Journal of Computer-Aided Molecular Design, 5 (1991) 599–616
ESCOM

J-CAMD 138

Reaction mechanisms in peptide synthesis. Part 2. Tautomerism of the peptide bond

Jerzy Ciarkowski*, Francis M.F. Chen and N. Leo Benoiton**

Department of Biochemistry, Faculty of Medicine, University of Ottawa, Ottawa, Ont., Canada K1H 8M5

Received 1 December 1990
Accepted 26 April 1991

Key words AM1; MOPAC, *N*-Acetylglycine amide; α-Hydroxyimine; Dimerization, Hydrogen exchange

SUMMARY

We had concluded in previous work that ring opening of a 2-alkyl-5(4*H*)-oxazolone by water or ammonia leads to transient high-energy imidol intermediates which instantly tautomerize to the native amides. Using the MOPAC molecular orbital program, detailed geometric and energetic characteristics of the tautomerism of a peptide bond have been determined on the AM1 level. The results demonstrate that tautomerism of a peptide bond comprises a three-stage process involving three successive transition states and a bimolecular mechanism: (i) $E \rightarrow Z$ peptide bond isomerization followed by dimerization, (ii) concerted double-hydrogen exchange leading to an α-hydroxyimine (imidic acid) followed by splitting of the dimer, and (iii) $Z \rightarrow E$ *N*-methylimine inversion. While pathway (iii→ii→i) is predicted as a feasible route terminating in the formation of a peptide bond, the inverse route (iii←ii←i) is excluded as a possible initial step in the generation of a 5(4*H*)-oxazolone intermediate.

INTRODUCTION

In previous work [1] we reported that the semiempirical molecular orbital AM1 method [2] seemed to be reliable enough for making semiquantitative predictions on the mechanisms of reactions involved in peptide synthesis. Specifically, by studying the possible ways of addition of simple oxygen or nitrogen nucleophiles to 2-methyl-5(4*H*)-oxazolone, which is representative of the activated form of the parent acid that is responsible for racemization in peptide synthesis [3–5], the AM1 method led us to the conclusion that the tautomeric imidic acid form of the amide may be a high-energy intermediate in a sequence of elementary reactions leading eventually to the elongation of a peptide chain by one amino acid residue:

*On leave from the University of Gdansk, Poland.
**To whom correspondence should be addressed.

0920-654X/$ 5.00

oxazolone + :XH → amide tautomer-X

where X = OH or NH_2.

To the best of our knowledge, a transient presence of the imidic acid as an intermediate in reactions of secondary amides has never been reported. This may be simply because the imidic acid has an elevated internal energy compared to that of the parent amide and/or because the energy barrier for its conversion to the amide is relatively low:

	RCO-NHR′ **1, 3, 5,** or **7**			↔	R(OH)C=NR′ **2, 4, 6** or **8**		
		R	R′			R	R′
where	**1, 2**:	H	H	;	**3, 4**:	CH_3	CH_2COOH
	5, 6:	CH_3	CH_2CONH_2	;	**7, 8**:	H	$CH_2COO(CO)OCH_3$

Only in the field of heterocyclic chemistry has prototropy involving an amide bond been recognized and well documented [6,7]. The process usually takes up the form of a lactam/lactim mutual proton exchange. Recent reports attribute a significant role to this kind of tautomerism in purine and pyrimidine bases as a possible cause of genetic mutations [8,9]. The apparent lack of prototropy in open-chain amides as compared to heterocyclic amides is not so surprizing when one considers that generation of the lactam bond in a heterocycle destroys formal aromacity while generation of a lactim usually implies a build-up of formal aromacity [6]. Thus, in heterocycles, both effects tend to diminish or even cancel the energy difference between the lactam and lactim states by increasing the energy of the former and simultaneously decreasing the energy of the latter. In this context it is interesting to note that the first suggestion of tautomerism as a type of chemical behavior was made for the amide/imidic acid isomerization of isatin in 1882 by A. Baeyer as part of his classical work on indigo [10].

Our current objective is to shed more light on open-chain amide/imidic acid tautomerism. While the above considerations suggest that conversion of an imidic acid into its parent amide is facile, and that the inverse does not occur, the following provides rigorous characterization of all the processes that can possibly be associated with such a tautomerism. The simplest case of formamide tautomerism (**1**/**2**) was analyzed in depth, compared with the evidence available in the literature [11], and the data were then used as a basis and reference for analyses of *N*-acetylglycine, Ac-Gly (**3**/**4**), *N*-acetylglycine amide, Ac-Gly-NH_2 (**5**/**6**), and *N*-formylglycine methylcarbonic acid anhydride, For-Gly-O-COOMe (**7**/**8**).

MATERIALS AND METHODS

Computing was done on the University of Ottawa IBM/370-compatible Amdahl 5860 mainframe computer that is managed by a VM/CMS operating system using the General Molecular Orbital Package MOPAC, version 4.01 [12]. All pertinent MOPAC options were explored for the formamide/formimidic acid case; only the PATH facility and subsequent gradient norm refinement [12] were used in the other cases. That the latter were sufficient for achieving the objectives issues from the first study. Initial geometries of flexible molecules were taken either directly as an

outcome of the reactions of 2-methyl-5(4*H*)-oxazolone with water and ammonia [1] for the tautomers of Ac-Gly and Ac-Gly-NH_2, respectively, or they were restricted to fulfill some constraints with specific goals in mind, such as for the **7**/**8** tautomerism (see below). SYMMETRY [12] was exploited to reduce the computer time when justified. Apart from restrictions arising from symmetry and/or a selected reaction coordinate, geometries were always fully optimized on the AM1 level. In particular, they were optimized for all the stationary points corresponding to either the energy minima or to the saddles (transition states, TrSts). Even though MOPAC allows for thermodynamic calculations no attempts were made to do them since the AM1 method is parametrized against experimental heat-of-formation (HoF) values at standard conditions [2] so it incorporates the zero-point-energy and thermal energy of vibrations and rotations.

The effect of a weakly polar solvent (μ = 1.6 D) not interacting with the solute was simulated in selected cases by means of a 'solvation shell' composed of dipoles (μ = 1.6 D) located in the vertices of a regular eicosahedron and surrounding their contents, the reacting species [1]. The mutual orientation of the dipoles making the solvation cavity, the cavity size and contents, and the relative positions of the reactants and cavity were always allowed to relax fully, as described in detail in the accompanying paper [1].

The molecular images presented were prepared by use of the PCMODEL or PCDISPLAY program, supplied by Serena Software (Bloomington, IN), on a PC AT-compatible desktop computer. Both programs can directly accommodate the data formats typical of MOPAC.

Only the geometry of the species pertinent to the **1**/**2** tautomerism (i.e. the changing geometry of the reaction sites in **3**/**4**, **5**/**6** and **7**/**8**) are provided (Table 1). The optimized geometries for all the remaining species and pertinent to Figs. 2, 3 and 8 are supplied as *supplementary material* *.

RESULTS AND DISCUSSION

All attempts to 'execute' the transfer of a hydrogen atom from an amide nitrogen to the carbonyl oxygen (or vice versa) of the same amide group revealed unreasonably high activation barriers. Thus, it became apparent that the tautomerism must be at least a bimolecular event. The obvious candidate for consideration was the cyclic amide (lactam) dimer [7,15a], that is analogous to the classical carboxylic acid dimer [15b]. Formamide, the simplest example, was examined first. This choice could offer the speed and versatility required for testing the method, and hopefully help avoid time-consuming calculations on the more elaborate models **3**–**8**. In addition, available experimental [16,17] and theoretical [11] data for this compound provided a suitable reference for assessing the results. Previous considerations had referred to the tautomeric forms in their energy minima, and neglected TrSts and associated reaction coordinates. We believe that the present study is the first attempt to define details of the *course* of amide/imidic acid tautomerism.

*Formamide (**1**)/formimidic acid (**2**) tautomerism*

Since each of the semiempirical methods included in MOPAC has been parametrized against experimental values of HoF at standard conditions (1 at 25°C), the results of the HoF calculations presented below refer to the gas phase. The results for the **1**/**2** system in vacuo as well as the litera-

*Contact: Depository of Unpublished Data, CISTI, National Research Council of Canada, Ottawa, Ont., Canada K1A 0S2.

TABLE 1
A CONCISE PROTOCOL FROM THE AM1 OPTIMIZATION OF THE REACTANTS[a]

No.	Final geometry							Connectivities N1	N2	N3	Formamide	Gas phase[b]
	AM1, T = 280, symmetry formamide											
1	C	0.000000	0	0.000000	0	0.000000	0	0	0	0	C1-N3	1.352
2	O	1.244471	1	0.000000	0	0.000000	0	1	0	0	C1-O2	1.219
3	N	1.365315	1	122.047713	1	0 000000	0	1	2	0	C1-H5	1.098
4	H	0.989646	1	120 708627	1	0.437064	1	3	1	2	N3-H4	1.002
5	H	1.114540	1	122.944956	1	−179.855174	1	1	2	3	N3-H6	1.002
6	H	0.986773	1	121 132248	1	−179.366240	1	3	1	2	H4-N3-H6	121.6
											H4-N3-C1	118.5
											N3-C1-O2	124.7
											N3-C1-H5	112.7
	AM1, T = 280, symmetry formamide tautomer											
1	C	0.000000	0	0.000000	0	0.000000	0	0	0	0		
2	O	1.373753	1	0.000000	0	0.000000	0	1	0	0		
3	N	1.283409	1	121.500391	1	0.000000	0	1	2	0		
4	H	0.971070	1	110.935686	1	−0.001141	1	2	1	3		
5	H	1.112043	1	107.831919	1	179.999981	1	1	2	3		
6	H	0.994649	1	115.725850	1	−179.999786	1	3	1	2		
	AM1, T = 280, symmetry formamide dimer											
1	XX	0 000000	0	0.000000	0	0.000000	0	0	0	0		
2	XX	1.000000	0	0.000000	0	0.000000	0	1	0	0		
3	C	2.165858	0	90.000000	0	0.000000	0	1	2	0		
4	O	1.249332	1	63.337151	1	179.996532	1	3	1	2		
5	N	1.359239	1	122.201361	1	−0.062515	1	3	4	1		
6	H	0.996455	1	121.256942	1	0.028873	1	5	3	4		
7	H	1.114836	1	122.120757	1	180.012443	1	3	4	5		
8	H	0.986936	1	120.793369	1	179.984421	1	5	3	4		
9	C	4.331716	1	0.000000	0	0.000000	0	3	1	2		
10	O	1.249332	0	63.337151	0	−0.003468	0	9	1	2		
11	N	1.359239	0	122.201361	0	−0.062515	0	9	10	1		
12	H	0.996455	0	121 256942	0	0.028873	0	11	9	10		
13	H	1 114836	0	122.120757	0	180.012443	0	9	10	11		
14	H	0.986936	0	120 793369	0	179.984421	0	11	9	10		
	AM1, T = 280, symmetry formamide tautomer dimer											
1	XX	0 000000	0	0.000000	0	0.000000	0	0	0	0		
2	XX	1.000000	0	0 000000	0	0.000000	0	1	0	0		
3	C	2.394497	0	90.000000	0	0.000000	0	1	2	0		
4	O	1.369545	1	65.027596	1	−179.904412	1	3	1	2		
5	N	1.285185	1	122.202116	1	−0.175220	1	3	4	1		
6	H	2.544636	1	125.110820	1	0.126853	1	5	3	4		
7	H	1.112785	1	107 960413	1	180.003370	1	3	4	5		
8	H	0.994622	1	115.425881	1	−179.997686	1	5	3	4		
9	C	4.788995	1	0 000000	0	0.000000	0	3	1	2		
10	O	1.369545	0	65.027596	0	0.095588	0	9	1	2		
11	N	1.285185	0	122.202116	0	−0.175220	0	9	10	1		
12	H	2.544636	0	125.110820	0	0.126853	0	11	9	10		
13	H	1.112785	0	107.960413	0	180.003370	0	9	10	11		
14	H	0 994622	0	115.425881	0	−179.997686	0	11	9	10		

TABLE 1 (continued)

No.	Final geometry						Connectivities N1	N2	N3
	AM1, T = 1600, symmetry NLLSQ restart formamide/tautomer dimer TrSt								
1	XX	0.000000	0	0.000000	0	0.000000	0	0	0	0
2	XX	1.000000	0	0.000000	0	0.000000	0	1	0	0
3	C	1.879617	0	90.000000	0	0.000000	0	1	2	0
4	O	1.312983	1	64.615253	1	179.999389	1	3	1	2
5	N	1.309983	1	123.545219	1	0.000435	1	3	4	1
6	H	1.322212	1	123.094271	1	−0.001596	1	5	3	4
7	H	1.115965	1	113.427328	1	−179.926931	1	3	4	5
8	H	0.988851	1	117.819438	1	180.002324	1	5	3	4
9	C	3.759234	1	0.000000	0	0.000000	0	3	1	2
10	O	1.312983	0	64.615253	0	−0.000611	0	9	1	2
11	N	1.309983	0	123.545219	0	0.000435	0	9	10	1
12	H	1.322212	0	123.094271	0	−0.001596	0	11	9	10
13	H	1.115965	0	113.427328	0	−179.926931	0	9	10	11
14	H	0.988851	0	117.819438	0	180.002324	0	11	9	10

[a] These are reproductions of 'archive files generated by MOPAC and subsequently submitted to a minor re-editing. In an archive file the molecular geometry is represented by an array of data fulfilling a certain standard format, which is called the Z-matrix. In this array, the three first real-number columns refer to bond lengths, bond angles and dihedral angles (i.e., *internal coordinates*) defined by means of the consecutive atoms (N0) and their connectivities (N1, N2 and N3 respectively). The 1 or 0 immediately following an internal coordinate indicates that the preceding value was or was not optimized, respectively, during the calculation. For example, the fourth line in the Z-matrix for $(\mathbf{1})_2$ indicates that the length of bond O(**4**)-C(**3**) = 1.249 Å, that valence angle O(**4**)-C(**3**)-XX(**1**) = 63.3° and that dihedral angle O(**4**)-C(**3**)-XX(**1**)-XX(**2**) = 180.0°. XX represents a *dummy atom* that is disregarded in the calculation, but XX is sometimes useful for imposing symmetry constraints or fixing the reference frame at a specified point of a system. In the current example it is clear that XX(**1**) and XX(**2**) are fictitious atoms serving to place the origin of the coordinate system at a specific point. This point (fictitious atom XX(**1**)) Å lies at the end of fictitious bond C(**3**)-XX(**1**) = 2.166 Å and enters fictitious valence angle O(**4**)-C(**3**)-XX(**1**) = 63.3° which serves to define the position of O(**4**) relative to the origin. The whole molecular system evolves in an inductive manner through the successive addition of atoms starting with atom 1, whether real or fictitious. It can be realized that the dimer $(\mathbf{1})_2$ is symmetrical by noting that the internal coordinates of atoms 9–14 have not been optimized and that they have adopted the values of their symmetry-related counterparts, atoms 3–8, respectively.
[b] Ref. 16, p. 660.

TABLE 2
SIMULATED EFFECTS OF 'SOLVATION' ON THE TAUTOMERISM OF FORMAMIDE

	2×**1**	Δ° **1**→$(\mathbf{1})_2$	$(\mathbf{1})_2$	Δ‡ $(\mathbf{1})_2$→TrSt	TrSt	Δ‡ $(\mathbf{2})_2$→TrSt	$(\mathbf{2})_2$	Δ° **2**→$(\mathbf{2})_2$	2×**2**
Solv. shell radius[a]	8.006	–	9.022	–	9.113	–	9.183	–	8.246
$E_{solvated}$[b]	−92.98	−5.88	−98.86	51.89	−46.97	19.57	−66.54	−1.56	−64.98
E_{vac}[b]	−89.52	−8.14	−97.65	51.27	−46.38	19.04	−65.42	−2.86	−62.56
Δ[b]	−3.46	2.26	−1.21	0.62	−0.59	0.53	−1.12	1.30	−2.42

[a] In Å.
[b] In kcal/mol.

ture reference data for the monomers are given in Table 1. It is seen that AM1 overestimates slightly (by $\approx$0.03–0.04 Å) the lengths of the C=O and C=N bonds of the amide and hydroxyimine forms, respectively, compared to the experimental [16] and/or ab initio HF/6-31G* [11] data. Also, the value of the AM1-optimized N-C-O angle in the amide is $\approx 3°$ lower than the experimental [16] and ab initio [11] values. These differences indicate that AM1 attributes slightly more electron delocalization than actually takes place in the NCO fragment in both forms **1** and **2**. The observed differences, however, are not critical and they fit well with the experimentally observed spread of the related values for amides and imines [11,16,18].

It is also seen that the geometries of the amide and hydroxyimine forms do not change appreciably upon dimerization; quite reasonably the N-H or O-H, as well as the formal C=O or C=N bonds, become minutely weaker (longer) while the formal C-N or C-O bonds become minutely stronger (shorter) in forms $(\mathbf{1})_2$ and $(\mathbf{2})_2$, compared to the respective monomers. The optimized structures of the dimers are effectively planar and possess the C_{2h}-symmetry. The N-H···O and O-H···N hydrogen bonds (H-bonds) are practically linear. The inter-amidic H-bond in $(\mathbf{1})_2$ is much

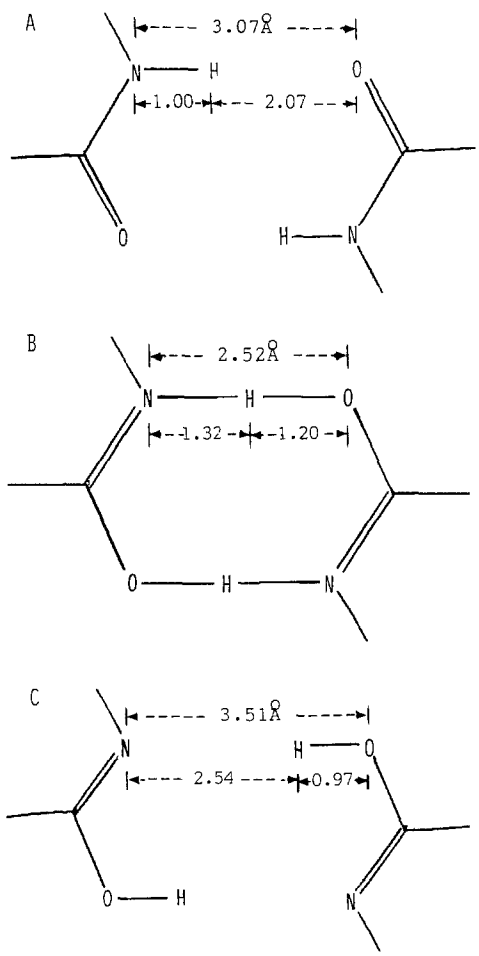
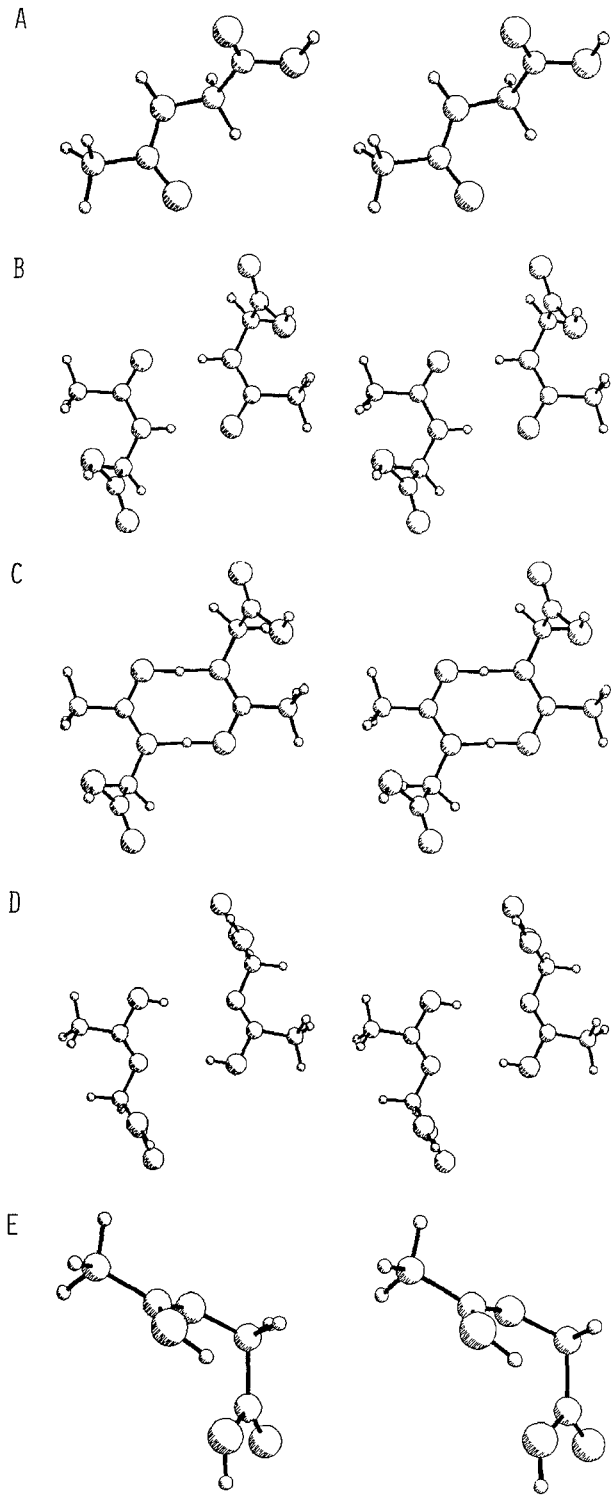


Fig. 1. Optimized geometries of: A, $(\mathbf{1})_2$; B, TrSt to the **1**/**2** tautomerism; C, $(\mathbf{2})_2$.



g. 2. Optimized geometries of: A, **3**; B, $(\mathbf{3})_2$; C, TrSt to the **3**/**4** tautomerism; D, $(\mathbf{4})_2$, E, **4**.

stronger than the inter-hydroxyimidic bond in $(\mathbf{2})_2$; this is inferred from the potential energy differences between the free and dimeric forms given in Table 2, as well as from the pertinent distances in Fig. 1. These results agree with the well-documented views on the H-bond donor/acceptor properties of amides and imines [19].

The hydrogen exchange within the dimer was examined by means of the SADDLE [20] and PATH facilities within MOPAC, the latter with or without SYMMETRY-imposed constraints [12]. The SYMMETRY-unconstrained PATH, executed only on one N···H···O atom set (Fig. 1), forced the other one to synchronously follow *exactly the same* changes, so that the study, after subsequent NLLSQ optimization, led to the definition of a unique TrSt of C_{2h}-symmetry, identical with that obtained when the C_{2h}- symmetry was imposed on the whole PATH search. Such a result indicated unequivocably that the double-hydrogen transfer had a *concerted* character. The most distinct features of dimers $(\mathbf{1})_2$ and $(\mathbf{2})_2$, as well as their common TrSt, are given in Fig. 1.

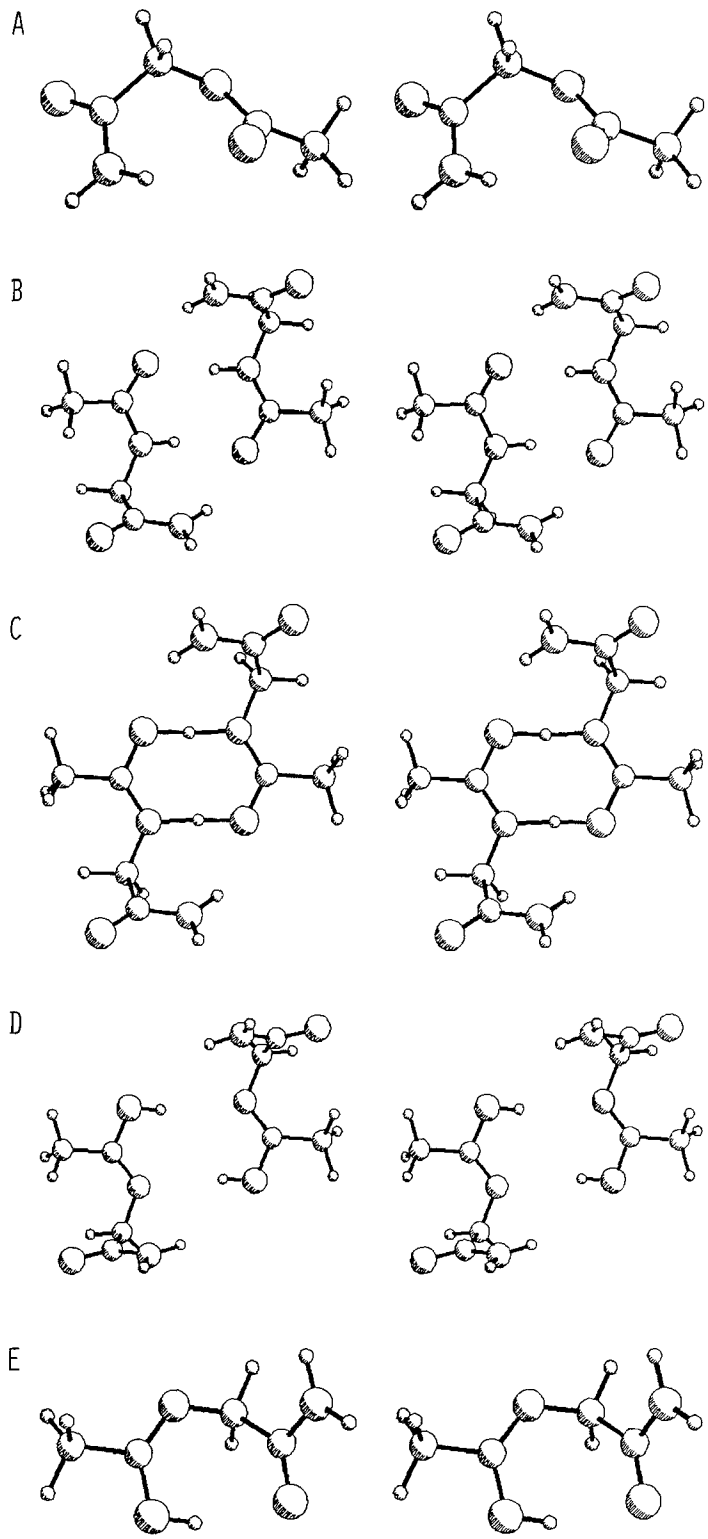
It is seen that the dimer shrinks considerably in the TrSt, as evidenced by the fact that the pair of equivalent *O*···H···*N* distances attain their common minimum (=2.521 Å) which is shorter than the corresponding distances in the amide dimer (3.069 Å) by 0.548 Å and shorter than the corresponding distances in the imidic acid dimer (3.518 Å) by 0.997 Å. The results of analysis by PATH [12] (not shown) indicate that the exchanging hydrogen atoms move essentially along the intermolecular lines connecting the O and N atoms so that either the O-H or N-H distances can make equally proper reaction coordinates. Since both hydrogen atoms move in a precisely concerted way, the studies described below involving dimers of glycine derivatives used SYMMETRY constraints whenever possible, in order to reduce the computational task.

While inclusion of the nonspecific solvation model [1] into these calculations imposed practically no changes on the respective geometries, it provided a moderating effect on the heats of dimerization of both the amide and amidic acid forms (Table 2). The obtained values of heats of solvation compare well with the experimental data [21,22]. The potential energy changes suggest that the conversion of **2** into **1** in both the gas phase and in solution is facile ($\Delta E^{\neq} = 19.04$ kcal/mol) but that the reverse is not ($\Delta E^{\neq} = 51.27$ kcal/mol) (see Table 2).

*Ac-Gly: 1-oxoethylaminoacetic acid (**3**) / 1-hydroxyethylidene iminoacetic acid (**4**) and Ac-Gly-NH_2: 1-oxoethylaminoacetamide (**5**) / 1-hydroxyethylideneiminoacetamide (**6**) tautomerism*

Both types of tautomerism were considered at the same level as that for formamide, except that the simulation of a nonspecific weakly polar solvent was not included. This omission should have no significant effect on the general conclusions that follow (see above and Ref. 1). In the case of tautomerism of the 'peptide' (acetamido group) bond in Ac-Gly and Ac-Gly-NH_2, two additional processes had to be considered, namely, peptide bond internal *cis*/*trans* rotation in **3** and **5**, and *N*-alkylimine *Z*/*E* inversion in **4** and **6**, both types of conformational conversions being prerequisites for either a secondary *cis*-amide or an *E*-1-hydroxy-*N*-alkylimine (*E*-imidic acid) dimerization. Optimized geometries for the imidic acid and peptide monomers and dimers and the TrSts for the **4**→**3** and the **6**→**5** tautomerisms are given in Figs. 2 and 3, respectively. In the dimers the C_2-symmetry is maintained and the double-hydrogen exchange occurs in a concerted way within a planar eight-atom framework that includes the exchanging hydrogen atoms. This is similar to the formamide/formimidic acid case. It is clear from the figures that in going from monomer to dimer, one of the stereomutations alluded to above must occur.

The numerical data characterizing most of the stationary points on the reaction pathways **4**→**3**



g. 3. Optimized geometries of: A, **5**; B. $(\mathbf{5})_2$, C, TrSt to the **5**/**6** tautomerism; D, $(\mathbf{6})_2$; E, **6**.

TABLE 3
POTENTIAL ENERGY DIFFERENCES IMPLICATED BY THE **3**/**4** TAUTOMERISM[a]

State	E	Reaction step[b]	ΔE
2 × **4** (*E*)	−257.96	2 × **4**(*E*)→(**4**)$_2$(*Z*)[c]	3.28
(**4**)$_2$ (*Z*)	−254.68	(**4**)$_2$(*Z*)→TrSt	19.61
		cumulative change	22.89
TrSt	−235.07	TrSt→(**3**)$_2$(*cis*)	−46.48
		cumulative change	−23.59
(**3**)$_2$ (*cis*)	−281.55	(**3**)$_2$ (*cis*)→2 × **3**(*trans*)[c]	4 16
		total balance	−19.43
2 × **3** (*trans*)	−277.39	tot.bal./1 molecule	−9.72

[a] Calculations (kcal/mol) refer to in vacuo.
[b] Free **4** has been taken as a reference.
[c] $Z \rightleftarrows E$ alkylimine inversion in **4** and *cis* $\rightleftarrows$ *trans* peptide rotation in **3** are characterized in Figs 4 and 6, respectively.

and **6**→**5** are given in Tables 3 and 4. The alkylimine inversions in **4** and **6** are given in Figs. 4 and 5, respectively; the same data for peptide bond rotations in **3** and **5** are given in Figs. 6 and 7.

It is clear from Figs. 4 and 5 that AM1 predicts the barriers to internal *N*-alkylimine inversion in **4** and **6** ($\Delta E^{+} \approx 21.5$ kcal/mol) in perfect agreement with the experimental data on inversion barriers obtained for these types of compounds by dynamic NMR [23]. The slight hysteresis of the curves indicates the presence of retardation forces arising from relatively weak intramolecular hydrogen-bonding interactions imposing C_7-type conformations on both **4** and **6** [1]. Apparently, these seven-membered-ring internal chelates can be disrupted with very little strain.

The values obtained indicate that imine inversions in **4** and **6** are feasible at room temperature, so the primary condition allowing them to form cyclic dimers (**4**)$_2$ and (**6**)$_2$ that could subsequently undergo a tautomeric double-proton exchange is met. Even if the conformational changes in Figs. 4 and 5 are described in terms of dihedral angles O-N=C-C$^{\alpha}$, the simultaneous record of accom-

TABLE 4
POTENTIAL ENERGY DIFFERENCES IMPLICATED BY THE **5**/**6** TAUTOMERISM[a]

State	E	Reaction step[b]	ΔE
2 × **6** (*E*)	−158 72	2 × **6**(*E*)→(**6**)$_2$(*Z*)[c]	1.24
(**6**)$_2$ (*Z*)	−157.48	(**6**)$_2$(*Z*)→TrSt	19.37
		cumulative change	22 61
TrSt	−138 11	TrSt→(**5**)$_2$(*cis*)	−49.16
		cumulative change	−28.55
(**5**)$_2$ (*cis*)	−187.27	(**5**)$_2$(*cis*)→2 × **5**(*trans*)[c]	8.25
		total balance	−20.30
2 × **5** (*trans*)	−179.02	tot.bal./1 molecule	−10.15

[a] Calculations (kcal/mol) refer to in vacuo.
[b] Free **6** has been taken as a reference
[c] $Z \rightleftarrows E$ alkylimine inversion in **6** and *cis* $\rightleftarrows$ *trans* peptide rotation in **5** are characterized in Figs 5 and 7, respectively.

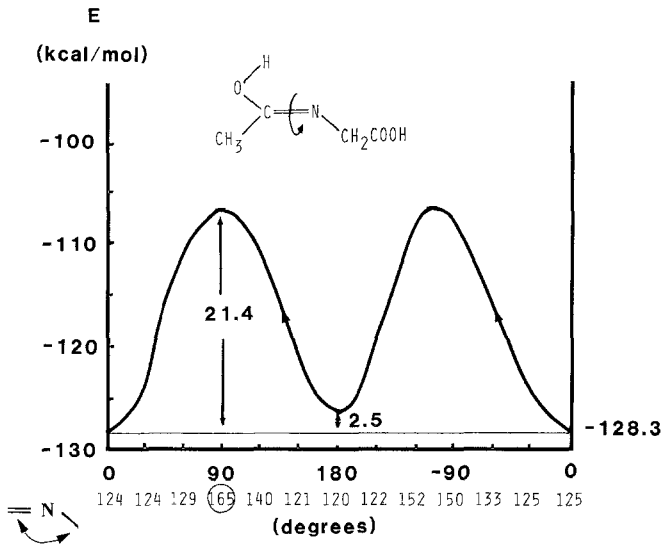


Fig. 4. Internal energy profile for *N*-alkylimine $Z \rightleftharpoons E$ inversion in **4**

panying changes imposed on the O=C-N valence angles leaves no doubt that these conversions have mainly inversive character. In fact, they are elliptical rotations flattened to fit the O-N=C plane, which is evident from the fact that the O-N=C valence angles attain values close to 170° near the TrSt (see Figs. 4 and 5). Finally, it is seen that the *E*-isomers are less stable than their *Z*-counterparts by 1.3–2.5 kcal/mol.

Inspection of Figs. 6 and 7 indicates that AM1 predicts a barrier to *cis* → *trans* rotation about the peptide bond of 18 kcal/mol for Ac-Gly-NH_2 (**5**), and 13 kcal/mol for Ac-Gly (**3**). The *trans*-

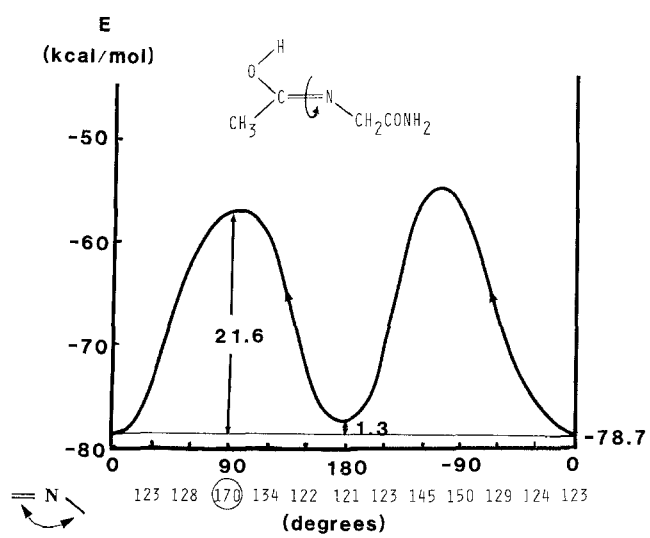


Fig 5. Internal energy profile for *N*-alkylimine $Z \rightleftharpoons E$ inversion in **6**

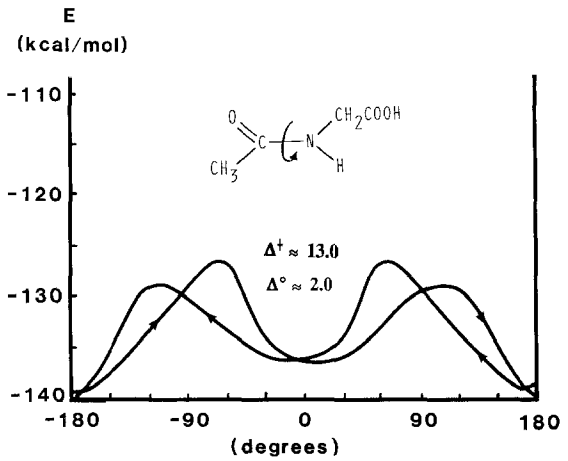


Fig 6. Internal energy profile for *N*-alkylacetamide rotation in **3**. Clearly, the shape of the curve depends on the sense of rotation.

amide bonds are more stable than the *cis*-bonds by 3.0 and 2.0 kcal/mol, respectively. While the results for **5** are perfectly compatible with the well-documented experimental [23] and high-level theoretical [24] evidence, the results for **3** show a *trans* → *cis* rotation barrier at least 4–5 kcal/mol lower than expected. This observation may mean failure of AM1 in this particular case, however, why has it not failed in the akin case of **5**? It more likely indicates the presence of a unique intramolecular screening furnished by the adjacent carboxylic group that gives rise to the enormous catalytic effect observed on rotation of the amide in **3**. That such a specific and strong intramolecular interaction exists is indicated by the strong retardation effects reflected in the severe disymmetry of the potential energy curves in Figs. 6 and 7.

Detailed analyses (not shown) of the changes in geometry accompanying the secondary amide rotations in **3** and **5** indicate that the hystereses of the curves are caused by different phenomena in the two cases. This may provide a clue to explain the strikingly different energy barriers to inter-

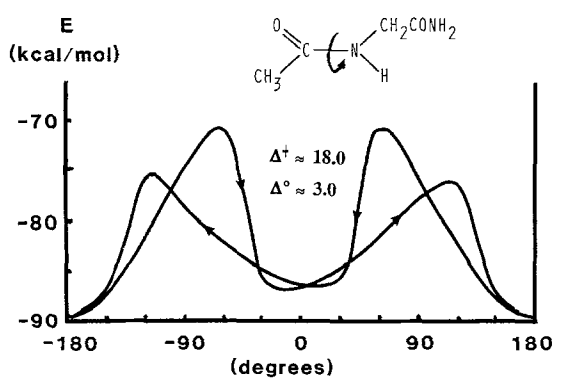


Fig. 7. Internal energy profile for *N*-alkylacetamide rotation in **5**. Clearly, the shape of the curve depends on the sense of rotation.

nal rotation in the compounds which are similar otherwise. Thus, the overall variations in molecular geometry accompanying rotation of the peptide bond in Ac-Gly-NH_2 indicate a sudden disruption (formation) of a C_7-type [25] intramolecular chelate during the *trans/cis* conversions. It should be noted that the disruption and formation take place at the different points of the interconversion coordinate, which is referred to as the retardation effect (see Figs. 6 and 7). On the other hand, similar inspection of the changes in geometry accompanying the peptide bond rotation in Ac-Gly indicates that the C_5-type [25] intramolecular chelate is maintained throughout the conversion cycle, even if there is strain at some stages of the rotation. In summary, AM1 delivers a picture of amide rotation and *N*-alkylimine inversion which is remarkably consistent with the experimental reference data that are available, and simultaneously admits symmetrical dimerizations of both a secondary amide and the corresponding amidic acid.

With regard to the overal energetics, it is seen that, allowing for minor modifications, the well-known pattern encountered in the study of the **1**/**2** tautomerism (Table 3) is repeated. The modifications concern mainly the dimerization step which for **3**/**4** and **5**/**6** tautomerism involves the conformational changes discussed above. Since the conformationally-converted species, immediately before dimerization, have potential energies $\approx$1.5–3.0 kcal/mol higher than those of their most stable parent conformers, the modification ranges from 3 to 6 kcal/mol on the bimolecular basis. In any event, the data indicate clearly that the **4** $\rightarrow$ **3** and **6** $\rightarrow$ **5** conversions are possible while the reverse, **4** $\leftarrow$ **3** and **6** $\leftarrow$ **5** are not. This remains in full compliance with the results for **1**/**2**.

*For-Gly-O-COOMe: 1-oxomethylaminoacetic methylcarbonic anhydride (**7**) / 1-hydroxymethylideneiminoacetic methylcarbonic anhydride (**8**) tautomerism*

We have found that the AM1-optimized monomer **8** assumes a seven-membered intramolecular chelate in its minimum-energy conformation with an intramolecular H-bond between the 'peptide' hydrogen and the electron pair of the methoxy oxygen as the key feature (Fig. 8). It appeared to us that if this kind of intramolecular chelation were sustained in the two dimeric forms of **7** and **8**, i.e., $(\mathbf{7})_2$ and $(\mathbf{8})_2$, then it would be of value to see whether the intermediate between $(\mathbf{7})_2$ and $(\mathbf{8})_2$, that is the TrSt pertinent to the **7**/**8** tautomerism, is also subjected to this kind of chelation, and if so, whether the chelation affects the energy barrier to tautomerization to any extent. A significant decrease (by $\approx$15 kcal/mol) of this energy barrier might suggest tautomerism as a possible process initiating a chain of elementary reactions leading to generation of the oxazolone from a mixed anhydride [25].

Although the standard [12] geometry optimizations of the minimum-energy states required 'only' 300–700 s of CPU time, the non-linear least-squares [12] full gradient minimization (NLLSQ) of the TrSts required as much as 5 h of CPU time. This was also the reason for choosing the formyl rather than acetyl derivative of glycine as model. The results are given in Table 5 and Fig. 8.

As the data show, the C_2-symmetry is maintained all along from $(\mathbf{7})_2$ to $(\mathbf{8})_2$, the double-hydrogen exchange occurs in a concerted manner, and the reacting centre is made of a symmetrical and effectively planar eight-membered ring. In other words, all significant features of the reaction considered above are maintained. The anticipated type of intramolecular chelation is indeed favored in the two dimers ($(\mathbf{7})_2$ and $(\mathbf{8})_2$) involved. The key features are the exchanging H-atoms which enter into peculiar bifurcated H-bonds because of the simultaneous dimerization (Fig. 8). Surprizingly, the bifurcation blows out just in the TrSt so that our hypothesis must be rejected.

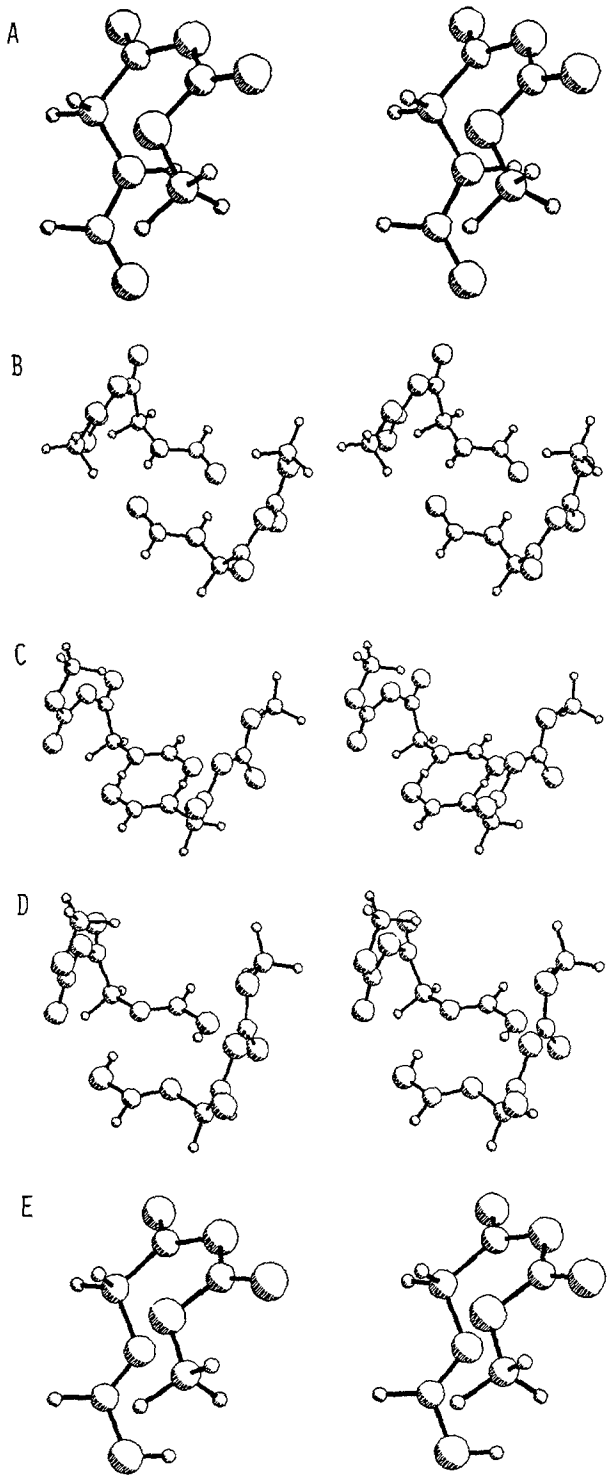


Fig. 8. Optimized geometries of: A, **7**; B, $(\mathbf{7})_2$; C, TrSt to the **7**/**8** tautomerism; D, $(\mathbf{8})_2$; E, **8**

TABLE 5
POTENTIAL ENERGY DIFFERENCES IMPLICATED BY THE **7**/**8** TAUTOMERISM[a]

State	E	Reaction step[b]	ΔE
2×**8**	−375.41	2×**8**→(**8**)$_2$	−6.41
(**8**)$_2$	−381.82	(**8**)$_2$→TrSt	25.08
		cumulative change	18.67
TrSt	−356.74	TrSt→(**7**)$_2$	−51.81
		cumulative change	−33 14
(**7**)$_2$	−408.55	(**7**)$_2$→2×**7**	13.61
		total balance	−19.53
		tot.bal./1 molecule	−9.77
2×**7**	−394.94		

[a] Calculations (kcal/mol) refer to in vacuo
[b] Free **8** has been taken as a reference

Hence, it is seen that in analogy with the previous cases, the conversion **8** → **7** is possible but the reverse is not. In addition, the heats of dimerization of both **7** and **8** are much more favorable than those of their counterparts presented above. This is presumably an effect reflecting replacement of the *N*-acetyl group by the *N*-formyl group. The total potential energy balance, $\Delta E^\circ = 9.77$ kcal/mol, is nearly the same as that for the two former derivatives of glycine.

General

The message to be gleaned from Tables 2–5 and Figs. 1–3 and 8 is that the double-hydrogen ex-

TABLE 6
HEATS OF DIMERIZATION, ENTHALPIES AND LENGTHS OF NH·· O H-BONDS IN SELECTED AMIDE DIMERS[a]

Amide	ΔE	ΔH	R_{ON}	Method	Ref.
Ac-Gly (**3**)	−8.36[b]	–	3.07	AM1	this work
Ac-Gly-NH_2 (**5**)	−14.35[b]	–	3.00	AM1	this work
For-Gly-O-$COOCH_3$ (**7**)	−13.61	–	3.00	AM1	this work
Formamide (**1**)	−8.14	–	3.07	AM1	this work
	−11.5	–	–	GTO: (7,3;3)→[2,1.1][c]	27
	−14.0	–	–	GTO: (7,3;3)→[2,1,1][d]	28
	–	–	2.93	X-ray	29
C_3–C_6 lactams	–	−7.1–−7.8	–	IR	30
Dioxopiperazine	–	7.0	–	Heat of sublimation	31
Propionamide[e]	–	7.85	–	IR	32

[a] ΔE and ΔH in kcal/mol, R_{ON} in Å.
[b] Net dimerization corrected for conformational conversion, see text.
[c] NH-bond length optimized.
[d] Stiff, i.e. assumed geometry.
[e] $\Delta F(298\ K) = -2.25$ kcal/mol.

change between an amide dimer and its tautomer, a related imidic acid dimer, occurs in a concerted way within an effectively planar eight-atom cyclic complex. This is essentially the same well-known phenomenon which is associated with carboxylic acid dimers [15b]. Both H-atoms move symmetrically along the straight lines connecting the O and N acceptors, each belonging to a different monomer, so that the H-exchange is *inter*-molecular within the dimer. The *N*···H···*O* distance attains its minimum at the TrSt where it is substantially smaller than in either of the tautomeric dimers. Whereas stabilization of transient imidic acids [1] into their parent amides is kinetically allowed, the reverse seems to be both kinetically and thermodynamically forbidden.

In this section, we consider quantitative features of selected results for which reliable experimental or theoretical data are available. Comparison of these data provides an assessment of the performance of the AM1 method in the current application and an indication of performance to be expected in similar types of applications. Though no theoretical or experimental data concerning the topology and energetics/thermodynamics of the prototropy described in this investigation are available, three aspects dealt with have been addressed elsewhere: (i) cyclic dimerization of formamide and formimidic acid and strictly related to it the energy (enthalpy) of H-bonding implicated in the former; and (ii) the relative stabilities and associated differences in energies of the tautomeric forms.

The enthalpies of dimerization of the amides studied here are compared with high-level theoretical [26,27] and experimental [28–31] data in Table 6. In contrast to the case of carboxylic acids, cyclic dimers of amides are far less common since primary amides can also take part in the poly-molecular H-bond-mediated association of chains [19,32]. On the other hand, secondary amides (i.e., peptides) with the notable exception of small-ring lactams, are confined to poly-molecular chain association because they can exist only in the *trans*-conformation [24] (Figs. 6 and 7) which is incompatible with cyclic dimerization. It is clear from Table 6 that our AM1 results are perfectly compatible with the limited experimental data available. The slightly shorter experimental N···O distances in formamide may be a reflection of the effect of crystal lattice forces. Unfortu-

TABLE 7
RELATIVE STABILITIES OF THE TAUTOMERS FOR SELECTED AMIDES[a]

Amide	ΔE	ΔH	ΔG[b]	Conditions	Ref.
Formamide (**1**)	13.48	–	–	AM1/gas phase	this work
	12.2(0.4)[c]	–	–	ab initio[c]	11
Ac-Gly (**3**)	9.72	–	–	AM1/gas phase	this work
Ac-Gly-NH_2 (**5**)	10.15	–	–	AM1/gas phase	this work
For-Gly-O-$COOCH_3$ (**7**)	9.77	–	–	AM1/gas phase	this work
δ-Valerolactam	–	–	7.3	gas/ICR spectroscopy/140 K	34
2-Pyridone	−1.0(2.5)	−0.3(0.3)	−0.8[d]	gas phase/UV spectroscopy	7
	−1.00	–	–	HF/6–31G**	35
	0.4(0.6)[c]	–	–	ab initio[c]	11

[a] In kcal/mol.

[b] At 300 K unless noted otherwise.

[c] A number of geometry-optimized ab initio calculations on levels varying from HT/STO 3–21G to $CISD_4$/−31G* the 'best theoretical estimate' [11] is quoted in the Table.

[d] 405 K. $\Delta S^\circ = 1.0(1.5)$ e.u

nately, there are no 'state-of-the-art' high-level ab initio results for the formamide dimer, so the numbers quoted as theoretical evidence in Table 6 may be of limited value. In particular the value of −14 kcal/mol for E [27] should be treated with caution since it was obtained for an assumed stiff geometry. It is also clear from Figs. 6 and 7 in Ref. 19 that poly-association should compete heavily with dimerization in the amides considered here.

High-level ab initio data [11,34] obtained for the tautomers of formamide and 2-pyridone may provide a link between our AM1 results for formamide and the experimental data for 2-pyridone (Table 7). The fact that the ab initio results for 2-pyridone display excellent agreement with the experimental data and that the ab initio data for formamide match ours provides an indirect proof that AM1 reliably estimates potential energy differences between the amide tautomers considered in this paper. Moreover, our secondary-amide data compare favorably with the unique experimental results for δ-valerolactam [33].

CONCLUSION

When the AM1 method is applied with appropriate caution, it provides valuable guidance for studying mechanisms of reactions involved in peptide synthesis. Its greatest asset as for any good theoretical method lies in its power to address topics not readily amenable to experimental verification. In the present context it has meant distinguishing among a few possible reaction pathways and in particular providing convincing geometrical characteristics of the TrSts.

ACKNOWLEDGEMENTS

This research was financially supported by the Medical Research Council of Canada of which N.L.B. is a Career Investigator. We are grateful to Bachem Inc., Torrance, CA for financial assistance and to the Polish Scientific Research Council (KBN) for a travel grant (BW/5-300-4-065-1) to J.C.

REFERENCES

1 Ciarkowski, J., Chen, F.M.F. and Benoiton, N L., J. Comput.-Aided Mol. Design, 5(1991) 585.
2 Dewar, M J.S , Zoebisch, E G , Healy, E F. and Stewart, J.J.P., J Am. Chem. Soc., 107(1985) 3902.
3 Benoiton, N.L. and Chen, F.M.F., Can J. Chem., 59(1981) 384.
4 Chen, F.M.F. and Benoiton, N.L., Can J. Chem , 65(1987) 619
5 Chen, F M.F. and Benoiton. N L., J. Chem Soc. Chem. Commun., (1981) 336
6 a. Katritzky, A R. and Lagowski, J.M., Adv. Heterocyclic Chem., (1963) 1 and 2.
b. Albert, A., Heterocyclic Chemistry. Athlone Press, London, 1968.
c. Elguero, J , Marzin. C., Katritzky, A.R. and Linda, P , Adv. Heterocyclic Chem., (1975) Suppl. No. 1.
7 Beak, P., Acc Chem. Res., 10(1977) 186
8 Kwiatkowski, J.S , Zielinski, T.J. and Rein, R., Adv. Quantum Chem . 18(1986) 85.
9 Szczesniak, M., Szczepaniak, K., Kwiatkowski, J S., KuBulat, K. and Person, W B., J. Am Chem. Soc., 110(1988) 8319, and refs. therein.
10 Baeyer, A. and Oekonomides, S., Ber. Dtsch. Chem. Ges., 15(1882) 2093.
11 Schlegel, H.B., Gund, P. and Fluder, E.M., J. Am Chem. Soc., 104(1982) 5347, and refs therein
12 Stewart, J.J.P., MOPAC. A General Molecular Orbital Package, modified by Eger, M , to suit IBM/370-compatible scalar computers utilizing VS/FORTRAN version 4 1.4 or higher compilers, 1988. QCPE 459.
13 Bingham,. R.C , Dewar, M.J S and Lo, D.H., J. Am. Chem. Soc., 97(1975) 1285

14 Dewar, M.J S and Thiel, W., J Am. Chem. Soc., 99(1977) 899.
15 a. Roberts, J.D and Caserio, M.C., Basic Principles of Organic Chemistry, W.A. Benjamin, New York, NY, 1965, pp. 675–676.
b. ibid, pp. 509–510.
16 Harmony, M.D , Laurie, V.W., Kuczkowski, R.L., Schwendeman, R H., Ramsay, D.A., Lovas, F.J., Lafferty, W.J. and Maki, A.G., J Phys Chem. Ref. Data, 8(1979) 619.
17 Lumbroso, H. and Pappalardo, G.C., J. Mol. Struct., 43(1978) 97.
18 Gross, E. and Meienhofer, J., In Gross, E. and Meienhofer, J., (Eds.) The Peptides, Vol. 1 (Analysis, Synthesis, Biology). Academic Press, New York, NY, 1979, pp. 1–64.
19 Pimentel, G.C. and McClellan, A.L., The Hydrogen Bond, W.H. Freeman & Co., San Francisco, CA, 1960.
20 Dewar, M J.S., Healy, E.F. and Stewart, J.J.P., J. Chem. Soc. Farad. Trans II, 80(1984) 227.
21 Jones III, F.M. and Arnett, E.M., Prog. Phys Org. Chem., 11(1974) 263.
22 Srivastava, R. and Smith, B.D , J. Phys. Chem. Ref Data, 16(1987) 219.
23 Jackman, L M., In Jackman, L.M. and Cotton, F.A. (Eds.) Dynamic Nuclear Magnetic Resonance, Academic Press, New York, NY, 1975. Chap. 7.
24 Zimmerman, S.S., In Udenfriend, S., Meienhofer, J. and Hruby, V. (Eds) The Peptides, Vol. 7 (Analysis, Synthesis, Biology), Academic Press, New York, NY, 1985, pp. 165–212.
25 a. Chen, F.M.F and Benoiton, N.L., Can. J. Chem., 65(1987) 619
b. Chen, F.M.F., Slebioda, M. and Benoiton, N.L., Int. J. Pept. Protein Res., 31(1988) 339.
26 Janoschek, R., Theoret. Chim. Acta, 32(1974) 49.
27 Dreyfus, M., Maigret, B. and Pullman, A., Theoret Chim. Acta, 17(1970) 109.
28 a. Ladell, J. and Post, B., Acta Crystallogr., 7(195) 559.
b Itoh, K. and Shimanouchi, T.J., J. Mol Spectrosc., 42(1972) 86.
29 Luck, W.A.P., Naturwissenschaften, 52(1965) 25, 49.
30 Seki, S , Suzuki, K. and Koide, T., J. Chem. Soc. Japan, Pure Chem. Sec., 77(1956) 346.
31 Badger, R M. and Rubalcava, M., Proc. Natl. Acad. Sci. USA, 40(1954) 12.
32 Schuster, P., Zundel, G. and Sandorfy, C., The Hydrogen Bond, Vols. 1–3, North-Holland, Amsterdam, 1976.
33 Cook, M.J., Katritzky, A R., Taagepera, M., Singh, T.D. and Taft, R.W., J Am. Chem Soc., 98(1976) 6048.
34 Scanlan, M.J., Hillier, I.H. and MacDowell, A.A., J. Am. Chem. Soc., 105(1983) 3568.