(i) the structures of the bases *as seen by water* are quite different from the structures of the base *as seen*, for example, *by x rays;* and (ii) concept like hydrophobicity (see, for example, Figure 3, insert a) can be put on more quantitative grounds because of the availability of potentials of the type given in Tables V and VI.

References and Notes

(1) This work has been partially presented at the International Symposium of Theoretical Chemistry, Boulder, Col., June 1975, and at the VI Simposio dei Chimici Teorici di Lingua Latina, Arles, France, September 1975.
(2) E. Clementi, F. Cavallone, and R. Scordamaglia, *J. Am. Chem. Soc.*, preceding paper in this issue.
(3) G. Bolis and E. Clementi, *J. Am. Chem. Soc.*, following paper in this issue.
(4) E. Clementi, "Lecture Notes in Chemistry", Vol. 2, Springer-Verlag, Heidelberg, W. Germany, 1976.
(5) B. Pullman and A. Pullman, *Q. Rev. Biophys.*, 7, 4 (1975).
(6) E. Clementi and D. R. Davies, *J. Comput. Phys.*, **1**, 223 (1966); E. Clementi, *Int. J. Quantum Chem., Symp.*, **No. 1**, 307 (1967). See in addition E. Clementi and J. Mehl, Special IBM Technical Reports: (a) IBMOL-V Program's System Manual (June 1975); (b) IBMOL-V Program's Users Manual (July 1971). E. Clementi, *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 2942 (1972).
(7) E. Clementi, J. M. André, M. Cl. André, D. Klint, and D. Hahn, *Acta Phys. Acad. Sci. Hung.*, **27**, 493 (1968). The same basis set has been used to study the hydrogen bond between base pairs: E. Clementi, J. Mehl, and W. von Niessen, *J. Chem. Phys.*, **54**, 508 (1971); in this work the basis set has been compared to a larger basis set.
(8) M. Spencer, *Acta Crystallogr.*, **12**, 59, 66 (1959).
(9) R. S. Mulliken, *J. Chem. Phys.*, **23**, 1883, 1841, 2338, 2343 (1955). The electronic charges we have computed for water are 8.6221 for the oxygens atom and 0.6889 for each one of the hydrogen atoms. The dependence of the NCH on the basis set has been analyzed for the case of water by E. Clementi and H. Popkie, *J. Chem. Phys.*, **57**, 1077 (1972).
(10) E. Clementi, *J. Chem. Phys.*, **46**, 3842 (1967); E. Clementi and A. Routh, *Int. J. Quantum. Chem.*, 6, 525 (1972); H. Popkie and E. Clementi, *J. Chem. Phys.*, **57**, 4870 (1972).
(11) The geometry of the complex "amino acid-water" is defined by giving the *x, y, z* coordinates of each atom (relative to a selected Cartesian frame of axes). The coordinates for the atoms in each amino acid are given in Table I-IV. The coordinates of the water molecule are given in a set of tables, where we report a code number for each computation, the *x, y, z* coordinates of the two hydrogen atoms and of the oxygen atom, and the computed total energy (in atomic units) for the amino acid-water complex. These tables are available as a Montedison Technical Report, DDC-766, February 1976; a copy of the report can be secured by request to the authors of this paper.

Analytical Potentials from "ab Initio" Computations for the Interaction between Biomolecules. 3. Reliability and Transferability of the Pair Potentials[1a]

G. Bolis[1b,c] and E. Clementi*[1d]

Contribution from the Università degli Studi di Milano, Facoltà di Fisica, 20100 Milan, Italy, and the Società Montedison, Istituto Ricerche "G. Donegani", 28100 Novara, Italy. *Received August 30, 1976*

Abstract: The pair potentials reported in the previous papers of this series have been analyzed using phenylalanine as a test case. For this amino acid (the last one of the naturally occurring amino acids to be studied) we have computed (in the SCF-LCAO-MO approximation) the interaction with a molecule of water placed at 75 different positions (or orientations) relative to phenylalanine. The computed interaction energies can be duplicated to a reasonable level of accuracy by making use of the previously reported pair potentials (with an error of about 1 kcal/mol); therefore, we conclude that our pair potentials are, to some extent, transferable to molecules that are chemically similar to those previously studied. We have discussed problems related to: (1) more refined selections of the points needed to sample the multi-dimensional interaction energy surface; and (2) the need to select a more physical form for the pair potential.

In the previous papers of this series we have presented a number of pair potentials to describe the interaction energy of a molecule of water (W) with several molecules (M); in particular we have considered 21 out of the 22 naturally occurring amino acids[2] and the four bases of DNA.[3]

In this paper we shall discuss some aspects related to the *reliability* of the pair potentials for those molecules we have considered; we shall, in addition, investigate the possibility of using the potentials to describe the interaction energy for molecules not explicitly considered in our work (i.e., *transferability* of the pair potentials). As a test case for the transferability we shall use the amino acid phenylalanine, previously not studied in this series. In Table I and in Figure 1 we report the geometry of phenylalanine; since the geometry given in ref 4 refers to the zwitterionic form, we have introduced in Table I and Figure 1 the necessary coordinate variations for the $-NH_2$ and -COOH groups. In Table I, we give a code name for each atom, a "class" specification (see ref 2 and 3 and following discussions), the x, y, z coordinates for each atom (in atomic units), the computed net charges,[5] NCH, and the computed molecular orbital valency state energies,[6] MOVS. The basis set used to compute the phenylalanine molecules, the water molecules, and the phenylalanine-water complexes has been discussed in ref 2. The total energy for the water molecules, E(W), and for the phenylalanine, E(M), are $E(W) = -75.733\,057$ and $E(M) = -549.373\,069$ au, respectively.

For the complex we have considered 75 different positions or orientations of the water relative to phenylalanine. In Table II we report a running index (1-75) to identify the computation of the complex, the x, y, z coordinates of the atoms for water (the phenylalanine is kept fixed in space at the position specified in Table I), the total energy of the complex E(M,W) expressed in atomic units, the interaction energy $I(M,W) = E(M,W) - E(M) - E(W)$, expressed in kilocalories per mole, and in the last two columns two different approximations to I(M,W), discussed later in this paper.

In the preceding papers of this series we have not reported the equivalent of Table II, for obvious reasons of space (about 2000 cases have been studied); however, such tables are available to the interested reader as specified in ref 2 and 3. For these computations we have used a computer program, named IBMOL.[7]

Table I. Geometry,[a] Charges, MO Valency State for Phenylalanine[b]

Atom	Class	x	y	z	Charge	MOVS
O(2)	9	4.499	11.490	3.621	−0.51	0.50
O(1)	10	2.391	8.694	1.328	−0.38	0.43
N	11	2.218	5.312	5.001	−0.52	0.91
C(B)	7	6.308	7.188	6.524	−0.38	1.06
C(G)	19	8.024	5.801	4.547	−0.01	1.16
C(D1)	18	9.253	7.151	2.671	−0.22	1.12
C(E1)	18	10.763	5.902	0.936	−0.21	1.13
C(2)	18	11.045	3.303	1.078	−0.21	1.12
C(E2)	18	9.816	1.954	2.954	−0.21	1.13
C(D2)	18	8.306	3.203	4.688	−0.21	1.12
C′	5	3.374	9.249	3.217	0.48	1.37
C(A)	8	3.588	7.676	5.559	−0.15	1.08
H(B1)	2	7.166	8.980	6.987	0.21	0.29
H(B2)	2	6.204	6.044	8.211	0.21	0.29
H(D1)	16	9.034	9.177	2.561	0.20	0.29
H(E1)	16	11.722	6.954	−0.527	0.21	0.30
H(2)	16	12.225	2.329	−0.274	0.21	0.30
H(E2)	16	10.033	−0.073	3.064	0.21	0.30
H(D2)	16	7.349	2.150	6.152	0.21	0.30
H(1)	1	3.081	4.386	3.401	0.24	0.28
H(2)	1	0.272	5.741	4.568	0.23	0.28
H(A)	2	2.741	8.637	7.149	0.22	0.29
H(O2)	4	4.303	12.412	2.114	0.37	0.31

[a] The geometry has been taken from G. V. Gurskaya, *Sov. Phys.-Crystallogr.*, **9,** 6 (1965). [b] Total energy −549.373 au.

Results and Discussion

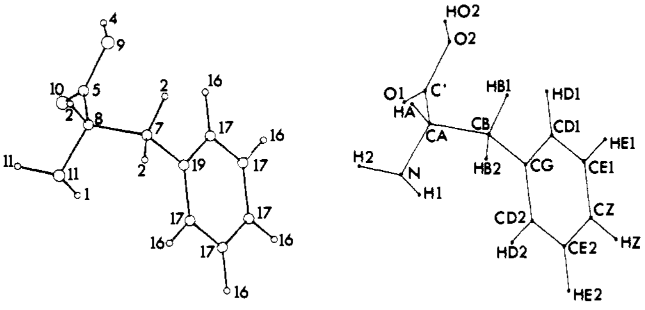
On the Form of the Pair Potentials. The computed interaction energies have been fitted with pair potentials of a very simple analytical form. If we designate with i (or k) an atom on M, and with j (or l) an atom on W, our simple form is a modified Lennard-Jones potential, namely

$$I(\mathrm{M,W}) = \sum_i \sum_j I_{ij} = \sum_i \sum_j [-A_{ij}(r_{ij})^{-6} + B_{ij}(r_{ij})^{-12} + C_{ij}q_iq_j(r_{ij})^{-1}]$$

In this expression A, B, and C are fitting constants; q_i and q_j are the computed net charges for the i and j atom in the separated molecules; and r_{ij} is the distance between the atoms of the pair i and j. In the above pair potentials the constants A, B, and C are different for *each pair*, i and j; such potential is quite flexible, computationally very fast even for Monte Carlo computations.[2] However, the use of a different set of constants for each pair is somewhat redundant (and would complicate the fitting to an unbearable limit for large molecular systems).

Indeed, it is reasonable to assume that the set of constants A, B, and C for a pair of atoms i and j can be used for other pairs, e.g., k and l, under fulfillment of some condition. Clearly, this would be the case for complexes with symmetry; however, most molecules of interest to us have no symmetry. As a first approximation we could state that if $q_i = q_k$ and if $q_j = q_l$ and if i and k as well j and l have the same atomic numbers Z, then both pairs could be described by the same set of constants A, B, and C. The drawback of these conditions lies in the practical difficulty to define the exact meaning of "*equality*" between net charges (equal to the first, or to the second, or to the third significant figure?).

"Chemical evidences" could be used to decide when a pair can be considered equal to another pair; however, this criterium presents the drawback that it is not easy to find the pertinent experimental data for the interaction of one biomolecule with one molecule of water in the gas phase and at low pressure. (Even if such data would be available it is not too clear, if we could make use of it: we wish to fit *our computed SCF-LCAO-MO energies* and such energies are far from being "exact" computations obtained from the "exact" wave function. In other words, the criteria should be selected from a representation that has the same "accuracy" as the set of data to be fitted.)



PHENYLALANINE

Figure 1. Geometry and identification code for the atoms in phenylalanine. This amino acid is described two times. The first time, the atom is identified with the same alphanumerical identification code adopted in Table I reporting the molecular geometry. The second time, each atom is identified by a numerical code which designates the class in which the atom belongs; such class code is given, in addition, in Table I reporting the molecular geometry.

These comments are presented to indicate that likely we must use a compromising attitude in selecting criteria for the applicability of a pair potential to different pairs of interacting atoms. In addition, since we intend to bring to conclusion a rather laborious working plan (see ref 2) we must find criteria that are amenable of applicability to the potential surfaces for much larger complexes than those examined up to now.[8]

To conclude, we have selected the following criteria, that are to be taken in a rather flexible way, mainly as indications, rather than dogma: (a) the net charges, NCH, provide an indication of the electronic density of an atom in the molecule; (b) the energies of the molecular orbital valency state, MOVS, provide an indication of the energy of an atom in the molecule; (c) "chemical evidences" should not be contradicted (but not necessary followed); and (d) the numerical values obtained by the fitting technique for the set of constants A, B, and C are

Table II. Cartesian Coordinates and Interaction Energies (in atomic units) of the Complex Water-Phenylalanine

Case	Atom	x^a	y^a	z^a	Energy^a	BE^b	GF^b	SF^b
1	O	0.853 26	7.867 07	−1.705 67	−625.027 60	49.32		
	H(1)	0.013 55	6.324 06	−2.136 98				
	H(2)	0.720 39	8.886 99	−3.193 67				
2	O	0.853 26	7.867 07	−1.705 67	−624.979 51	79.45		
	H(1)	0.986 12	6.847 16	−0.217 67				
	H(2)	1.692 96	9.410 08	−1.274 35				
3	O	0.633 65	7.748 96	−2.139 05	−625.074 44	19.89		
	H(1)	−0.206 05	6.205 95	−2.570 37				
	H(2)	0.500 78	8.768 87	−3.627 05				
4	O	0.633 65	7.748 96	−2.139 05	−625.068 60	23.53		
	H(1)	0.766 51	6.729 04	−0.651 05				
	H(2)	1.473 35	9.291 97	−1.707 74				
5	O	0.633 65	7.748 96	−2.139 05	−625.073 74	20.33		
	H(1)	−0.676 08	6.884 20	−1.239 70				
	H(2)	0.030 76	9.447 13	−2.296 38				
6	O	0.633 65	7.748 96	−2.139 05	−624.074 58	19.83		
	H(1)	1.236 54	6.050 78	−1.981 72				
	H(2)	1.943 37	8.613 71	−3.038 41				
7	O	−0.244 78	7.276 49	−3.872 59	−625.102 66	2.17	1.97	2.34
	H(1)	−1.084 48	5.733 48	−4.303 90				
	H(2)	−0.377 64	8.296 41	−5.360 58				
8	O	−0.244 78	7.276 49	−3.872 59	−625.110 41	−2.69	−2.63	−2.48
	H(1)	−0.111 92	6.256 58	−2.384 59				
	H(2)	0.594 92	8.819 51	−3.441 27				
9	O	−0.244 78	7.276 49	−3.872 59	−625.104 24	1.18	0.82	1.15
	H(1)	−1.554 51	6.411 74	−2.973 23				
	H(2)	−0.847 67	8.974 67	−4.029 91				
10	O	0.244 78	7.276 49	−3.872 59	−625.104 77	0.84	0.12	0.45
	H(1)	0.358 11	5.578 32	−3.715 26				
	H(2)	1.064 95	8.141 25	−4.771 94				
11	O	−6.393 77	3.969 25	−16.007 33	−625.106 10	0.01	0.05	0.07
	H(1)	−7.233 47	2.426 25	−16.438 65				
	H(2)	−6.526 64	4.989 16	−17.495 33				
12	O	−6.393 77	3.969 25	−16.007 33	−625.106 03	0.05	−0.09	−0.08
	H(1)	−6.260 91	2.949 34	−14.519 33				
	H(2)	−5.554 07	5.512 25	−15.576 01				
13	O	7.523 59	11.603 84	10.028 43	−625.096 41	6.086		
	H(1)	5.881 66	11.223 56	9.372 03				
	H(2)	8.634 98	10.481 01	9.147 81				
14	O	7.523 59	11.603 84	10.028 43	−625.106 03	0.06	−0.32	−0.29
	H(1)	6.412 14	12.726 69	10.909 06				
	H(2)	9.165 59	11.984 11	10.684 82				
15	O	7.643 39	11.943 25	10.375 49	−625.103 10	1.90	2.20	1.86
	H(1)	6.001 38	11.562 98	9.719 10				
	H(2)	8.754 85	10.820 40	9.494 86				
16	O	7.643 39	11.943 25	10.375 49	−625.106 82	−0.43	0.05	−0.81
	H(1)	6.531 73	13.066 15	11.256 14				
	H(2)	9.285 58	12.323 47	11.031 87				
17	O	7.643 39	11.943 25	10.375 49	−625.107 10	−0.61	0.01	−0.16
	H(1)	6.125 94	11.554 07	11.279 88				
	H(2)	8.879 66	10.811 42	11.055 62				
18	O	7.643 39	11.943 25	10.375 49	−625.107 41	−0.80	−0.97	−0.79
	H(1)	6.407 11	13.075 08	9.695 36				
	H(2)	9.160 83	12.332 44	9.471 10				
19	O	8.122 58	13.300 88	11.763 74	−625.106 30	−0.11	0.52	0.38
	H(1)	6.480 52	12.920 62	11.107 35				
	H(2)	9.234 10	12.178 01	10.883 10				
20	O	8.122 58	13.300 88	11.763 74	−625.106 25	−0.07	−0.62	−0.57
	H(1)	7.011 00	14.423 76	12.644 38				
	H(2)	9.764 71	13.681 11	12.420 12				
21	O	8.122 58	13.300 88	11.763 74	−625.106 19	−0.03	0.07	−0.03
	H(1)	6.605 14	12.911 69	12.668 12				
	H(2)	9.358 86	12.169 03	12.443 88				
22	O	8.122 58	13.300 88	11.763 74	−625.106 37	−0.15	−0.34	−0.42
	H(1)	6.886 30	14.432 71	11.083 61				
	H(2)	9.640 02	13.690 07	10.859 34				
23	O	11.476 93	22.804 25	21.481 46	−625.106 06	0.04	0.04	0.02
	H(1)	9.834 94	22.423 98	20.825 07				
	H(2)	12.588 40	21.681 40	20.600 83				
24	O	11.476 93	22.804 25	21.481 46	−625.106 15	−0.01	−0.01	−0.02
	H(1)	10.365 48	23.927 10	22.362 09				
	H(2)	13.118 94	23.184 52	22.137 86				

Table II (*continued*)

Case	Atom	x^a	y^a	z^a	Energy[a]	BE[b]	GF[b]	SF[b]
25	O	7.510 74	5.354 25	−2.144 71	−625.090 39	9.91		
	H(1)	9.050 28	4.673 40	−2.806 75				
	H(2)	6.472 75	5.627 81	−3.600 65				
26	O	7.510 74	5.354 25	−2.144 71	−625.052 89	33.44		
	H(1)	8.548 74	5.080 69	−0.688 77				
	H(2)	5.971 21	6.035 09	−1.482 66				
27	O	7.623 99	5.262 28	−2.622 95	−625.096 80	5.83		
	H(1)	9.163 53	4.581 44	−3.285 00				
	H(2)	6.586 00	5.535 85	−4.078 89				
28	O	7.623 99	5.262 28	−2.622 95	−625.076 08	18.88		
	H(1)	8.661 98	4.988 72	−1.167 01				
	H(2)	6.084 46	5.943 13	−1.960 90				
29	O	7.623 99	5.262 28	−2.622 95	−625.097 31	5.52		
	H(1)	8.502 97	3.761 40	−2.126 19				
	H(2)	5.925 44	4.715 80	−2.920 09				
30	O	7.623 99	5.262 28	−2.622 95	−625.085 62	12.86		
	H(1)	9.322 55	5.808 75	−2.325 81				
	H(2)	6.745 02	6.763 17	−3.119 71				
31	O	8.076 99	4.894 42	−4.535 93	−625.104 32	1.13	1.88	1.47
	H(1)	9.616 53	4.213 58	−5.197 97				
	H(2)	7.039 00	5.167 99	−5.991 86				
32	O	8.076 99	4.894 42	−4.535 93	−625.108 26	−1.34	−0.86	−0.95
	H(1)	9.114 98	4.620 86	−3.079 99				
	H(2)	6.537 46	5.575 27	−3.873 88				
33	O	8.076 99	4.894 42	−4.535 93	−625.106 44	−0.19	−0.17	−0.54
	H(1)	8.955 97	3.393 56	−4.039 17				
	H(2)	6.378 44	4.347 96	−4.833 06				
34	O	8.076 99	4.894 42	−4.535 93	−625.104 45	1.04	1.54	1.16
	H(1)	9.775 54	5.440 88	−4.238 79				
	H(2)	7.198 01	6.395 30	−5.032 68				
35	O	11.247 97	2.319 42	−17.926 75	−625.106 13	−0.00	−0.01	0.02
	H(1)	12.787 51	1.638 57	−18.588 80				
	H(2)	10.209 98	2.592 98	−19.382 69				
36	O	11.247 97	2.319 42	−17.926 75	−625.106 06	0.03	−0.08	−0.03
	H(1)	12.285 95	2.045 86	−16.470 82				
	H(2)	9.708 45	3.000 25	−17.264 70				
37	O	15.637 92	4.175 01	4.147 61	−625.105 26	0.54	0.30	0.47
	H(1)	14.464 87	2.814 56	3.935 21				
	H(2)	14.651 73	5.668 96	3.887 66				
38	O	15.637 92	4.175 01	4.147 61	−625.101 04	3.18	3.56	3.42
	H(1)	16.624 11	2.681 07	4.407 57				
	H(2)	16.810 98	5.535 48	4.360 01				
39	O	16.125 48	4.144 87	4.254 27	−625.107 48	−0.85	−1.68	−1.24
	H(1)	14.952 46	2.784 93	4.041 87				
	H(2)	15.139 26	5.638 32	3.994 32				
40	O	16.125 48	4.144 87	4.254 27	−625.102 91	2.01	2.78	2.36
	H(1)	17.111 68	2.650 95	4.514 23				
	H(2)	17.298 54	5.505 33	4.466 67				
41	O	16.125 48	4.144 87	4.254 27	−625.104 84	0.80	0.60	0.29
	H(1)	15.797 52	2.751 05	5.359 58				
	H(2)	15.984 38	5.605 45	5.312 02				
42	O	16.125 48	4.144 87	4.254 27	−625.103 88	1.40	1.90	1.52
	H(1)	16.266 59	2.684 30	3.196 52				
	H(2)	16.453 45	5.538 70	3.148 97				
43	O	18.075 73	4.024 31	4.680 91	−625.107 25	−0.70	−1.32	−1.01
	H(1)	16.902 76	2.665 17	4.468 49				
	H(2)	17.089 45	5.516 95	4.420 97				
44	O	18.075 73	4.024 31	4.680 91	−625.104 87	0.78	0.87	0.89
	H(1)	19.061 98	2.531 35	4.940 85				
	H(2)	19.248 72	5.383 80	4.893 33				
45	O	18.075 73	4.024 31	4.680 91	−625.106 14	−0.00	−0.17	−0.17
	H(1)	17.747 76	2.630 48	5.786 22				
	H(2)	17.934 62	5.484 89	5.738 66				
46	O	18.075 73	4.024 31	4.680 91	−625.105 48	0.40	0.46	0.44
	H(1)	18.216 83	2.563 74	3.623 16				
	H(2)	18.403 69	5.418 14	3.575 60				
47	O	31.727 42	3.180 37	7.667 37	−625.105 83	0.18	−0.05	0.02
	H(1)	30.554 46	1.821 22	7.454 95				
	H(2)	30.741 15	4.673 01	7.407 43				
48	O	31.727 42	3.180 37	7.667 37	−625.105 83	0.18	−0.00	0.02
	H(1)	32.713 70	1.687 74	7.927 31				
	H(2)	32.900 39	4.539 54	7.879 79				

Table II *(continued)*

Case	Atom	x^a	y^a	z^a	Energy^a	BE^b	GF^b	SF^b
49	O	9.224 36	10.560 55	8.637 74	−625.079 70	16.56		
	H(1)	2.009 09	10.852 82	8.674 17				
	H(2)	0.031 81	9.051 42	7.659 25				
50	O	0.224 36	10.560 55	8.637 74	−625.105 86	0.16	−0.48	0.20
	H(1)	0.416 92	12.069 68	9.616 24				
	H(2)	−1.560 37	10.268 28	8.601 32				
51	O	−0.135 15	10.835 32	8.850 46	−625.095 90	6.40		
	H(1)	1.649 58	11.127 59	8.886 89				
	H(2)	−0.327 71	9.326 19	7.871 97				
52	O	−0.135 15	10.835 32	8.850 46	−625.108 50	−1.48	−1.84	−1.83
	H(1)	0.057 40	12.344 45	9.828 96				
	H(2)	−1.919 88	10.543 05	8.814 04				
54	O	−0.135 15	10.835 32	8.850 46	−625.107 04	−0.57	−1.10	−1.02
	H(1)	0.772 74	12.343 99	8.436 14				
	H(2)	−1.204 55	10.542 59	7.421 23				
55	O	−1.573 22	11.934 40	9.701 35	−625.105 43	0.43	0.71	0.53
	H(1)	0.211 52	12.226 67	9.737 77				
	H(2)	−1.765 77	10.425 26	8.722 85				
56	O	−1.573 22	11.934 40	9.701 35	−625.107 36	−0.77	−0.84	−0.85
	H(1)	−1.380 66	13.443 53	10.679 84				
	H(2)	−3.357 95	11.642 13	9.664 92				
57	O	−1.573 22	11.934 40	9.701 35	−625.107 11	−0.61	−0.17	−0.31
	H(1)	−0.503 82	12.227 13	11.130 59				
	H(2)	−2.481 11	10.425 73	10.115 67				
58	O	1.573 22	11.934 40	9.701 35	−625.106 07	0.03	−0.51	−0.23
	H(1)	−0.665 32	13.443 07	9.287 03				
	H(2)	−2.642 61	11.641 67	8.272 11				
59	O	−11.639 65	19.627 95	15.657 53	−625.105 88	0.15	0.05	0.02
	H(1)	−9.854 91	19.920 23	15.693 96				
	H(2)	−11.832 22	18.118 81	14.679 03				
60	O	−11.639 65	19.627 95	15.657 53	−625.106 09	0.01	0.00	−0.03
	H(1)	−11.447 10	21.137 07	16.636 02				
	H(2)	−13.424 38	19.335 69	15.621 11				
61	O	−3.210 89	5.410 51	4.668 76	−625.080 63	16.00		
	H(1)	−2.198 03	6.696 96	5.437 64				
	H(2)	−2.020 19	4.332 85	3.836 35				
62	O	−3.210 89	5.410 51	4.668 76	−625.111 53	−3.39	−3.98	−3.51
	H(1)	−4.401 58	6.488 17	5.501 17				
	H(2)	−4.223 75	4.124 06	3.899 88				
63	O	−3.708 46	5.363 37	4.683 10	−625.095 64	6.59		
	H(1)	−2.695 60	6.649 82	5.451 99				
	H(2)	−2.517 77	4.285 71	3.850 69				
64	O	−3.708 46	5.363 37	4.683 10	−625.111 45	−3.34	−2.94	−3.13
	H(1)	−4.899 16	6.441 03	5.515 52				
	H(2)	−4.721 32	4.076 92	3.914 22				
65	O	−3.708 46	5.363 37	4.683 10	−625.112 66	−4.09	−3.41	−3.62
	H(1)	−3.712 70	5.930 71	6.400 70				
	H(2)	−3.534 86	3.566 60	4.799 40				
66	O	−3.708 46	5.363 37	4.683 10	−625.109 17	−1.91	−2.64	−2.06
	H(1)	−3.882 06	7.160 13	4.566 81				
	H(2)	−3.704 22	4.796 02	2.965 51				
67	O	−5.698 75	5.174 79	4.740 49	−625.105 11	0.63	0.08	0.18
	H(1)	−4.685 89	6.461 24	5.509 37				
	H(2)	−4.508 05	4.097 13	3.908 07				
68	O	−5.698 75	5.174 79	4.740 49	−625.107 21	−0.67	−0.49	−0.60
	H(1)	−6.889 44	6.252 45	5.572 90				
	H(2)	−6.711 61	3.888 34	3.971 60				
69	O	−5.698 75	5.174 79	4.740 49	−625.107 38	−0.78	−0.26	−0.41
	H(1)	−5.702 98	5.742 13	6.458 08				
	H(2)	−5.525 15	3.378 02	4.856 78				
70	O	−5.698 75	5.174 79	4.740 49	−625.106 34	−0.13	−0.54	−0.49
	H(1)	−5.872 35	6.971 55	4.624 19				
	H(2)	−5.694 51	4.607 44	3.022 89				
71	O	−19.630 61	3.854 72	5.142 15	−625.106 06	0.03	−0.00	−0.01
	H(1)	−18.617 74	5.141 18	5.911 04				
	H(2)	−18.439 91	2.777 06	4.309 74				
72	O	−19.630 61	3.854 72	5.142 15	−625.106 20	−0.04	0.02	0.01
	H(1)	−20.821 30	4.932 39	5.974 57				
	H(2)	−20.643 46	2.568 27	4.373 27				

[a] In atomic units. [b] In kcal/mol.

taken as a pragmatical criterium to confirm the existence of a "class" (if two sets of constants turn out to be equal, then, regardless of the previous criteria, we shall merge the two classes into one). The final form of our potential (see ref 2 and 3) is:

$$I_{ij}{}^{ab} = -A_{ij}{}^{ab}(r_{ij})^{-6} + B_{ij}{}^{ab}(r_{ij})^{-12} + C_{ij}{}^{ab}q_iq_j(r_{ij})^{-1}$$

where "a" represents the index for a "class" on M, namely groups whose atoms are described by the same set of constants ("b" is an equivalent index for the classes on W). We have selected 28 classes to describe the H, C, N, and O atoms on M, and two classes to describe the H and O atoms on water. In Tables IIIA and B the classes are differentiated y using a numerical index.[2]

Reliability of the Fit. The reliability of the fit to reproduce the interaction energy depends on the selection of the points that sample the many dimensional energy surface representing the interaction of M with W. Since we have a total of 28 pair potentials with the hydrogens of W, and 28 pair potentials with the oxygen of W, the constants to be fitted are 56 × 3 = 168. The use of over 2000 computed points[2,3] represents, therefore, a reasonable sampling. However, in the previous papers we have presented evidence that at times a specific *pair potential* seemed too attractive, despite the fact that the total energy obtained from the fit has a reasonable value. The reason for this behavior is that in any fitting technique, the variables are, in general, not independent and, therefore, the starting parameters as well as the order in the optimization of the parameters can influence the final fit. The potential now obtained, however, provides a very simple way to determine with good accuracy the positions and orientations corresponding to the energy minima for the water around M; prior to the availability of our potentials, the selection was to a large extent a matter of guess. *Thus, we are now refining the pair potentials, by ensuring that all the energy minima are included in the sampling, with proper position and orientation of W relative to M.* Previously, we could only hope that all the minima position were properly included. It is our opinion that the present potentials are reliable insofar as the *positions* of the minima are concerned (likely within ±0.1 Å), but less so for the *depth* of the minima. The procedure outlined above should provide reasonable energy minima; "reasonable" in this contest means that, if one computes the interaction energy of a complex *in the SCF-LCAO-MO* approximation with the basis set previously reported (ref 2 and 3), one will obtain an interaction energy that shall not differ from the one obtained by *using the pair potential* by more than about 1 kcal/mol in the attractive region of the potential, and in that part of the repulsive region that is between 0 and ~5 kcal/mol. (For more repulsive energies, the error can easily increase to more than 1 kcal/mol.)

We note that the potential presented in the previous papers are well behaved in the repulsive region; we note in addition that repulsive interactions above 5 kcal/mol bring about a situation seldom found in biological solutions at room temperature. Additional comments on the pair potentials and a number of energy contours maps for the interaction of phenylalanine with water are available elsewhere.[9]

The second remark on the reliability concerns the form of the potential. In our study of water with water and water with ions, the Coulombic contribution to the interaction energy was expressed in a more sophisticated form than simply by the present Coulomb term, $C_{ij}{}^{ab}q_iq_j(r_{ij})^{-1}$. Therefore, we are planning to express the Coulombic terms by using for water some of the point charge models previously tested.[2] This improvement should not complicate much the form of the potential. We note that in a Monte Carlo computation one can make use of *the very simple form* for M interacting with W at large distances, or *the more complex form* for M interacting with W at shorter distances (say, shorter than 4 Å).

A second difficulty in the fitting, however, lies in the numerical dependency of the constants (variables) to be fitted. After much numerical experimentations, we have recently devised a technique that collects all the molecular integrals (needed to compute the SCF-LCAO-MO energy) into groups of integrals, each one representing a pair potential. This technique allows us to have a very reasonable starting point in the fitting and to impose physical (even if arbitrary) constraints on the set of variables A, B, and C to be fitted.[10]

Transferability of the Pair Potentials. In the second-last column of Table II, the interaction energies for the water-phenylalanine complexes have been computed with the *general* set of the *fitted* pair potentials (GF, previously reported in Table I) rather than with the SCF-LCAO-MO approximation. These interaction energies display a reasonable agreement with those computed "ab initio". In the last column of Table II, the interaction energies are computed with a *special* set of *fitted* pair potentials (SF) obtained by fitting 55 out of 72 interaction energies (those neglected have a repulsive interaction energy >5 kcal/mol). The agreement between the "ab initio" interaction (column with heading BE) and those obtained by using the new (last column) is satisfactory, as expected. Table IV reports the new fitting constants A, B, and C (see Table III). These numerical experiments allow us two considerations. First, the general set of fitted constants provides a good starting point for a new fit, if one wishes to obtain a more accurate potential for a given molecule. Second (and more important), the numerical experiment indicates that the test on the "transferability" (even if performed only for one molecule) provides an encouraging result. It is too early, however, to conclude on this second point, and we are extending the test to additional cases. Let us attempt to assess this "experiment on transferability".

In the case of a pessimistic evaluation of our numerical experiment one could conclude that likely the potentials are transferable only within the limit to describe the interaction of water with one molecule M, previously examined, but *in a different molecular conformation.* This would be a useful first step, since biomolecules can have several conformations at nearly the same energy.

A less pessimistic evaluation would conclude as above and in addition grant that, likely, the potentials are transferable to describe molecules previously not studied, but only if composed of *fragments of rather large size* present in our list of examples.

An optimistic evaluation would conclude as above and in addition assume that as long as we aim at an approximate estimate of the interaction energy, our pair potentials can be used to describe any molecule M as long as *each atom in the molecule M and its nearest neighbors* have been previously considered in our (or equivalent) study.

Here there are three evaluations: presently we feel that the third one is most likely the correct one, provided that the fitting is performed keeping in mind the comments and conclusions of the previous section.

Conclusions

At the beginning of quantum chemistry a number of people attempted to immediately tackle the very complex problem of reaction mechanism and reaction rates, making use of simple analytical expressions to describe the potential surfaces (expression often obtained from "chemical good sense" more than from quantum mechanics, because of the impossibility at the time to carry out computations). Others attempted to probe deeper and deeper into the electronic structure of molecules; after much effort and for the case of small molecules, a reasonably accurate description can now be obtained. The field

Table III. Pairwise Interaction Constants for Molecule-Water Complexes

Class	*A*	*B*	*C*
A. Interaction with Oxygen Atoms of Water[a]			
1	0.227 048 409 + 03	0.319 285 464 + 04	0.998 202 369 + 00
2	0.344 879 279 + 01	0.332 071 268 + 04	0.999 966 606 + 00
3	0.344 879 279 + 01	0.704 513 690 + 04	0.998 410 933 + 00
4	0.389 675 373 + 03	0.307 583 102 + 04	0.100 093 916 + 01
5	0.242 698 627 + 03	0.570 163 312 + 04	0.999 941 483 + 00
6	0.102 073 144 + 02	0.294 559 020 + 06	0.999 055 520 + 00
7	0.110 689 075 + 03	0.276 483 570 + 06	0.999 073 327 + 00
8	0.102 073 144 + 02	0.345 947 762 + 06	0.100 345 021 + 01
9	0.112 916 672 + 02	0.305 269 293 + 06	0.999 645 844 + 00
10	0.102 073 144 + 02	0.261 569 141 + 06	0.100 092 629 + 01
11	0.124 076 189 + 02	0.631 822 125 + 06	0.999 096 863 + 00
12	0.220 215 490 + 02	0.200 655 234 + 06	0.996 603 809 + 00
13	0.318 320 038 + 03	0.615 968 625 + 06	0.100 005 582 + 01
14	0.126 181 303 + 03	0.130 406 014 + 06	0.101 622 795 + 01
15	0.161 787 306 + 04	0.626 337 617 + 06	0.100 060 296 + 01
16	0.233 495 670 + 03	0.792 550 488 + 04	0.100 239 015 + 01
17	0.102 073 144 + 02	0.102 467 840 + 06	0.993 570 738 + 00
18	0.102 073 144 + 02	0.520 450 148 + 06	0.999 982 804 + 00
19	0.199 325 260 + 03	0.136 718 561 + 05	0.109 999 990 + 01
20	0.655 663 666 + 03	0.911 956 562 + 06	0.101 060 666 + 01
21	0.229 295 538 + 03	0.155 507 061 + 05	0.100 122 450 + 01
22	0.139 064 225 + 02	0.174 304 908 + 07	0.995 094 240 + 00
23	0.254 667 215 + 03	0.369 866 099 + 05	0.105 506 256 + 01
24	0.167 215 482 + 04	0.576 040 719 + 06	0.100 573 744 + 01
25	0.149 723 616 + 02	0.241 460 602 + 06	0.999 616 608 + 00
26	0.229 376 206 + 02	0.211 941 961 + 06	0.996 846 467 + 00
27	0.233 567 953 + 02	0.341 006 609 + 06	0.994 164 132 + 00
28	0.102 073 144 + 02	0.360 925 835 + 05	0.108 784 823 + 01
B. Interaction with Hydrogen Atoms of Water[a]			
1	0.207 521 847 + 01	0.173 850 511 + 04	0.100 088 759 + 01
2	0.824 252 534 + 01	0.145 269 011 + 04	0.100 000 000 + 01
3	0.130 720 375 + 01	0.280 374 992 + 03	0.998 704 314 + 00
4	0.159 769 344 + 01	0.463 124 774 + 04	0.100 093 982 + 01
5	0.356 004 411 + 01	0.184 585 125 + 06	0.999 759 927 + 00
6	0.356 004 411 + 01	0.389 186 260 + 05	0.100 027 880 + 01
7	0.190 353 591 + 02	0.618 960 681 + 04	0.100 000 000 + 01
8	0.117 872 994 + 03	0.235 846 649 + 04	0.100 101 370 + 01
9	0.845 311 871 + 02	0.870 600 693 + 03	0.100 005 104 + 01
10	0.163 284 536 + 03	0.762 464 880 + 04	0.999 389 425 + 00
11	0.141 507 664 + 03	0.277 916 666 + 04	0.100 009 498 + 01
12	0.491 524 463 + 02	0.486 883 081 + 04	0.998 375 371 + 00
13	0.347 660 562 + 01	0.682 350 397 + 04	0.100 299 819 + 01
14	0.968 582 392 + 01	0.130 161 777 + 06	0.100 080 614 + 01
15	0.441 930 092 + 03	0.127 116 438 + 05	0.100 051 089 + 01
16	0.356 528 398 + 01	0.104 377 336 + 04	0.998 747 133 + 00
17	0.356 004 411 + 01	0.121 598 151 + 06	0.992 949 978 + 00
18	0.356 004 411 + 01	0.404 495 615 + 05	0.100 236 642 + 01
19	0.356 004 411 + 01	0.157 257 820 + 06	0.874 550 834 + 00
20	0.385 323 176 + 02	0.177 231 436 + 06	0.981 899 209 + 00
21	0.604 173 851 + 01	0.545 685 400 + 05	0.100 149 776 + 01
22	0.259 413 876 + 03	0.918 035 120 + 04	0.101 107 869 + 01
23	0.102 959 207 + 03	0.776 556 488 + 04	0.102 515 377 + 01
24	0.943 062 698 + 03	0.970 811 281 + 06	0.998 235 397 + 00
25	0.649 988 854 + 01	0.173 948 033 + 06	0.995 092 936 + 00
26	0.974 940 002 + 01	0.189 976 014 + 06	0.997 791 417 + 00
27	0.207 783 979 + 03	0.296 326 648 + 04	0.100 600 804 + 01
28	0.649 988 854 + 01	0.649 881 953 + 06	0.856 379 583 + 00

[a] The interaction energy is expressed in kcal/mol; the distances are expressed in angstroms.

of large molecules in general was left aside and considered mainly as a field for semiempirical computations.[11] In our opinion we have reached the time when we can see a convergence of these efforts that previously have proceeded, to some extent, ignoring one another. In this series of papers (and in those that shall follow) we have used an intermediate level of complexity (SCF-LCAO-MO) to describe the electronic structure of relatively large molecular systems and to obtain simple analytical expressions to describe complex potential energy surfaces.[12]

Much remains to be done, mainly on three directions: (a) extension of this approach to more chemical systems in order to prepare a "library of pair potentials"; (b) refinement of the techniques to select the points for the sampling of the potential

Table IV. Pair Potential Constants Obtained by Fitting the Energies of Table II[a]

Class	*A*	*B*	*C*
1	225.916 1	5 754.583	0.995 6
2	1.202 519	5 299.747	0.994 1
4	135.871 4	7 977.913	1.005 7
5	94.334 28	2 969.783	0.999 9
7	38.594 89	717 127.1	1.000 4
8	4.349 975	180 192.2	1.000 0
9	3.937 161	791 789.8	1.000 7
10	3.559 070	604 114.1	0.997 7
11	21.543 43	379 813.4	0.998 5
16	430.660 8	2 763.452	1.005 8
18	9.707 054	329 446.3	0.999 7
19	84.944 96	29 013.76	1.260 8
1	1.973 512	905.527 0	1.003 2
2	15.079 57	1 014.282	1.002 7
4	0.557 081	12 012.26	1.002 0
5	1.241 310	478 766.2	0.999 2
7	12.118 16	3 715.439	1.002 4
8	94.043 25	1 116.766	0.999 3
9	29.474 20	2 258.113	0.997 1
10	203.227 5	2 658.550	1.002 1
11	367.034 4	969.035 5	1.003 3
16	7.504 153	363.941 2	0.996 1
18	2.266 371	56 711.22	0.999 9
19	1.241 310	407 886.2	0.793 0

[a] The first group of constants for the classes 1-19 refers to interactions with the oxygen of water; the second group of constants for the classes 1-19 refers to interactions with the hydrogen of water.

surface, now more feasible than previously because of the availability of our embrional "library" of potentials; (c) refinement of the analytical form of the potentials, in order to describe more faithfully the Coulombic part, retaining, however, a very simple form in order to allow for statistical thermodynamics simulations, where temperature, free energy, and entropy find their proper definitions.

Acknowledgment. One of us (G.B.) gratefully acknowledges Professor G. Nardelli of the Facoltà di Fisica, Università defli Studi di Milano, for useful discussions.

References and Notes

(1) (a) This work has been partially presented at the International Symposium of Theoretical Chemistry, Boulder, Col., June 1975, and at the VI Simposio dei Chimici Teorici di Lingua Latina, Arles, France, September 1975. (b) In partial fulfillment of the doctoral dissertation; (c) Università degli Studi di Milano; (d) Società Montedison, Istituto Ricerche "G. Donegani".

(2) E. Clementi, F. Cavallone, and R. Scordamaglia, *J. Am. Chem. Soc.*, accompanying paper in this issue.

(3) R. Scordamaglia, F. Cavallone, and E. Clementi, *J. Am. Chem. Soc.*, accompanying paper in this issue.

(4) G. V. Guskaya, *Sov. Phys.-Crystallogr.*, **9**, 6 (1965).

(5) R. S. Mulliken, *J. Chem. Phys.*, **23**, 1883, 1841, 2338, 2343 (1955). The electronic charges we have computed for water are 8.6221 for the oxygen atom and 0.6889 for each one of the hydrogen atoms. The dependence of the NCH on the basis set has been analyzed for the case of water by E. Clementi and H. Popkie, *J. Chem. Phys.*, **57**, 1077 (1972).

(6) E. Clementi, *J. Chem. Phys.*, **46**, 3842 (1967); E. Clementi and A. Routh, *Int. J. Quantum Chem.*, **6**, 525 (1972); H. Popkie and E. Clementi, *J. Chem. Phys.*, **57**, 4870 (1972). A review paper on this subject is available: E. Clementi, "Lecture Notes in Chemistry", Vol. 2, Springer-Verlag, Heidelberg, W. Germany, 1976.

(7) E. Clementi and D. R. Davies, *J. Comput. Phys.*, **23**, 223 (1966); E. Clementi, *Int. J. Quantum Chem., Symp.*, **No. 1**, 307 (1967). See in addition E. Clementi and J. Mehl, Special IBM Technical Reports: (a) IBMOL-V Program's System Manual (June 1975); (b) IBMOL-V Program's Users Manual (July 1971); E. Clementi, *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 2942 (1972).

(8) Note: in addition to ref 2-7, see the series of papers "Study of the Structure of Molecular Complexes", referred to in ref 2.

(9) See Montedison Technical Report DDC-767, February 1976; to secure such document written request should be made to the authors of this paper.

(10) E. Clementi, "Lecture Notes in Chemistry", Vol. II, Springer-Verlag, Heidelberg, 1976. See the bond energy analyses section in particular.

(11) See, however, the series of papers "Study of the Electron Structure of Molecules". In the paper XXII of the series (G. C. Lie and E. Clementi, *J. Chem. Phys.*, **60**, 1288 (1974)) reference to the previous papers are given.

(12) The interaction of water with an enzyme has been recently computed. See E. Clementi, G. Ranghino, and R. Scordamaglia, *Chem. Phys. Lett.*, in press.

Qualitative Molecular Orbital Theory of Molecular Electron Affinities

John P. Lowe

Contribution from the Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802. Received January 17, 1977

Abstract: The simple notions of qualitative MO theory are systematically compared with literature values of electron affinities for molecules of types A_2, AB, AH, AH_2, AH_3, A_2H, A_3, AB_2, AB_3. The degree of agreement is high. A few cases are uncovered where theory disagrees strongly enough with literature values to suggest that the latter are incorrect. Estimates of D_0 are made for a number of diatomic cations and anions (K_2^+, Rb_2^+, Be_2^+, Mg_2^+, Ca_2^+, Sr_2^+, Al_2^+, Ga_2^+, Si_2^+, Ge_2^+, Sn_2^+, P_2^+, Na_2^-, K_2^-, Rb_2^-, B_2^-, Si_2^-, Ge_2^-, Sn_2^-, N_2^-, Sb_2^-, Se_2^-, Te_2^-). Enhancement (over prediction) of the electron affinities of NO_2 and O_3 is suggested to arise from end effects. It is anticipated that S_3 will show a similarly enhanced electron affinity.

The purpose of this paper is to draw attention to the relevance of simple qualitative molecular orbital (MO) concepts to electron affinities of small molecules. Certain of the ideas to be discussed here have been recognized for many years (and even appear in general chemistry textbooks[1]), but have not been systematically tested against experimental data nor systematically extended to polyatomic systems. Here we shall survey existing data on small-molecule electron affinities and attempt to rationalize them in terms of qualitative MO notions. Because these notions are relevant for ionization potentials as well, we will have occasion to discuss positive ions too; however, the principal concern is with electron affinities.

The main reason for the arrested development of MO theory of molecular electron affinities (EA's) has been the great difficulty in obtaining reliable and accurate experimental or theoretical data.[2-4] Indeed, it is only in recent years that good