

# Comparison of the structures of the endothelin A receptor antagonists BQ123 and *N*-methyl leucine BQ123 with the crystal structure of the C-terminal tail of endothelin-1

Catherine E. Peishoff<sup>a</sup>, Robert W. Janes<sup>b</sup>, B.A. Wallace<sup>b,\*</sup>

<sup>a</sup>Department of Physical and Structural Chemistry, SmithKline Beecham Pharmaceuticals, King of Prussia, PA 19406, USA

<sup>b</sup>Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, UK

Received 1 September 1995

**Abstract** The functionally important regions of the cyclic pentapeptide endothelin A receptor antagonist BQ123 are shown to correlate with the structure of the C-terminal tail of endothelin-1, as found in the recently-determined X-ray crystal structure. Residues 18 and 21 of endothelin-1 are spatially juxtaposed such that they superpose extremely well with *D*-Asp and *D*-Trp of the antagonist, consistent with the residues on this surface of the endothelin helix being important for binding. This study provides new information on the three-dimensional nature of the endothelin A receptor binding site which may prove useful for rational drug design.

**Key words:** Endothelin; X-Ray crystallography; Three-dimensional structure; Drug design; Molecular modeling; G-Protein coupled receptors

## 1. Introduction

Endothelin (ET) is a 21-amino acid polypeptide (Fig. 1), which is the most potent natural vasoconstrictor yet identified [1]. It binds to G-protein coupled receptors in a wide range of tissue types and has been linked with a number of diseases, including hypertension, myocardial infarction, atherosclerosis and renal failure. As such, it has been a target for drug discovery and considerable effort has been expended to develop small molecule analogues of it which may prove to be effective antagonists [2].

To rationally design such an antagonist, one would optimally have the three-dimensional structure of both the receptor and its native ligand. To date, however, there is no detailed structural information available on the ET receptors nor, indeed, on any of the G-protein coupled receptors. There is, however, a wealth of structure/activity relationship (SAR) data on the native ligand and on many synthetic analogues of ET. In addition there exists detailed three-dimensional structural information on human ET and a number of its analogues and on small molecule antagonists. These data, provided they can be rationalized in terms of common structural features, should be useful in developing an image of the receptor binding pocket features which are crucial for ligand binding. In turn, this would provide essential data for the rational design of new antagonists.

A number of structures have been proposed for endothelin:

several from NMR studies [3–14] and, recently, one from X-ray crystallography [15]. The NMR structures have been from organic solutions or aqueous/organic solvent mixtures; the X-ray structure was from a purely aqueous solution. The NMR and X-ray structures differ primarily in the conformation of the C-terminal part of the molecule [16]. The C-terminus is essentially a disordered structure in all of the NMR models, but forms an irregular, but well-ordered helix in the crystal structure. The presence of such a helix is also consistent with fluorescence studies of the molecule in aqueous solutions [17]. Thus, the question has arisen as to which of the NMR or X-ray structures can be most useful in providing information on the bioactive conformation of the molecule. As the C-terminal residues have been identified as being crucial for receptor binding, the importance of the detailed structure of this region of the molecule is obvious. One approach to assessing whether the structure is a bioactive conformation is to compare it with an analogue molecule which is conformationally restricted and determine if there is any sensible overlap of crucial chemical moieties between the two.

In 1991, a cyclic pentapeptide isolated from a fermentation broth of *Streptomyces misakiensis* was found to selectively inhibit the binding of ET-1 to the endothelin A receptor subtype and to functionally antagonize ET-1 induced vasoconstriction [18]. This molecule, cyclo(*D*-Trp-*D*-Glu-*L*-Ala-*D*-Alloile-*L*-Leu), has served as the lead structure from which a number of more potent analogues have been derived, most notably BQ123, cyclo(*D*-Trp-*D*-Asp-*L*-Pro-*D*-Val-*L*-Leu), a potent and highly selective endothelin receptor antagonist (Fig. 2) [19,20]. While the sequence of BQ123 does not match any part of the ET-1 sequence, it contains features similar to those at the carboxy-terminal end, e.g. tryptophan preceded by two hydrophobic residues and an aspartic acid, and thus is generally thought to mimic this region [20].

While no X-ray structure of BQ123 exists, NMR studies of this molecule and analogues in a variety of solvents and solvent mixtures [22–26] show a remarkably consistent structure for the polypeptide backbone (which is defined by a type II  $\beta$ -turn containing the leucine and *D*-tryptophan at positions  $i + 1$  and  $i + 2$ , respectively, and an inverse  $\gamma$ -turn at the proline). The side chain positions display strong preferences in hydrophilic environments; in particular, a close contact is observed between leucine and *D*-tryptophan. However, in the non-polar solvent CDCl<sub>3</sub>, no such preference is observed for the equipotent analogue *N*-methyl-leucine BQ123 [25]. The collective results suggest a strongly preferred backbone structure with environmentally-dependent sidechain positions.

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

The goal of this study was to determine if there is any correspondence between the three-dimensional structures of BQ123 and its analogues and the C-terminal tail of endothelin-1, particularly between residues in both molecules known for being important either for binding or function. It was anticipated that if a good correlation could be found, then this might give an idea of the region of the structure that was involved in the binding site and insight into the geometric nature of the receptor binding pocket, and could thus be useful in the rational design of new antagonists. Further it might provide evidence as to which of the many structures of the ET C-terminal tail was relevant to the endothelin A receptor-bound conformation.

## 2. Materials and methods

The X-ray structure of human ET in crystals prepared from an aqueous solution has been previously published [15], and the coordinates deposited in the Brookhaven Protein Data Bank (acquisition code 1EDN). The structure of *N*-methyl-leucine BQ123 is a low energy structure determined by NMR spectroscopy from solution in CDCl<sub>3</sub> [25]. *N*-Methyl-leucine BQ123 is a highly active and specific antagonist [27] and was chosen for this comparative study because its methyl group imposes additional steric constraints, which must be allowed for in any binding site model. The residues numbers for BQ123 are denoted with *a'* to distinguish them from the residue numbers in ET.

Alignments were done graphically to achieve maximal superimposition of the residue pairs D-Asp<sup>2</sup>/Asp<sup>18</sup>, Lnm<sup>5</sup>/Ile<sup>20</sup>, and D-Trp<sup>1</sup>/Trp<sup>21</sup>, respectively, while maintaining good backbone correspondence. The structure of ET-1 was held rigid as was the backbone conformation of the peptide antagonist. Only the sidechains of the antagonist were allowed to rotate to achieve the overlay. The residue pairs were elected on the basis of SAR data highlighting the Asp [28] and Trp [29] residues in the peptide antagonists as important to binding affinity, and Asp<sup>18</sup>, Ile<sup>20</sup>, and Trp<sup>21</sup> in ET-1 as important to either binding or function [30–32]. Many other potential overlaps were examined, but none produced as strong a functional correlation as the one discussed here.

Fig. 3 was produced using the software SETOR [33]. The *N*-methyl-leucine residue is designated Lnm.

## 3. Results

On first inspection, the composition of *N*-methyl-leucine BQ123 and endothelin appear very different: not only is the small molecule cyclic in nature, as opposed to the linear C-terminus of the native peptide, it also includes three amino acids with different (D-) chirality, and two conformationally-challenged imino acids (proline and *N*-methyl-leucine) not present in ET. Despite this, their three-dimensional structures exhibit very similar features. In the superposition in Fig. 3, it can be seen that the position of the carboxylate group of Asp<sup>18</sup> in ET corresponds closely to that of D-Asp<sup>2</sup> in BQ123, whilst the sidechains of the pairs Ile<sup>20</sup>/Lnm<sup>5</sup> and Trp<sup>21</sup>/D-Trp<sup>1</sup> also overlap well (note: in the figure, the  $\beta$ - $\gamma$  bond of Asp<sup>18</sup> is rotated to clarify the correspondence between the carboxylate groups).

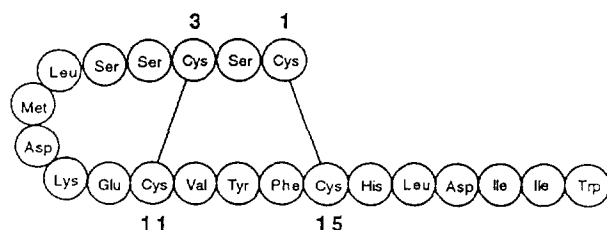


Fig. 1. Amino acid sequence of human endothelin-1.

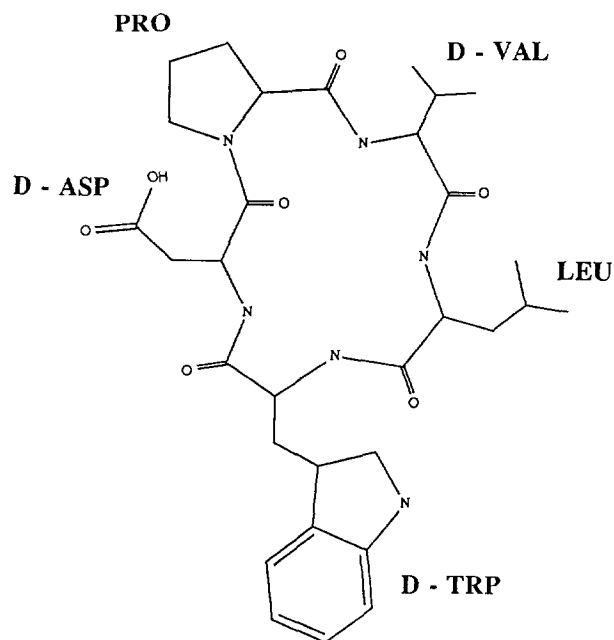


Fig. 2. Chemical structure of the endothelin receptor antagonist BQ123.

With this correspondence, several antagonist backbone elements can also be aligned, notably the carbonyl groups of Val<sup>3</sup> and Lnm<sup>5</sup> to those of Ile<sup>19</sup> and Ile<sup>20</sup> of ET-1, respectively. This may be, in part, because the ET helix, as stated earlier, is irregular, with some distorted  $\phi$ ,  $\psi$  angles, and not all the expected hydrogen-bonds formed. Had the helix been more 'perfect' the overlap with the small molecule may not have been as good. The overlap between the structures of BQ123 and the C-terminal five amino acids of ET is excellent as found in the crystal structure, but the same cannot be said for any of the NMR structures, because they do not contain helical C-termini.

The side chains that are believed to be important for binding based on SAR studies, notably residues 18 and 21, overlap very well even though the backbone conformations of the two molecules are different, and hence their overlap not particularly good. The backbones of the two molecules veer off in different directions in the region of the molecule between residues 18 and 21. This allows both the D-amino acid sidechains of Asp<sup>2</sup> and Trp<sup>1</sup> in BQ123 to line up with the L-amino acids of Asp<sup>18</sup> and Trp<sup>21</sup> in ET, whilst the L-Ile<sup>20</sup> sidechain can still point in the same direction as L-Lnm<sup>5</sup>. The consequence of this is that there is not a one-to-one correspondence of the Pro<sup>3</sup> and Val<sup>4</sup> residues of BQ123 with residues in ET. Similarly, there is no obvious counterpart in ET for the *N*-methyl group in *N*-methyl-leucine. The surface features of the two molecules, which are expected to be those recognized by the receptor, are very similar despite the differences in the buried parts of the molecules.

Other superpositions (especially when only Ca atoms are used in alignment algorithms) produce better correspondences of the backbone, but in those cases, none of the sidechains overlap, and the surface contours are entirely different.

It should be noted that obtaining the superposition of these molecules only required the rotation of the side chains in the smaller molecule. It was not necessary to invoke molecular

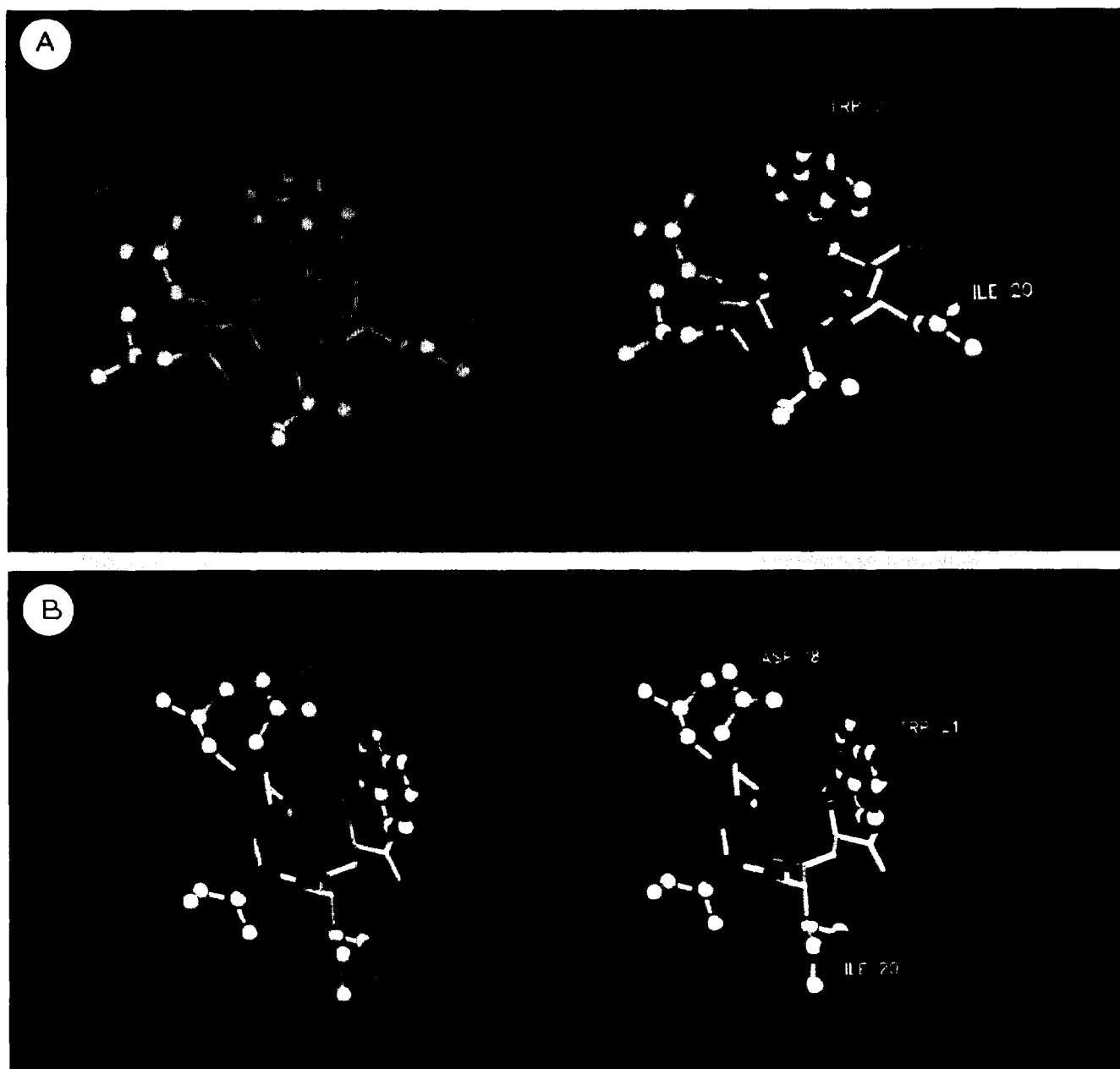


Fig. 3. Two stereo-views showing the overlay of *N*-methyl-leucine BQ123 (in green) and residues 17–21 of human endothelin-1 (as found in the crystal structure [15]) (in grey).

dynamics simulations for either peptide. In fact, no alteration of the ET-1 X-ray coordinates was required.

The advantage of dealing with the structure of a small molecule in drug design is that it may be much more conformationally restrained than is the native peptide, so there is a limit to the range of possible structures that it may adopt in order to fit into a receptor binding site. In the case of BQ123, the constraints are imposed by its small size, the presence of the conformationally restricting amino acid proline, the stereochemical variations at the *Ca* positions, and the cyclic nature of the peptide. As a consequence, all the NMR studies have produced very similar structures for this small molecule. Thus there is little freedom in the choice of possible structures to be used for the comparisons.

#### 4. Discussion

Three naturally-occurring isoforms of endothelin have been identified, ET-1, ET-2, ET-3, which are highly homologous, differing only in the nature of 2–5 amino acids in the N-terminal part of the molecule. The C-terminal seven amino acids are identical in all isoforms, and amino acid substitutions as well as alanine scan studies have indicated the importance of the C-terminal amino acids in receptor binding. SAR studies on peptide and non-peptide small molecule antagonists have pointed to the importance of a juxtaposition of an acidic group (presumably Asp<sup>18</sup> in the native endothelin) and at least one aromatic group (in native ET this would likely be Trp<sup>21</sup>). Thus, it was postulated that these types of side chains should be in

similar relative positions in both the small molecule and in the native peptide [34]. In BQ123 the corresponding groups would appear to be D-Asp<sup>2</sup> and D-Trp<sup>1</sup>. Although there are a number of possible acidic/aromatic combinations in ET-1, SAR studies for both molecules support a correspondence of BQ123 and its analogues to the C-terminal tail of ET-1.

The overlay studies reported here would suggest that the endothelin A receptor binding site has a geometry that would accommodate the tryptophan side chain as well as a closely juxtaposed carboxylate from the aspartic acid sidechain, but that the residues on the opposite side of the molecule (i.e. residues 17 and 19 of ET-1 and Pro<sup>3</sup> and Val<sup>4</sup> of BQ123) are not the primary sources of contact with the receptor.

The endothelin A receptor (ET<sub>A</sub>) is selective in its binding of the endothelin isoforms such that ET-1 = ET-2 >> ET-3; the endothelin B receptor (ET<sub>B</sub>) binds all isoforms with nearly equal affinity. There is an apparent paradox in suggesting that BQ123, an ET<sub>A</sub> selective antagonist, overlays a region of the endothelins found through analogue studies to be important for the binding of both ET<sub>A</sub> and ET<sub>B</sub>. Recent mutagenesis data suggest that a residue located in transmembrane segment 2 of ET<sub>A</sub> affects the selectivity of ET<sub>A</sub> for agonist peptides as well as its affinity for BQ123. However, modification of the corresponding residue in ET<sub>B</sub> does not alter the non-selective profile of this receptor subtype, nor does it confer affinity for BQ123. Selectivity, therefore, does not appear to arise in this system from an easily identified set of interactions. If, however, as this study suggests, BQ123 and its analogues overlay the part of the isoforms that have a common sequence, then the differences in isoform binding affinities may arise from the regions where their sequences differ, notably residue 2 and the region from residue 4 to residue 7. The far lower binding affinity of ET-3 for the endothelin A receptor may well be caused by steric hindrances that lie outside the binding site indicated in this study.

Even though it is now widely believed that BQ123 mimics the C-terminus of ET [21], earlier studies by Satoh and Barlow [35] which had compared model structures for BQ123 with the NMR structure of endothelin, suggested an alternative correspondence might exist between BQ123 and residues 6–8 of the ET polypeptide. However, that work was done prior to the availability of the ET crystal structure, the NMR structures of BQ123, and the wealth of SAR data now in hand. The conformation of the central turn region of ET which was their proposed region of correspondence, now appears in both the NMR and crystal structures to be somewhat different from the one used in their comparison study. In addition, their overlap with the N-terminus of endothelin does not take into account the crucial nature of the C-terminal residues.

Future work will entail comparisons of the ET structure with other synthetic peptide and non-peptide antagonists for the endothelin receptors, and in this way build up a more complete view of the nature of the binding pocket. Initial attempts to design small molecule antagonists from the BQ peptide series alone has produced compounds with only moderate affinity [29], indicating the need for a more complete description of the cyclic pentapeptide binding site. Our comparison of BQ123 and the C-terminus of ET has provided important information on the nature of the binding site geometric requirements, and given support to the notion that the crystal structure of ET may represent a bioactive receptor-bound conformation. In addition,

it emphasizes the notion previously suggested [15,36] that one surface of the ET C-terminal helix, which includes residues 18 and 21 and arises from consecutive turns of the helix, is very important for receptor binding.

*Acknowledgements:* This work was supported, in part, by a grant from the British Heart Foundation (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] Elliott, J.E., Lago, M.A. and Peishoff, C.E. (1995) in: *Endothelin Receptors from the Gene to the Human* (Ruffolo Jr., R.R. ed.) pp. 79–107, CRC Press, Boca Raton.
- [3] Endo, S., Inooka, H., Ishibashi, Y., Kitada, C., Mizuta, O. and Fujino, M. (1989) *FEBS Lett.* 257, 149–154.
- [4] Saudek, V., Hoflack, J. and Pelton, J.T. (1989) *FEBS Lett.* 257, 145–148.
- [5] Krystek Jr., S.R., Bassolino, D.A., Novotny, J., Chen, C., Marschner, T.M. and Andersen, N.H. (1991) *FEBS Lett.* 281, 212–218.
- [6] Munro, S., Craik, D.J., McConville, D., Hall, J.G., Searle, M., Bicknell, W., Scanlon, D. and Chandler, C. (1991) *FEBS Lett.* 278, 9–13.
- [7] Aumelas, A., Chiche, L., Mahe, E., L-Nguyen, D., Sizun, P., Berthault, P. and Perly, B. (1991) *Int. J. Peptide Protein Res.* 37, 315–324.
- [8] Reily, M.D. and Dunbar Jr., J.B. (1991) *Biochem. Biophys. Res. Commun.* 178, 570–577.
- [9] Saudek, V., Hoflack, J. and Pelton, J.T. (1991) *Int. J. Peptide Protein Res.* 37, 174–179.
- [10] Tamaoki, H., Kobayashi, Y., Nishimura, S., Ohkubo, T., Kyogaku, Y., Nakajima, K., Kumagaya, S., Kimura, T. and Sakakibara, S. (1991) *Protein Eng.* 4, 509–518.
- [11] Dalgarno, D.C., Slater, L., Chackalamannil, S. and Senior, M.M. (1992) *Int. J. Peptide Protein Res.* 40, 515–523.
- [12] Andersen, N.H., Chen, C., Marschner, T.M., Krystek Jr., S.R. and Bassolino, D. (1992) *Biochemistry* 31, 1280–1295.
- [13] Coles, M., Munro, S.L.A. and Craik, D.J. (1994) *J. Med. Chem.* 37, 656–664.
- [14] Andersen, N.H., Chen, C. and Lee, G.M. (1995) *Protein Peptide Lett.* 1, 215–222.
- [15] Janes, R.W., Peapus, D.H. and Wallace, B.A. (1994) *Nature Struct. Biol.* 1, 311–319.
- [16] Wallace, B.A., Janes, R.W., Bassolino, D.A. and Krystek, S.R. (1995) *Protein Sci.* 4, 75–83.
- [17] Cowley, D.J. and Pelton, J.T. (1995) *Int. J. Peptide Protein Res.* 46, 56–64.
- [18] Ihara, M., Fukuroda, T., Saeki, T., Masuru, N., Kojiri, K., Suda, H. and Yano, M. (1991) *Biochem. Biophys. Res. Commun.* 178, 132–137.
- [19] Ihara, M., Noguchi, K., Saeki, T., Fukuroda, T., Tsuchida, S., Kimura, S., Fukami, T., Ishikawa, K., Nishikibe, M. and Yano, M. (1992) *Life Sci.* 50, 247–255.
- [20] Ishikawa, K., Fukami, T., Nagase, T., Fujita, K., Hayama, T., Niiyama, K., Mase, T., Ihara, M. and Yano, M. (1992) *J. Med. Chem.* 35, 2139–2142.
- [21] Moreland, S. (1994) *Cardiovascular Drug Rev.* 12, 48–69.
- [22] Atkinson, R.A. and Pelton, J.T. (1992) *FEBS Lett.* 296, 1–6.
- [23] Krystek Jr., S.R., Bassolino, D.A., Brucoleri, R.E., Hunt, J.T., Porubcan, M.A., Wandler, C.F. and Andersen, N.H. (1992) *FEBS Lett.* 299, 255–261.
- [24] Reily, M.D., Thanabal, V., Omecinsky, D.O., Dunbar, J.B., Doherty, A.M. and DePue, P.L. (1992) *FEBS Lett.* 300, 136–140.
- [25] Bean, J.W., Peishoff, C.E. and Kopple, K.D. (1994) *Int. J. Peptide Protein Res.* 44, 223–232.
- [26] Gonnella, N.C., Zhang, X., Jin, Y., Prakash, O., Paris, C.G., Kolossvary, I., Guida, W.C., Bohacek, R.S., Vlattas, I. and Sytwu, T. (1994) *Int. J. Peptide Protein Res.* 43, 454–462.

- [27] Ishikawa, K., Fukami, T., Hayama, T., Niiyama, K., Nagase, T., Mase, T., Fujita, K., Kumagai, U., Urakawa, Y., Kimura, S., Ihara, M. and Yano, M. (1992) in: *Peptides: Chemistry and Biology* (Smith, J. and Rivier, J. eds.) pp. 812–813, Escom, Leiden.
- [28] Bogusky, M.J., Brady, S.F., Sisko, J.T., Nutt, R.F. and Smith, G.M. (1993) *Int. J. Peptide Protein Res.* 42, 194–203.
- [29] Murugesan, M., Gu, Z., Lee, V., Webb, M.L., Liu, E. C.-K., Hermsneuer, M. and Hunt, J.T. (1995) *Bioorg. Med. Chem. Lett.* 5, 253–258.
- [30] Hunt, J.T., Lee, V.G., Stein, P.D., Hedberg, A., Liu, E.C.-K., McMullen, D. and Moreland, S. (1991) *Bioorg. Med. Chem. Lett.* 1, 33–38.
- [31] Hunt, J.T. (1992) *Drug News Perspect.* 5, 78–92.
- [32] Tam, J.P., Liu, W., Zhang, J.-W., Galantino, M., Bertolero, F., Cristiani, C., Vaghi, F. and DeCastiglione, R.D. (1994) *Peptides* 15, 703–708.
- [33] Evans, S.V. (1993) *J. Mol. Graphics* 11, 134–138.
- [34] Ohlstein, E.H., Nambi, P., Douglas, S.A., Edwards, R.M., Gellai, M., Lago, A., Leber, J.D., Cousins, R.D., Gao, A., Frazee, J.S., Peishoff, C.E., Bean, J.W., Eggleston, D.S., Elshourbagy, N.A., Kumar, C., Lee, J.A., Yue, T.L., Loudon, C., Brooks, D.P., Weinstock, J., Feuerstein, G., Poste, G., Ruffolo Jr., R.R., Gleason, J.G. and Elliott, J.D. (1994) *Proc. Natl. Acad. Sci. USA* 91, 8052–8056.
- [35] Satoh, T. and Barlow, D. (1992) *FEBS Lett.* 257, 83–87.
- [36] Wallace, B.A. and Janes, R.W. (1995) *J. Cardiovascular Pharm.* 26, s250–s253.