EXPERIMENTAL STUDY OF THE INFLUENCE OF SPECIFIC INTRAMOLECULAR INTERACTIONS ON THE CONFORMATION OF MODEL MOLECULES. (PEPTIDES AND OLIGOPEPTIDES)

J. NÉEL

Laboratoire de Chimie-Physique Macromoléculaire
E.N.S.I.C., 1, rue Grandville, 54 Nancy, France

ABSTRACT

Several experimental methods have been used for studying the conformations taken by dipeptidic molecules when dissolved, at a very low concentration, in an inactive solvent such as carbon tetrachloride.

Through systematic investigations performed by infrared and nmr spectroscopy, it has been possible to obtain accurate data about these conformations and to compare with the results of various theoretical treatments.

Several typical examples are reported which concern aminoacid derivatives containing a paraffinic, an aromatic or a polar side substituent. It appears that the conformations taken by such molecules are mainly determined by specific intramolecular forces.

These results agree only with the conclusions of the theoretical calculations in which such interactions have been considered.

It is known that the properties of peptides and polypeptides are partly determined by their conformations. For this reason many attempts have been made during the last few years[1], to compute the energy of these molecules as a function of their deformation parameters. Usually, these approaches have been by partitioning the potential energy of the system into several discrete contributions such as electrostatic and non-bonded interactions, barriers to internal rotation around single bonds, hydrogen bonding and so on. The formulae for these various contributions are generally deduced from physico-chemical data on model compounds of low molecular weight.

Such a scheme, however, is not entirely satisfactory because the fundamental formulae and parameters used to define some of the components of the total molecular potential energy are not well established and differ, often appreciably, from one author to another.

For instance, in the case of the simplest polypeptides, the great stability of the α-helix can as well be explained either by taking into account mainly the internal hydrogen bonds, or only the non-bonded interactions. So, if the absolute values of the various components of the potential energy are not exactly known, it is not possible to get a clear idea about the actual

nature of the driving forces which mainly determine this peculiar conformation.

Owing to the numerous degrees of freedom which allow the deformation of polypeptidic molecules, it appears that such compounds are not a good choice for checking the validity of theoretical results. It is better to work with simpler molecules which contain only two peptidic linkages and which can be prepared by the chemical transformation of natural α-aminoacids. It is quite obvious that any theoretical treatment which disagrees with experimental data obtained with these very simple models will probably be inadequate to explain the conformational behaviour of more sophisticated systems like high molecular weight polypeptides or proteins.

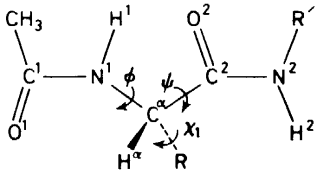


Figure 1. Dipeptidic species I (R, R').

For these reasons we have paid some attention to compounds similar to those which are drawn in *Figure 1* where the two angles ϕ and ψ measure the rotations around N^1–C^α and C^α–C^2 and are conventionally taken equal to zero in the fully-stretched planar conformation.

A fine illustration of the discrepancy which can arise between the results of different theoretical calculations is shown in *Figure 2* which shows the conformational maps of acetylglycine *N*-methylamide (a, b) and acetylalanine *N*-methyl amide (c, d). For these two compounds, the potential energy has been computed by Brant and coworkers[2], first without taking into consideration the dipolar intramolecular forces and secondly after the introduction of an additional electric potential component. We can easily see that, in each case, the two maps so computed are quite different and that the representative points which correspond to the energy minima are appreciably shifted.

In such a situation, it seems to me that experimental data are quite useful to appreciate the relative weight of the various components contained in the energy function. So, we have tried to find out accurate experimental methods to determine the conformations taken by such molecules when surrounded by an inactive solvent such as carbon tetrachloride and I will report some results which have been obtained in my own laboratory and in Professor Lascombe's in Bordeaux[3].

Although they are not the only experimental tools available for such an investigation, our work has been performed mainly by infrared and nmr spectroscopy.

By studying a great number of dipeptides similar to the compounds shown in *Figure 1*, we have been able to prove that such molecules can be in different conformations when dissolved, at a very low concentration, in carbon tetrachloride.

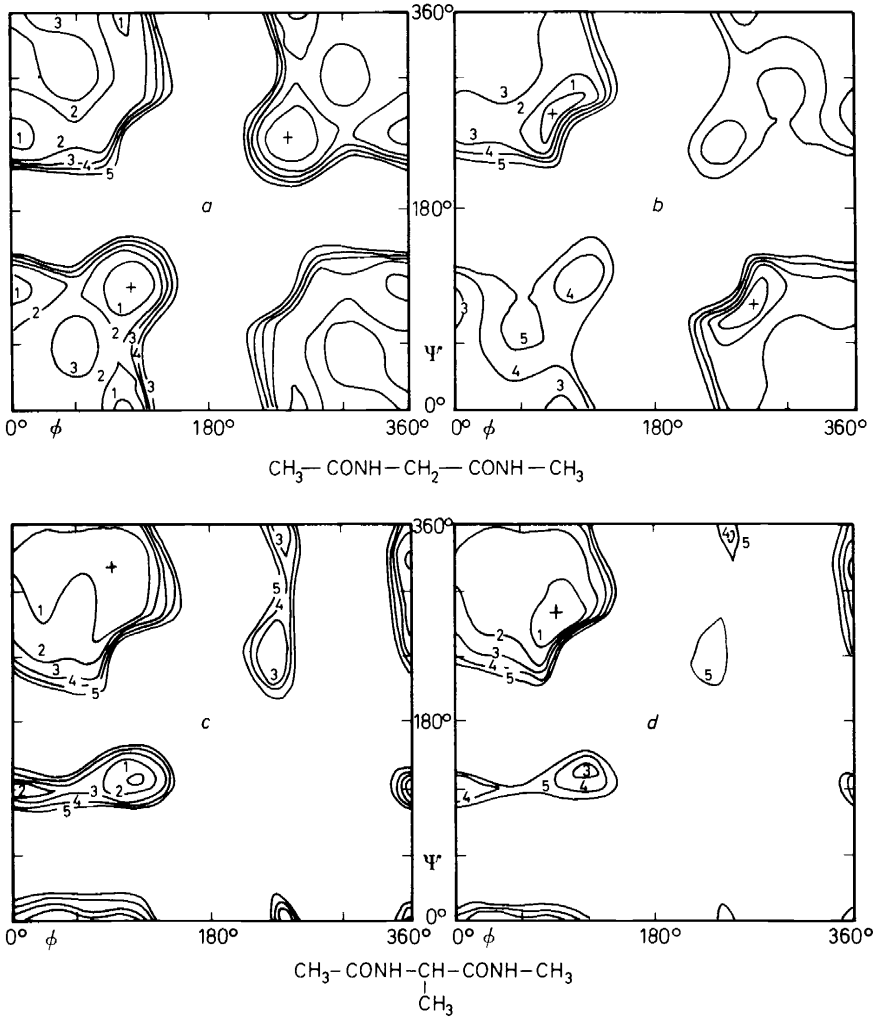


Figure 2. Conformational maps of glycine and alanine derivatives computed (a,c) without taking into consideration the intramolecular dipolar interactions; and (b,d) after the introduction of an additional electric component. After Brant, Miller and Flory[2].

The first drawn in *Figure 3* is quite similar to the fully-stretched form. The relative disposition of the two dipoles N^1–H^1 and C^2–O^2 is such that there is obviously some interaction between them. These two sites form with the C^α atom a pentagonal ring and, for this reason, we call this conformation the C_5 structure. In fact, it can differ from the fully-stretched one by a slight warping of the molecule due to an unbalanced steric repulsion between the side substituent and the atoms H^2 and O^1. So, the ϕ and ψ rotational angles can be slightly different from zero.

The second one, labelled the C_7 conformation, is a ring structure which is folded by an intramolecular hydrogen bond between H^2 and O^1 as shown

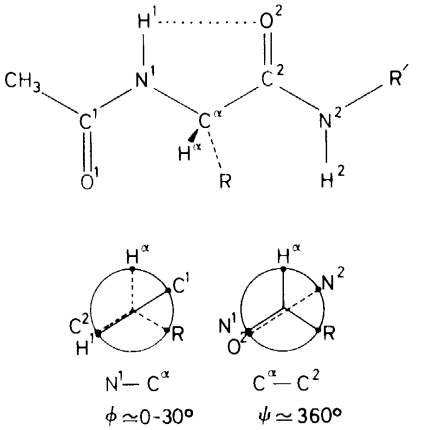


Figure 3. C_5 conformation of molecules I (R, R').

in *Figure 4*. The amide groups lie in two planes which make an angle of 115°. When R is not a hydrogen atom, two different C_7 conformations can exist according to the inclination of the C^{α}—R bond with respect to the intersection line of these two planes. These two forms, which are called equatorial and axial, are represented on the usual conformational map by two centrosymmetric points the coordinates of which are respectively $\Phi = 105°$, $\Psi = 230°$ (for the equatorial structure) and $\Phi = 255°$, $\Psi = 130°$ (for the axial one).

Finally it is possible that open forms can also be present. Up to this date, we are not able to say whether they are random conformations or a sequence of several discrete favourable ones. Nevertheless our own observations prove that the relative number of these open forms is quite low under the conditions in which we have performed our experiments. In fact, in carbon

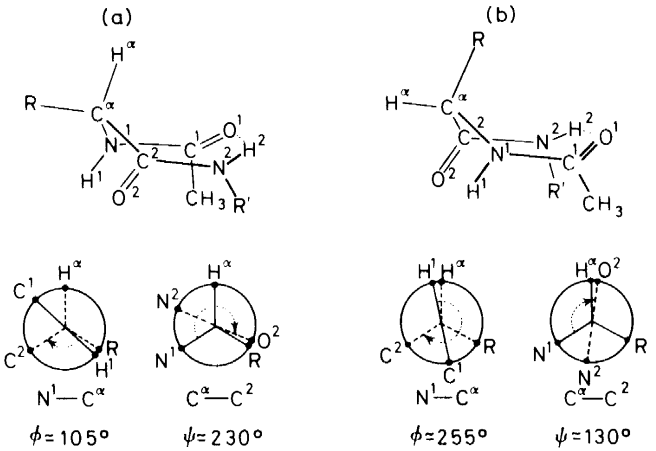


Figure 4. C_7 conformations of molecules I (R, R').
(a) equatorial C_7 conformation; (b) axial C_7 conformation.

tetrachloride, these peculiar conformations appear only when the temperature exceeds 70°C.

All these conclusions are deduced mainly from infrared spectroscopic measurements[3].

We have operated with a great number of slightly different model compounds which can be classified into four groups according to the formulae which are collected in *Figure 5*.

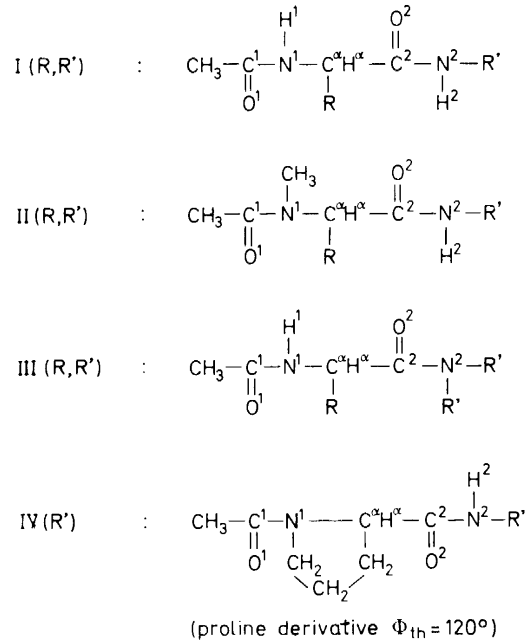


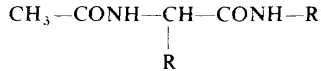
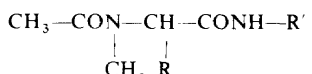
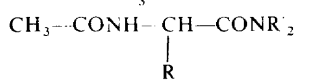
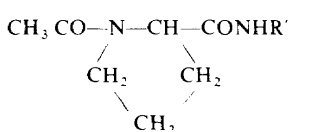
Figure 5. General formulae of the model compounds which have been studied.

It is known that when the vibration of an N–H oscillator is disturbed by some interaction or by its inclusion in a hydrogen bond, its stretching frequency is lowered. Since an intramolecular hydrogen-bonding is hardly ever complete the corresponding absorption band is generally split into two components. If one operates in a very dilute solution in an inactive solvent, the solute molecules are neither self-associated nor solvated and the measurement of the band splitting can be used to investigate the conformations taken by the solute molecules in this medium.

Under such conditions we have studied many compounds, some of them being listed in *Table 1*. The experiments have mainly been performed in the frequency range included between 3 250 and 3 500 cm^{-1} which contains the N–H stretching absorption bands.

In this range, each spectrum generally presents several maxima that have been classified into four groups labelled C_7, C_5, L' and L, corresponding to an order of increasing frequency. Some of these spectra have been drawn in *Figure 6* where, for instance, we can successively see, starting from the

Table 1. Compounds examined for N–H ir absorption spectra.

Type	Representation	R	R′	Amino-acid
	I(H, Me)	—H	—CH_3	glycine
	I(Me, Me)	—CH_3	—CH_3	alanine
	I(*i*Pr, Me)	—$CH(CH_3)_2$	—CH_3	valine
	I(*i*Bu, Me)	—CH_2—$CH(CH_3)_2$	—CH_3	leucine
	I(Bu, Me)	—CH_2—CH_2—CH_2—CH_3	—CH_3	norleucine
	I(Et—S—Me, Me)	—CH_2—CH_2—S—CH_3	—CH_3	methionine
	I(Bz, Me)	—CH_2—C_6H_5	—CH_3	phenylalanine
	I(Bz—O—Bz, Me)	—CH_2—C_6H_4—O—CH_2—C_6H_5	—CH_3	*O*-benzyltyrosine
	I(H, Et)	—H	—C_2H_5	glycine
	I(Me, Et)	—CH_3	—C_2H_5	alanine
	I(*i*Pr, Et)	—$CH(CH_3)_2$	—C_2H_5	valine
	I(Bu, Et)	—CH_2—CH_2—CH_2—CH_3	—C_2H_5	norleucine
	I(Et—S—Me, Et)	—CH_2—CH_2—S—CH_3	—C_2H_5	methionine
	I(Bz, Et)	—CH_2—C_6H_5	—C_2H_5	phenylalanine
	I(Bz—O—Bz, Et)	—CH_2—C_6H_4—O—CH_2—C_6H_5	—C_2H_5	*O*-benzyltyrosine
	II(H, Me)	—H	—CH_3	sarcosine
	II(H, Et)	—H	—C_2H_5	sarcosine
	II(Me, Me)	—CH_3	—CH_3	*N*-methylalanine
	III(H, Me)	—H	—CH_3	glycine
	III(Me, Me)	—CH_3	—CH_3	alanine
	III(Et—S—Me, Me)	—CH_2—CH_2—S—CH_3	—CH_3	methionine
	III(*i*Bu, Me)	—CH_2—$CH(CH_3)_2$	—CH_3	leucine
	III(*i*Pr, Me)	—$CH(CH_3)_2$	—CH_3	valine
	III(Bz, Me)	—CH_2—C_6H_5	—CH_3	phenylalanine
	III(Bz—O—Bz, Me)	—CH_2—C_6H_4—O—CH_2—C_6H_5	—CH_3	*O*-benzyltyrosine
	III(H, Et)	—H	—C_2H_5	glycine
	III(Me, Et)	—CH_3	—C_2H_5	alanine
	III(Et—S—Me, Et)	—CH_2—CH_2—S—CH_3	—C_2H_5	methionine
	III(*i*Pr, Et)	—$CH(CH_3)_2$	—C_2H_5	valine
	III(Bz—O—Bz, Et)	—CH_2—C_6H_4—O—CH_2—C_6H_5	—C_2H_5	*O*-benzyltyrosine
	IV(Me)		—CH_3	proline
	IV(Et)		—C_2H_5	proline

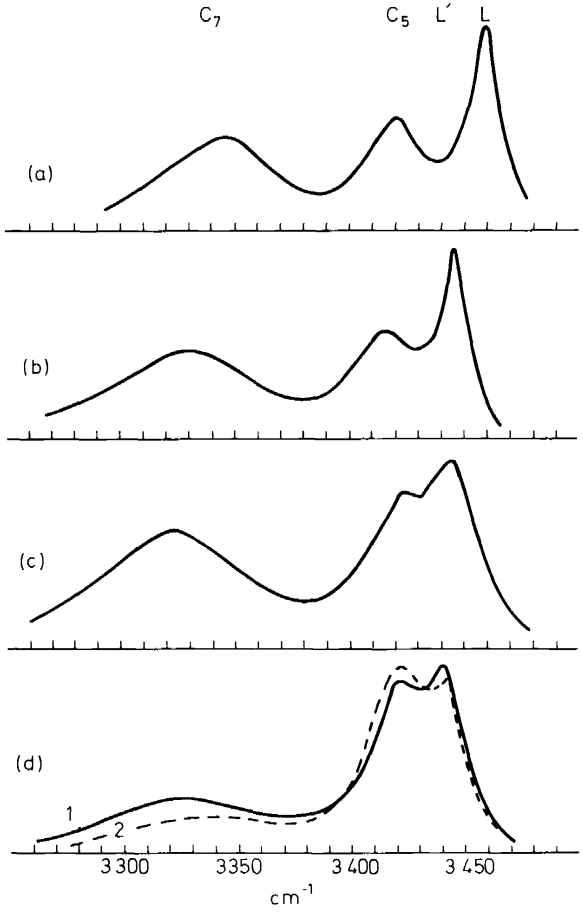


Figure 6. Infrared spectra of *N*-acetylated derivatives of methylglycinamide (a), ethylalaninamide (b), ethylvalinamide (c), ethylphenylalaninamide (d_1), and ethyl *O*-benzyltyrosinamide (d_2).

upper part of the figure, the spectra of the *N*-acetylated derivatives of methylglycinamide (a), ethylalaninamide (b), ethylvalinamide (c), ethylphenylalaninamide (d_1) and ethyl *O*-benzyltyrosinamide (in dotted line) (d_2).

By comparing with the absorption of simple *N*-methyl and *N*-ethyl amides, it is possible to prove that the two high frequency maxima L and L′ are due to free N–H oscillators when differently surrounded. The highest one (L) is attributed to free N–H vibrators contiguous to a single tetrahedral carbon atom. It is the case for N^1–H^1 when R is a hydrogen atom and for N^2–H^2 when R′ is a methyl substituent, and these two conditions are satisfied only in the model compound which corresponds to the upper spectrum. The L′ maximum appears when the nitrogen atom is next to a sequence containing at least two consecutive tetrahedral carbon atoms, that is to say N^1–H^1 when R is not a hydrogen atom and N^2–H^2 when R′ is an ethyl substituent.

Table 2. ν_{N-H} frequencies of dipeptides listed in *Table 1* (dilute solution in CCl_4, $T = 25°C$)

Compound	ν_{NH} Frequencies cm^{-1}					
	C_7	C_5	L′	L	$(\Delta\nu)_5$ cm^{-1}	$(\Delta\nu)_7$ cm^{-1}
I((H, Me)	3346	3421	—	3461	40	115
I(Me, Me)	3340	3414	3443	3460	29	120
I(*i*Pr, Me)	3330	(3424)sh.	(3437)sh.	3461	13	131
I(*i*Bu, Me)	3325	(3425)sh.	3442	3460	17	135
I(Bu, Me)	3330	(3422)sh.	3442	3461	20	131
I(Et—S—Me, Me)	3342	3420	3440	3457	20	115
I(Bz, Me)	3340	3420	3441	3452	21	112
I(Bz—O—Bz, Me)	(3355)*	3419	3442	3452	23	(97)*
I(H, Et)	3332	3417	3446	3460	43	114
I(Me, Et)	3332	3416	3445	—	29	113
I(*i*Pr, Et)	3322	3423	3443	—	20	121
I(Bu, Et)	3322	(3425)sh.	3445	—	20	123
I(Et—S—Me, Et)	3332	3422	3442	—	20	110
I(Bz, Et)	3325	3421	3441	—	20	116
I(Bz—O—Bz, Et)	(3343)*	3422	3442	—	20	(99)*
II(H, Me)	3367	—	—	3455	—	88
II(H, Et)	3355	—	3442	—	—	87
II(Me, Me)	3375	—	—	3452	—	77

III(H, Me)	—	3412	—			
III(Me, Me)	—	3416	(3433)sh.	—	17	—
III(Et—S—Me, Me)	—	3410	(3433)sh.	—	23	—
III(*i*Bu, Me)	—	3416	(3430)sh.	—	14	—
III(*i*Pr, Me)	—	(3415)sh.	3431	—	16	—
III(Bz, Me)	—		3424	—		—
III(Bz—O—Bz, Me)	—		3426	—		—
III(H, Et)	—	3412				—
III(Me, Et)	—	3415	(3433)sh	—	18	—
III(Et—S—Me, Et)	—	3413	(3430)sh	—	17	—
III(*i*Pr, Et)	—	(3415)sh.	3431	—	16	—
III(Bz—O—Bz, Et)	—		3429	—		—
IV(Me)	3325	—	—	(3460)*	—	135
IV(Et)	3313	—	(3446)*	—	—	133

* Very weak absorption band.
$(\Delta\nu)_5$ = Lowering of the stretching frequency of the $N^1—H^1$ oscillator due to the formation of the C_5 structure.
$(\Delta\nu)_7$ = Lowering of the stretching frequency of the $N^2–H^2$ oscillator consecutive to its inclusion in the hydrogen bond which locks the C_7 conformation.

Taking into consideration the experimental infrared data which are collected in *Table 2*, it appears quite obvious that the C_7 and C_5 absorption bands are to be respectively attributed to the C_7 and C_5 conformations previously described. In accordance with this assumption, the C_7 band does not appear in the spectra of compounds III where the second nitrogen atom is fully substituted.

On the other hand, the C_5 maximum is lacking in the case of models II and IV because of the total substitution of the first amide group.

So, infrared spectroscopy proves that dipeptidic species I can exist, when dissolved at a very low concentration in carbon tetrachloride, in two different conformations labelled C_5 and C_7 which can be identified by their absorption maxima according to the attributions shown in *Figure 7*.

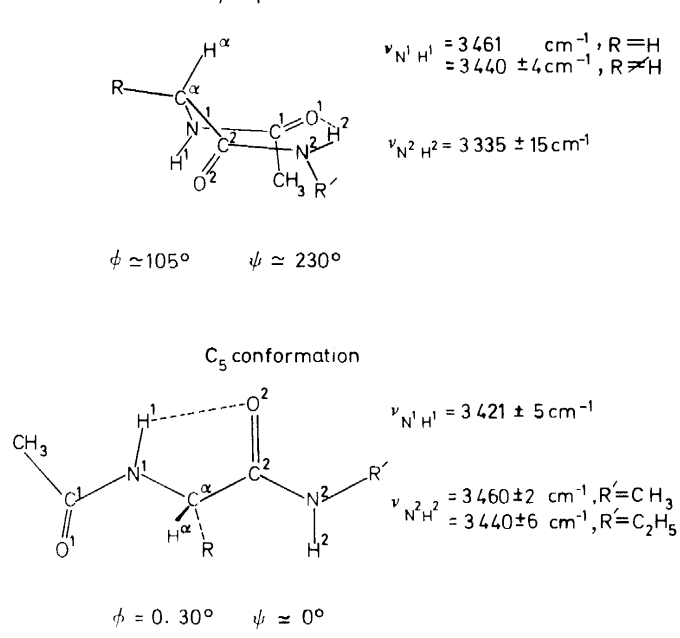


Figure 7. Attribution of the various maxima observed on the infrared spectra of compounds I(R,R').

NMR spectroscopy is also a very powerful technique to get accurate data about the conformation of these molecules.

In fact, the ϕ angle is directly related to the spin-spin coupling constant J between the vicinal protons H^1 and H^α. So, the ϕ values can be obtained by measuring the magnitude of this coupling. First, we have to set up experimentally the correlation between J and ϕ.

This subject has been studied by Bystrov and coworkers[4] in Moscow and by ourselves[5]. The proceeding consists in parametrizing the usual Karplus function with experimental J values measured with model amides in which the ϕ angle is well known.

We have performed this calibration with *N*-methylacetamide and several rigid heterocyclic lactams like pyrrolidone, isoquinuclidone and some piperidinone and dihydrouracil derivatives. Our results agree with an empirical function which is represented, taking the experimental errors into account, by the area comprised between the two continuous lines drawn in *Figure 8*.

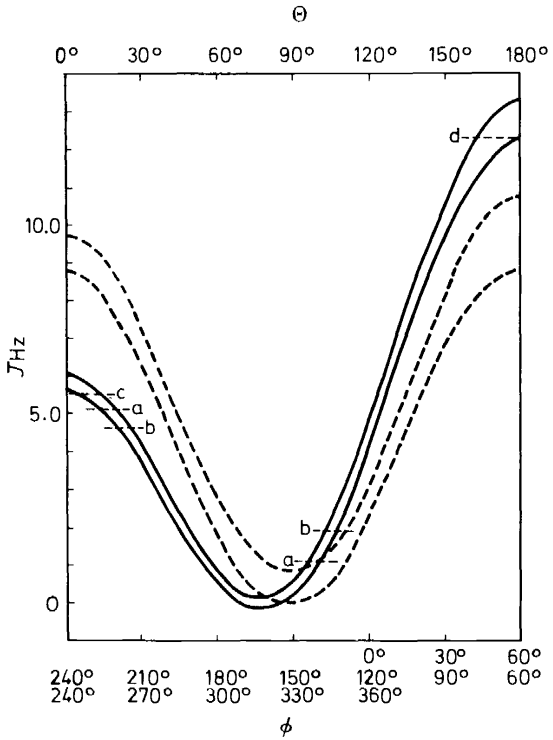


Figure 8. Correspondence between the rotational angle ϕ and the coupling constant J.
After Néel *et al.*[5]: area comprised between the continuous lines.
After Bystrov *et al.*[4]: area comprised between the dotted lines.

Finally, the third available experimental technique we have used, is the measurement of dipole moments. It is obvious that the relative orientation of the two amide linkages determines the magnitude of the dipole moment. The experimental difficulty lies in the necessity to operate in an inactive solvent at a very low concentration to avoid any solvation and any self-association of the solute. Professor Bergman in Jerusalem has succeeded in performing such measurements with some of our compounds. These new data can be interpreted by plotting them on the map drawn in *Figure 9* which has been computed in such a way that the level lines join points which are representative of conformations having the same dipole moment.

The results obtained by the systematic use of all these experimental methods enable us to advance some considerations about the conformations of the molecules we have investigated.

A favourable situation is found in the case of models III where the C_7 forms are forbidden because of the full substitution of the terminal nitrogen atom. So, we merely have to take into consideration the possibility of an equilibrium between the C_5 conformation and the open forms. For instance, examination of the infrared spectra collected in *Figure 10* shows a quite progressive increase in the relative number of open forms when the side substituent is made more and more bulky by operating successively with dimethylamides of acetylglycine (R = H, curve 1), acetylmethionine (R = CH_2CH_2—S—CH_3, curve 2), acetylleucine (R = *i*Bu, curve 3), acetylvaline (R = *i*Pr, curve 4), and acetyl-*O*-benzyltyrosine (curve 5).

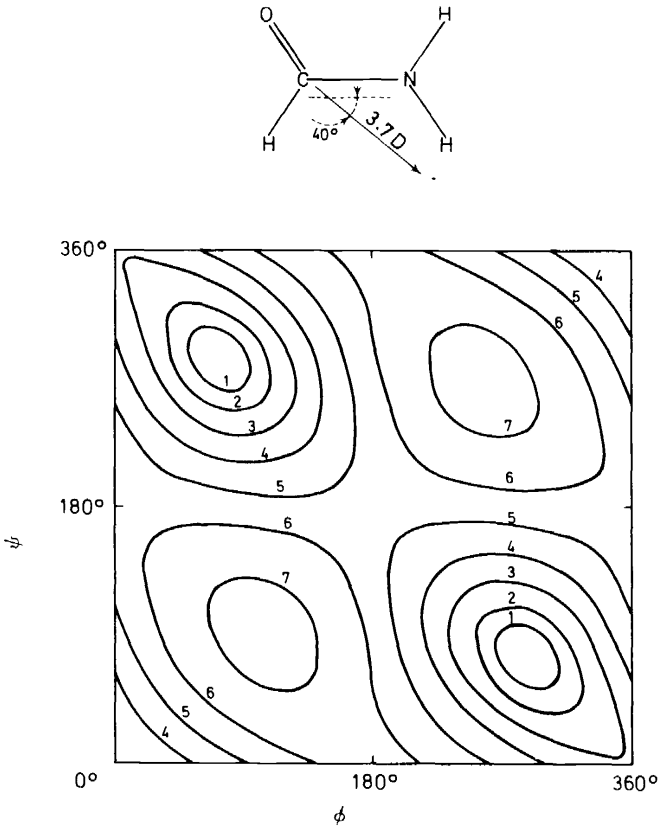


Figure 9. Dipole moment (in Debye units) of compounds I as a function of ϕ and ψ dihedral angles.

On the glycine derivative spectrum we observe a single band located at 3412 cm^{-1} and we therefore conclude that, in this case, every solute molecule is in the C_5 conformation.

In the case of the other compounds, a shoulder appears on the high frequency edge of the C_5 band and is attributed to free N–H oscillators due to the presence of open structures.

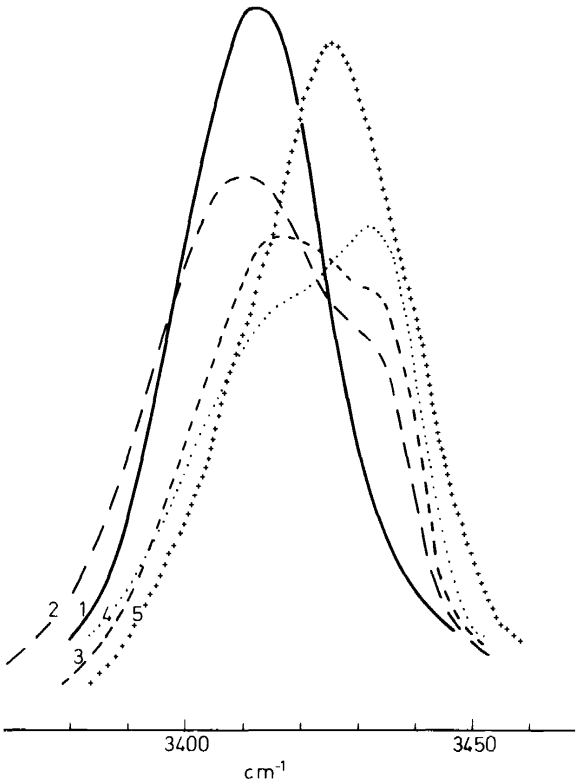


Figure 10. Infrared spectra of compounds III (R,Me)

1	R = H	(glycine)
2	R = C_2H_4—S—Me	(methionine)
3	R = *i*-Bu	(leucine)
4	R = *i* Pr	(valine)
5	R = CH_2—C_6H_4—O—$CH_2C_6H_5$	(benzyltyrosine)

Finally, the spectrum of the *O*-benzyltyrosine derivative (curve 5) is a quite peculiar one. All the molecules are in an open form and the single maximum which is seen at 3431 cm^{-1} must obviously be attributed to free N–H vibrators.

The influence of the bulkiness of the side substituent can easily be explained if we consider the intramolecular non-bonded interactions as shown in *Figure 11*.

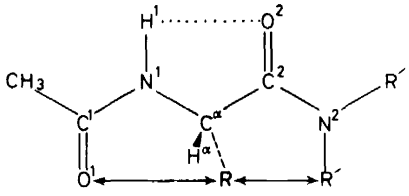


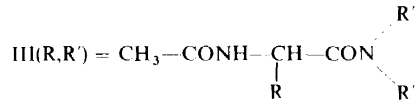
Figure 11. Internal steric repulsions which induce the warping of the C_5 conformations of model molecules III (R,R′).

Due to the peculiar symmetry of this molecule the C_5 conformation of the glycine derivative is very near to the fully-stretched structure, the average values of the rotational angles ϕ and ψ being equal to zero.

On the contrary the dissymmetry introduced by the side substituent in the other aminoacids induces a warping of these molecules.

Our nmr observations are quite in accordance with this assumption and the experimental results collected in *Table 3* show the correlative variation of the azimuthal ϕ angle.

Table 3. Determination of the ϕ angle by nmr spectroscopy



Compound	J_{obs} Hz	J_{corr} Hz	$\Theta°$	$\Phi°$
III(H, Et)	4.0	4.2	120 ± 2	0 ± 2
III(Me, Et)	8.2	9.1	144 ± 3	24 ± 3
III(Bz—O—Bz, Et)	8.7	9.7	147 ± 3	27 ± 3
III(Et—S—Me, Et)	9.0	10.0	148 ± 3	28 ± 3
III(Bu, Et)	8.9	9.8	148 ± 3	28 ± 3
III(*i*Pr, Et)	9.5	10.5	152 ± 4	32 ± 4
III(Me, Me)	7.8	8.6	141 ± 3	21 ± 3
III(Et—S—Me, Me)	8.5	9.4	145 ± 3	25 ± 3
III(*i*Bu, Me)	8.9	9.9	148 ± 3	28 ± 3
III(Bz, Me)	9.0	10.0	148 ± 3	28 ± 3
III(*i*Pr, Me)	9.0	10.0	148 ± 3	28 ± 3

The ϕ value for the *O*-benzyltyrosine derivative is worthy of notice because the infrared spectrum of this compound obtained in pretty similar conditions, shows that all those molecules are in the open conformation. The ϕ angle remaining below 30° we conclude that the releasing of the intramolecular interaction is likely to be attributed to an increase of the ψ-angle.

Similar measurements have been done with models I which contain two monosubstituted amide groups. Some experimental difficulties arise due to the more intense self-association of these compounds and to the complexity of their nmr spectra. Nevertheless, it has been possible to estimate the ϕ angles which characterize the C_5 conformations taken by these molecules and in *Table 4*, these values are compared to the frequency shifts resulting when the N–H oscillator is included in the C_5 ring structure. There is a pretty good correlation between these two values when the side substituent is a paraffinic one. The frequency shift drops when the ϕ angle increases which means that the interaction between the N–H and CO sites is progressively weakened.

Using the same experimental methods, it is possible to get some quantitative data about the C_7 conformations.

Infrared examination of models II in which the C_5 structure is forbidden due to the complete substitution of the first nitrogen atom has enabled us to

Table 4. Correlation between the frequency shift $(\Delta\nu)_5$ and the ϕ dihedral angle which characterizes the C_5 conformation of the model compounds I(R, CH_3)

CH_3—CONH—CHR—CONH –CH_3

R	$C_5\phi^\circ$ (nmr)	$(\Delta\nu)_5$ (ir) cm^{-1}
H (glycine)	0	40
Me (alanine)	21	29
Bu (norleucine)	24	20
*i*Bu (leucine)	26	17
*i*Pr (valine)	28	13

assign the absorption maximum located at 3 335 cm^{-1} to the N–H oscillator when included in the seven-membered coordinated ring.

As for nmr investigations, the model I deriving from glycine is a more favourable case since there is no cause to distinguish the axial and equatorial C_7 forms. So, in this peculiar case, we have merely to take into consideration the equilibrium between the C_5 structure and a single C_7 conformation.

Using infrared spectroscopy, our colleagues Avignon and Huong[6] in Bordeaux have succeeded in measuring the relative amounts of the C_5 and C_7 conformations in the case of the *N*-methylamide of acetylglycine. They have found that the molar ratio of the C_7 form is equal to 60 %.

In addition to our own results obtained under very similar conditions by nmr spectroscopy, these data have enabled us to estimate to 6.2 Hz the average spin coupling constant which characterizes the C_7 form of the glycine derivative. This value corresponds to a ϕ angle equal to 105° and this conclusion agrees quite well with the description previously given for the C_7 structure.

In the case of models I deriving from other aminoacids we have to take into account the two possible axial and equatorial conformations.

In most cases, there is experimental evidence that the equatorial form is the more stable one.

Our first argument arises from some considerations about the infrared spectra of models IV and II.

In the first case the pyrrolidine ring of the proline residue is such that the C_7 form can only be an equatorial one, and the corresponding frequency shift, which is indicated in *Figure 12* is found equal to 135 cm^{-1}.

The same measurement has been performed with the compound II derived from *N*-methylalanine in which the steric hindrance between the two adjacent methyl groups allows the axial conformation only. The frequency shift is then lowered to 77 cm^{-1}.

Taking into consideration the experimental values which are gathered in *Table 2* and in the lower part of *Figure 12*, and which concern dipeptidic species I in which the side substituent is a paraffinic one, it appears clearly that the equatorial form is the most probable one.

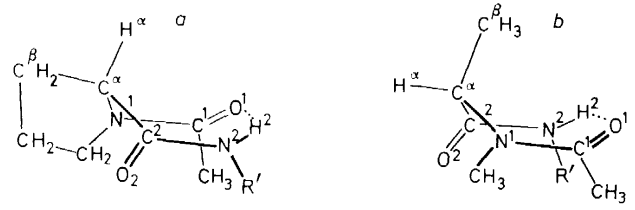


Figure 12. $(\Delta\nu)_7$ shifts which characterize the equatorial and axial C_7 conformations.

Aminoacid	Compound	$(\Delta\nu)_7$ exp cm^{-1}	Aminoacid	Compound	$(\Delta\nu)_7$ exp cm^{-1}
Glycine	I(H, Me)	115	Leucine	I(*i*Bu, Me)	135
Alanine	I(Me, Me)	120	Valine	I(*i*Pr, Me)	131
Norleucine	I(Bu, Me)	131	Methionine	I(Et—S—Me, Me)	115

A second reason can be found in a systematic study of the free N–H absorption bands.

In *Figure 13*, one can see how the frequency of the first N–H oscillator, which is free in the C_7 conformation, depends on its orientation relative to the adjacent C–H covalent bond.

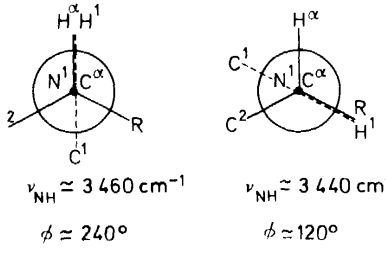


Figure 13. Influence of the orientation of the C^7 contiguous carbon atom on the ν_{N-H} stretching frequency.

In the axial conformation, the situation is such that we should observe a free band at 3 460 cm^{-1}. On the contrary, in the case of an equatorial one, we may expect a maximum at 3 440 cm^{-1}. So, the data collected in *Table 5* are quite in agreement with this last assumption.

Table 5. (ν_{N-H}) free frequency of the C_7 conformation of compounds I(R, Et)

	Expected frequencies	
$(C_7)_{ax}$	$(\nu_{NH})_{free} = 3460$ cm^{-1}	$\phi = 240°$
$(C_7)_{eq}$	$(\nu_{NH})_{free} = 3440$ cm^{-1}	$\phi = 120°$
	Experimental frequencies $(\nu_{NH})_{free}$	
Alanine	I(Me, Et)	3445
Norleucine	I(Bu, Et)	3445
Valine	I(*i*Pr, Et)	3443
Methionine	I(Et—S—Me, Et)	3442
Phenylalanine	I(Bz, Et)	3441
O-benzyltyrosine	I(Bz—O—Bz, Et)	3442

So, the experimental results we have obtained with the dipeptidic models I when the side-substituent is a paraffinic one, lead to the following conclusions.

Under the conditions of our measurements, we have mainly identified the C_5 and the equatorial C_7 conformations. At room temperature we did not observe any appreciable amount of open structures.

It has been possible to evaluate the molar ratio of the two forms and the results are reported in *Table 6*.

Table 6. Relative amounts of C_5 and C_7 forms for several dipeptidic models I(R, CH_3) (dilute solution in CCl_4, $T = 25°C$)

Side substituent	$\%C_5$	$\%C_7$
H(glycine)	40	60
Me(alanine)	30	70
*n*Bu(norleucine)	30	70
*i*Bu(leucine)	25	75
*i*Pr(valine)	25	75
Bz(phenylalanine)	55	45
$-C_2H_4-S-$Me(methionine)	43	57

It appears that the relative number of the C_5 forms slightly decreases when the side substituent becomes more bulky. This can be explained by the correlative warping of this structure.

On the contrary, when the side chain is not completely inactive, anomalies are observed as the abnormal stabilization of the C_5 form in the case of phenylalanine and methionine residues.

In these two cases, this peculiar behaviour is due to a second intramolecular attractive interaction as shown in *Figure 14* for the phenylalanine derivative.

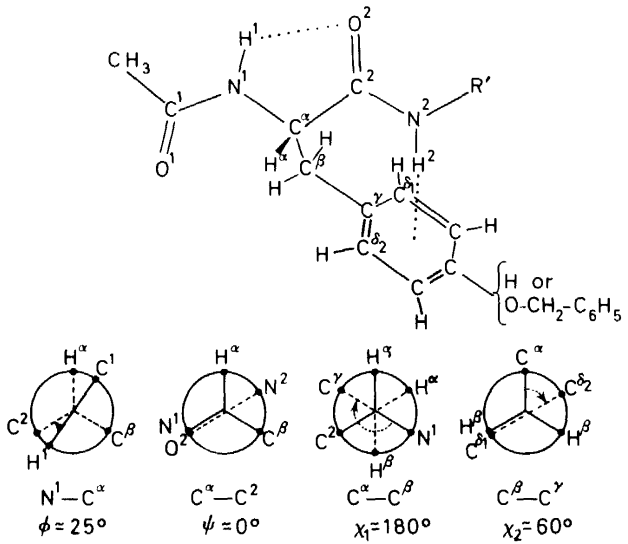


Figure 14. Stabilization of the C_5 conformation in the case of models I derived from phenylalanine. After Marraud[7].

More specific effects appear when operating with a model compound corresponding to a trifunctional aminoacid. In such a case it is necessary to proceed by studying many models of increasing complexity. For instance, in the case of the serine residue, systematic spectroscopic examinations have been performed with the various chemicals shown in *Figure 15*.

Using this method, Marraud[7] in Nancy, has succeeded in elucidating the conformations of several trifunctional dipeptides.

For instance, he has proved that *S*-ethylcysteine, though it is an isomer of methionine, has a quite opposite conformational behaviour. In its case, nearly all the solute molecules are in the equatorial C_7 form due to an additional interaction between the sulphur atom and the second peptide group as shown in *Figure 16*.

Although it is included in a five-membered coordination ring, this particular hydrogen bond involving a sulphur atom is strong enough to lower

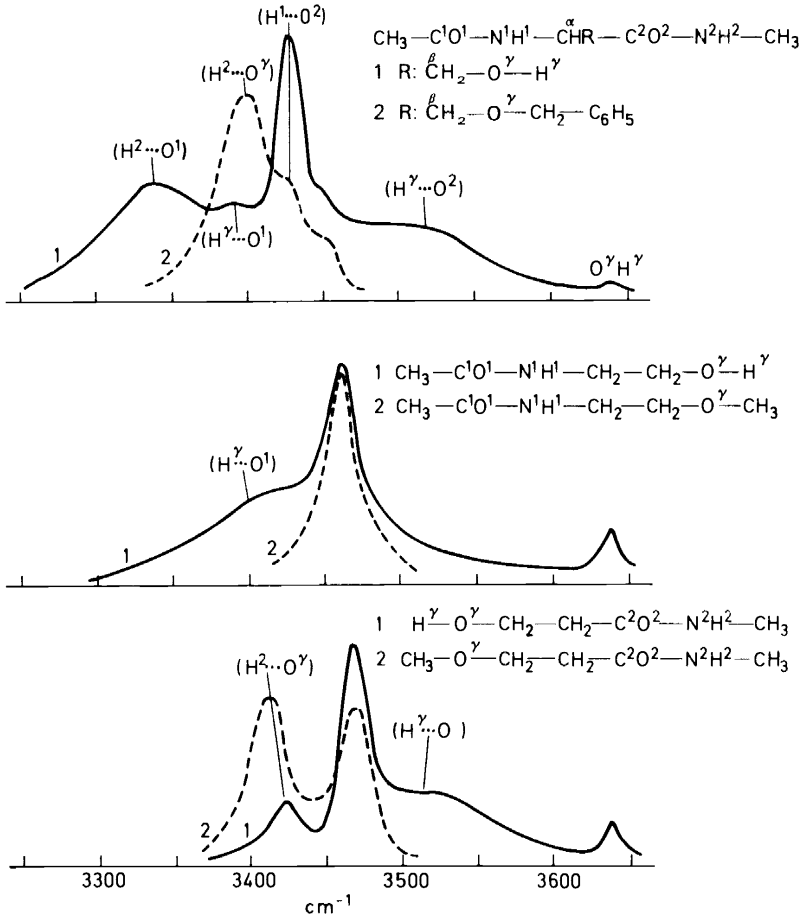


Figure 15. Infrared investigations[7] performed on several model compounds in order to elucidate the conformation of *N*-acetyl-*N'*-methyl serinamide I (CH_2OH, Me).

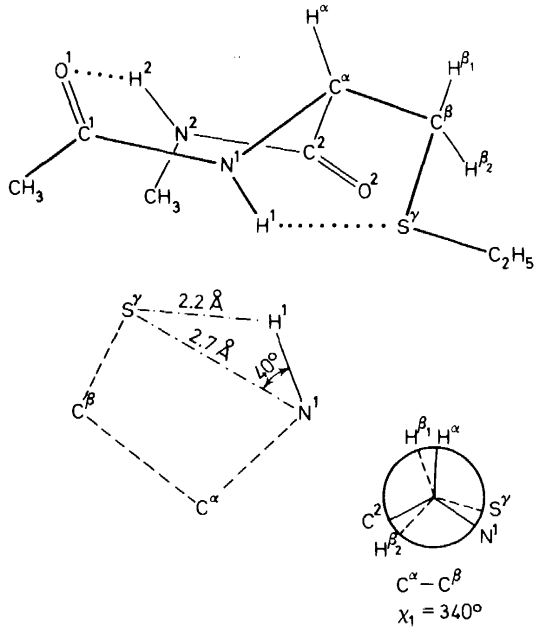


Figure 16. Conformation of the model compound I (—CH_2—S—Et, Me) derived from S-ethylcysteine.

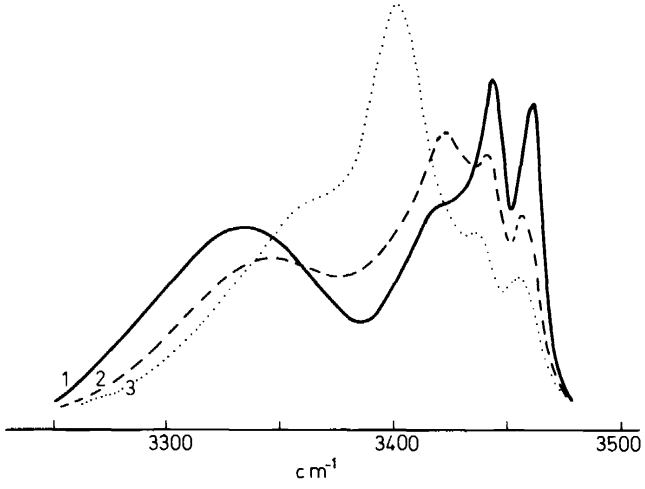


Figure 17. Infrared spectra of:

1. I(Bu, Me) (norleucine)
2. I(—CH_2—CH_2—S—Me, Me) (methionine)
3. I(—CH_2—S—Et, Me) (S-ethylcysteine)

the N–H stretching frequency to 3 402 cm^{-1} and to induce a slight distortion of the C_7 adjacent ring so that the corresponding absorption band is slightly shifted. Thus, the infrared spectra of the two isomeric models are quite different as shown in *Figure 17*.

In *Figure 18*, the two upper patterns represent the two conformations which are in equilibrium in the case of the serine derivative while the third is the only form observed after benzylation of the hydroxyl group[8].

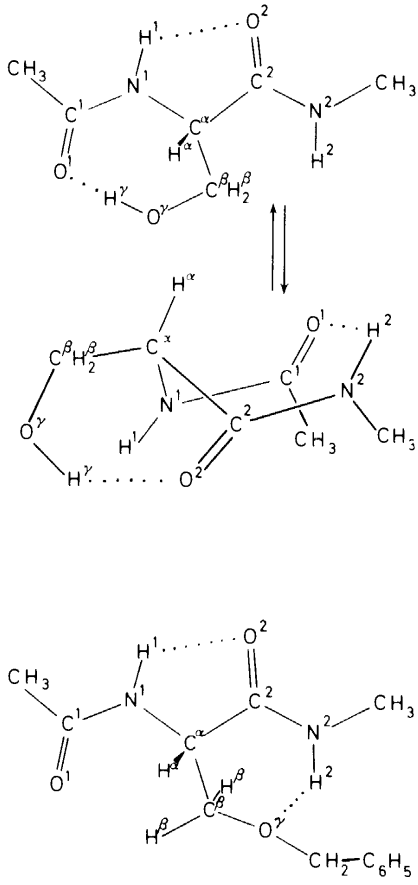


Figure 18. Conformations of serine derivatives after Marraud and Néel[8].

Homoserine, shown in *Figure 19* is a special case[7]. Some of these dipeptidic molecules are in a new conformation characterized by the rotational angles $\phi = 60°$, $\psi = 300°$, $\chi_1 = 30°$, $\chi_2 = 30°$, $\chi_3 = 60°$. The peculiar behaviour of this molecule is due to the fact that the orientation of the side functional group, with respect to the main peptidic chain, depends upon three degrees of freedom. So, the hydroxyl group can be in such an arrangement that it can act, in the same conformation, as an acceptor as well as a donor site in the stabilization of two adjacent coordinated rings.

The experimental data that have just been reported can be compared with the computed conformational maps resulting from different theoretical examinations.

The less elaborate one is the hard sphere approximation according to which each atom is considered as a an incompressible particle. Such a treatment allows us to settle the boundaries of the sterically forbidden areas on the conformational map.

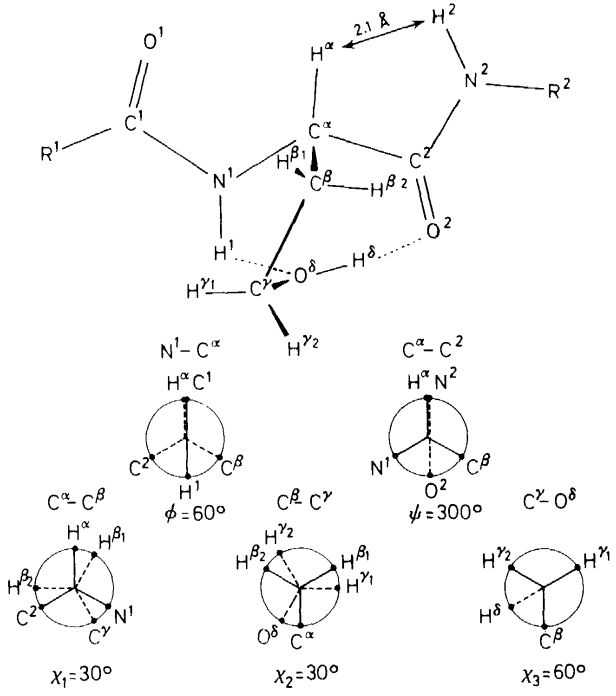


Figure 19. Specific conformation of some of the molecules of the homoserine derivative.

A better approach consists in computing the free enthalpy for each point of the square diagram by using a multicomponent potential function taking into account several contributions such as intramolecular steric and electrostatic effects and internal hydrogen bonds.

In the last two years, more systematic attempts have been made in omitting the usual partitioning of the potential energy into empirically determined components and in proceeding within the more general scheme of the molecular orbital method.

Thus, the conformational maps of the simplest dipeptides have been computed by Hoffmann and Imamura[9] with the extended Huckel method.

More recently, Pullman and coworkers[10] have carried out other treatments with their new method; Perturbative Configuration Interaction using Localized Orbitals (PCILO). This is an analysis which goes beyond the

self-consistent field approximation in the calculation of the ground state energy by incorporating an appreciable fraction of the correlation energy.

It is of some interest to compare the results of these various theories with some of our own experimental conclusions.

I have selected the alanine and serine derivatives which are quite different typical examples.

The former corresponds to the simplest asymmetric peptidic residue and the latter, which derives from a trifunctional aminoacid, is worthy of notice because of its specific conformational behaviour.

In the case of the alanine derivative, we had found that 70% of the molecules are in the equatorial C_7 form, the others being in a slightly warped C_5 conformation where the ϕ and ψ angles are respectively equal to 21° and 350°.

Table 7 enables us to compare these data with different theoretical results obtained by various authors who have used either the hard sphere approach (values indicated under the letter A), the multicomponent potential function (letter B), or quantum-mechanical procedures (letter C).

The hard sphere analysis shows that the C_5 conformation is sterically allowed. On the contrary, it is generally concluded that the C_7 forms are forbidden and this discrepancy with our experimental observations proves that such a treatment undervalues the influence of attractive intramolecular forces.

When using the second method the conclusions are somewhat ambiguous. The best agreement is found when the potential function used contains a component which takes into account the dipolar interactions or the possibility of hydrogen bonding as in the more recent calculations published by Scheraga[11] and Popov[12].

As for the quantum-mechanical methods, it appears that our experimental data completely disagree with the results obtained by the extended Huckel type treatment.

On the contrary, by PCILO calculations, Pullman and coworkers[13] were able to draw a free enthalpy map which reveals three minima corresponding to the C_5 and the two C_7 conformations. The relative stability order which results from their calculations is different from our own. A complete agreement has been obtained recently after they have introduced in the computer programme an entropy correction which can be evaluated from the free enthalpy surface. In the definitive map so obtained, which is called the probability map, the three more favourable conformations are located as shown in the last column of *Table 7* and their relative probabilities are quite consistent with our own experimental conclusions.

Due to the greater complexity of this model, the serine derivative is a more specific example. Let us remember that we have concluded from our experimental measurements that this molecule can also exist in the C_5 and the equatorial C_7 forms. These two conformations are stabilized by additional hydrogen bonds which involve the side hydroxyl group so that the two rotational angles χ_1 and χ_2 which characterize the side-substituent have well-defined values.

The free enthalpy map computed by the PCILO method[14] is shown in *Figure 20*. One can see four minima and the two deepest correspond to the

Table 7. Conclusions of the various theoretical calculations for the dipeptidic model I(Me, Me) derived from alanine, to be compared to the experimental data: $(C_7)_{eq}$,$\phi = 105°$,$\psi = 230°$; 70 %; C_5,$\phi = 20°$,$\psi = 340\text{-}360°$,30 %

Method of calculation	A[a]	A[b]	A[c]	B[d]	B[e]	B[f]	B[g]	C[h] (EHT)	C[i] (PCILO)	C[j] (PCILO)
C_5 $\begin{cases}\phi(°)\\ \psi(°)\end{cases}$				—	21.8 345.1	—	—	60 330	0 0	9 344
C_{7eq} $\begin{cases}\phi(°)\\ \psi(°)\end{cases}$	—		—	(95) (250)	105 252	(100) (275)	108.2 249.7	—	100 220	102 220
C_{7ax} $\begin{cases}\phi(°)\\ \psi(°)\end{cases}$	—	—	—	—	249.1 111.2	—	240.9 113.1	—	260 140	255 140
Order of decreasing stability of the conformations ↓				C_{7eq} C_5, C_{7ax}	C_{7eq} C_5 C_{7ax}	C_{7eq} C_5, C_{7ax}	C_{7ax} C_{7eq} C_5	C_5 C_{7eq}, C_{7ax}	C_{7ax} C_{7eq} C_5	C_{7eq} C_5 C_{7ax}

Dashed boxes: theoretically forbidden conformations.
Empty boxes: theoretically allowed conformations (without any further data given by the author concerning the ϕ and ψ values).
a: after G. N. Ramachandran *et al. Biophys. J.* **5**, 911 (1965).
b: after G. N. Ramachandran *et al. Biophys. J.* 6, 849 (1966).
c: after H. A. Scheraga *et al. Biopolymers* **4**, 369 (1966).
d: after H. A. Scheraga *et al. J. Chem. Phys.* **45**, 2091 (1966).
e: after G. M. Crippen and H. A. Scheraga: see ref. 11.
f: after P. J. Flory *et al.*: see ref. 2.
g: after E. M. Popov *et al.*: see ref. 12.
h: after R. Hoffmann and A. Imamura: see ref. 9.
i: after B. Pullman *et al. J. Theoret. Biol.* **26**, 321 (1970).
j: after B. Pullman *et al.*: see ref. 13.

conformations which have been experimentally identified. The others are so sharp that they should probably disappear on a probability map.

As a general conclusion, we consider that there is now good agreement between the more elaborate theoretical calculations and the experimental observations in the case of very simple peptidic compounds when the measurements are performed at a very low concentration in an inactive solvent.

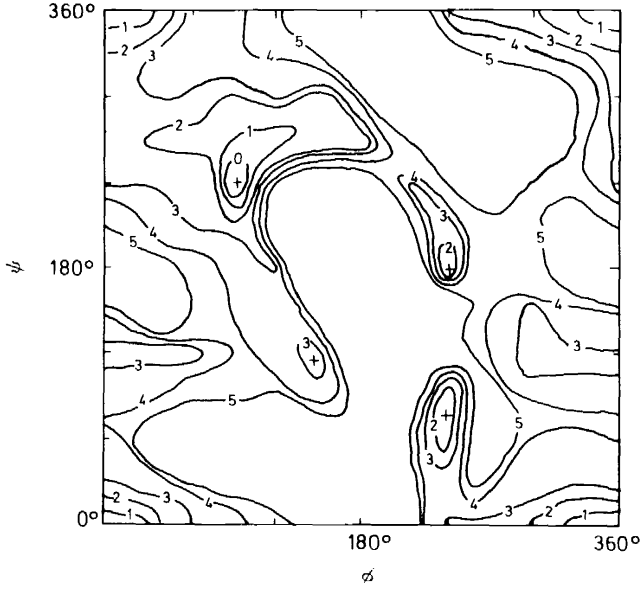


Figure 20. Calculated energy map[14] for the ground state of *N*-acetyl-*N'*-methyl serinamide I(CH_2OH, Me)

Such conditions are very peculiar and it is highly probable that there is no simple correlation between the observations I have reported and the actual conformation of the same residues when included in a protein or in a high molecular weight polypeptide.

It is quite obvious that in these more complex species longer range interactions should be taken into account as well as the specific influence of the solvent in which they can be dissolved.

ACKNOWLEDGEMENTS

The results reported in this lecture have been obtained through a co-operative research carried out by the Laboratory of Macromolecular Chemistry in Nancy (Prof. J. Néel) and the Laboratory of Infrared Spectroscopy in Bordeaux (Prof. J. Lascombe).

Experiments have been performed by Drs. M. Avignon, M. T. Cung, P. V. Huong and M. Marraud.

The author wishes to thank Prof. M. L. Josein, B. Pullman and Dr. C. Garrigou-Lagrange for helpful discussions.

REFERENCES

[1a] P. de Santis, E. Giglio, A. M. Liquori and A. Ripamonti, *Nature* **206**, 456 (1965).
[b] P. R. Schimmel and P. J. Flory, *J. Mol. Biol.* **34**, 105 (1968).
[c] H. A. Scheraga, *Advan. Phys. Org. Chem.* **6**, 103 (1968).
[d] V. Madison and J. Schellman, *Biopolymers* **9**, 511 (1970).
[2] D. A. Brant, W. G. Miller and P. J. Flory, *J. Mol. Biol.* **23**, 47 (1967).
[3a] M. Avignon, P. V. Huong, J. Lascombe, M· Marraud and J. Néel, *Biopolymers* **8**, 69 (1969).
[b] M. Marraud, J. Néel, M. Avignon and P. V. Huong, *J. Chim. Phys.* **67**, 959 (1970).
[4a] V. F. Bystrov, S. L. Portnova, V. I. Tsetlin, V. T. Ivanov and Y. A. Ovchinnikov, *Tetrahedron* **25**, 493 (1969).
[b] V. F. Bystrov, S. L. Portnova, T. A. Balashova, V. I. Tsetlin, V. T. Ivanov, P. V. Kostetsky and Y. A. Ovchinnikov, *Tetrahedron Letters* **60**, 5283 (1969).
[5] M. T. Cung, D. Canet, P. Granger, M. Marraud and J. Néel, *Compt. Rend.* **C269**, 580 (1969).
[6] M. Avignon and P. V. Huong, *Biopolymers* **9**, 427 (1970).
[7] M. Marraud, *Etude des conformations prises par des molécules dipeptidiques en solution dans un solvant inerte. Rôle des interactions spécifiques dues au substituant latéral.* Thesis. Nancy (1971).
[8] M. Marraud and J. Néel, *Conformational study of serine dipeptides by using infrared spectroscopy.* International Symposium on Macromolecules IUPAC. Joint Symposium J3. Boston (1971).
[9] R. Hoffman and A. Imamura, *Biopolymers* **7**, 207 (1969).
[10] B. Maigret, B. Pullman and M. Dreyfus, *J. Theoret. Biol.* **26**, 321 (1970).
[11] G. M. Crippen and H. A. Scheraga, *Proc. Nat. Acad. Sci. U.S.* **64**, 42 (1969).
[12] E. M. Popov, V. G. Dashevskii, G. M. Lipkind and S. F. Arkhipova, *Mol. Biol.* **2**, 622 (1968).
[13] B. Maigret, B. Pullman and D. Perahia, *J. Theoret. Biol.* **31**, 269 (1971).
[14] D. Perahia, B. Maigret and B. Pullman, *Theoret. Chim. Acta* **19**, 121 (1970).