THE CONFORMATIONAL ENERGY MAP OF AN ALANYL RESIDUE PRECEDING PROLINE : A QUANTUM-MECHANICAL APPROACH.

M. Maigret, B. Pullman and J. Caillet

Institut de Biologie Physico-Chimique, 13, rue P. et M. Curie, Paris 5è, France.

Received June 29, 1970

Summary. Quantum-mechanical computations on the conformation energy map of a C^{β}-containing amino acid residue preceding proline in a polypeptide chain indicate that the right-handed α-helical region is an allowed one, provided that account is being taken of the flexibility of the pyrrolidine ring. The results of the computations are compared with X-ray data for nine globular proteins.

It is known that in distinction to other amino acid residues, the presence of a prolyl residue in a polypeptide chain induces limitations on the conformations permitted to the L-amino acid immediately preceding it in the chain. On the basis of theoretical investigations on the model dipeptides GLY-L-PRO and L-ALA-L-PRO, Schimmel and Flory (1) concluded that the right-handed α-helix region (R_{α}) is disallowed for an amino acid residue with a C^{β} followed by a prolyl residue. The conformational space allowed for such an amino acid residue occurred only for $\Psi > 240°$. This limitation was introduced by steric repulsions between C^{β} of the ALA residue and the $C^{\delta}H_2$ group of the pyrrolidine ring (Fig. 1). More recently, however, Damiani et al. (2) guided by the existence of LYS 87 of myoglobin (preceding its PRO 88) in the R_{α} conformation, have shown by new calculations that the occurrence of the R_{α} conformation for a C^{β} containing residue followed by PRO cannot be completely excluded, the steric hindrance defined above being possibly compensated by other factors such as folding or packing.

The above quoted calculations were all carried out by empirical methods (3) and used the geometry of the pyrrolidine ring obtained from X-ray study of L-Leu-L-Pro-Gly (4). Their analysis shows that the availability of the R_{α} conformation for the residue preceding PRO depends essentially on the mode of description of the $C^{\delta}H_2$ group of PRO (considered as a bulky "methyl" atom or as built up of

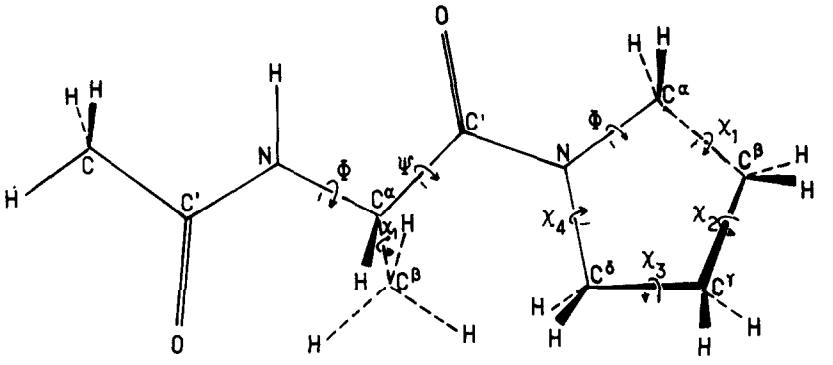

Figure 1. The model compound and notations.

individual atoms). Their conclusion is thus artificial, because linked to the theoretical method used.

We have undertaken a more fundament investigation of the problem by quantum-mechanical calculations, using the PCILO method,following our general stu-

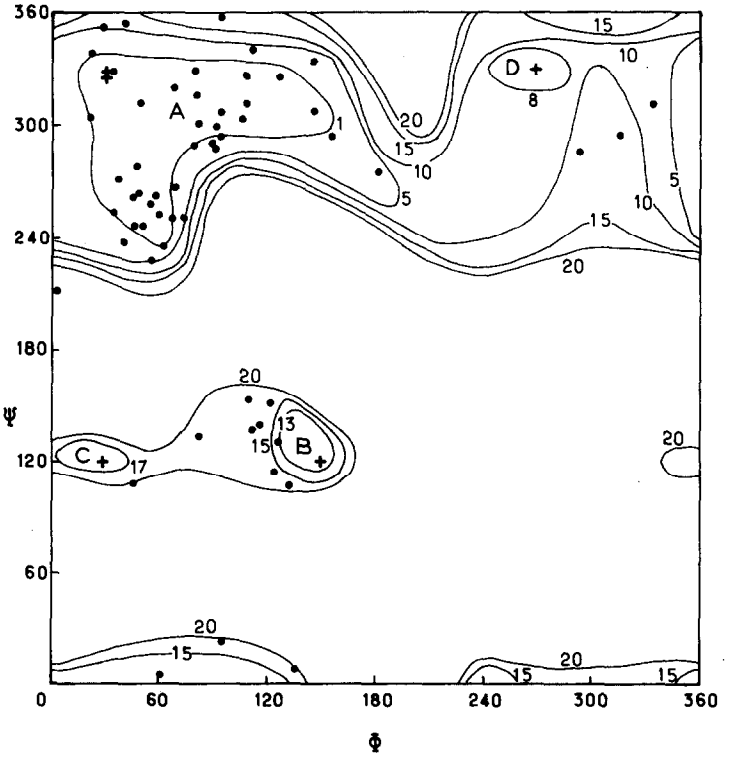

Figure 2. Conformational energy map for ALA proceding PRO with a fixed geometry of the pyrrolidine ring.
* global minimum, + local minima • experimental conformations (table I).

TABLE I

The conformation of aminoacid residues preceding a prolyl residue in globular proteins. The underlined residues occur in or near B_α region.

Protein	Residue	Φ	Ψ	Protein	Residue	Φ	Ψ
Lysozyme[9]	Thr 69	58	263	Carboxypeptidase A[15]	His 29	40	239
	Ile 78	35	329		Arg 45	94	297
Myoglobin[10]	His 36	49	264		Arg 59	146	337
	Lys 87	133	107		Asn 93	80	318
	Ile 99	90	290		Asn 112	68	252
	His 119	36	254		Ser 159	91	300
α-Lactalbumin[11]	Leu 25	110	155		Tyr 204	96	23
	Asp 69	58	263		Ile 213	127	327
Ribonuclease-S[12]	Lys 41	81	300		Leu 281	79	290
	Tyr 92	146	307		Ile 287	125	115
	Asn 113	48	279	α-chymotrypsin[16]	Val 3	69	267
	Val 116	62	5		Glu 7	107	327
Erythrocruorin[13]	Asp 19	62	236		Val 23	94	307
	Asp 31	46	263		Try 27	46	246
	Ala 53	123	153		Leu 123	80	329
	Val 88	116	141		Thr 151	90	287
Subtilisin-BPN'[14]	Val 4	294	266		Leu 160	29	351
	Ala 13	112	138		Gly 197	94	359
	His 39	156	294		Thr 224	49	311
	Val 51	37	272	Oxyhaemoglobin, α chain[17]	Phe 36	57	229
	Thr 55	46	108		Phe 43	56	260
	Ala 85	73	253		Leu 76	128	131
	Gly 128	317	295		Asp 94	107	304
	Tyr 167	4	213		Leu 113	46	245
	Tyr 171	21	305		Thr 118	112	342
	Gly 193	23	339	β chain	Tyr 35	61	253
	Ala 200	40	356		Asp 50	336	312
	Leu 209	68	320		Asn 57	183	276
	Ser 224	83	133		Asp 99	110	314
	His 238	50	246		Thr 123	137	7

dies in this field (5)-(8). The more so as the investigation of X-ray data on 9 globular proteins (Table I) indicates that out of 60 residues preceding PRO, 9 (15°/$_o$) occur in the R_α conformation or close to it. Our computations involved two steps. In the first one, the conformational map in the (Φ, Ψ) plane has been constructed, in increments of 30°, with a fixed geometry adopted for the pyrrolidine ring (following 2, with C_γ exo or χ_1 negative) and the alanyl residue (3). The results are shown in fig. 2. They indicate the existence of four zones of local energy minima, labelled A, B, C and D. The global minimum lies in A at (Φ, Ψ) = 30°, 330°. The minimum B in the R_α region is about 13 Kcal/mole above the global one and this region appears therefore as a disallowed one.

In the second step, a refined calculation has been performed, taking into account the possible flexibility of the pyrrolidine ring as demonstrated recently by the work of Ramachandran et al. (18) and involving moreover an optimization procedure (19)(20). Because of the localization of the most important steric hindrance around the $C^\beta H_3$ of ALA and $C^\delta H_2$ of PRO only these groups have been considered in the optimization procedure. Thus, for each (Φ, Ψ) point of the alanyl conformational map, this process involved the set of five variables : χ_1 of the ALA residue, χ_3 and χ_4 of the PRO residue, the $N-C^\delta$ bond length and the $C'-N-C^\delta$ bond angle of the pyrrolidine ring. In the areas of stability a further optimization with respect to Φ and Ψ and the above quoted variables has finally been performed in order to fix all the possible minima.

The results of this refined computations are indicated in fig. 3. The general contours of the conformational map are similar to those of fig. 2. The global minimum still lies in the A zone (at Φ, Ψ = 31°, 327°) and there appear in that zone several other local minima at about 2 Kcal/mole above the global one. The local minima in zones B, C and D undergo only small displacements and a new local minimum appears in zone E. The minimum in zone B (Φ, Ψ = 147°, 131°) now occurs, however, at only 6.1 Kcal/mole above the global minimum and the R_α conformation appears thus as a possible one for a C^β containing residue preceding PRO, in as much as such a residue is represented by ALA. The minimum in the neighbour zone C is a relatively high one, located at 14,4 Kcal above the global one.

The distortion of the geometrical data are of importance. We find :

$$48° < \chi_1^{ALA} < 65°, \qquad -20° < \chi_3^{PRO} < -15°,$$

$$4° < \chi_4^{PRO} < 13°, \qquad 1.44\ \text{Å} < N-C^\delta < 1.41\ \text{Å} \text{ and}$$

$$125° < \tau_{C'NC} < 128°,$$

in very good agreement with the suggestions of (18).

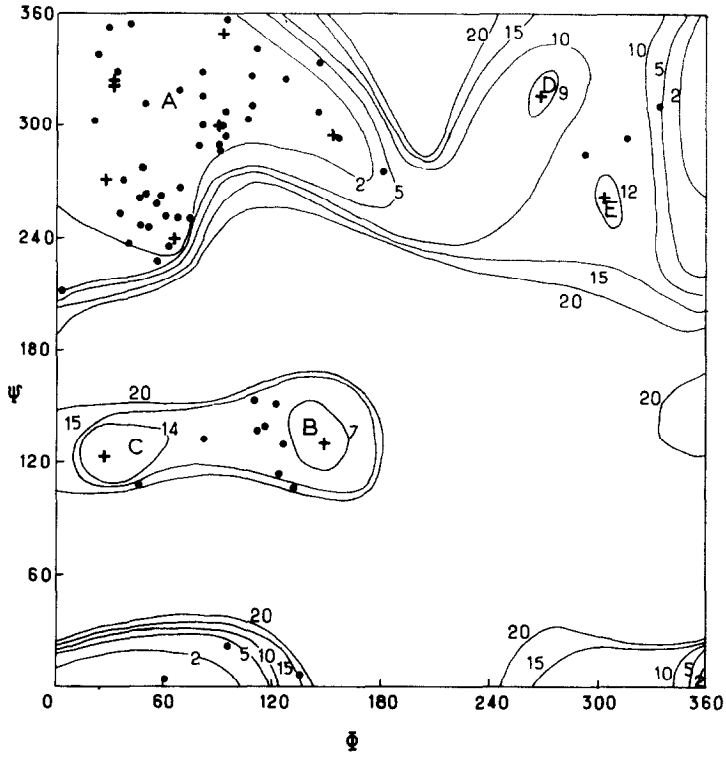

Figure 3. Conformational energy map for ALA proceeding PRO with a flexible geometry of the pyrrolidine ring.
‡ global minimum, + local minima • experimental conformations (table I).

It may thus be considered that the occurrence of a C^{β} containing residue proceding PRO in the R_{α} conformation is greatly favored by the departure of the C^{β}-side chain from its rotational minimum and by the distortion of the pyrrolidine ring. It may be added that in 8 cases out of 9 the residues preceding PRO which occur in or near the R_{α} region are followed by a PRO which has itself the R_{α} conformation (while the PRO's are distributed nearly equally in the regions $(\Phi, \Psi) \sim 120°$, 150° and 120°, 330°). Finally it may be useful to point out the overall agreement between the computed zones of stability and the localization of the conformation of all the residues listed in table I. Only a few residues from subtilisin fall somewhat outside the limits of these zones. In view of the approximate nature of the X-ray data for that and some other of the proteins quoted, presently in the stage of refinement, the general agreement may be considered very satisfactory.

References.

1. Schimmel P.R. and Flory P.J., J. Mol. Biol. 34, 105 (1968).
2. Damiani A., De Santis P. and Pizzi A., Nature 226, 542 (1970).
3. See e.g. Scheraga H.A., Adv. Phys. Org. Chem. 6, 103 (1968).
4. Leung Y.C. and Marsh R.E., Acta Cryst. 11, 17 (1958).
5. Maigret B., Pullman B. and Dreyfus M., J. Theoret. Biol. 26, 321 (1970).
6. Pullman B., Maigret B. and Perahia D., Compt. Rend. Acad. Sci. Paris, 270, 1396 (1970).
7. Maigret B., Pullman B. and Perahia D., Biopolymers, in press.
8. Pullman B., Maigret B. and Perahia D., Theoret. Chim. Acta, in press.
9. Phillips D.C., Proc. Natl. Acad. Sci. U.S., 57, 484 (1967).
10. Watson H.C., Progress in Stereochemistry, 4, 299 (1969).
11. Browne W.J., North A.C.T., Phillips D.C., Brew K., Vanaman T.C. and Hill R.L., J. Mol. Biol. 42, 65 (1969).
12. Private communication from Dr. Wyckoff. See also Wyckoff H.W., Tsernoglou D., Hanson A.W., Knox J.R., Lee B. and Richards F.M., J. Biol. Chem. 245, 305 (1970)
13. Private communication from Dr. Steigemann. See also Huber R., Epp O. and Formanek H., Naturwissenschaften, 56, 362 (1969).
14. Private communication from Dr. Wright. See also Wright C.S., Alden R.A. and Kraut J., Nature, 221, 235 (1969).
15. Private communication from Dr. Rieke. See also Rieke G.N., Hartsuck J.A., Ludwig M.L., Quiocho F.A., Steitz T.A. and Lipscomb W.N., Proc. Natl. Acad. Sci. U.S., 58, 2220 (1967).
16. Birktoft J.J., Mathews B.M. and Blow D.M., Biochem. Biophys. Research Comm. 36, 131 (1969).
17. Calculated from X-ray data kindly provided by Dr. Perutz the values are very approximate. See also Perutz M.F., Muirhead H., Cox J.M., Goaman L.C.G., Mathews F.S., McGandy E.L. and Webb L.E., Nature 219, 29 (1968).
18. Balasubramanian R., Lakshminarayanan A.V., Sabesan M.N., Teguni G., Venkatesan, K. and Ramachandran G.N., in press.
19. Nelder J.A. and Mead R., Computer J. 7, 308 (1968).
20. Chandler J.P., Simplex Method. Quantum Chemistry Program Exchange (QCPE), Dept. of Chemistry, Indiana University, Bloomington, Indiana.