FEBS Letters 355 (1994) 263–266

FEBS 14830

Conformation of the C-terminus of endothelin-1 in aqueous solution studied by Monte-Carlo simulation

Masataka Kuroda, Kazuto Yamazaki, Tooru Taga*

Faculty of Pharmaceutical Sciences, Kyoto University, Sakyo-ku, Kyoto 606, Japan

Received 8 September 1994

Abstract The conformation of the C-terminus of endothelin-1 in an aqueous solution has been analyzed by a Monte-Carlo simulation including the hydration energy term. The C-terminus may adopt multiple or flexible conformations in the solution, but the classification of the conformations shows that a comparably large number of conformers take a similar folding form in which the C-terminal chain is extended along the α-helix of the N-terminal core, and hydrophobic clusters are formed between the side groups of the C- and N-terminus. The aromatic ring of Trp^{21} is located at a certain distance from the two charged side groups of Asp^8 and Glu^{10}, and the functional groups of the key residues in the N-terminus are uncovered with the C-terminus tail.

Key words: Endothelin; Monte-Carlo simulation; Hydration; Conformation

1. Introduction

Endothelin is a well-known endogenous 21 amino acid peptide which has been isolated from aortic endothelial cells [1]. It has a very potent and sustained vasoconstrictor activity. Three isomers of endothelin have been shown so far. The original endothelin is called endothelin-1 (ET-1), and ET-2 and ET-3 were identified by human genomic analysis [2]. ET-1 has a globular N-terminus which is crosslinked by the two disulfide bridges, and a hydrophobic C-terminus consisting of 6 amino residues from His^{16} to Trp^{21} (Fig. 1). Many structure–activity studies of endothelin have shown that the two disulfide bonds and the Trp^{21} residue are important for functional activity [3]. It has been reported that the C-terminal fragment binds to the receptor with a certain affinity [4], and the conformation of the C-terminus is functionally important. Various ^1H NMR studies of the ET-1 structure in aqueous solution have shown that the globular N-terminus forms a rigid core, including a β-turn between Ser^5 and Asp^8, and an α-helix between Lys^9 and Cys^{15} [5–7], although there is a small variance in the residues from Cys^3 to Asp^8 [8]. On the other hand, in most of the NMR models, the C-terminus is described as an undetermined or flexible region, and does not have a dominant conformation. Recently, the entire conformation of ET-1 in the solid state was determined by X-ray crystallography [9]. In the structure, the C-terminal 16–21 residues form an ordered α-helix, which is quite different from the NMR results. The difference between the solid- and solution-phase structures will be a consequence of the different environments around the molecule, and the conformation of the C-terminus in solution will be largely affected by the solute–solvent interactions. In the present work, a Monte-Carlo simulation of the C-terminus, taking account of the aqueous solvent effect, was carried out in order to elucidate its specific solution-phase structure which was undetermined by the NMR methods.

2. Materials and methods

2.1. Molecular model

To carry out the simulation near to physiological conditions, the molecular model of ET-1 in a zwitterionized form was assumed. The N-terminus was fixed in a folding form, and the coordinates of this core portion were obtained from the NMR study of ET-1 in aqueous solution [8]. The variable region of the C-terminus in the model had 23 rotatable ϕ and ψ bonds of the main chain, including the ψ bond of Cys^{15}, and the 12 χ^n bonds of the side chains except for χ^3 of Leu^{18}, and χ^2 and χ^3 of Ile^{19} and Ile^{20}.

2.2. Monte-Carlo simulation

The Monte-Carlo (MC) simulation was undertaken using the standard Metropolis algorithm [10]. The conformation energy included the non-bonding, electrostatic, bond-torsional, hydrogen-bonded energies, which were calculated by the use of the modified ECEPP force field parameters [11]. The hydration energy was computed from the peptide–water accessible surface area. The surface area for each atom was calculated by the method developed by Perrot et al. [12] and multiplied by the characteristic hydration coefficients which had been evaluated by Ooi et al. [13]. The hydration energy was added to the conformation energy.

The MC sampling of a number of conformations of the peptide chain was performed in separate series of the MC iteration processes, each series including three stages: annealing, energy minimizing and sampling stages. In the annealing stage, the acceptance of the new conformation was artificially increased by assuming a small dumping factor (0.1) for ΔE. In the energy minimizing stage, the system was led into the local energy minimum. In the sampling stage, the first 200 steps before the energetic equilibrium of the system were rejected, and the following 1000 steps were adopted into the statistical calculations. A total of 1000 series of the MC processes were calculated.

2.3. Classification of conformers

Many types of conformations were involved in 182,922 independent conformations which were found in a total of 10^6 samplings through the MC procedure. The generated structures were spread with a large root mean square deviation of 7.7 Å. In order to extract a specific structure from these various conformations, at the start, the conformers were classified into groups with characteristic types defined by the ϕ, ψ and χ^1 dihedral angles. The conformation of each residue was assigned to one of the nine types indicating α-helix ($-180 < \phi < 0$, $-120 < \psi < 0$), β-strand ($-180 < \phi < 0$, $0 < \psi < 240$) and the other, each having three probable rotamers about the χ^1 bond. As a result, the conformers were classified into 508 groups. In the next step, a minimum energy structure representing one group was compared with the other representatives, and when the root mean square distance between the

---

*Corresponding author. Fax: (81) (75) 753-4544.

SSDI 0014-5793(94)01217-2

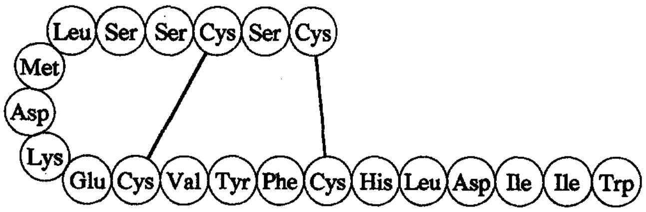

Fig. 1. Amino acid sequence of endothelin-1.

two structures was less than 2.5 Å, the two groups were gathered into one group. Through this process, the number of groups were reduced to 392.

All computations were performed on an IBM RS/6000 using the FORTRAN program written by the authors [14].

3. Results and discussion

The NMR studies of ET-1 in solutions, undertaken so far, could not determine the conformation of the C-terminus because of the high flexibility in solution [5–8]. Our present MC study also shows that the C-terminus does not have a greatly dominant conformation, but through the procedure of classification of the conformations, some specific features of the C-terminus chain have been revealed. Fig. 2 is a bar chart of the number of conformations vs. their minimum energy in the classified group. It shows that the energy variation is large, and the number of conformations in one group is small, the percentages being less than 1% for most groups. However, obviously there are a limited number of groups encompassing a comparably large number of conformations, which are in the middle range of the wide energy distribution in Fig. 2. The percentages, minimum energies, and conformation types of ten groups, including the large number of conformers, are listed in Table 1. Fig. 3 shows stereo views of four conformers representing the high percentage groups. Interestingly, these conformers have a similar folding form in the overall molecular view, and the conformations of the ten groups in Table 1 have the common features that the main chain turns at Cys^{15} and His^{16}, and the side chain of His^{16} has the same orientation. The conformers in the #1–#10 groups, corresponding to 18% of the whole conformers, have this particular structure, i.e. the C-terminus chain is approximately extended along the α-helix of the N-terminus core. The major conformations that the C-terminal tail interacts with the core of the head group are like Saudek et al. predicted [15]. Judging from the conformation types of the residues in Table 1, the major conformers do not have a characteristic secondary structure such as a helix form observed in the crystal structure [9], and the main chains take multiple conformations in a similar molecular form. The lowest energy conformation is not involved in the major conformations; the energy is lower than that of group #1 by 18.6 kcal/mol, and the conformation belongs to the group that includes a small number of conformations with a low percentage of 0.2%. A stereo

Table 1
The groups including large number of conformations

Group (#)	Percentage[a]	ΔE[b] (kcal/mol)	Conformation[c]					
1	3.3	0.00	S3	S2	S3	H3	O3	S/H2
2	2.5	−0.31	S3	S3	S3	H3	S3	S/H2
3	2.3	2.60	H3	O3	H2	S3	S1	S/H2
4	2.1	−11.15	H2	S3	S2	S1	H3	S/H2
5	2.1	−0.60	S3	S3	S3	S3	S1	S/H3
6	1.4	−2.40	S3	H3	S3	S1	S1	S/H3
7	1.2	7.34	S3	S3	S3	S1	S2	S/H3
8	1.1	5.43	S3	S2	S2	S1	S1	O3
9	1.1	3.11	S2	S3	S3	S3	H3	S/H3
10	1.1	8.09	S1	S3	S3	S1	S1	S/H3

[a] The percentage of the conformers involved in the group.
[b] The energy of the representative conformer.
[c] The letters represent the types of chain defined by the ϕ and ψ region for the six residues 16–21 of the C-terminus; H, α-helix form; S, β-strand form; O, other form. The numerical numbers distinguish the χ^1 rotamers; 1, 60 ± 60°; 2, 180 ± 60°; 3, −60 ± 60°.

Table 2
Hydrophobic contact residues between the N- and the C-terminus

Group (#)	N-Terminus	C-Terminus
1	Leu^{6},Val^{12}	Leu^{17}
	Lys^{9}	Trp^{21}
	Val^{12}	Ile^{19}
2	Leu^{6}, Val^{12}	Leu^{17}
	Leu^{6}, Val^{12}	Trp^{21}
	Lys^{9}	Ile^{20}
3	Leu^{6}, Val^{12}	Leu^{17}
	Val^{12}	Ile^{20}
	Leu^{6}, Met^{7}	Trp^{21}
4	Leu^{6}	Ile^{19}, Trp^{21}
	Met^{7}	Trp^{21}
	Val^{12}	Ile^{20}
5	Leu^{6}, Val^{12}	Leu^{17}
	Leu^{6}, Val^{12}	Ile^{20}
6	Val^{12}	Leu^{17}
	Tyr^{13}	Ile^{20}
7	Leu^{6}, Val^{12}	Leu^{17}
	Val^{12}	Trp^{21}
8	Met^{7}	Leu^{17}, Ile^{19}
	Lys^{9}	Trp^{20}
9	Met^{7}	Leu^{17}, Ile^{19}
10	Met^{7}	Ile^{19}, Ile^{20}
	Val^{12}	Trp^{21}

The alkyl or aromatic groups of the residues in both termini participate in the hydrophobic interactions.

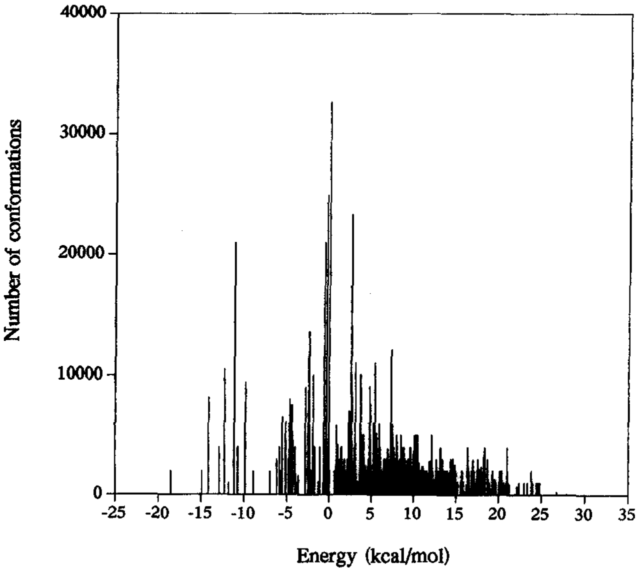

Fig. 2. Number of conformations vs. relative energy for the classified groups.

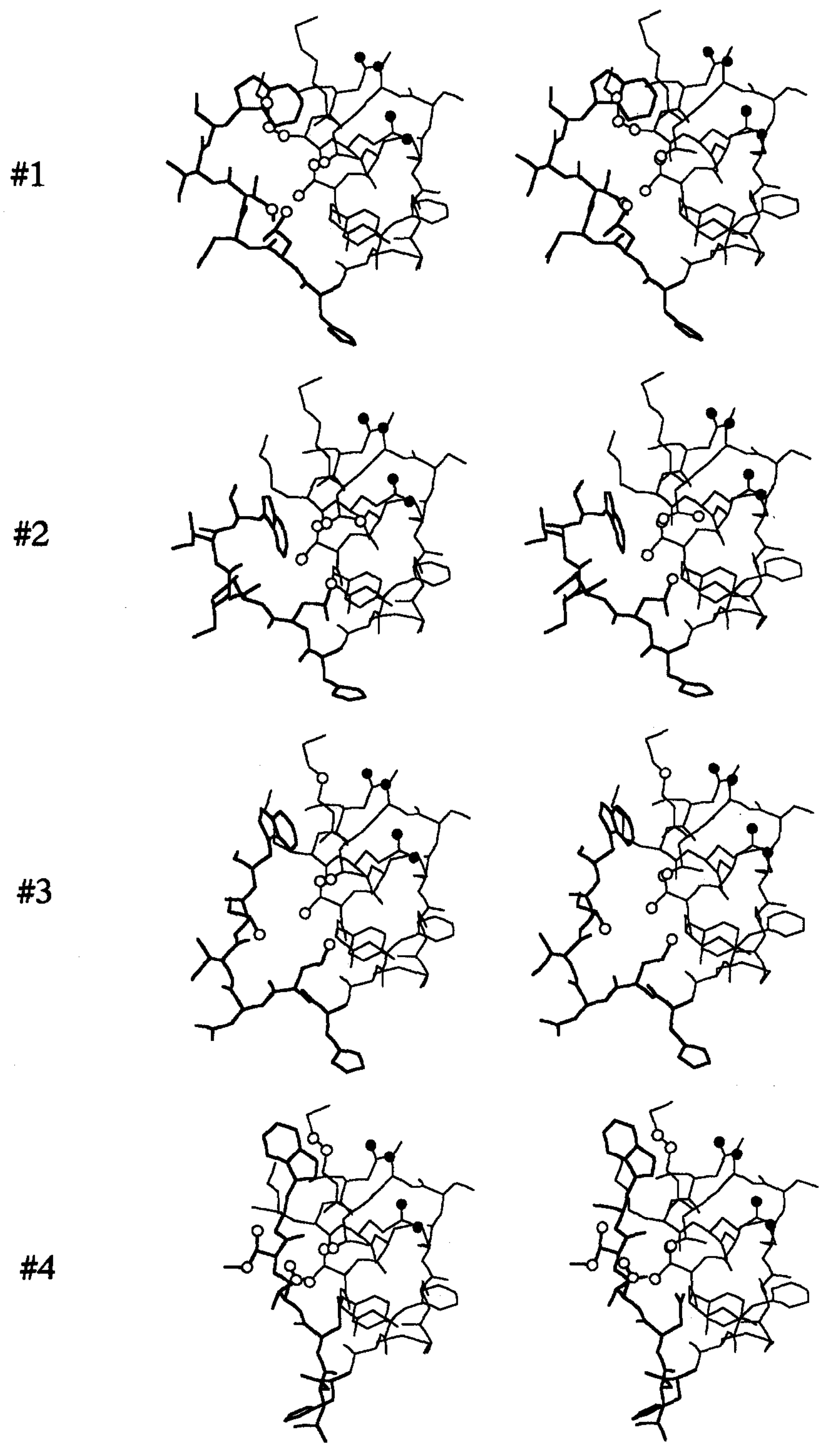

Fig. 3. Stereo views of four structures representing the major conformer groups. The C-terminus is shown in thick lines and the N-terminus in thin lines. The open circles indicate the alkyl carbon atoms forming hydrophobic clusters within the inter-atominc distance of 5Å. The aromatic ring of the C-terminus residue Trp[21] also participates in the clusters. The filled circles indicate the carboxyl oxygens of Asp[8] and Glu[10] which are functionally important residues.

view of the lowest energy structure is given in Fig. 4, in which the charged amino groups of the N-terminus is located near the deprotonated carboxyl groups of Trp[21] and Asp[18], their electrostatic interactions stabilizing the structure. An obvious difference between the major conformations and the lowest energy conformation is found in the ψ rotation of Cys[15]: the *trans* form for the major conformations and the *gauche* form for the lowest energy conformation. The notable peculiarity in the major conformations is that the side groups of the C-terminus form some hydrophobic clusters with the side groups of the N-terminal

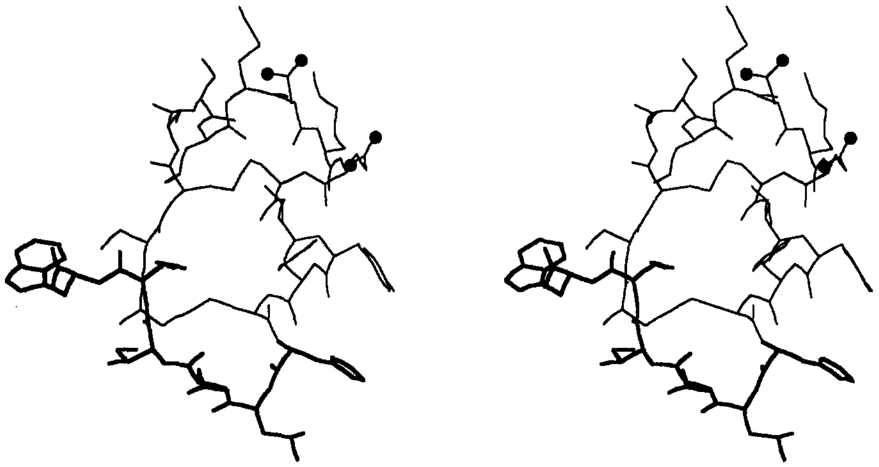

Fig. 4. A stereo view of the lowest energy structure. The C-terminus is shown in thick lines and the N-terminus in thin lines. The filled circles indicate the carboxyl oxygens of Asp^8 and Glu^{10} which are functionally important residues.

core, and that no hydrogen bonds are formed between the groups. Table 2 lists the residues having the hydrophobic contacts between the C-terminus and the head core. In most conformers, Leu^{17} of the C-terminus is in close proximity to Leu^6 and Val^{12} of the core, and in some cases, Ile^{19}, Ile^{20}, and Trp^{21} of the C-terminus locate near the alkyl side chains of Met^7 and Lys^9. Andersen et al. proposed that the C-terminus interconverts between at least two discrete conformations [8]. Such interconversion may occur between the conformations found here, accompanied by changing combinations of the hydrophobic contacts between the residues given in Table 2.

From the structure–function studies it has been reported that deletion of Trp^{21} decreases the constrictor response by three orders of magnitude [3], and that four residues, Asp^8, Glu^{10}, Tyr^{13}, and Phe^{14} in the core region are key residues for ET-1 binding to the receptor and exhibiting the biological function [16]. The functional importance of Trp^{21} may be related to the present specific conformations that the aromatic ring of Trp^{21} is located at a certain distance from the two charged side groups of Asp^8 and Glu^{10} (Fig. 3). As described above, the flexible C-terminus is extended along the core helix, and locates near the key residues. However, the functional groups of these residues are not covered with the C-terminus chain, and are exposed on the molecular surface. Such a conformation may be a desirable form for receptor recognition. Considering the flexibility of the C-terminus, the specific conformation would not be kept constantly under the physiological circumstance. However, the present study has led to the conclusion that at least the receptor-favorable conformations exist in a fairly high ratio in aqueous solution.

References

[1] Yanagisawa, M., Kurihara, H., Kimura, S., Tomobe, Y., Kobayashi, M., Mitsui, Y., Yazaki, Y., Goto, K. and Masaki, T. (1988) Nature 322, 411–415.

[2] Inoue, A., Yanagisawa, M., Kimura, S., Kasuya, Y., Miyauchi, T., Goto, K. and Masaki, T. (1989) Proc. Natl. Acad. Sci. USA 86, 2863–2867.

[3] Kimura, S., Kasuya, Y., Sawamura, T., Shinmi, O., Sugita, Y., Yanagisawa, M., Goto, K. and Masaki T. (1988) Biochem. Biophys. Res. Commun. 156, 1182–1186.

[4] Doherty, A.M., Cody, W.L., Leitz, N.L., DePue, P.L., Taylor, M.D., Rapundalo, S.T., Hingorani, G.P., Major, T.C., Panek, R.L. and Taylor, D.G. (1991) J. Cardiovasc. Pharmacol. 17, S59–S61

[5] Krystek, S.R., Jr., Bassolino, D.A., Novotny, J., Chen, C., Marshner, T.M. and Andersen, N.H. (1991) FEBS lett. 281, 212–218.

[6] Tamaoki, H., Kobayashi, Y., Nishimura, S., Ohkubo, T., Kyogoku, Y., Nakajima, K., Kumagaye, S., Kimura, T. and Sakakibara, S. (1991) Protein Eng. 4, 509–518.

[7] Reily, M.D. and Dunbar Jr., J.B. (1991) Biochem. Biophys. Res. Commun. 178, 570–577.

[8] Andersen, N.H., Chen, C., Marshner, T.M., Krystek Jr., S.R. and Bassolino, D.A. (1992) Biochemistry 31, 1280–1295.

[9] Janes, R.W., Peapus, D.H. and Wallace, B.A. (1994) Nature Struct. Biol. 1, 311–319.

[10] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) J. Chem. Phys., 21, 1087–1092.

[11] Momany, F.A., McGuire, R.F., Burgess, A.W. and Scheraga, H.A. (1975) J. Phys. Chem. 79, 2361–2381.

[12] Perrot, G., Cheng, B., Gibson, K.D., Vila, J., Palmer, K.A. Nayeem, A., Maigret, B. and Scheraga, H.A. (1992) J. Comp. Chem. 13, 1–11.

[13] Ooi, T., Oobatake, M., Nemethy, G. and Scheraga, H.A. (1987) Proc. Natl. Acad. Sci. USA 84, 3086–3090.

[14] Kuroda, M., Yamazaki, K., Kobayashi, T. and Taga, T. (1994) Bull. Chem. Soc. Jpn. 67, 648–652.

[15] Saudek, V., Hoflack, J. and Pelton, J.T. (1989) FEBS Lett. 257, 145–148.

[16] Nakajima, K., Kubo, S., Kumagaye, S., Nishino, H., Tsunemi, M., Inui, T., Kuroda, H., Chino, H., Watanabe, T., Kimura, T. and Sakakibara, S. (1989) Biochem. Biophys. Res. Commun. 163, 424–429.