Amino Acids (1994) 7: 223–230

Amino
Acids

Structure-energy relationship in ω-amino acids and related compounds

U. Seebacher and **M. Ramek**

Institut für Physikalische und Theoretische Chemie, Technische Universität Graz, Graz, Austria

Accepted October 27, 1993

Summary. The potential energy surfaces of various ω-amino acids, ω-hydroxy acids, and ω-amino alcohols have been investigated by ab-initio RHF calculations in the recent years. In order to find a common scheme for these molecules an energy function was developed, in which the energy is defined as a function of only those dihedral angles that are essential for the description of the nuclear framework. A least square fit procedure is performed to yield the parameters. Minimum geometries with unfavourable contributions are sorted out, as long as the correlation coefficient is below a predefined threshhold. The differences between the given (quantum chemical) and the estimated energies are small for minimum geometries which were used to fit the parameters, and larger for those which have been omitted in the fit. For conformers with intramolecular hydrogen bonding these energy differences turn out to be in the range of the interaction energy predicted for the H-bond contribution. The method can therefore be used as a semi-quantitative measure of the influence of additional interactions to the conformational energy.

Keywords: Amino acids – Interaction – Hydrogen bond – Structure – Ab-initio – Energy

Abbreviations: GABA: γ-aminobutyric acid; PES: potential energy surface; RHF: Roothaan-Hartree-Fock.

Introduction

The potential energy surfaces of ω-amino acids, ω-hydroxy acids, and ω-amino alcohols with 2 to 4 carbon atoms have been investigated by ab-initio RHF calculations (Ramek, 1990a, b; Kelterer and Ramek, 1991, 1993; Flock and Ramek, 1992, 1993; Kelterer et al., 1992; Ramek and Flock, 1993). These molecules share some structural features due to the two functional groups, which are located at each end of the molecule. The most important of these features is that these molecules tend to build ring structures, which are stabilized by an

intramolecular H-bond. In the case of β-alanine the structure-energy relationships become obvious when the conformations are compared with each other, and the H-bond influence can be estimated in a straightforward manner from the energy differences between similar conformations (Ramek, 1990a). This is, however, not generalizable: in the case of GABA no such comparison is possible due to the greater flexibility of the molecule. The same is true for γ-hydroxybutyric acid, 4-aminobutanol, β-hydroxypropionic acid, and 3-aminopropanol. The goal of the present study was to develop an energy function, which is suitable to describe this class of molecules and which allows the quantification of intramolecular interaction energies.

Method

The energy of a molecule is usually described as a function of bond lengths R, bond angles Θ and torsional angles Φ:

$$E_{\text{bonds}} = \sum_{\text{bonds}} K_R(R - R_0)^2 + \sum_{\text{angles}} K_\Theta(\Theta - \Theta_0)^2 + \sum_n \sum_{\text{dihedrals}} V_{n,\Phi} \cos(n\Phi)$$

The first two terms use a quadratic potential, since the harmonic oscillator application is valid near equilibrium geometries. The potential for an internal rotation, the torsion of the molecule, must be periodic and is modelled by a cosine or a sum over cosines. R_0 and Θ_0 are the equilibrium values of bond lengths and valence angles.

Since most of the molecules discussed here have more local minima in the PES than variables of the energy function (e.g. GABA: 62 minima, 42 internal coordinates; γ-hydroxybutyric acid: 64 minima, 39 internal coordinates; 4-aminobutanol: 110 minima, 48 internals), it is a convincing idea to optimize the parameters K_R, K_Θ, and V_n in a least-square fit. There is, however, an essential difference between such a fit and the usual energy function mentioned above: K_R, K_Θ, and V_n are parameters for individual equilibrium geometries in the usual energy function, but global values in the case of the least-square fit. Similarily, R_0 and Θ_0 are no longer equilibrium values of individual conformations, but mean values in the latter case. Indeed such a least square fit can be performed in a straightforward manner; for numerical reasons it is necessary to take the relative energies of the equilibrium conformations (these are the absolute energies minus the energy of the global minimum, transformed to kJ/mol or kcal/mol). Since there are always some minimum geometries that do not fit into the regular pattern that is common to most of the conformations, the correlation of the least square fit will be poor. The correlation coefficient r is defined as

$$r = \frac{\left(\sum_i E_{\text{fit},i} - E_{\text{mean}}\right)^2}{\left(\sum_i E_{\text{given},i} - E_{\text{mean}}\right)^2}$$

with E_{mean} being the mean value of all given (ab-initio) energies.

We therefore used a procedure, which in an iterative manner excludes each geometry successively from the fit. The geometry, whose omission yields the most significant improvement of the correlation, is left out for the next loop. The procedure stops, when enough minima have been sorted out to yield a correlation coefficient better than some predefined threshold. Then for each geometry the energy is calculated. For the minima sorted out the difference between the ab-initio energy and the estimated energy is a measure for additional contributions, which result from other than the steric parameters used for the energy function.

Since the goal is to catch those features of the molecules that are essential for their conformational behaviour, we performed an assay scanning a range of energy functions. First a function giving a full description of the molecule by bond lengths, bond angles and

torsional angles was tested:

$$E(R, \Theta, \Phi) = E_0 + \sum_{\text{bonds}} K_R(R - R_0)^2 + \sum_{\text{angles}} K_\Theta(\Theta - \Theta_0)^2 + \sum_{\text{dihedrals}} V_{1,\Phi} \cos(\Phi).$$

It should be pointed out that the reference energy E_0 in this function has no physical meaning. The most important requirement for a function to be successful with regard to the present purpose, is the "production" of reasonable energy differences between given (ab-initio) and calculated energies for geometries with extra intramolecular interactions. From former estimates (Ramek, 1990a, b) approximate values of the H-bond contribution for β-alanine and GABA are known. Therefore these values were taken as a first performance measure for the function under investigation. The "full" function $E(R, \Theta, \Phi)$ did not yield satisfactory results. Since it is obvious that the inclusion of redundant information in the energy function makes it more difficult for the fitting algorithm to discern between dominant and minor contributions, various simplifications of this ansatz were tested. It turned out that omission of the valence angles contributions, i. e. an energy function

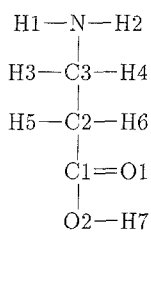
$$E(R, \Phi) = E_0 + \sum_{\text{bonds}} K_R(R - R_0)^2 + \sum_{\text{dihedrals}} V_{1,\Phi} \cos(\Phi)$$

improved the result considerably. From this it may be inferred that these angles are not very specific for the structure-energy relationship.

A further improvement was achieved by omitting the bond lengths too. This is a somehow striking result, since the bond lengths in the equilibrium structures considered exhibit a significant amount of variation. However, most of this scattering is reflected in the values of the dihedral angles (cf. Fig. 4 in (Ramek, 1990a) and Fig. 4 in (Ramek et al., 1992)), so that the bond lengths can be excluded from the energy function. (As a matter of fact, the torsional angle values are sufficient to reconstruct the structure of a molecule, assuming standard values for bond lengths and valence angles.)

This limitation of the energy function to terms describing the dihedral angles only greatly reduces the number of variables: for a molecule with n atoms only $n - 3$ variables are needed instead of $3n - 6$. This has the remarkable side effect that smaller compounds can be treated too. The PES of β-alanine, e.g., has 21 symmetry unique local minima, a number which is smaller than 33, the number of internal coordinates that are necessary for an exact description of the nuclear framework of β-alanine, but larger than 10, the number of dihedral angles.

With a complete set of torsional angles the molecule is still overdetermined; for saturated hydrocarbons, e.g., it is sufficient to define the orientation of the carbon atoms. It turned out that our method works best when the energy is represented only as a sum over the cosines of the essential dihedral angles, i.e. the ones related to the main chain and the terminal groups. Fig. 1 exemplifies this for β-alanine.



H1–N–C3–C2 **H2–N–C3–C2**
H3–C3–C2–C1 H4–C3–C2–C1
H5–C2–C3–N H6–C2–C3–N
O1–C1–C2–C3
H7–O2–C1–C2
N–C3–C2–C1
C3–C2–C1–O2

Fig. 1. Nuclear framework of β-alanine together with a list of all dihedral angles. The dihedral angles related to the terminal groups and those describing the main chain (given in bold face type) are essential for the orientation of the molecule

The final energy function, which captures all the essential features of the molecules under investigation, therefore is

$$E = E_0 + \sum_{\text{essential}} V_\Phi \cos \Phi.$$

Other variations of the energy function such as expansion to a series of cosines or cosines plus sines always lead to a deterioration of the results. Inclusion of saddle point geometries was first thought to be a further possibility to enlarge the equation pool for statistical reasons, but their occurence in the equation system only confuses the determination of the important structural contributions.

The choice of the threshold for the correlation coefficient is not critical, and with the omission of all unusual conformations (i.e., H-bonded and sterically hindered ones) the correlation soon becomes better than 0.99. Tuning the threshhold to a precision of 0.999 or even 0.9999 does not alter the results significantly, but only effects the execution time. Furthermore, the value of 0.99 is in accordance with the restricted precision of the description by the essential dihedral angles only. This value therefore was chosen as threshhold for all subsequent calculations.

Results

To further exploit the suitability of the method described above, calculations were performed for the three series glycine, β-alanine, and GABA; 2-aminoethanol, 3-aminopropanol, and 4-aminobutanol; glycolic acid, β-hydroxypropionic acid, and γ-hydroxybutyric acid. Table 1 gives a list of the essential dihedral angles of these compounds. The energy function

$$E = E_0 + \sum_{\text{essential}} V_\Phi \cos \Phi$$

was used throughout; geometries with non-covalent interatomic distances less than 0.98% of the sum of the van der Waal's radii were excluded from the fit in an initial step.

Table 2 summarizes the coefficients V_Φ as well as the reference energy term E_0. Obviously the most prominent torsional angle is the one that describes the conformation of the carboxy group (dihedral 5) in the ω-amino and ω-hydroxy acids. From the value of this coefficient the energy difference between syn-periplanar- and anti-periplanar-configurations of the carboxy group may be inferred as shown in Table 3.

As already mentioned, the difference between the given (ab-initio) energies and the ones calculated via the least square fit procedure may be expected to be a good measure for the intramolecular interactions in those conformers that have been excluded during the fit. Table 4 lists these energies for the minima with the strongest hydrogen bonds. For β-alanine the value in Table 4 (35.7 kJ/mol) compares well with the earlier estimate of 37.5 $\pm$ 1.5 kJ/mol (Ramek, 1990a).

Discussion

The method described above works quite well for all of the test molecules. Of course, better results are likely if more data of local minima are available. These limitations become obvious in the case of the C2-compounds (glycine, ethanol-

Table 1. Essential dihedral angles of the set of test molecules

	Dihedral 1	Dihedral 2	Dihedral 3	Dihedral 4	Dihedral 5	Dihedral 6
Glycine	O2 C1 C2 N	H1 N C2 C1	H2 N C2 C1		H O2 C1 C2	
2-Aminoethanol	O2 C1 C2 N	H1 N C2 C1	H2 N C2 C1		H O2 C1 C2	
Glycolic acid	O2 C1 C2 O3	H O3 C2 C1			H O2 C1 C2	
β-Alanine	C1 C2 C3 N	H1 N C3 C2	H2 N C3 C2	O1 C1 C2 C3	H O2 C1 C2	
3-Aminopropanol	C1 C2 C3 N	H1 N C3 C2	H2 N C3 C2	O1 C1 C2 C3	H O2 C1 C2	
β-Hydroxypropionic acid	C1 C2 C3 O	H O C3 C2		O1 C1 C2 C3	H O2 C1 C2	
GABA	C2 C3 C4 N	H1 N C4 C3	H2 N C4 C3	O1 C1 C2 C3	H O2 C1 C2	C1 C2 C3 C4
4-Aminobutanol	C2 C3 C4 N	H1 N C4 C3	H2 N C4 C3	O1 C1 C2 C3	H O2 C1 C2	C1 C2 C3 C4
γ-Hydroxybutyric acid	C2 C3 C4 O	H O C4 C3		O1 C1 C2 C3	H O2 C1 C2	C1 C2 C3 C4

Table 2. Coefficients and reference energy (kJ/mol) of the least square fit of the set of test molecules

	Coeff. 1	Coeff. 2	Coeff. 3	Coeff. 4	Coeff. 5	Coeff. 6	E_0
Glycine	3.7454	5.2806	−4.6170		17.9170		39.6280
2-Aminoethanol	−2.0655	1.5760	1.0546		2.1833		20.9463
Glycolic acid	2.4653	−12.1500			20.1700		61.0496
β-Alanine	−3.4488	1.4381	1.4805	4.4647	17.6108		43.6532
3-Aminopropanol	−0.2893	−2.1284	3.1378	−1.8924	1.6046		33.6992
β-Hydroxypropionic acid	−1.6011	2.6127		3.2029	20.1822		63.4228
GABA	−2.2123	0.6406	1.8123	2.9681	16.9285	−0.2323	47.8540
4-Aminobutanol	0.9456	−1.4159	−1.0022	−0.1277	0.9973	3.7109	42.1489
γ-Hydroxybutyric acid	−4.0244	1.2981		4.0670	17.6735	−0.5456	51.9209

Table 3. The energy difference (kJ/mol) between the syn-periplanar and anti-periplanar orientation of the carboxy group of the tested acids can easily be obtained by multiplying the corresponding coefficient with (cos 0° − cos 180°) = 2

	Coeff. 5	ΔE
Glycine	17.9170	35.8340
β-Alanine	17.6108	35.2216
GABA	16.9285	33.8570
Glycolic acid	20.1700	40.3400
β-Hydroxypropionic acid	20.1822	40.3644
γ-Hydroxybutyric acid	17.6735	35.3470

Table 4. Hydrogen bond energies (kJ/mol) for the most stable H-bonded conformers in the set of test molecules

Molecule	Reference	$E_{\text{ab initio}}$	E_{calc}	ΔE
Glycine	Ramek and Cheng, 1992	11.896	42.830	30.934
Glycolic acid	Flock and Ramek, 1992	13.380	67.771	54.391
Ethanolamine	Kelterer and Ramek, 1991	0.000	11.281	11.281
β-Alanine	Ramek, 1990a	6.477	42.207	35.730
3-Aminopropanol	Kelterer et al., 1992	0.000	14.596	14.596
β-Hydroxypropionic acid	to be published	16.631	52.232	36.601
GABA	Ramek and Flock, 1993	6.960	42.411	35.451
4-Aminobutanol	Kelterer and Ramek, 1993	0.000	24.870	24.870
γ-Hydroxybutyric acid	Flock and Ramek, 1993	10.250	45.230	34.980

amine, and glycolic acid), which have only few local minima in their potential energy surfaces (7, 10, and 6, respectively). For glycine only the minima with overlapping van der Waal's radii have been omitted in the fitting procedure, for glycolic acid all minima except the one with the strongest H-bonding were taken in the fit. Nevertheless, the values calculated for the C2-compounds are not too bad. Fig. 2 shows the energies of the intramolecular O–H···O and O–H···N interactions as a function of the the ring size for all acid conformers with this hydrogen bond. In these displays, the values of the C2-compounds, especially the one for glycolic acid, differ noteably from those of the C3- and C4-compounds. The latter indicate a fairly constant interaction energy for the conformers with the strongest H-bond. For GABA, β-hydroxypropionic acid, and γ-hydroxybutyric acid more than one conformer exists with this specific interaction. The lesser interaction energy, which is predicted for these conformers, is in good agreement with the ab-initio data (bond lengths, vibration frequencies, and bond orders).

Fig. 3 compares the estimated interaction energy for the ω-amino-*n*-alkanoles with the O–H distances, which also reflect the amount of intramolecular interaction. The trends in both quantities are essentially the same; here the value of the C2-compound is in better agreement with those of the

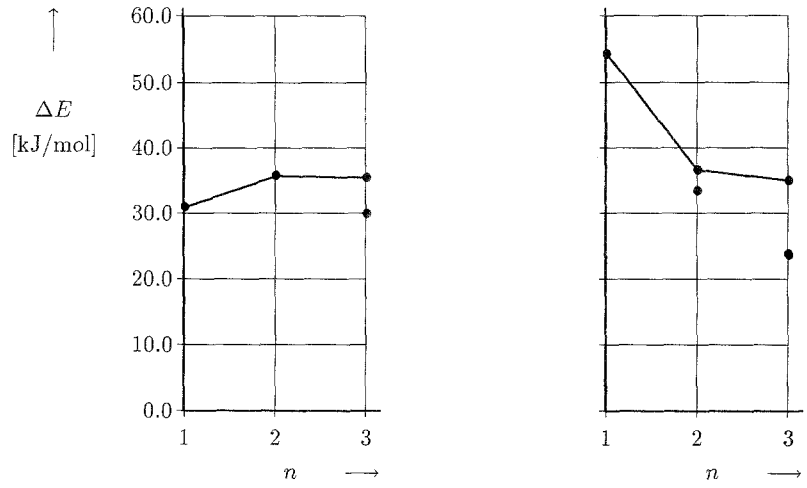


Fig. 2. Estimated energies of the intramolecular hydrogen bonds of the ω-amino acids $H_2N-(CH_2)_n-COOH$ (left) and the ω-hydroxy acids $HO-(CH_2)_n-COOH$ (right). Glycine, glycolic acid, and β-alanine form one symmetry unique H-bonded conformer, GABA and β-hydroxypropionic acid form two H-bonded conformers. γ-Hydroxybutyric acid forms three H-bonded conformers, two of which cannot be distinguished in the display

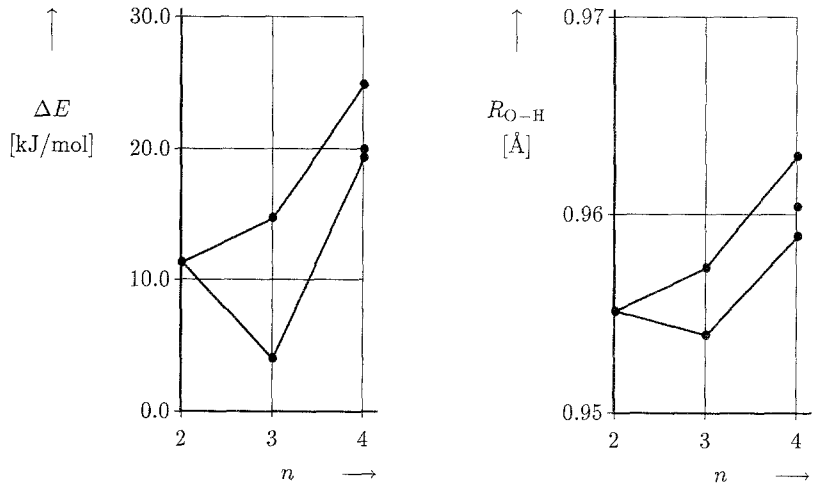


Fig. 3. Estimated energies of the intramolecular hydrogen bonds in the ω-amino-n-alkanols $H_2N-(CH_2)_n-OH$, compared with the equilibrium O–H distances. 2-Aminoethanol ($n = 2$) forms one symmetry unique H-bonded conformer, 3-aminopropanol forms two H-bonded conformers, and 4-aminobutanol forms three H-bonded conformers

C3- and C4-compounds, which seems to be due to the slightly larger number of local minima in the PES of ethanolamine.

It can be seen from Table 4 that the interaction energy is slightly larger for the hydroxy acids than for the amino acids. This is remarkable with respect to the fact that the hydrogen bond is stronger in the amino acids, but it may be explained by considering that the intramolecular O–H···O and O–H···N interactions also include electrostatic contributions, which are larger in the hydroxy acids due to the higher electronegativity of the oxygen atom.

The method in its current form is limited to unbranched compounds and has been tested only for few compounds, the data of which were readily accessible to us. The results are encouraging enough to justify an extension to branched compounds also, which would allow the treatment of such important compounds as α-amino acids or cyclic systems.

References

Flock M, Ramek M (1992) Ab-initio SCF investigation of glycolic acid. Int J Quant Chem: Quant Chem Symp 26: 505–515

Flock M, Ramek M (1993) Ab-initio SCF investigation of γ-hydroxybutyric acid. J Mol Struct (Theochem) (accepted)

Kelterer A-M, Ramek M (1991) Intramolecular hydrogen bonding in 2-aminoethanol, 3-aminopropanol and 4-aminobutanol. J Mol Struct (Theochem) 232: 189–201

Kelterer A-M, Flock M, Ramek M (1992) Ab initio SCF investigation of 3-aminopropanol and 3-aminopropanal. J Mol Struct (Theochem) 276: 35–59

Kelterer A-M, Ramek M (1993) Ab initio SCF investigation of the potential energy surface of 4-aminobutanol. Int J Quant Chem: Quant Chem Symp 27: 479–490

Ramek M (1990a) Ab-initio SCF investigation of β-alanine. J Mol Struct (Theochem) 208: 301–355

Ramek M (1990b) Intramolecular hydrogen bonding in neutral glycine, β-alanine, γ-aminobutyric acid, and δ-aminopentane acid. Int J Quant Chem: Quant Biol Symp 17: 45–53

Ramek M, Cheng VKW (1992) On the role of polarization functions in SCF calculations of glycine and related systems with intramolecular hydrogen bonding. Int J Quant Chem: Quant Biol Symp 19: 15–26

Ramek M, Flock M (1993) Ab-initio SCF investigation of γ-aminobutyric acid. Amino Acids (accepted)

Ramek M, Flock M, Kelterer A-M, Cheng VKW (1992) Intramolecular interactions in β-alanine, 3-aminopropanal and 3-aminopropanol. J Mol Struct (Theochem) 276: 61–81

Authors' address: Doz. Dr. M. Ramek, Institut für Physikalische und Theoretische Chemie, Technische Universität Graz, Brockmanngasse 27/4, A-8010 Graz, Austria.

Received April 5, 1993

Verleger: Springer-Verlag KG, Sachsenplatz 4–6, A-1201 Wien. – Herausgeber: Prof. Dr. Gert Lubec, Universitäts-Kinderklinik, Währinger Gürtel 18–20, A-1090 Wien. – Redaktion: Sachsenplatz 4–6, A-1201 Wien. – Satz und Umbruch: Asco Trade Typesetting Limited, Hong Kong; Offsetdruck: Novographic, Ing. W. Schmid, Maurer-Lange-Gasse 64, A-1238 Wien. – Verlagsort: Wien – Herstellungsort: Wien. Printed in Austria.