BULLETIN OF THE CHEMICAL SOCIETY OF JAPAN VOL. 42 3118—3123 (1969)

The Electronic Structures of Glycine in the Isolated State and in Water

Akira IMAMURA, Hiroko FUJITA and Chikayoshi NAGATA

Biophysics Division, National Cancer Center Research Institute, Tsukiji, Chuo-ku, Tokyo

(Received May 20, 1969)

The electronic structures of glycine have been calculated for the dipolar-ion form and the neutral form by means of the CNDO method. The relative stability of both the dipolar-ion form and the neutral form in the gas state was reproduced by our calculation. The geometry of the glycine was also reproduced by the CNDO method, indicating the validity of our calculation. In addition, the relative stability of the two tautomeric forms of glycine was proved to be independent of its geometrical modifications. The experimental fact that the dipolar-ion form is more stable than the neutral form in an aqueous solution was well explained by the calculation of the interaction energy between the solute and the solvent.

Glycine is the simplest α-amino acid and constitutes the backbone of the protein, namely, polypeptide. Therefore, the electronic structure of glycine is of interest since the general features of the electronic structures of the amino acid and the backbone of protein can be derived from that of glycine.

Glycine is well known to exist as a dipolar ion in water and crystal, and, on the other hand, to exist in the neutral form in the isolated state.[1] Therefore, the interaction between glycine and water makes the dipolar-ion form more stable than the neutral form.

The electronic structures of various α-amino acids have already been calculated by Del Re's method;[2] a good parallelism between the electron densities of the proton and the NMR chemical shift[3] was found. Moreover, the dissociation constants of α-amino acids were found to be correlated with the calculated energy differences between the dissociated and non-dissociated forms.[4] Recently, the CNDO method was developed by Pople *et al.*;[5,6] it has proved to be very useful for the prediction of the electronic structures of many organic molecules, especially polar molecules, in the ground state.[7,8]

In the present paper, the total electronic energies of glycine in its neutral and dipolar-ion forms were calculated by the CNDO method.[5,6] The most stable geometry of glycine obtained by the calculations agrees well with the results of X-ray diffraction.[9,10] Moreover, the relative stabilities of the neutral form and of the dipolar-ion form were calculated in water and in gas, and the solvent effect considered.

---

[1] K. Biemann, J. Seibl and F. Gapp, *J. Am. Chem. Soc.*, **83**, 3795 (1961); G. Junk and H. Sevec, *ibid.*, **85**, 839 (1963); J. P. Greenstein and M. Winitz, "Chemistry of the Amino Acids," Vol. 1, John Wiley & Sons, New York (1961), Chap. 4.

[2] G. Del Re, *J. Chem. Soc.*, **1958**, 4031.

[3] G. Del Re, B. Pullman and T. Yonezawa, *Biochim, Biophys. Acta*, **75**, 153 (1963).

[4] T. Yonezawa, G. Del Re and B. Pullman, This Bulletin, **37**, 985 (1964).

[5] J. A. Pople, D. P. Santry and G. A. Segal, *J. Chem. Phys.*, **43**, s 129 (1965); J. A. Pople and G. A. Segal, *ibid.*, **43**, s135 (1965).

[6] J. A. Pople and G. A. Segal, *ibid.*, **44**, 3289 (1966).

[7] J. A. Pople and M. Gordon, *J. Am. Chem. Soc.*, **89**, 4253 (1967).

[8] J. E. Bloor and D. L. Breen, *ibid.*, **89**, 6835 (1967).

[9] G. Albrecht and R. B. Corey, *ibid.*, **61**, 1087 (1939).

[10] R. E. Marsh, *Acta Cryst.*, **11**, 654 (1958); Y. Iitaka, *ibid.*, **11**, 225 (1958).

TABLE 1. THE PARAMETERS USED IN THE CALCULATION

Atom	H	C	N	O
Extended Hückel method				
$Z_{\text{eff.}}$	1.2	3.25	3.90	4.55
I_p[a] s orbital	−13.6	−21.43	−27.50	−35.30
p orbital		−11.42	−14.49	−17.76
CNDO method				
$Z_{\text{eff.}}$	1.2	3.25	3.90	4.55
$\frac{1}{2}(I+A)$[a] s orbital	− 7.176	−14.051	−19.316	−25.390
p orbital		− 5.572	− 7.275	− 9.111
β[a]	− 9.0	−21.0	−25.0	−31.0

a) in eV.

Methods

The calculations were carried out in the framework of the extended Hückel method developed by Hoffmann[11] and the CNDO/2 method developed by Pople *et al.*[5,6] All the parameters included in these methods are collected in Table 1; these values are the same as those usually employed. In the CNDO method the total energy is the sum of the total electronic and core-repulsion energies. The core-repulsion energy was calculated by the following equation:

$$E_{\text{core rep.}}=\frac{1}{2}\sum_{A}\sum_{\substack{B\\(\neq A)}}\frac{Z_A Z_B e^2}{R_{AB}} \qquad (1)$$

where R_{AB} is the distance between the atoms A and B and where Z_A is the effective nuclear charge on the atom A. In order to carry out the population analysis by the CNDO method, the molecular orbitals on the CNDO basis were transformed to those on the Slater function basis by using the overlap matrix. Then the new set of the molecular orbitals thus obtained was employed to calculate the total electron densities using the population analysis[12] which was also employed in the extended Hückel method.

The convergency of the iteration procedure in the CNDO method was checked against the total electronic energy; that is, if the ratio of the total electronic energy to the energy difference between the former iteration and the present one was less than 10^{-6}, the convergency was considered to be accomplished. The usual number of the iteration was eight or nine.

Results and Discussions

The Electronic Structures of Glycine in the Neutral and the Dipolar-ion Forms. The electronic structures of glycine in the neutral and the dipolar-ion forms were calculated by the extended Hückel[11] and CNDO/2 methods.[5,6] The geometries of glycine assumed[9,10] are indicated in Fig. 1, while the total energies obtained are listed in Table 2. As is shown in Table 2, the extended Hückel method predicts that the dipolar-ion form will be more stable than the neutral form, in

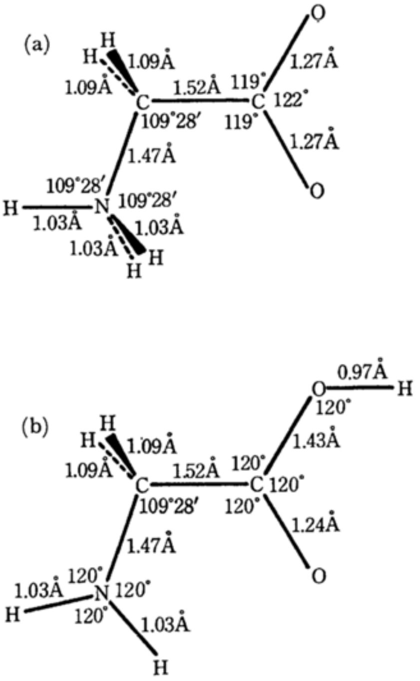

Fig. 1. The geometry of glycine used for calculations.
(a) the dipolar ion form
(b) the neutral form

TABLE 2. THE TOTAL ENERGY OF GLYCINE CALCULATED BY THE EXTENDED HÜCKEL AND CNDO METHODS (in eV)

	Extended Hückel	CNDO
Neutral form	−630.81	−1807.58
Dipolar ion form	−632.36	−1804.44

---

11) R. Hoffmann, *J. Chem. Phys.*, **39**, 1397 (1963).

12) R. S. Mulliken, *ibid.*, **23**, 1833, 1841, 2338, 2343 (1955).

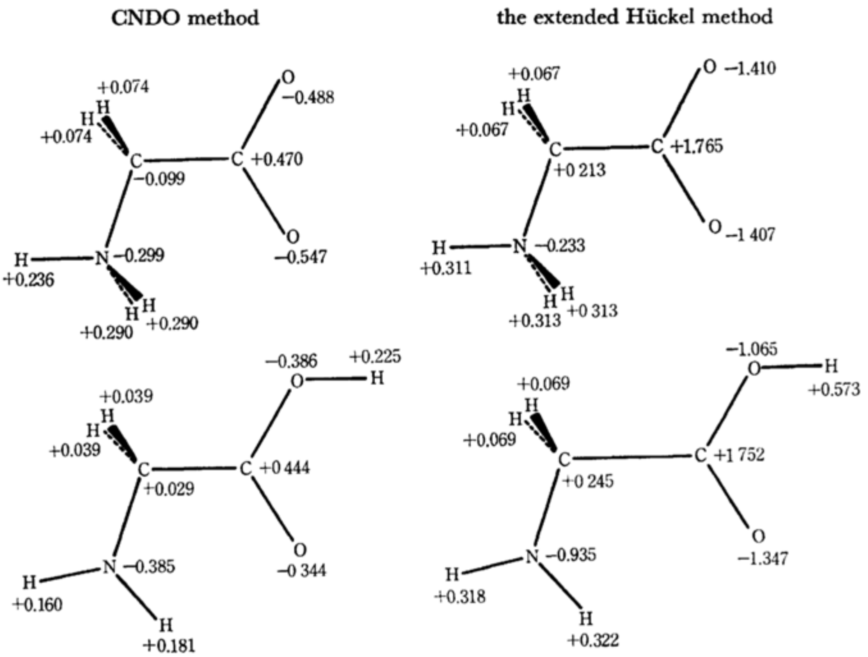

Fig. 2. The electron distribution of glycine calculated by CNDO and the extended Hückel methods.

agreement with the previous calculation;[13] on the contrary, the neutral form is indicated to be more stable by the CNDO method. This discrepancy between the two methods may be attributable to the explicit neglect of the electron-electron repulsion in the extended Hückel method. The above-mentioned inference was supported by the electron distributions of glycine calculated by the two methods; that is, the electron distributions calculated by the extended Hückel method deviate from neutral more than those calculated by the CNDO method, as is shown in Fig. 2. In other words, the electron densities of glycine calculated by the extended Hückel method exaggerate its polar nature too much; for example, the oxygen net negative charges in glycine exceed one, and the net positive charge of one of the carbons is approximately two, because of the neglect of the explicit effect of electron-electron repulsion. Glycine in water or in the solid state is well known to exist as a dipolar ion.[1] However, glycine in the isolated state, namely, in the gas state, is in the neutral form;[1] this is in accordance with the results calculated by the CNDO method.

The Geometry of Glycine in the Dipolar-ion Form. In order to investigate the validity of the CNDO method for such a polar molecule as glycine and estimate the variation in the total energy with the bond lengths and the bond angles, the calculations were carried out for various geometries of the molecule in order to establish the energy minimum. First of all, the energy dependences upon the C–C and C–N bond lengths were investigated, with the other geometrical parameters fixed as indicated in Fig. 1. The calculated results are listed in Table 3. In Table 3-A, the C–C bond length is fixed at 1.52 Å, while the C–N bond lengths range from 1.20 to 1.60 Å. Thus we obtain the energy minimum at 1.43 Å. Therefore, in Table 3-B the C–N bond length is fixed at three points, namely at 1.39, 1.43, and 1.47 Å, and the energy dependences upon the change in the C–C bond length was investigated. In the table, the energy minimum was found at the C–C bond length of 1.48 Å for all C–N bond lengths. That is to say, equilibrium C–C and C–N bond lengths were determined independently of each other. The equilibrium bond lengths calculated were approximately in accordance with the data obtained by X-ray diffraction analysis; that is, from X-ray diffraction analysis the C–C bond length was determined as 1.52 Å,[9,10] and for the C–N bond two different values, of 1.47[10] and 1.39 Å,[9] have been given by different authors. With regard to the bond angle, the energy variation with the angle between the C–C and C–N bonds was calculated, fixing the C–C bond at 1.48 Å and the C–N bond at 1.43 Å. In this case, the bond angle between the two C–H bonds was changed in harmony with the alteration of the CCN bond angle; that is, the directions of the four bonds were determined in order to construct four orthonormally-hybridized

---

[13] T. Yonezawa, "Seitai Ryôshi Kagaku" (Biological Quantum Chemistry), ed. by K. Fukui, Kyoritsu, Tokyo (1967), p. 29.

Table 3. The energy dependence on the change of the bond length[a)]

(A) C–C bond length = 1.52 Å

C–N bond length (Å)	1.20	1.30	1.39	1.43
Total energy (eV)	−1800.03	−1803.21	−1804.36	−1804.50
C–N bond length (Å)	1.47	1.51	1.55	1.60
Total energy (eV)	−1804.45	−1804.25	−1803.95	−1803.45

(B) C–N bond length = 1.39 Å

C–C bond length (Å)	1.44	1.48	1.52
Total energy (eV)	−1804.29	−1804.40	−1804.36

C–N bond length = 1.43 Å

C–C bond length (Å)	1.44	1.48	1.52
Total energy (eV)	−1804.46	−1804.55	−1804.50

C–N bond length = 1.47 Å

C–C bond length (Å)	1.44	1.48	1.52
Total energy (eV)	−1804.41	−1804.50	−1804.44

a) Observed C–C bond length = 1.52 Å[9,10]
Observed C–N bond length = 1.39 Å[9] or 1.47 Å[10]

Table 4. The energy dependence on the NCC angle[a)]

NCC angle	106°	109°28′	112°	115°	118°
Total energy (eV)	-1804.51_7	-1804.52_4	-1804.51_7	−1804.47	−1804.41

a) Observed = 112°[9] or 118.8°[10]

Table 5. The energy dependence on the rotational angle around the C–C bond[a)]

Rotational angle	0°	30°	60°	90°
Total energy (eV)	−1804.55	−1804.48	−1804.35	−1804.30

a) Observed ≒ 20°[10]

orbitals for any values of the bond angle. The total energies calculated at various angles between the C–C and C–N bonds are listed in Table 4. The energy minimum was found at the tetrahedral configuration, which agrees fairly well with the experimental data of 112°[9] and 118.8°.[10] Moreover, the variation in the energy with the change in the angle is so small that the determination of the equilibrium bond angle is probably very difficult. Judging from this fact, the small discrepancy between the calculated bond angle and the observed one is possibly not serious. The last geometrical parameter investigated was the rotational angle around the C–C bond, with the C–C bond length of 1.48, with the C–N bond length of 1.43 Å, and with a tetrahedral configuration of the central carbon. The results in Table 5 indicate that the planar configuration is most stable. From X-ray diffraction analysis, glycine has been shown to deviate slightly from the planar configuration.[9,10] On the whole, the geometry of glycine predicted by means of the CNDO method is in good agreement with the experimental results. In addition, the energy change of the molecule with the alterations in the geometry is rather small. Therefore, it is quite impossible to expect the reversion of the relative stability of the neutral and the dipolar-ion forms by means of the variation in the geometry.

The Electronic Structure of Glycine in Water. As has already been mentioned, glycine exists in the dipolar-ion form in water, but in the neutral form in gas.[1] These facts suggest that the interaction between glycine and water stabilizes the dipolar-ion form more than the neutral form. The exact calculation of the interaction energy between the solute and the solvent is quite impossible because of the diversity of the relative conformations of glycine and water in solution. The interaction energy can be estimated roughly, however, by such a simple model as the cavity in a dielectric solution.[14] The alternative approach is treatment by

---

[14] For example, H. Fröhlich, "Theory of Dielectrics," The Clarendon Press, Oxford (1958), Chap. 2.

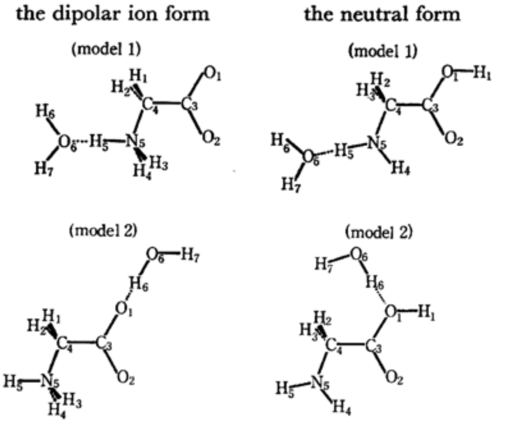

Fig. 3. The models of the hydrogen bond between glycine and water.

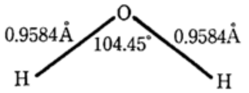

Fig. 4. The geometry of water.

TABLE 6. THE STABILIZATION ENERGY OF THE HYDROGEN BONDS BETWEEN GLYCINE AND WATER

N–O distance (model 1) O–O distance (model 2)	Total energy (eV)	
	Model 1	Model 2
Dipolar ion form		
2.0	−2341.79	−2343.44
2.3	−2345.39	−2345.59
2.4	−2345.61	−2345.67
2.5	−2345.65	−2345.63
2.75	−2345.54	−2345.46
3.0	−2345.37	−2345.31
3.5	−2345.21	−2345.17
∞	−2345.09	−2345.09
Stabilization energy	0.56	0.54
Neutral form		
2.0	−2343.85	−2345.49
2.4	−2348.14	−2348.33
2.5	−2348.29	−2348.39
2.6	−2348.34	−2348.38
2.7	−2348.34	−2348.35
3.0	−2348.26	−2348.24
3.5	−2348.17	−2348.16
∞	−2348.12	−2348.12
Stabilization energy	0.22	0.27

the molecular orbital theory, in which the representative relative conformations of the solute and the solvent are tentatively assumed and in which the calculation is carried out for the conformation thus assumed. In this method, however, only a part of the interaction energy can be obtained; in other words, the molecular orbital approach can be used only to estimate the order of magnitude of the interaction energy. From this point of view, it is of interest to calculate the stabilization energies of the hydrogen bonds between water and glycine in the dipolar-ion form or in the neutral form in order to estimate the magnitude of the solvent effect. Although the stabilization energy by the hydrogen-bond formation is a part of the interaction energy, its contribution to the energy may reasonably be the main one; thus, this treatment can be employed for the rough estimation of the interaction energy. The models of hydrogen bonds calculated are shown in Fig. 3. The model 1 in Fig. 3 had already been accepted for the study of the protolysis kinetics of amino acid analogs.[15] The geometry of water is shown in Fig. 4. The total energies calculated by the CNDO method are listed in Table 6 for the model system of the hydrogen bonds with various intermolecular distances. For any models, the energy minimum was obtained between the intermolecular distances of 2.4 and 2.7 Å; this is in good agreement with the results obtained by means of the X-ray diffraction method used on usual hydrogen bonds. It is noteworthy that the stabilization energies by hydrogen-bond formation in the dipolar-ion form are greater than those in the neutral form by about 0.3 eV. Since many hydrogen-bond formations and other types of interaction between glycine and water can be expected in a water solution, the differences in the stabilization energies calculated are large enough to change the relative stability of the neutral and the dipolar-ion forms of glycine. In other words, the dipolar-ion form of glycine is more stable than the neutral form in water, probably because of such a solute-solvent interaction as the hydrogen bonds.

The electron densities obtained by the population analysis for various intermolecular distances are listed in Table 7. From this table, the charge transfer between water and glycine in the dipolar-ion form may be found to be larger than those in the neutral form, corresponding to the relative magnitudes of the stabilization energies. More remarkable is the fact that the charge redistribution by the hydrogen-bond formation occurs predominantly not between molecules, but within molecules; this is in good agreement with the non-empirical calculation of the hydrogen bond between water molecules[16] and the CNDO calculation of the dimer of the formic acids.[17] Moreover, when hydrogen bonds are formed, the proton-accepting

---

15) M. Sheinblatt, *J. Chem. Phys.*, **36**, 3103 (1962).

16) K. Morokuma and L. Pedersen, *ibid.*, **48**, 3275 (1968).

17) A. Ocvirk, H. Ažman and D. Hadži, *Theoret. Chim. Acta*, **10**, 187 (1968); A. Pullman and H. Berthod, *ibid.*, **10**, 461 (1968).

TABLE 7. THE CHARGE DISTRIBUTION FOR VARIOUS HYDROGEN BOND SYSTEMS

(A) Dipolar ion form—model 1[a)]

N$_5$–O$_6$ distance	Atomic population					
	H$_1$	H$_3$	H$_5$	H$_6$	H$_7$	CT[b)]
2.0	0.970	0.765	0.462	0.700	0.708	−0.069
2.5	0.941	0.726	0.683	0.755	0.761	−0.033
3.0	0.931	0.713	0.741	0.780	0.784	−0.007
3.5	0.928	0.710	0.755	0.789	0.793	−0.001
∞	0.927	0.707	0.767	0.804	0.804	0.0

N$_5$–O$_6$ distance	Atomic population					
	O$_1$	O$_2$	C$_3$	C$_4$	N$_5$	O$_6$
2.0	6.525	6.556	3.527	4.065	5.461	6.523
2.5	6.502	6.555	3.528	4.084	5.345	6.451
3.0	6.505	6.548	3.525	4.090	5.309	6.429
3.5	6.493	6.553	3.527	4.093	5.302	6.417
∞	6.491	6.552	3.527	4.095	5.298	6.392

a) The numbering of the atoms is shown in Fig. 3.
b) The charge transfer quantity from water to glycine.

(B) Dipolar ion form—model 2[a)]

O$_1$–O$_6$ distance	Atomic population					
	H$_1$	H$_3$	H$_5$	H$_6$	H$_7$	CT[b)]
2.0	0.910	0.703	0.755	0.526	0.905	−0.020
2.5	0.921	0.706	0.762	0.709	0.842	−0.018
3.0	0.925	0.707	0.765	0.760	0.820	−0.004
3.5	0.927	0.707	0.766	0.778	0.813	−0.001
∞	0.927	0.707	0.767	0.804	0.804	0.0

O$_1$–O$_6$ distance	Atomic population					
	O$_1$	O$_2$	C$_3$	C$_4$	N$_5$	O$_6$
2.0	6.576	6.516	3.504	4.097	5.303	6.589
2.5	6.518	6.537	3.514	4.096	5.299	6.467
3.0	6.511	6.543	3.519	4.094	5.298	6.424
3.5	6.507	6.545	3.522	4.094	5.298	6.410
∞	6.491	6.552	3.527	4.095	5.298	6.392

a) The numbering of the atoms is shown in Fig. 3.
b) The charge transfer quantity from glycine to water.

(C) Neutral form—model 1[a)]

N$_5$–O$_6$ distance	Atomic population						
	H$_1$	H$_2$	H$_4$	H$_5$	H$_6$	H$_7$	CT[b)]
2.0	0.787	0.995	0.872	0.540	0.767	0.767	0.007
2.5	0.778	0.972	0.835	0.761	0.789	0.788	−0.011
3.0	0.776	0.964	0.824	0.810	0.799	0.799	−0.003
3.5	0.775	0.962	0.821	0.830	0.802	0.802	−0.001
∞	0.775	0.961	0.819	0.840	0.804	0.804	0.0

N$_5$–O$_6$ distance	Atomic population					
	O$_1$	O$_2$	C$_3$	C$_4$	N$_5$	O$_6$
2.0	6.400	6.343	3.571	3.962	5.529	6.473
2.5	6.393	6.338	3.566	3.968	5.425	6.412
3.0	6.391	6.337	3.564	3.972	5.394	6.399
3.5	6.390	6.336	3.564	3.973	5.387	6.395
∞	6.386	6.344	3.556	3.971	5.385	6.392

a) The numbering of the atoms is shown in Fig. 3.
b) The charge transfer quantity from water to glycine.

(D) Neutral form—model 2[a)]

O$_1$–O$_6$ distance	Atomic population						
	H$_1$	H$_2$	H$_4$	H$_5$	H$_6$	H$_7$	CT[b)]
2.0	0.756	0.949	0.817	0.832	0.578	0.882	0.006
2.5	0.764	0.959	0.817	0.837	0.750	0.829	−0.009
3.0	0.770	0.962	0.818	0.839	0.790	0.812	−0.002
3.5	0.773	0.963	0.818	0.839	0.799	0.808	0.0
∞	0.775	0.961	0.819	0.840	0.804	0.804	0.0

O$_1$–O$_6$ distance	Atomic population					
	O$_1$	O$_2$	C$_3$	C$_4$	N$_5$	O$_6$
2.0	6.467	6.332	3.550	3.970	5.385	6.535
2.5	6.407	6.342	3.549	3.971	5.385	6.430
3.0	6.394	6.346	3.551	3.970	5.385	6.400
3.5	6.390	6.347	3.552	3.971	5.385	6.393
∞	6.386	6.344	3.556	3.971	5.385	6.392

a) The numbering of the atoms is shown in Fig. 3.
b) The charge transfer quantity from glycine to water.

oxygen becomes more negative, and the hydrogen-bonding proton, more positive; this is in contrast with the other atoms, the charge of which does not change considerably except in the case of a few atoms adjacent to the hydrogen bond. Therefore, the mutual electrostatic interaction can be considered to determine predominantly the charge redistribution of the molecule by the hydrogen-bond formation. The electronic structures at the intermolecular distance of 2.0 Å show a characteristic feature; that is, the charge distribution of the molecules at 2.0 Å is quite different from that at 2.5 Å because of the strong solute-solvent interaction; this is parallel with the sudden increase in the total energy indicated in Table 6.

The authors are grateful to Miss Fumiko Kitazumi for her assistance in the programming.