the comparison of the dipole as a vector value rather than a scalar one.
3. The large errors of the HF/6-31G* level of theory in the prediction of molecular electrostatic potential and the dipole moment of amino acids justify that the accurate prediction of the electrostatic properties of proteins requires effects of electron correlation to be taken into account.

Acknowledgements. The authors are thankful to Jerzy M. Rudzinski and James J.P. Stewart for their support.

References

1. Scuseria GE (1999) J Phys Chem A 103: 4782–4790
2. Stewart JJP (2001) 86. Optimization of parameters for NDDO methods. Book of abstracts. ACS National Meeting, Chicago, IL, 26–30 August 2001

3. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107: 3902–3909
4. Stewart JJP (1989) J Comput Chem 10: 209–220
5. Stewart JJP (1990) J Comput-Aided Mol Des 4: 1–105
6. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JJ, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14: 1347–1363
7. Hehre WJ, Ditchfield R, Pople JA (1972) J Chem Phys 56: 2257–2261
8. Pietro WJ, Francl MM, Hehre WJ, DeFrees DJ, Pople JA, Binkley JS (1982) J Am Chem Soc 104: 5039–5048
9. Pople JA, Binkley JS, Seeger R (1976) Int J Quantum Chem Symp 10: 1–19
10. Sponer J, Hobza P, Leszczinski J (1996) Hydrogen bonding and stacking interactions of DNA bases: Nonadditivity of interactions, importance of higher-order electron correlation contributions, and comparison with empirical potentials. Book of abstracts. 5th conference on current trends in computational chemistry (CCTCC-5), Jackson State University, Vicksburg, MS, 1–2 November 1996
11. Becke AD (1993) J Chem Phys 98: 5648–5642
12. Stephens PJ, Devlin FJ, Chablowski CF, Frisch MJ (1994) J Phys Chem 98: 11623–11627
13. Hertwig RH, Koch W (1997) Chem Phys Lett 268: 345–351
14. Kuyper LF, Hunter RN, Ashton D, Merz KM Jr, Kollman PA (1991) J Phys Chem 95: 6661–6666
15. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) J Phys Chem B 102: 3586–3616