Reaction pathway of proton transfer from the neutral to zwitterionic forms of amino acids. Support for a water molecule-mediated mechanism

Sławomir J. Grabowski,[1]* Tadeusz M. Krygowski[2] and Beata Stępień[2]

[1]Institute of Chemistry, University of Białystok, Al. J. Piłudskiego 11/4, 15-443 Białystok, Poland
[2]Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

ABSTRACT: The analysis of bond valences and precise neutron diffraction data taken from the Cambridge Structural Database were used to model the reaction path for the proton transfer of amino acids. The correlations confirm the mechanism with the support of water molecules mediating the proton transfer and also show that intramolecular proton transfer is much less probable owing to the strains within amino acid molecules. A concept of the basicity and acidity in the crystalline state is presented. Copyright © 2000 John Wiley & Sons, Ltd.

KEYWORDS: bond valences; amino acids; proton transfer; reaction path; basicity and acidity in crystals

INTRODUCTION

Proton transfer reactions are of fundamental importance in chemistry, biochemistry and biology. They participate in many chemical reactions (including some industrial processes), in catalysis and in particular in enzyme catalysis, they determine the function of membrane protein channels, the strength of hydrogen bonds, etc.[1,2] Reactions of this kind are also associated with two forms of existence of amino acids: the neutral and the zwitterionic form. Moreover, amino acids as constituents of proteins belong to the most important molecules in biological systems. These two forms of existence of amino acids are related to each other by the proton transfer reaction. Since the zwitterionic form of amino acids is predominant for their molecules existing in water solutions, and hence also in their native forms, the studies on the proton transfer from the neutral to the zwitterionic form may be very helpful in understanding life processes.

The importance of these processes and their universal character are a reason why studies on the reaction path for the proton transfer of amino acids have been undertaken and are the subject of investigation in this paper. For the analysis of the reaction pathway the accurate neutron diffraction data of organic crystals taken from the Cambridge Structural Database (CSD)[3] were used.

Correspondence to: S. J. Grabowski, Institute of Chemistry, University of Białystok, Al. J. Piłudskiego 11/4, 15-443 Białystok, Poland.
E-mail: slagra@noc.uwb.edu.pl

THE MODEL AND CHOICE OF DATA SAMPLE

It has been pointed out that the shape of correlation curves between geometrical parameters of molecules may be reminiscent of structural changes occurring along the pathways of the respective chemical reactions.[4,5] In other words, the points of the correlation curve correspond to different states of the investigated system during the reaction. For example, this method was used to study the reaction pathways for nucleophilic substitution, the addition and elimination reactions of carbonyl derivatives and proton transfer within the O—H···O systems.[6–9]

The direct intramolecular proton transfer in amino acids (Scheme 1) is connected with suitable changes of geometrical parameters, including the lengthening of the O—H and N—H bonds, the changes in H···N and H···O contacts, respectively, and the possibility of rotation around the CN bond. This, however, needs energy and hence seems to be much less favorable.
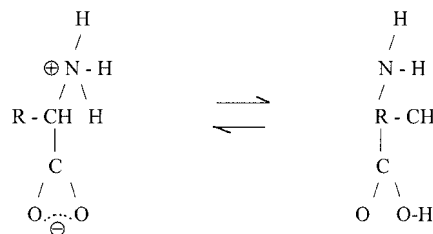
Similar geometrical changes, except bond angles deformations, occur for the proton transfer mediated by water molecules. This may be described in terms of the following two-step equilibria processes:

first step: NH_2—R—COOH + H_2O =
H_3O^+ + NH_2—R—COO^-

second step: H_3O^+ + NH_2—R—COO^- =
H_2O + NH_3^+—R—COO^-
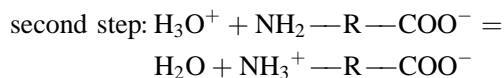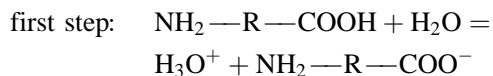
In order to study the respective changes of geometrical parameters during the proton transfer, the interatomic distances in the —C—O—H···O— and —C=O···H—O— as well as in —O—H···NH_2— and —NH_3^+···O systems were taken from the CSD. Only the accurate neutron diffraction data were taken into account (e.s.d.s ≤ 0.005 Å and $R \leq 8\%$).

ANALYSIS OF NEUTRON DIFFRACTION DATA

In order to study the H-bond geometry, or more generally the geometrical parameters containing the H atoms, it is useful to consider neutron diffraction data since for this method the positions of the H-atoms are well determined. Forty-four accurate neutron diffraction O—H···O=C systems and 98 O···H—O—C systems were found in the CSD. For these two data sets the scatter plot is presented in Fig. 1, showing the changes in geometrical parameters during the proton transfer in amino acids in aqueous solutions.

The circles correspond to the neutron diffraction data and the continuous line represents the relation between the O—H bond length and O···H distance within the
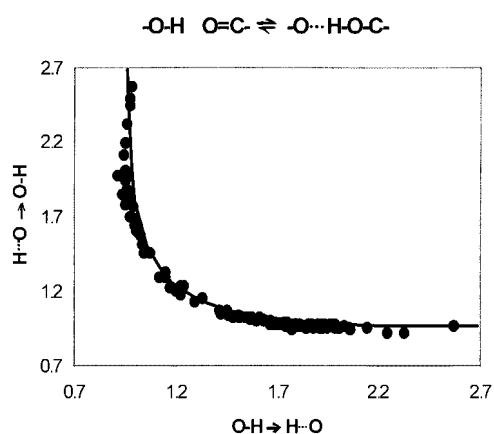

-O-H O=C- ⇌ -O···H-O-C-

Figure 1. Relation between O—H and H···O distances for data sets taken from the CSD and containing C—O—H···O and C=O···H—O systems. All distances in Å

O—H···O bond obtained from the bond valence sum[10] applied for the H atom within O—H···O system.[4,8,9]

$$n_{OH} + n_{H···O} = 1 \qquad (1)$$

where n_{OH} and $n_{H···O}$ are the bond valences of the O—H bond and H···O contact, respectively. The idea of bond valence has been used in many problems in structural chemistry.[4–16] For many types of bonds, the bond valence n correlates inversely with bond length. In order to describe bond valence, two relations are usually used:

$$n_{ij} = (r_{ij}/r_0)^{-m} \qquad (2)$$

and

$$n_{ij} = \exp\left[(r_0 - r_{ij})/B\right] \qquad (3)$$

where r_{ij} is the length of the bond between atoms i and j and n_{ij} is the corresponding bond valence. B, m and r_0 are constants and r_0 is usually the length of the reference single bond not perturbed by intermolecular interactions.

Equations (2) and (3) are a consequence of the character of intermolecular forces.[4,17] Equation (3) was used in this study for calculations of bond valences. For the O—H···O systems $r_0 = 0.957$ Å (the single O—H bond length for water molecules in the gaseous state).[4] The constant value B may be determined from the definition of bond valence [i.e. directly from Eqn. (3)] if we know the value of n_{ij} for a given bond length or contact length r_{ij}. The length of the contact is usually known for the value of n_{ij} equal to 0.5:

$$0.5 = \exp\left[(0.957 - r_{1/2})/B\right] \qquad (4)$$

where $r_{1/2} = 1.22$ Å is based on the averaged data of Speakman.[18] The constant value B obtained from Eqn. (4) corresponding to the H···O pair of atoms is equal to 0.38.

As shown in Fig. 1, the bond valence sum corresponding to the O—H···O systems [eqn (1)] is in very good agreement with the experimental data (obtained from the accurate neutron diffraction measurements) and properly describes the geometrical changes for the amino acid molecules during the proton transfer.

In the original definition of the bond number it was stated that n is the number of electron pairs involved in the bond (the bond valence model described and studied extensively by Brown[10] has a similar meaning to the bond number idea introduced by Pauling,[17] the terminology being in accordance with the bond valence model is applied in this paper). Additionally, according to the bond valence model[10] the oxidation state for a given atom is equal to the sum of bond valences of all its bonds (intermolecular contacts and typical covalent bonds). Then it is possible to express the bond valence sum
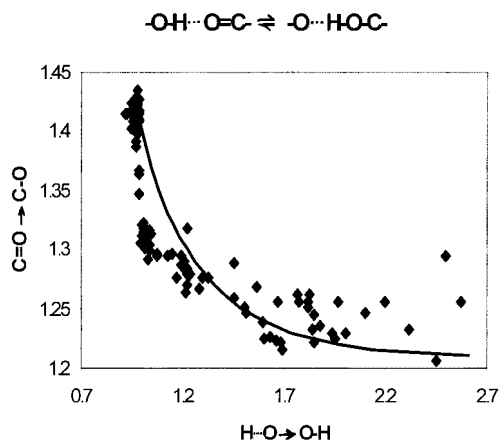
Figure 2. Scatter plot of the C=O bond lengths vs H···O distances and C—O bond lengths vs O—H bond lengths for C—O—H···O and C=O···H—O systems. All distances in Å

Figure 3. N—H vs H···O (or H···N vs O—H), distances in Å

for the oxygen atom within the C=O···H—O and C—O—H···O systems in the following way;

$$n_{C=O} + n_{H···O} = 2 \qquad (5a)$$

or

$$n_{C—O} + n_{O—H} = 2 \qquad (5b)$$

Equation (5a) was used earlier[8] to study the influence of intermolecular forces in the crystals on the C=O bond length.

$n_{C=O}$ (and $n_{C—O}$) in Eqn. (5) is defined according to Eqn. (3);

$$n_{C=O} = \exp\left[(r_{C—O}{}^0 - r_{C=O})/B^*\right] \qquad (6)$$

and the constant B^* can be calculated knowing the lengths of the single and double C—O and C=O bonds (1.428 and 1.209 Å, respectively) and it is equal to 0.32.

Figure 2 shows the theoretical relation (continuous line) based on the bond valence sum applied for the oxygen atom of C=O (and C—O) bond [Eqns (5)]. The squares in Fig. 2 show the neutron diffraction data for the C—O—H···O and C=O···H—O systems. The samples are the same as those used for the plot presented in Fig. 1. The experimental results are in good qualitative agreement with the bond valence sum.

The process of proton transfer within amino acids is also connected with the geometrical changes of the NH_2 groups. Then the —H_2N···H—O and $NH_3{}^+$···O systems were taken into account to show the proton transfer reaction path. The appropriate accurate neutron diffraction geometrical parameteres were those taken from the CSD. Forty-two $NH_3{}^+$···O and two NH_2···H—O systems were found. Only those systems were considered for which the H···O and H···N distances fulfil the criteria of the existence of H-bonds. This means that only the
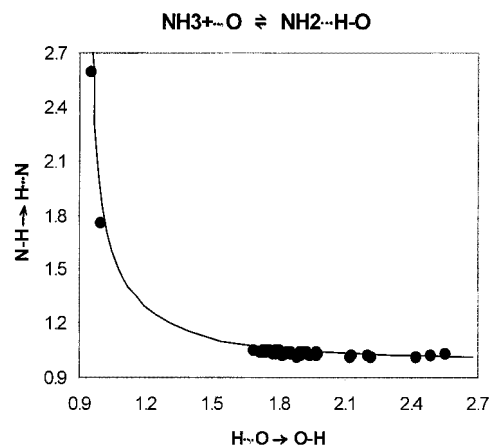
distances not exceeding 2.6 Å were included. Thus the small number of the NH_2···H—O systems is due to rather strict conditions of choosing the systems (accurate data and H···N less than 2.6 Å). The NH_2 group is a weak Lewis base and this may also be the reason why not many such systems have been found.[19,20] The potential role of the amino nitrogen atom as a hydrogen bond acceptor has been described in detail.[20] The authors hold that the amino group can accept hydrogen bonds under special circumstances in macromolecules; additionally, the amino–hydrogen interaction is attractive but weaker than a hydroxyl–hydroxyl hydrogen bond.

The neutron diffraction data are presented in Fig. 3 and compared with the relation (continuous line) obtained from the bond valence sum:

$$n_{N—H} + n_{H···O} = 1 \qquad (7)$$

The way of calculating $n_{H···O}$ (or $n_{O—H}$) was described earlier in this paper. In order to calculate $n_{N—H}$ ($n_{H···N}$) the exponential definition of bond valence was also used [Eqn. (3)]. Then appropriate constant values $r_0 = 1.01$ Å (the single N—H bond length of methylamine[21]), B calculated from the value of $r_{1/2}$ (1.269 Å[22]) and equal to 0.37 were used to calculate the $n_{N—H}$ and $n_{H···N}$ bond valences.

The agrement between the experimental data and the curve obtained from Eqn. (7) is excellent in spite of the simplicity of the model used here.

THE ANALYSIS OF α-AMINO ACIDS

In order to understand better the important role played by the proton transfer and hydrogen bonding for the interactions occurring in polypeptides and proteins, and their conformations, the geometry of α-amino acids has been investigated here. Hence a search of the CSD was performed for ordered, non-deuterated and precise

Table 1. The data set of α-amino acids used for the investigation of their acidities and basicities

Refcode	Structure	R value (%)
DLASPA02	DL-Aspartic acid	4.2
DLSERN11	DL-Serine	2.0
GLUTAM01	L-Glutamine	3.2
GLYCIN03	Glycine (α-form)	3.2
GLYCIN05	Glycine (α-form)	4.5
GLYCIN15	Glycine (γ-form)	2.5
GLYCIN16[a]	Glycine (γ-form)	2.4
LALNIN12	L-Alanine	2.2
LCYSTN12	L-Cysteine	4.6
LGLUAC03	L-Glutamic acid (α-form)	2.1
LGLUAC11	L-Glutamic acid (β-form)	2.6
LHISTD13	L-Histidine	5.8
LTHREO02	L-Threonine	6.8
LTYROS11	L-Tyrosine	4.0

[a] Temperature of the measurement 83 K; for all remaining structures the temperature was 295 K

(e.s.d.s ≤ 0.005 Å and $R \leq 8\%$) neutron crystal structures of α-amino acids. To avoid unfavorable effects from the inhomogeneity of the H-bonds as strong intermolecular interactions, the hydrochloride complexes and hydrated structures were excluded from the investigations. Table 1 presents the structures of the sample chosen for the investigations.

Figure 4 shows the dependence of the aciditiy of N—H bonds of the NH_3^+ groups on bond valences of the corresponding intermolecular contacts. If the N—H bond is not involved in any H-bonding interaction (as an H-donor), then its bond valence n_{NH} should be equal to 1.00. Any N—H···acceptor interaction will extend this bond leading to a decrease of the bond valence value, n_{NH}, thus the quantity $A = 1 - n_{NH}$ will increase. This may be associated with the acidity of the proton involved. The N—H bond lengths were taken from the sample presented in Table 1. The sum of bond valences of intermolecular contacts (H···X) should be a measure of basicity of Xs. The stronger these interactions, the shorter
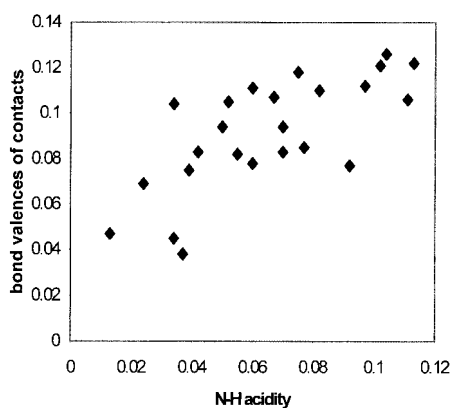
Figure 4. N—H bond acidity plotted against the sum of bond valences of the respective intermolecular contacts

are the H···X contacts and the larger is the value of the N—H acidity, A. The sum of bond valences of intermolecular contacts (H···X) should be equal to the N—H acidity. Figure 4 confirms approximately this relation. In consequence, the bond valence sum for this kind of systems may be described in the following way:[10]

$$n_{N—H} + \Sigma n_{H···X} = 1 \qquad (8)$$

Only the strong H-bonds were taken into account, which means that only short H···O and H···N contacts were included (for greater distances the values of bond valences are negligible).

A MEDIATION ROLE OF A WATER MOLECULE DURING THE PROCESS OF PROTON TRANSFER

In earlier studies, the influence of water on the intramolecular proton transfer in 8-hydroxyquinoline (oxine) and glycine was investigated (J. Langlet, E. Kassab, G. Coquerel, and S. Petit, presented at the 8th International Congress of Quantum Chemistry, 19–23 June 1994, Prague) using *ab initio* calculations (6–31G* basis set, HF and MP2 levels). The authors of that work held that glycine appears to be more sensitive to the solvation effect than oxine since for glycine the zwitterionic form is more stable than the neutral form by 4.3 kcal mol^{-1} and the energy barrier height between these two forms in an aqueous solution is 3.7 kcal mol^{-1} (calculations within the MP2 level). For both isolated and solvated oxine the neutral form is more stable.

Recent investigations on glycine and water[23,24] show that in the gas phase the neutral glycine structure is preferred by about 20 kcal mol^{-1}, while the zwitterion is preferred by about 10 kcal mol^{-1} in water solution. According to the authors it should be possible to observe the transition from the gas phase neutral form to the zwitterionic form in the aqueous phase by successively increasing the number of water molecules that surround the glycine molecule. According to their investigations, two dihydrated zwitterionic glycine structures are possible and hence two mechanisms of the intramolecular transfer; within the first mechanism the transfer of the proton is possible directly from the COO^- group to the NH_3^+ group, and for the second mechanism the transfer is possible via water molecules.

The calculations performed on glycine in the gas phase and in water solution are in agreement with the other investigations. For example, in order to examine the influence of the solvent environment on protein H-bridges, high-level *ab initio* calculations have been performed in the gas phase and in three different solvents on a salt bridge as modeled by formate and guanidinium ions.[25] The authors held that according to the results of calculations the neutral H-bonded complex is favored,
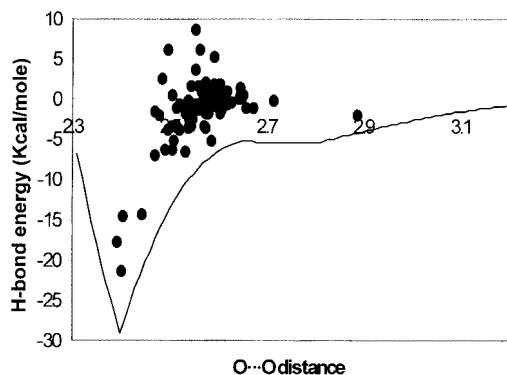
Figure 5. Intramolecular x-ray H-bonds for six-membered systems taken from the CSD

For the curve obtained from the modified Lippincott–Schroeder model[9] and presented in Fig. 5 (solid line), there is the minimum for 2.44 Å, which agrees very well with the MP2 and B3LYP results obtained for the simple system.

When the mean energy of all 89 six-membered H-bonding systems is compared with the mean value for seven-membered systems (seven cases), -1.4 and -15.7 kcal mol^{-1}, we find that the difference is dramatic, indicating a significant role of strain in six-membered rings. The situation in the case of intramolecular proton transfer in amino acids is still worse energetically—the strain in five-membered rings is even stronger and hence this mechanism is much less probable.

whereas in polar solvents such as water and DMSO, the zwitterionic form becomes much more stable. Similar conclusions were presented by Scheiner and Kar,[26] who examined the H-bonds and proton transfer for the complex between formic acid and methylenimine. The results of their calculations showed that the ion pair is more stable than the neutral pair for large values of the dielectric constant.

The correlations shown in this paper confirm the role of water molecules in the mechanism of proton transfer. The mechanism without the mediation of water molecules has also been discussed in the literature (for which the intramolecular H-bonds are the conformations of the reaction path). However, such a mechanism requires considerable changes of the molecular geometry. Apart from bond length variations present in both mechanisms, the changes in bond angle may be energetically very costly. To confirm this point, calculations of the H-bond energies for six-membered intramolecular H-bonds taken from the CSD were performed, applying a modified[9] Lippincott-Schroeder model.[27] In these cases the O—H⋯O bond is forced to approach the O—H⋯O angle close to 120°, thus far away from the optimal value of 180°. As a result, the experimental points in Fig. 5 are biased by this form of energy (strain energy) and that is why they are 5–10 kcal mol^{-1} over the theoretical line (continuous line), which corresponds to the unstrained situation. That situation, for unstrained systems, is presented in detail in Ref. 9, where the points lie exactly on the line. Figure 5 is in agreement with the previous investigations, the hydrogen bond strengths greater than 20 kcal mol^{-1} correspond to low barrier hydrogen bonds[28] for which the O⋯O distance is in the range 2.4–2.5 Å. Additionally, the solid line in Fig. 5 is in agreement with the hydrogen bond energies calculated at three levels of theory as a function of the O⋯O distance.[29] For two levels of theory, MP2/6–31+G(d,p) and B3LYP/6–31+G(d,p), the minimum of the H-bond energy exists for an O⋯O distance of 2.43 Å. The HF curve (the same basis set) shows the minimum for 2.52 Å.

REFERENCES

1. Limbach HH, Manz J. *Ber. Bunsenges. Phys. Chem.* 1998; **102**: 289–291, and references cited therein
2. Tuckerman ME, Marx D, Klein ML, Parrinello M. *Science* 1997; **275**: 817–820.
3. Allen FH, Davies JE, Galloy JE, Johnson JJ, Kennard O, Macrae CF, Mitchell EM, Mitchell GF, Smith JM, Watson DG. *J. Chem. Inf. Comput. Sci.* 1991; **31**: 187–204.
4. Dunitz JD. *X-Ray Analysis and the Structure of Organic Molecules*. Cornell University Press: Ithaca, NY, 1979.
5. Bürgi HB, Dunitz JD. *J. Am. Chem. Soc.* 1987; **109**: 2924–2926.
6. Bürgi HB. *Angew. Chem., Int. Edit. Engl.* 1975; **14**: 460–473.
7. Cieplak AS. In *Structure Correlation*, Bürgi H-B, Dunitz JD (eds). VCH: Weinheim, 1994, 205–302.
8. Grabowski SJ. *Tetrahedron* 1998; **54**: 10153–10160.
9. Grabowski SJ, Krygowski TM. *Tetrahedron* 1998; **54**: 5683–5694;; Grabowski SJ, Krygowski TM. *Chem. Phys. Lett.* 1999; **305**: 247–250.
10. Brown ID. *Acta Crystallogr., Sect. B* 1992; **48**: 553–572, and references cited therein.
11. Dziembowska T, Szczodrowska B, Krygowski TM, Grabowski SJ. *J. Phys. Org. Chem.* 1994; **7**: 142–146.
12. Gilli P, Bertolasi V, Ferretti V, Gilli G. *J. Am. Chem. Soc.* 1994; **116**: 909–915.
13. Gilli G, Bertolasi V, Veronese AC. *Acta Crystallogr., Sect. B* 1983; **39**: 450–456.
14. Grabowski SJ. *Croat. Chem. Acta* 1988; **61**: 805–814.
15. Krygowski TM, Cyrański M. *Tetrahedron* 1996; **52**: 10255–10264.
16. Palenik GJ. *Inorg. Chem.* 1997; **36**: 4888–4890.
17. Pauling L. *J. Am. Chem. Soc.* 1947; **60**: 542.
18. Speakman JC. In *Molecular Structure by Diffraction Methods*, vol. 3, Sim GA, Sutton LE (eds). Chemical Society: London, 1975, 89.
19. Etter MC. *Acc. Chem. Res.* 1990; **23**: 120–126.
20. Luisi B, Orozco M, Sponer J, Luque FJ, Shakked Z. *J. Mol. Biol.* 1998; **279**: 1123–1136.
21. Legon AC, Millen DJ. *Chem. Phys. Lett.* 1988; **147**: 484–489.
22. Saupe T, Kreger, Staab HA. *Angew. Chem. Int. Ed. Engl.* 1986; **25**: 451.
23. Jensen JH, Gordon MS. *J. Am. Chem. Soc.* 1995; **117**: 8159–8170.
24. Gordon MS, Jensen JH. *Acc. Chem. Res.* 1996; **29**: 536–543.
25. Zheng Y, Ornstein LR. *J. Am. Chem. Soc.* 1996; **118**: 11237–11243.
26. Scheiner S, Kar T. *J. Am. Chem. Soc.* 1997; **117**: 6970–6975.
27. Lippincott ER, Schroeder R. *J. Chem. Phys.* 1955; **23**: 1099–1106.
28. Garcia-Viloca M, González-Lafont A, Lluch JM. *J. Am. Chem. Soc.* 1997; **119**: 1081–1086.
29. Smallwood CJ, McAllister MA. *J. Am. Chem. Soc.* 1997; **119**: 11277–11281.