Vol. 112, No. 1, 1983 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 15, 1983 Pages 339-346

MODELS FOR THE CONFORMATIONAL BEHAVIOUR OF ANGIOTENSIN-II IN ACIDIC AQUEOUS SOLUTIONS

C. MARCHIONINI, B. MAIGRET & S. PRÉMILAT

Laboratoire de Biophysique Moléculaire
E.R.A. C.N.R.S. N° 828
UNIVERSITE DE NANCY I
Faculté des Sciences
B.P. N° 239
54506 VANDOEUVRE LES NANCY
FRANCE

Received February 3, 1983

Clustering analysis is applied to a sample of Angiotensin-II conformations obtained from a Monte-Carlo procedure. It is shown that 5 families of conformations can describe entirely the sample. Each family can be represented by a typical conformer. Two of these last models are predominant in the description of Angiotensin-II in acidic solutions. These theoretical models are in agreement with recent exeperimental results.

The linear octapeptide Angiotensin-II (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe) has been the subject of extensive investigations owing to the broad range of biological activities presented by this hormone (1). As the three dimensional structure of this potent pressor substance might be the key for the understanding of its interactions with its receptors in various tissus, numerous conformational studies (see 2-6 and references therein) have been devoted to this peptide and to analog molecules. As concerns the behaviour of the peptide hormone in solution, the conformations obtained from theoretical or experimental works present large differences, and no single model is consistent with all available data. In such a situation, it is unavailing to believe that only few conformations are present in solution as no particular one gives general agreement with experiments. The hormone ought to be considered as an extremely flexible molecule presenting a large amount of possible conformational states. Such a large number of conformations in equilibrium can explain the range of observations and the modifications of the conformational properties from one solvent to another. In such a situation, the usual computational approaches to molecular conformations using minimization of the conformational energy are irrelevant and only statistical or molecular dynamic computations can satisfactorily be used to tackle such problems.

0006-291X/83 $1.50

We have recently proposed and used a Monte-Carlo sampling method for calculations of molecular conformations (7,8). Applied to the simulation of the conformational behaviour of Angiotensin-II in aqueous solution at several pH, this procedure allows us to get statistical samples of conformations presenting averages in good accordance with experimental ^{1}H-NMR data. But as it is difficult to depict molecular models from these averages, it is necessary to get a usable description from the information contained in the whole Monte-Carlo sample . This can be achieved by clustering analysis, searching for groups of conformers with similar three-dimensional properties. These conformational families are easier to handle in further analysis as their number is limited and their properties more general than that of single individual conformers. The present paper deals with such an attempt applied to the determination of valuable conformational models of Angiotensin-II in acidic aqueous solutions (pH range from about 1 to 4).

METHOD

The method which has been used here is roughly similar to the procedure presented in a previous paper (9). The sample of molecular conformations which has been analysed comes from the Monte-carlo algorithm applied to polypeptide chains (8) and is made of $n = 2184$ *different conformations. To each of these conformations a population number* n_i *is associated in such a way that* $\Sigma n_i = 10000$ *the total number of conformations generated in the Monte-Carlo calculations (one can note that these 10000 conformations constitute a good representative sample in the case of the unimodal distributions obtained for short polypeptide chains at acidic pH). If the parameters describing the molecular geometry (bond lengths and bond angles) are considered as fixed, a molecular conformation can be defined wether by its variable dihedral angles, or by a set of variable interatomic distances. It has been shown (10) that the use of dihedral angles in clustering analysis of molecular conformations introduces difficulties, thus the only relevant parameters considered here to describe conformations are inter-atomic distances. This means that the conformational space of Angiotensin-II is considered in the clustering analysis as* m *dimensional, where* m *represents the number of distances between selected pairs of atoms. As in a medium-size molecule such as Angiotensin-II the amount of possible variable inter-atomic distances in too large, preliminary analysis are necessary to determine the usefull distances which describe the conformational space with a good accuracy. For this purpose, usual statistics (means, correlations,...), discriminant analysis and several hierachical clustering methods were first applied on the initial sample of* n *conformations taking into account all the variable interatomic distances in the molecule. As a result of these calculations, it appears that a subset of 23 distances carry the major part of the conformational information (the 21* C^{α}_{i}-C^{α}_{i+2} *distances, the N- and C-terminals end-to-end distance, the* Tyr$_4$-Phe$_8$ *aromatic side-chain centroïds distance). Among these 23 distances, the latest two are the most discriminant and the most independent ones.*

No clustering program ensures the user to reach the best possible partition of the analysed data, thus the application and interpretation of clustering algorithms are dependent on the experience and perspicacity of the user (11-12). To avoid such difficulties and to measure the validity of the partitions obtained, several clustering methods based on different algorithms (13-18) where used with the same data. The final *clusters of conformations retained are stable for at least 3 different approaches. When the "best" partition of the conformational space is obtained, each family of conformations is analysed so that comparison with experimental data can be done.*

RESULTS AND DISCUSSION

The sample of molecular conformations built from the Monte-Carlo procedure is well separated into five stable clusters. Table I gives the mean values of the end-to-end and Tyr_4-Phe_8 side-chain centroīds distances in each group of conformers. Compact conformations with distant aromatic rings pointing outside are found in cluster N°1 which represents only a small amount (5 %) of all the analysed structures. Cluster n°2 includes 25 % of the conformations and is characterized by extended conformations and by the larger distances between the Tyr_4-Phe_8 rings. The features of the conformations found in cluster n°3 (15 %) are the relative positions cf the aromatic side-chains which are close together with almost parallel rings while the backbone remains in a quite extended form. The conformations found in cluster N°4 (25 %) present a compromise between extended and folded chains : one can find in this group a small amount of conformations folded at the N-terminal with the remaining residues extended, and conformations folded at the C-terminal with the other residues extended. Cluster n°5 is the most important with 30 % of the sample and is built from the most extended chains with distant aromatic rings.

Looking now back to the usual $(\phi,\psi,\{\chi\})$ dihedral angles distributions in each of the families defined above, one can calculate their means and standard deviations using a procedure adapted to periodic variables.

TABLE I

Mean end-to-end and Tyr_4-Phe_8 side-chain centroīds distances in the clusters.

Clusters	End-to-end Distance (Å)	Tyr_4-Phe_8 side-chain centroid distances (Å)
1	9.8	14.9
2	21.3	15.8
3	19.6	6.9
4	17.8	12.9
5	23.4	12.6
in the whole sample	20.3	12.8

TABLE II

(ϕ,ψ,χ) Dihedral angles (IUPAC conventions) which characterize the typical three dimensional structure in each cluster of conformations.

Clusters	Asn_1	Arg_2		Val_3		Tyr_4		Ile_5		His_6		Pro_7	Phe_8
	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ψ	ϕ
N°1	135	- 145	120	- 70	150	- 65	- 45	- 90	150	50	90	110	- 140
N°2	180	- 165	130	- 70	155	- 85	145	- 65	135	- 160	90	- 30	- 100
N°3	180	- 160	125	- 90	180	- 85	145	- 100	55	- 170	170	135	- 135
N°4	180	- 165	130	- 75	150	- 85	145	- 80	130	70	90	130	- 100
N°5	175	- 165	115	-135	150	- 85	145	- 140	150	- 155	90	150	- 40

These values can be used to associate with each group of conformations a typical three-dimensional structure representative of most of the features of the family and defined in a suitable and simple way. These typical conformations represent usable conformational models which mimic the behaviour of Angiotensin-II in acidic aqueous solutions, and can be compared to experimental data or to other proposed conformations. The (ϕ,ψ) values which define these typical models are given in table II and the associated conformations are pictured on figure 1. Using the statistical weights n_i associated to each conformation and the weight of the family to which it partakes (the percentages given above), one can associate to each of these typical conformations a weight describing its representativity in the sample. From these calculations on can easily see that the most probable conformations are in clusters 3 and 4. One can also calculate NMR coupling constants for the amino-acid residues using as usually (8) a Karplus-type formulae and compare calculated values with experimental data. The $J(H_N-H_C\alpha)$ calculated for each family of conformers as well as those obtained from the whole sample (8) or from the mixture of the typical conformations are given in table III. For the last case the agreement between calculated and observed coupling constants is very good and even the larger differences noted for Ile and Phe are in the range of experimental precision. If the (ϕ,ψ) angles as well as their standard deviations for each residue in each family are plotted on the usual Ramachandran (ϕ,ψ) map, it appears that :

- conformations belonging to cluster n°1 are characterized by the fact that the Tyr_1-(ϕ,ψ) angles are in the α_R region and the His_6-(ϕ,ψ) angles in the α_L one,
- conformations in cluster n° 2 are defined by (ϕ,ψ) angles in the extended β region, except for the Pro_7 residue which is in the α_R one,
- the Ile_5 residue of the conformations in cluster n° 3 is in a C_7^{eq}

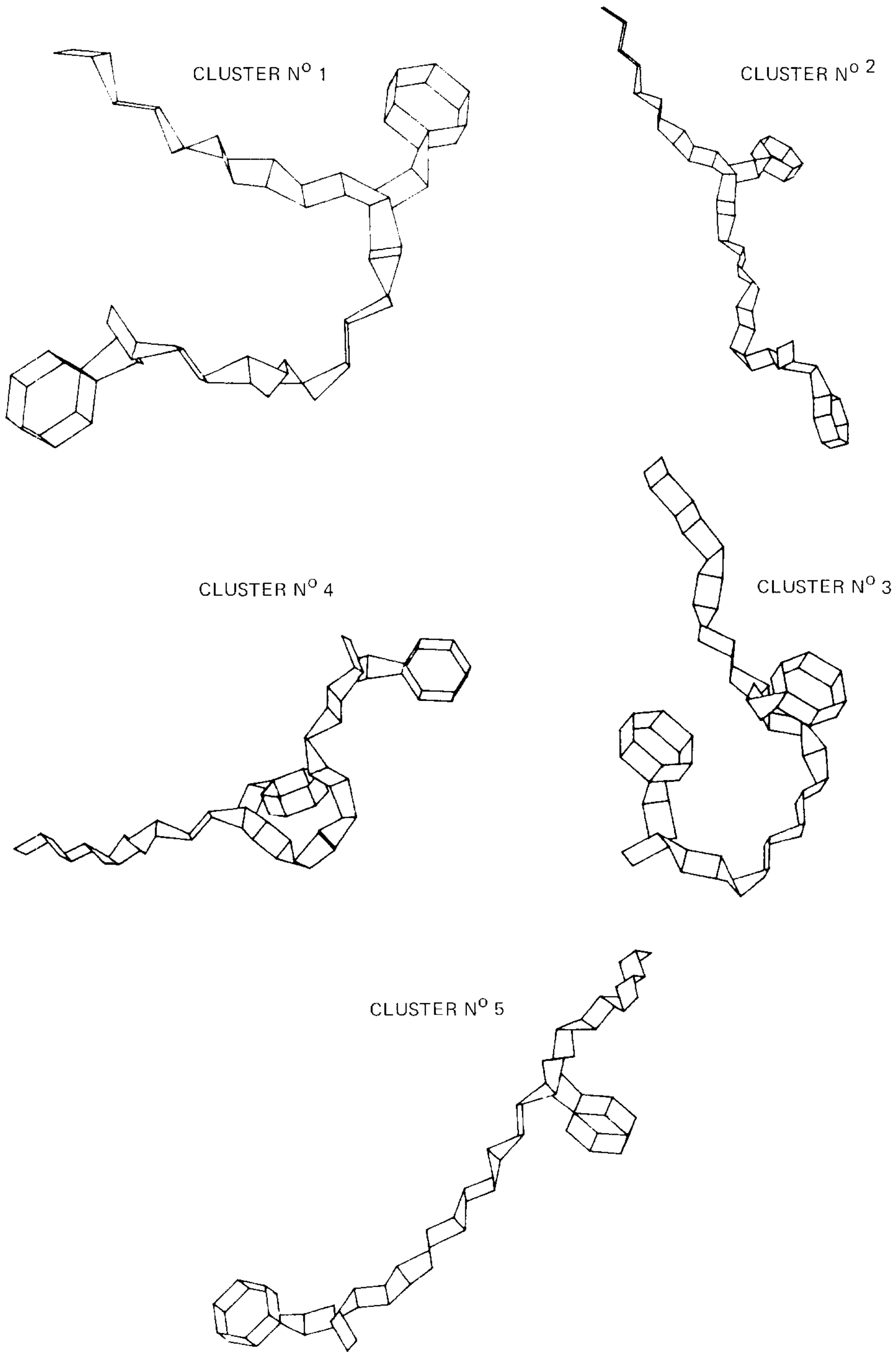

Figure 1

Typical conformations in each cluster.
Only the N-C$^{\alpha}$ - C' atoms of the peptide backbone are ribbon-like represented, as well as the Tyr$_1$ and Phe$_8$ aromatic side-chains.

TABLE III

$J(H_N-H_{C\alpha})$ coupling constants calculated with the clusters of conformations.

	$J_{H_N-H_{C\alpha}}$							
	Asn_1	Arg_2	Val_3	Tyr_4	Ile_5	His_6	Pro_7	Phe_8
Calculated in clusters :								
N°1	-	9.7	5.4	4.5	9.0	5.5	-	10.3
N°2	-	6.4	5.4	8.2	4.5	7.3	-	10.3
N°3	-	7.3	9.0	8.2	10.3	5.4	-	10.8
N°4	-	6.4	6.4	8.2	7.3	5.5	-	10.3
N°5	-	6.4	10.8	8.2	10.3	8.2	-	1.0
Calculated from the whole Monte-Carlo sample (8)	-	6.6	7.9	7.6	7.2	6.1	-	8.0
Calculated from the weighted typical conformation	-	6.6	7.8	8.0	8.0	6.7	-	7.5
From experimental data :								
Ref.(19)	-	6.5	7.9	7.2	8.0	6.0	-	7.3
Ref.(20)	-	6.7	7.6	8.0	8.5	7.0	-	7.5
Ref.(21)	-	-	-	8.1	8.5	6.6	-	-

structure, all the others being in the β region, with a fully extended (C_5) conformation for the His_6 residue,

- all residues have their (φ,ψ) angles in the β region for conformers in cluster n° 4 except for the His_6 amino-acid which has an α_L position as in cluster n° 1,

- in conformations of cluster n°5 the backbones of all the residues are extended, with two possibilities for the Phe_8 residue which can also take an α_R conformation.

When these results are compared with available experimental data, it appears that a good agreement is obtained with the most recent of them :

. Fluorescence-energy transfer studies which have been carried out recently by Schiller (2) on analog molecules indicates that Tyr_4-Trp_8 distances are of about 12 Å, and end-to-end distances of about 17 Å, which is in good agreement with the results obtained in cluster n° 4 which contains most of the probable conformations. Nevertheless, it should be noted that conclusions presented in a previous paper (8) were already acceptable as they were based on a mean end-to-end distance of about 20 Å and a Tyr_4-Phe_8 side-chain distance of about 13 Å. These results were interpreted as due to an equilibrium between several conformations and not to a time-averaged prefered conformation or to a random coil in which no defined group of conformers can be depicted.

. The weakly folded conformations in the C-terminal part of the molecule proposed by Schiller (2) and other authors can be explained by the α_L conformation of His_6 residue. The small amount of conformers folded in the C-terminal observed in the present study is not indicative (at acidic pH) of the situation described by Schiller with folded conformations in the N-terminal and, to a lesser degree, in the C-terminal part of the molecule. This difference can be due to the introduction of Trp_1 instead of Asn_1 which can modify the equilibrium between the different families and increase the amount of folded conformations in cluster n°1.

. The recent proposal of Lenkinski et al. (3-4) who interpret their results in term of a dominant three-dimensional structure presenting a C_7^{eq} - C_5 conformation for the Ile_5-His_6 fragment can also agree with our findings as such structures correspond to the typical conformation of cluster n°3.

It can thus be concluded that Angiotensin-II in acidic solutions presents an equilibrium mainly between 2 different families of conformers and that experimental data are mean values. The conformational groups obtained from the clustering of Monte-Carlo samples allow the determination of molecular models which can well represent these conformers as averages calculated with the models are in very good accordance with experimental measurements.

REFERENCES

1. PAGE I.N., & BUMPUS F.M. Editors, (1979), "Handbook of experimental pharmacology : Angiotensin-II", Springer-Verlag, vol.37
2. SCHILLER P.W.,(1980), Int.J.Peptides Protein Res., 16, 259
3. LENKINSKI R.E., STEPHENS R.L. & KRISHNA N.R., (1981), Biochemistry, 20, 3122
4. LENKINSKI R.E., STEPHENS R.L. & KRISHNA N.R.,(1981), Biochim.Biophys. Acta, 667, 157
5. PIRIOU F., LINTNER K., FERMANDJIAN S., FROMAGEOT P., KHOSLA M.C., SMEBY R.R. & BUMPUS F.M., (1980), Proc.Natl.Acad.Sci.USA, 77, 82
6. FERMANDJIAN S., PIRIOU S., SAKARELLOS C., LINTNER K., KHOSLA M.C., SMEBY R.R. & BUMPUS F.M., (1981), Biopolymers, 20, 1971
7. PREMILAT S. & MAIGRET B., (1977), J. Chem. Phys., 66, 3418
8. PREMILAT S. & MAIGRET B., (1980), J. Phys. Chem., 84, 293
9. MAIGRET B. & PREMILAT S., (1982), Biochem.Biophys.Res.Comm., 104, 971

10. FLAMENBAUM G., ABILLON E. & BENZECRI J.P., (1979), Cahiers de l'analyse des données, 4, 339

11. DUBES R. & JAIN A.K.,(1976), Pattern Recogn., 8, 247

12. DUBES R. & JAIN A.K.,(1979), Pattern Recogn., 11, 235

13. WOLFE J.H., (1970), Multivar.Behav.Res., 5, 329

14. WOLFE J.H., (1967), NORMIX 360 computer program., Rep. n° SRM-68-2 and 72-4, U.S. Naval Personnel Research Activity, San Diego, USA.

15. MACRAE D.J., (1971), Behav. Sciences, 16

16. VASSEUR C. & POSTAIRE J.G., (1980), IEEE Trans. Systems man cyber., 10, 145

17. DIDAY E. (1974), Pattern Recogn., 6, 1

18. DIDAY E. Editor (1980), INRIA,FRANCE "Optimization an classification automatique", vol. 1,2.

19. GLICKSON J.D., DADOK J. & MARSHALL G.R., (1974), Biochemistry, 13, 11

20. BLEICH H.E., GALARDY R.E., PRINTZ M.P., CRAIG L.C., (1973), Biochemistry, 20, 3122

21. LENKINSKI R.E., STEPHENS R.L. & KRISHNA N.R., (1981), Biochemistry, 20, 3122