A molecular model for human Big-Endothelin-1 (Big ET-1)

Heather Peto[a], Roger Corder[b], Robert W. Janes**[,a], B.A. Wallace[a,*]

[a]*Department of Crystallography, Birkbeck College, University of London, London, WC1E 7HX, UK*
[b]*The William Harvey Research Institute, St. Bartholomew's Medical College, Charterhouse Square, London, EC1M 6BQ, UK*

Received 5 August 1996

Abstract A molecular model has been developed for human Big Endothelin-1, which is the immediate precursor to the potent vasoconstrictor polypeptide endothelin-1 and the target of the highly specific endothelin converting enzyme. This model is produced by a threading algorithm protocol and is consistent with all the currently available structural and biochemical data for this molecule.

Key words: Threading algorithm; Endothelin converting enzyme; Crystal structure; Rational drug design; Endothelin precursor

1. Introduction

Human endothelin-1 (ET-1), the most potent vasoconstrictor yet identified, is a 21 amino acid disulphide-crosslinked polypeptide [1]; its biosynthesis involves a multistep process. The human ET-1 gene encodes for a 212 amino acid precursor, preproendothelin [2]. Removal of the signal sequence generates proendothelin-1 (195 amino acids). This is followed by both N-terminal and C-terminal cleavages, which remove 35 and 122 amino acids, respectively, to release an intermediate referred to as Big ET-1 (human form, 38 amino acids). The mature and active form of ET-1 is formed by an unusual but selective hydrolysis of the Trp-21–Val-22 bond of Big ET-1 by an endothelin converting enzyme (ECE) [3]. A number of structurally homologous isoforms of the endopeptidase called ECE-1 have been cloned [4–9]. They are derived from the same gene by differential splicing of mRNA transcripts [10].

Since Big ET-1 exhibits virtually no activity at any of the endothelin receptor sites, this means that molecules which can inhibit ECE, preventing the formation of the mature ET-1, are potential candidates for therapeutic drug design. To assist in the development of specific ECE inhibitors, a knowledge of the structure of Big ET-1 is necessary.

Although the 3-dimensional structure of the mature human ET-1 is known from crystallographic studies [11], there is little direct evidence for either the structure of Big ET-1 or for the structural requirements for specific processing of Big ET-1. There is, as yet, no X-ray structure available for Big ET-1. NMR [12,13] and circular dichroism (CD) [14] spectroscopic studies have concluded that the core residues of Big ET-1 must adopt similar conformations to those of the corresponding residues in ET-1. However, because the C-terminus of Big ET-1 appears to be relatively flexible, NMR studies have been unable provide any information on the structural nature of the C-terminal extension (CTE), i.e. residues 22–38, of Big ET-1. The CD studies have suggested qualitatively that the CTE contains some residues in helical and some residues in sheet conformations, with perhaps a single additional tight turn, but not which residues might be involved in such structures [14].

A number of biochemical experiments have shed some light on the nature of Big ET-1. Incorporation of bulky groups onto Lys-9 decreased markedly the vasopressor activity of Big ET-1 in vivo even though the same modifications to ET-1 had no effect on its hypertensive or in vitro vasoconstrictor activity [15]. The rate of conversion of these Lys-9 modified Big ET-1 molecules by the ECE activities of cultured smooth muscle or endothelial cells was substantially reduced. A similar decrease in the rate of hydrolysis occurred after reduction and *S*-carboxyamidomethylation of the polypeptide to generate a linear Big ET-1 molecule. These experiments led to the conclusion that a specific conformation of Big ET-1, potentially stabilized by an interaction between the C-terminal sequence and Lys-9 or an adjacent amino acid, as well as by the presence of one or both of the disulphide bonds, is important for the specific hydrolysis of the Trp-21–Val-22 bond [15].

In this paper, we describe the development of a model for Big ET-1 based on the crystal structure of ET-1 using a threading algorithm protocol. As will be shown, this model is consistent with all the currently available data on the structure and function of Big ET-1. Only one other model for Big ET-1 has appeared in the literature [16]. It was based on homology modeling with scorpion neurotoxin [17], and was proposed prior to the availability of the crystal structure of ET-1. However, since the ET-1 portion of that earlier Big ET-1 model [16] is unlike the now known structure of ET-1, it does not fulfil one of the structural requirements for the Big ET-1 molecule, and is thus unlikely to represent the structure of the actual molecule.

2. Materials and methods

2.1. Modeling procedure

The 212 residue sequence of human preproendothelin was threaded through all the known 3-dimensional structures of proteins held in the Brookhaven Protein Data Base up to January 1995 using the THREADER algorithm developed by Jones et al. [18–20]. The optimum aligned structure identified was 1gdh (d-glycerate dehydrogenase – apo form) [21]. The Epair value was −3.85, indicating a good alignment. Cys-53 to Val-74 of preproendothelin (equivalent to Cys-1 to Val-22 of Big ET-1) align with Lys-147 to Phe-168 of 1gdh; Asn-75 to Pro-77 (Asn-23 to Pro-25 in Big ET-1) form an insertion, and finally Glu-78 to Ser-90 (Glu-26 to Ser-38 in Big ET-1) align with Asp-169 to Ser-181 in 1gdh.

This procedure primarily provides a guide to the backbone conformation, without any requirement for corresponding sequence similarities. Flexibility observed in the threading procedure allowed replacement of the regular helix in 1gdh with the irregular helix (residues 10–

*Corresponding author. Fax: (44) (171) 631-6803.
E-mail: ubcg91c@ccs.bbk.ac.uk

**Present address:* Department of Biochemistry, Queen Mary and Westfield College, University of London, London E1 4NS, UK.

PII S0014-5793(96)00951-9

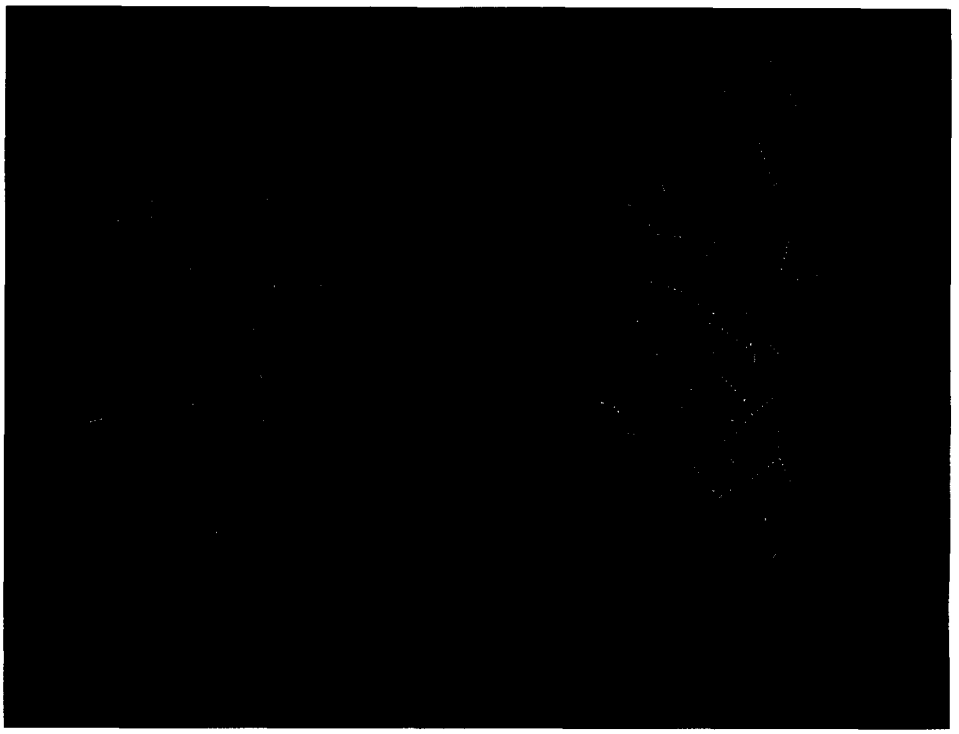

Fig. 1. Stereo view of the model of Big ET-1. The red coil traces the path of the polypeptide. Hydrogen bonds are shown in yellow.

20) of ET-1. Using the molecular modeling package SYBYL [22], ET-1 was fitted into the thread using its beta-strand as the target. The residues of 1gdh were mutated to correspond to the Big ET-1 sequence. The model was then minimized using the Kollman all atom force field and Gateiger-Marsili partial charge parameters in SYBYL. Dynamic analyses at 300 K indicated conformational flexibility of the region between Pro-30 and Ser-38. To examine the allowed conformations about the flexible Pro-30–Tyr-31 bond, a 360° psi torsion rotation of Tyr-31 at 5° intervals with minimization was undertaken, which indicated a strongly favourable electrostatic interaction when Ser-38 came into close proximity to Lys-9. The final model was the result of 1000 iterations of the minimization procedure.

2.2. Analyses of the model

The stereochemical quality of the model was evaluated using the PROCHECK software [23]. Analyses of the model including determinations of surface area, secondary structure, hydrogen bonds, etc. were performed using the CCP4 suite of programs (SERC Collaborative Computer Project N. 4, Daresbury Laboratory, UK). Graphics displays utilized PREPI [24] and SETOR software [25].

Fig. 2. Ribbon diagrams comparing the crystal structure of ET-1 [10] with the model for Big ET-1 (this work). The disulphide bonds are shown in yellow.

3. Results

3.1. Rationale for the method used

Big ET-1 is a relatively small polypeptide and hence we considered the conventional sequence homology modeling schemes used for larger proteins unlikely to be applicable in this case. Whilst we were able to find some fragments of structures in the Brookhaven Protein Data Base with modest homology to the CTE, there were none which exhibited sufficient overlap with the C-terminal part of ET-1 to provide information on how the CTE would attach onto and interact with the remainder of the molecule in three dimensions. Hence, we needed a method that would effectively model the whole of Big ET-1 rather than its parts, while at the same time being consistent with the known (ET-1) part of the structure.

Another consideration for the type of modeling utilized was that ET-1, and ultimately Big ET-1, are biosynthesized as a larger precursor, and it may be that the presence of the entire precursor protein is necessary to fold Big ET-1 efficiently into its correct structure. Therefore, we decided to test a threading algorithm [18] strategy to examine whether we could find a favourable fold for the whole precursor molecule, and then extract the folded Big ET-1 structure from the context of the larger molecule.

Table 1
Surface accessibilities of ET-1 and Big ET-1 structures (expressed as fractions of the maximal surface accessible area for each type of residue)

Residue	ET-1	Big ET-1
Cys-1	0.89	0.45
Ser-2	0.96	0.74
Cys-3	0.59	0.03
Ser-4	0.19	0.23
Ser-5	0.79	0.01
Leu-6	0.95	0.73
Met-7	0.62	0.80
Asp-8	0.89	0.47
Lys-9	0.84	0.45
Glu-10	0.52	0.42
Cys-11	0.02	0.04
Val-12	0.26	0.38
Tyr-13	0.84	0.83
Phe-14	0.42	0.41
Cys-15	0.10	0.02
His-16	0.52	0.45
Leu-17	0.51	0.44
Asp-18	0.86	0.39
Ile-19	0.68	0.31
Ile-20	0.55	0.59
Trp-21	0.93	0.61
Val-22		0.65
Asn-23		0.84
Thr-24		0.37
Pro-25		0.67
Glu-26		0.34
His-27		0.63
Val-28		0.60
Val-29		0.55
Pro-30		0.42
Tyr-31		0.91
Gly-32		0.68
Leu-33		0.55
Gly-34		0.27
Ser-35		0.70
Pro-36		0.90
Arg-37		1.00
Ser-38		0.67

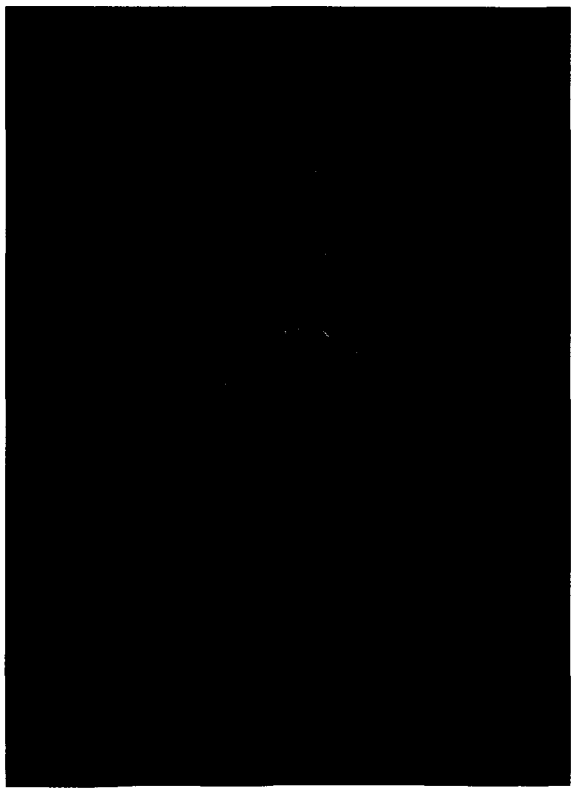

Fig. 3. Model of Big ET-1 showing the residues adjacent to the bonds cleaved by proteases (21–22 and 31–32) in red and those adjacent to bonds cleaved in ET-1 but not Big ET-1 (5–6, 16–17, and 18–19) in green.

An essential criterion for success was that the secondary structures of the ET-1 portion of the aligned structure must contain the same types of secondary structures found in the ET-1 crystal structure. The threading strategy was successful in that it found an optimum aligned structure with a low Epair and in which, most importantly, the helix and beta-strand alignments matched the equivalent structures in ET-1. The final model produced is shown in Fig. 1.

3.2. General description of the model

ET-1 consists of an N-terminal beta-strand linked by a loop region to a long irregular helix that extends from residue 10 to residue 20 near the C-terminal end of the molecule [11]. The molecule is cross-linked via two disulphide bonds between residues 1 and 15 and 3 and 11, respectively (Fig. 2). The structure of the endothelin core in the Big ET-1 model is essentially the same as in the ET-1 crystal structure, with the exception that the C-terminal helix is extended by 3 residues to Asn-23 in Big ET-1. Additionally, the CTE contains a hydrogen-bonded loop (between residues 18 and 24) that causes the chain to fold back on itself, creating a parallel beta-sheet that is hydrogen-bonded between residues 1 and 5 of the first strand and between residues 27 and 29 of the second strand, thereby producing a relatively compact structure. Finally, amino acids 35 and 38 in the C-terminus form a hydrogen-bonded tight turn which brings the side chain of

residue Ser-38 back to form a hydrogen-bonded contact with the side chain of residue Lys-9. Consequently, the Big ET-1 model is a globular structure.

The surface accessibilities (Table 1) of residues 1–5, which form one strand of the beta-sheet present in Big ET-1, are greatly reduced relative to their counterparts in ET-1. The surface accessibilities of residues at the beginning of the helical region (Glu-10–Cys-15) are essentially unchanged in Big ET-1 relative to ET-1, while those residues that form the C-terminal end of ET-1 are less accessible in Big ET-1 due to the presence of the loop in the CTE which folds back on residues 16–21 in Big ET-1. One other notable difference in Big ET-1 is the reduced accessibility of residue 9, whose side chain hydrogen bonds with the side chain of residue 38 in the CTE.

4. Discussion

The most important test of any molecular model is whether it is consistent with all the experimental data, both structural and biochemical, available for that molecule.

NMR studies [12,13] have shown that the conformations of residues 1–16 of Big ET-1 are very similar to those of the same residues in ET-1. However, the NMR studies have been unable to define accurately a structure for the part of the Big ET-1 molecule beyond residue 16, so the nature of the all-important C-terminal extension remains unknown. Thus, as our modeling only effectively changed the conformation of the last two residues (20 and 21) in ET-1, it is entirely consistent with the available NMR data.

Like the NMR studies, the only CD spectroscopic data for Big ET-1 also suggested that its core residues have very similar conformations to those in ET-1 [14]. That study concluded that it was difficult to quantitate the secondary structure for such a small flexible polypeptide, but estimated that the C-terminal extension would contain some additional residues in helical and sheet conformations and an additional turn-type structure. Residues 21–23 of the Big ET-1 model fulfil the additional helical content requirement, and residues 27–29 provide the additional sheet component, with the hydrogen-bonded turn between residues 35 and 38 producing the new turn component. Thus, this model is compatible with all the (albeit rather qualitative) information available on the secondary structure of Big ET-1.

While the spectroscopic studies have provided only a limited amount of structural information against which the model can be tested, there is some very specific experimental information available on the Big ET-1 structure from proteolysis studies, which has important implications for the structure. Firstly, the site of cleavage by ECE is between residues 21 and 22, so it is expected that these residues should have a high degree of surface accessibility to enable interaction with the protease. ECE is highly specific for the Trp–Val bond, so the side chains for those residues should be surface exposed. Indeed, these residues are more than 50% surface accessible, and the peptide bond itself is exposed in the Big ET-1 molecule, so cleavage by ECE would be expected (Fig. 3). Cleavage of Big ET-1 between bonds 31 and 32 by several other proteases has also been reported [26–28], so it is anticipated that these residues should also exhibit significant surface exposure. In the model, they do indeed have a surface accessibility of greater than 50%, again consistent with this being a potential cleavage site.

Lack of cleavage can also provide information on inaccessible sites in a molecule which can be useful in evaluating a structure, especially if there is evidence for cleavage of a related molecule by the same enzyme (Fig. 3). Bonds that are cleaved by various proteases in ET-1 but not Big ET-1, include Ser-5–Leu-6, His-16–Leu-17, and Asp-18–Ile-19 [29–32]. In all these cases, at least one residue of the pair is less than 50% exposed in Big ET-1, and also in all cases, there is a notable reduction in accessibilities from those in ET-1 (where all residues are $>50\%$ exposed). This is precisely the type of structure that would be expected in a correct model of Big ET-1.

Likewise, because Big ET-1 exhibits essentially no activity at any of the endothelin receptors, one would anticipate that its structure should reflect a decrease in accessibility relative to ET-1 for at least one of the residues primarily involved in the interactions with its receptors. A wide range of structure-activity relationship studies point to the residues on one face of the helix, notably 10, 14, 18, and 21 as being involved in the binding interaction [33]. Whilst residues 10 and 14 are only slightly altered in their accessibilities in Big ET-1, residues 18 and 21 have substantially reduced accessibilities, indicating the model is also consistent with these data.

The studies which showed that chemical modification of Lys-9 in Big ET-1 decreased the vasopressor activity in vivo, whilst the same modifications to ET-1 did not affect its activity, suggest that the conformation of Big ET-1 is altered by this modification in a way that makes it less susceptible to proteolysis by ECE [15]. This would be expected for the model proposed since Lys-9 is involved in a conformational-stabilizing side chain hydrogen bond that could not be present in a Lys-9-modified molecule. Hence, while the Lys-9 side chain is not apparently essential for interactions or activity of the mature ET-1, it does appear necessary for maintaining a proteolytically competent form of Big ET-1.

Finally, it has been observed that the cross-reactivity of Big ET-1 to antibodies produced against residues 1–15 of ET-1 is decreased relative to ET-1, suggesting that Big ET-1 must be a globular molecule with a CTE that blocks antibody accessibility to some of these amino terminal residues [15]. That is certainly true for the present model, where residues 1–5 and 8 and 9 are considerably more buried in the Big ET-1 structure than in the ET-1 structure.

In summary, we have developed a model for Big ET-1, the immediate precursor to ET-1, which is normally cleaved in vivo by the endothelin converting enzyme. While there is little direct structural information available on Big ET-1, there is a large amount of biochemical data against which any model of Big ET-1 must be tested. The model we have presented in this paper is stereochemically reasonable and consistent with all the structural and biochemical information currently available on this molecule. While a model such as this will obviously not be fully correct in all its details, it may serve a useful purpose for the development of further experiments and the design of new molecules which may act as inhibitors of ECE and thus provide the basis for useful therapeutic developments.

Acknowledgements: We thank R.T. Miller of the Biomolecular Structure Group at UCL for help in running the THREADER program. This work was supported by British Heart Foundation Grant 95009 (to B.A.W.).

References

[1] Yanigisawa, M., Kurihara, H., Kimura, S., Tomobe, Y., Kobayashi, M., Mitsui, Y., Yazaki, Y., Goto, K. and Masaki, T. (1988) Nature 332, 411–415.
[2] Bloch, K.D., Friedrich, S.P., Lee, M.E., Eddy, R.L., Show, T.B. and Quertermous, T. (1989) J. Biol. Chem. 264, 10851–10857.
[3] Turner, A.J. and Murphy, L.J. (1995) Biochem. Pharmacol. 51, 91–102.
[4] Shimada, K., Takahashi, M. and Tanzawa, K. (1994) J. Biol. Chem. 269, 18275–18278.
[5] Xu, D., Emoto, N., Giaid, A. Slaughter, C., Kaw, S., deWit, D. and Yanagisawa, M. (1994) Cell 78, 473–485.
[6] Ikura, T., Sawamura, T., Shiraki, T., Hosokawa, H., Kido, T., Hoshikawa, H., Shimada, K., Tanzawa, K., Kobayashi, S., Miwa, S. and Masaki, T. (1994) Biochem. Biophys. Res. Commun. 203, 1417–1422.
[7] Schmidt, M., Kroger, B., Jacob, E., Seulberger, H., Subkowski, T., Otter, R., Meyer, T., Schmalzing, G. and Hillen, H. (1994) FEBS Lett. 356, 238–243.
[8] Shimada, K., Matsushita, Y., Wakabayashi, K., Takahashi, M., Matsubara, A., Iijima, Y. and Tanzawa, K. (1995) Biochem. Biophys. Res. Commun. 207, 807–812.
[9] Yorimitsu, K., Moroi, K., Inagaki, N., Saito, T., Masuda, Y., Masaki, T., Seino, S. and Kimura, S. (1995) Biochem. Biophys. Res. Commun. 208, 721–727.
[10] Valdenaire, O., Rohrbacher, E. and Mattei, M.-G. (1995) J. Biol. Chem. 270, 29794–29798.
[11] Janes, R.W., Peapus, D.H. and Wallace, B.A. (1994) Nat. Struct. Biol. 1, 311–319.
[12] Donlan, M.L., Brown, F.K. and Jeffs, P.W. (1992) J. Biomol. NMR 2, 407–420.
[13] Inooka, H., Endo, S. Kikuchi, T., Wakimasu, M., Mizuta, E. and Fujino, M. (1990) Peptide Chem. 28, 409–414.
[14] Wallace, B.A. and Corder, R. submitted.
[15] Corder, R. (1996) Biochem. Pharmacol. 51, 259–266.
[16] Menziani, M.C., Cocchi, M., DeBenedetti, P.G., Gilbert, R.G., Richards, W.G., Zamai, M. and Caiolfa, V.R. (1991) J. Chim. Phys. Phys.-Chim. Biol. 88, 2687–2694.
[17] Almassy, R.J., Fontecilla-Camps, J.C., Suddath, F.L. and Bugg, C.E. (1983) J. Mol. Biol. 170, 497–527.
[18] Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992) Nature 358, 86–89.
[19] Jones, D.T., Miller, R.T. and Thornton, J.M. (1995) Proteins: Struct. Funct. Genet. 23, 387–397.
[20] Miller, R.T., Jones, D.T. and Thornton, J.M. (1996) FASEB J. 10, 171–178.
[21] Goldberg, J.D., Yoshida, T. and Brick, P. (1994) J. Mol. Biol. 236, 1123–1140.
[22] Sybyl, v. 6.1, Tripos Associates, St. Louis, MO, USA.
[23] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) J. App. Crystallogr. 26, 283–291.
[24] Islam, S.A. and Sternberg, M.J.E. (1991) PREPI Manual, Imperial Cancer Research Fund, London, UK.
[25] Evans, S.V. (1993) J. Mol. Graphics 11, 134.
[26] McMahon, E.G., Fok, K.F., Moore, W.M., Smith, C.E., Siegel, N.R. and Trapani, A.J. (1989) Biochem. Biophys. Res. Commun. 161, 406–413.
[27] Kaw, S., Hecker, M. and Vane, J.R. (1992) Proc. Natl. Acad. Sci. USA 89, 6886–6890.
[28] Wypij, D.M., Nichols, J.S., Novak, P.J., Stacy, D.L., Berman, J. and Wiseman, J.S. (1992) Biochem. Pharmacol. 43, 845–853.
[29] Sokolovsky, M., Galron, R., Kloog, Y., Bdolah, A., Indig, F.E., Blumberg, S. and Fleminger, G. (1990) Proc. Natl. Acad. Sci. USA 87, 4702–4706.
[30] Vijayaraghavan, J., Scicli, A.G., Carretero, O.A., Slaughter, C., Moomaw, C. and Hersh, L.B. (1990) J. Biol. Chem. 265, 14150–14155.
[31] Fagny, C., Michel, A., Leonard, I., Berkenboom, G., Fontaine, J. and Deschodt-Lanckman, M. (1991) Peptides 12, 773–338.
[32] Murphy, L.J., Corder, R., Mallet, A.I. and Turner, A.J. (1994) Br. J. Pharmacol. 113, 137–142.
[33] Wallace, B.A. and Janes, R.W. (1995) J. Cardiovasc. Pharmacol. 26, S250–S253.