

## DESIGN AND CONSTRUCTION OF BIOLOGICALLY ACTIVE PEPTIDES, INCLUDING ENZYMES

E. T. Kaiser

Laboratory of Bioorganic Chemistry and Biochemistry, The Rockefeller  
University, New York, New York 10021, USA

**Abstract** - Amphiphilic secondary structural regions have been found to be of great importance to the biological and physical properties of a variety of peptides and proteins that act at membrane surfaces. For systems ranging from apolipoproteins through toxins to peptide hormones we have prepared peptide models that have minimal homology to the corresponding regions of the natural peptides and proteins but simulate their biological and physical behavior. Most of our peptides have been synthesized by the stepwise solid phase method. However, in the synthesis of two of our models for apolipoprotein A-I we have utilized an oxime polymer that we have developed for a fragment synthesis-condensation approach. The latter approach will probably find considerable application, particularly in the preparation of large peptide and protein models.

### INTRODUCTION

A major goal of research in our laboratory is the design of enzymes, including both their active and binding sites. The problem of design of protein structure from the constituent amino acids, however, is complicated by the uncertainties which still exist in our ability to predict tertiary structure from primary amino acid sequence. For this reason, in much of the work which is to be described in this article in which we focus on the design of important binding regions of a variety of proteins and peptides, systems have been chosen where one can ignore, to a first approximation, tertiary structure. As will be described at the very end, we have also made progress along another line. In particular, we have taken existing enzyme systems with defined tertiary structures, and we have introduced new active sites in them. In a sense, this work is complementary to the studies that will be described in the greater part of this article which focuses on the design of binding sites involving principally secondary structural features.

If one sets out to look for systems where secondary structural features dominate the biological properties and physical properties of peptides and proteins, then one is immediately drawn to examples of peptides and proteins which bind at membrane surfaces. Membrane surfaces provide an amphiphilic environment. It seems reasonable that when peptides and proteins would bind to such surfaces they would adopt local structural features, secondary structures which are amphiphilic in nature and complementary to the amphiphilic character of the membrane surface. The secondary structures in the peptides and proteins which can be amphiphilic in character will include the commonly encountered types such as  $\alpha$ -helices,  $\beta$  strands, etc.

The first system to which we were drawn in our search for suitable models for biologically active peptides and proteins binding at membrane surfaces was apolipoprotein A-I (apo A-I), the principal polypeptide in high density lipoprotein (HDL) (Ref. 1). Structural analyses performed independently by Fitch (Ref. 2) and McLachlan (Ref. 3) were consistent with a picture of this protein being comprised in large part of  $\alpha$ -helical units, approximately twenty-two amino acids in length, in which the  $\alpha$ -helices showed a pronounced amphiphilic character. That is to say, in a prototypic helix one face, approximately covering one-third of the surface, was hydrophobic, whereas the remaining face consisted of positively and negatively charged hydrophilic residues. When we considered how the apolipoprotein might bind on the surface of HDL, an attractive hypothesis was that the amphiphilic  $\alpha$ -helices might bind with their axes roughly tangential to the surface of the HDL particle and the hydrophobic faces penetrating into the particle in contact with the hydrocarbon-like portion of the surface phospholipids. The hydrophilic face of the helices would be oriented towards the aqueous solution just as would be the case for the hydrophilic head groups of the surface phospholipids. Such amphiphilic  $\alpha$ -helices would be connected by "hinged" regions containing the helix breaking residues Gly or Pro. While this model for the binding of apo

A-I to the HDL surface looked appealing, evidence for the model was clearly needed. The approach we decided to undertake to test the proposed picture for the binding of apo A-I to the HDL surface is one which we have used throughout our work on the design of surface active proteins and peptides and is worth emphasizing here therefore.

The main thesis on which our approach is based is that if the biological and physical properties of a peptide or protein are dependent on the secondary structural features of a particular region then in building a model for the region in question, it should be possible to design a system which has minimum homology to the natural system but has the potential to form a very closely related secondary structure (Ref. 4). In other words, if it is thought that for a particular region of a peptide or protein a specific amino acid sequence is not crucial but rather that a certain type of secondary structural feature is necessary, then it should be possible to reproduce that secondary structure using an amino acid sequence which is relatively non-homologous to the natural sequence. For this reason in building our peptide models we have tried as much as possible to design systems which possess relatively little homology to the natural sequences in the amphiphilic secondary structural regions. We have not been able to eliminate sequence homology because in the work done to date we have confined ourselves to naturally occurring amino acids. If a Leu residue is encountered in a particular place in an amphiphilic  $\alpha$ -helix, for example, then we are not likely to replace this residue with another hydrophobic aliphatic amino acid since the Leu residue has a rather high potential for going into an  $\alpha$ -helix compared to other possible choices of residues. It is probable that if we were to replace Leu residues which are encountered in the hydrophobic areas of amphiphilic  $\alpha$ -helices with other aliphatic amino acids, we would obtain a less rather than a more effective amphiphilic  $\alpha$ -helix which would defeat the purposes of our modelling.

#### Apolipoproteins and peptide toxins

To illustrate our approach it is well to consider what we did in developing a model apolipoprotein A-I. In brief, we constructed a twenty-two amino acid containing peptide 1 which had the potential to form an amphiphilic  $\alpha$ -helix corresponding to a prototypic amphiphilic  $\alpha$ -helix which appears to be an underlying structural feature of apo A-I (Ref. 5) (see Fig. 1). The amphiphilic  $\alpha$ -helical peptide 1 which we constructed consisted primarily of three amino acids, Leu, as the aliphatic hydrophobic residue, Glu, as the hydrophilic negatively charged residue and Lys as the positively charged residue (assuming a pH near neutrality). The amino acids chosen were among the best  $\alpha$ -helix formers in their respective classes. The amino acids were distributed throughout the sequence in a way such that approximately one-third of the surface of the amphiphilic  $\alpha$ -helix was hydrophobic and the remaining two-thirds hydrophilic with a clustering of the charges in a positive and in a negative group (Fig. 2). The physical and biological properties of the twenty-two amino acid model peptide have been summarized extensively elsewhere (Ref. 5, 6 and 7). Suffice it to say that in all respects this peptide which is only one-eleventh the length of the whole apolipoprotein A-I molecule can mimic essentially all of the important biological and physical properties of the whole protein. What this indicates is that the crucial structural principles affecting the biological and physical properties of apo A-I are adequately reproduced by the amphiphilic  $\alpha$ -helical twenty-two amino acid peptide which has been constructed with a view toward minimizing the sequence homology with any particular region of the whole apolipoprotein.

1	10	
PRO-LYS-LEU-GLU-GLU-LEU-LYS-GLU-LYS-LEU-LYS-		
	20	<u>1</u>
GLU-LEU-LEU-GLU-LYS-LEU-LYS-GLU-LYS-LEU-ALA		
1	10	
PRO-LYS-LEU-GLU-GLU-LEU-LYS-GLU-LYS-LEU-LYS-		
	20	
GLU-LEU-LEU-GLU-LYS-LEU-LYS-GLU-LYS-LEU-ALA-		
	30	<u>2</u>
PRO-LYS-LEU-GLU-GLU-LEU-LYS-GLU-LYS-LEU-LYS-		
	40	
GLU-LEU-LEU-GLU-LYS-LEU-LYS-GLU-LYS-LEU-ALA		

Fig. 1 Amino acid sequences of peptide model 1 and peptide model 2 for apolipoprotein A-I.



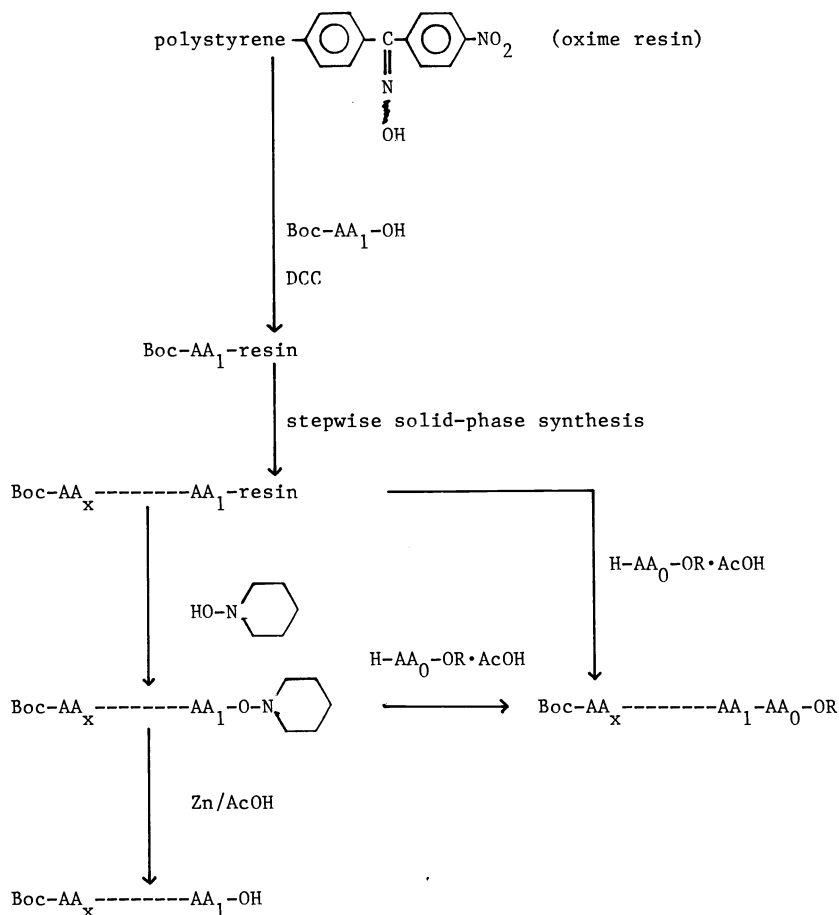


Fig. 3 Synthetic route leading to protected peptide segment employing the *p*-nitrobenzophenone oxime resin.

As outlined in Figures 4, 5 and 6, we have utilized peptide fragments synthesized on the oxime ester support to prepare two peptide models for apolipoprotein A-I. The first peptide which we prepared by a combination of fragment synthesis with fragment condensation on the oxime resin (Ref. 9) was the twenty-two amino acid containing model peptide which we had already synthesized by stepwise solid phase method and which we had already shown is an effective model of apo A-I. As shown in Fig. 4, by appropriate assembly of peptide fragments which themselves were synthesized on the oxime ester support, the apolipoprotein model peptide 1 could be prepared in a protected form very readily. An advantage of the approach used is that by judicious choice of protecting groups it was possible to synthesize the peptide itself without the need to use HF in deprotection. Furthermore, the individual fragments prepared were readily characterized as to their purity. Preparation of the twenty-two amino acid peptide from peptide fragments which themselves are known to be pure gives a product at the end which needs relatively little purification as compared to the product obtained by stepwise solid phase method for which more extensive purification is needed. Very recently, we have succeeded in preparing a forty-four amino acid containing model peptide which mimics apolipoprotein A-I and which corresponds to a system in which two units of model peptide 1 are linked by having the carboxyl of the C-terminal Ala residue of one unit connected to the amino terminal residue Pro of the other unit (Ref. 8). In considering the preparation of a peptide with a length such as this one, an important feature of the fragment synthesis-condensation approach is that one can demonstrate the purity of the fragments themselves, and that, therefore, the probability is fairly high that the result of fragment condensation will be the generation of a peptide which is largely the right product. As we build larger model peptides, we may feel generally more comfortable with their synthesis by the fragment synthesis-condensation approach than by the stepwise method despite the fact that the stepwise method is generally much more rapid. In any event, the forty-four amino acid containing peptide 2 was prepared so that we could examine the results of putting in the same molecule two amphiphilic  $\alpha$ -helical domains which model the basic secondary structural unit which we believe to be of fundamental importance for the case of apo A-I. We would like to examine the effects of having a potential "turn" region



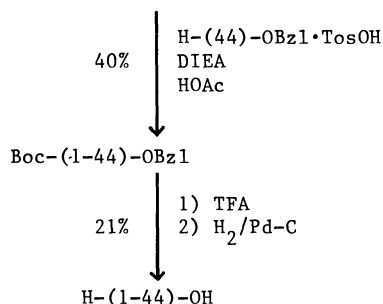


Fig. 5 Solid phase fragment coupling route to model peptide 2.

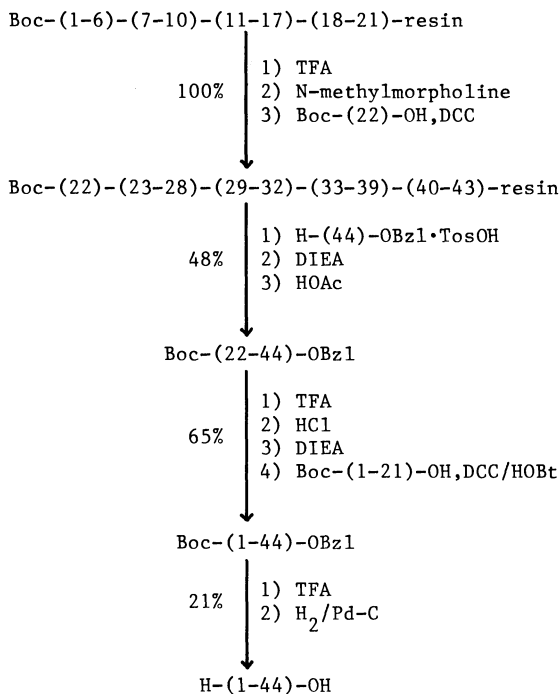


Fig. 6 Solution phase fragment coupling route to model peptide 2.

While our work on peptide 2 is still incomplete, some very interesting results have been obtained. For example, from circular dichroism studies the  $\alpha$ -helicity of this model peptide was estimated to be 54% in neutral buffered solution over the peptide concentration range between  $6 \times 10^{-5}$  M and  $6 \times 10^{-7}$  M. In a mixture of 1:1 trifluoroethanol/buffered solution (v/v) this value was increased to approximately 65%. We believe that the peptide exists in an aggregated state over the concentration range for which the aqueous measurements were made, since gel permeation chromatography on a column of Sephadex G-50 (fine) using pH 7.0<sub>5</sub> phosphate buffer containing 0.15 M NaCl and with the peptide at a concentration of  $5 \times 10^{-5}$  M gave a single peptide peak with an apparent molecular weight of 14,500, as calibrated employing globular proteins. This would suggest that at this concentration the peptide exists largely as a trimer. Furthermore, monolayer studies at the air-water interface indicate that the peptide forms a trimer in that amphiphilic environment (Ref. 12).

Besides the work on the apolipoproteins, the design of amphiphilic secondary structures has been extended in our laboratory to systems where in addition to the secondary structural regions there are active sites. The objective of this extension of our concepts is not only to improve our understanding of these systems but also to develop our ability to model molecules of increasing complexity enroute to the design of enzymes which contain active sites and also consist of complex tertiary structures. Among the systems we have examined is the peptide toxin melittin (Ref. 13 and 14). This peptide which causes the lysis of erythrocytes is in its principal form twenty-six amino acids in length (Ref. 15). We have analyzed the structure of melittin in terms of the existence of a segment extending from residue one to twenty which is an amphiphilic  $\alpha$ -helical region with a much more pronounced hydrophobic character than the amphiphilic  $\alpha$ -helices present in apo A-I. In addition to the

amphiphilic  $\alpha$ -helical region of melittin, there is a C-terminal cluster of positively charged residues in the C-terminal hexapeptide region. It is known that removal of the C-terminal hexapeptide region results in a twenty amino acid fragment which still binds to the red cell membranes but does not lyse them (Ref. 16). We have constructed a model peptide in which we have built an amphiphilic  $\alpha$ -helix with a very hydrophobic face and in which we have replaced the Pro residue which causes a kink in the helix with a Ser residue which does not. In other words, we have in a sense "amplified" the amphiphilic  $\alpha$ -helical properties of the 1-20 fragment. We have maintained the same C-terminal hexapeptide region as is present in melittin, and for the model peptide which we have so constructed, we have found that we can observe an even greater lytic activity than is seen for the naturally occurring peptide itself.

#### Peptide hormones

As discussed elsewhere, we have found also that there are many peptide hormones for which sizeable regions of the molecule appear to be forming amphiphilic secondary structures in appropriate environments (Ref. 17 and 18). For example, in the case of calcitonin, a peptide possessing potent hypocalcemic activity, we have proposed that the structure consists of an "active site" amino terminal heptapeptide which has a disulfide loop between the Cys-1 and Cys-7 residues, an amphiphilic  $\alpha$ -helical domain extending from residues eight through twenty-two, and a hydrophilic COOH-terminal region which reaches from residue twenty-three through thirty-two. In a stringent test of this proposal, we have synthesized two thirty-two amino acid peptide models for calcitonin (Ref. 18, 19 and 20). In these models while we have attempted to retain the structural characteristics of the amphiphilic  $\alpha$ -helical region, we have designed an amino acid sequence from residues eight through twenty-two in which we have tried to minimize sequence homology to any naturally occurring calcitonin. For the second of our models, studied very recently, we found that both the ability of the synthetic model to bind to calcitonin receptors and its hypocalcemic activity were approximately equivalent to the corresponding properties of the salmon hormone, although the model had been designed to have substantial non-homology with the naturally occurring peptide in the  $\alpha$ -helical region.

In recent studies the tendency of corticotropin releasing factor (CRF) to assume amphiphilic secondary structures when it is exposed to amphiphilic environments such as the air-water interface and the surfaces of unilamellar phospholipid vesicles has been demonstrated (Ref. 21). Our structural analysis for CRF suggests that this peptide hormone consists of two regions of high helical potential in which the hydrophilic and hydrophobic residues are segregated on opposite sides of the cylindrical helix. On the basis of our structural analysis and of the data on the binding of CRF to model amphiphilic environments, we believe that it is the predominately helical form which is the biologically active conformation of the peptide when it binds to cell membranes.

Another type of amphiphilic structure can be seen in the case of  $\beta$ -endorphin. We have proposed that for this hormone three basic structural units can be distinguished: a highly specific opiate recognition site at the N-terminus (residues 1-5) which is connected via a hydrophilic link (residues 6-12) to a potential amphiphilic helix in the C-terminal residues 13-31 (Ref. 22). For the C-terminal region there is an ambiguity as to whether we are dealing with an amphiphilic  $\alpha$ -helix or an amphiphilic  $\pi$ -helix. In the  $\alpha$ -helical form the hydrophobic domain would twist along the length of the helix. In the  $\pi$ -helical form the hydrophobic domain would lie straight along the length of the helix. Currently, we cannot distinguish between these two alternative proposals with confidence. Because we have had considerable experience in building  $\alpha$ -helices, it is primarily these types of systems that we have been testing for the C-terminal region of our model hormones. A number of peptide models for  $\beta$ -endorphin containing amphiphilic helical regions in their C-termini have been constructed (Ref. 22, 23, 24 and 25). A particularly striking model is a peptide in which, in addition to the construction of an amphiphilic  $\alpha$ -helical region from residues thirteen to thirty-one with a sequence having minimal homology to that of  $\beta$ -endorphin, a hydrophilic spacer region was included from residue six to twelve which had minimum homology to the equivalent region of the natural hormone (Ref. 23). Additionally, we have designed a peptide which is a "negative" model (Ref. 25). In this peptide many of the important features of  $\beta$ -endorphin are retained, with the principal difference from the natural system being that in an  $\alpha$ - or  $\pi$ -helical conformation of the region from residues thirteen to thirty-one, the amphiphilic character present in other model peptides is absent. From our studies with the "negative" model as well as the various "positive" models for the amphiphilic helical region, we have obtained much evidence for the importance of an amphiphilic helical structure in the region from residues thirteen to thirty-one, determining the resistance to proteolysis of the natural hormone and making some contribution to the interaction of the hormone with the opiate  $\delta$  and  $\mu$  receptors. In order to observe high opiate activity on the rat vas deferens the amphiphilicity of the C-terminal helical structure region was found to be essential whereas, in contrast, interaction with the opiate receptors on the guinea pig ileum did not appear to have a similar requirement (Ref. 25). Most recently, we have constructed a peptide model in which the whole C-terminal segment from residues 13-31 has been built using only D-amino acids (Ref. 26). In aqueous buffered

solutions as well as in an aqueous mixture with the structure-forming solvent trifluoro-ethanol, the circular dichroism spectra of this peptide show the presence of a considerable amount of left-handed structure. Using rat brain homogenate enzymatic degradation studies indicate that the peptide is stable in this medium. Furthermore, the model peptide shows a slightly higher affinity than  $\beta$ -endorphin in  $\delta$ - and  $\mu$ -opiate receptor binding assays while retaining the same  $\delta/\mu$  selectivity. The potency of the left-handed helical peptide in opiate assays on the guinea pig ileum is twice that of  $\beta$ -endorphin. It is particularly interesting that in the rat vas deferens assay which shows a high specificity for the C-terminal region of  $\beta$ -endorphin the model peptide shows mixed agonist-antagonist activity. It is clearly an analog of  $\beta$ -endorphin, as shown by its high potency as an agonist. Further, at low doses its antagonist effect on  $\beta$ -endorphin action shows that the model peptide interacts with the same receptor as the natural peptide hormone. Perhaps the most interesting finding is that this model peptide containing a left-handed amphiphilic helix displays a potent opiate analgesic effect when interjected intracerebroventricularly into mice. Together with earlier studies, the findings which we have now made with the new model show that the amphiphilic helical structure in the C-terminus of  $\beta$ -endorphin is of primary importance with respect to activity in the rat vas deferens and analgesic assays. Most intriguingly, in view of the drastic change in chirality of the C-terminal segment of the model the similarity of the *in vitro* and *in vivo* opiate activities of  $\beta$ -endorphin and the model peptide demonstrates that even a left-handed amphiphilic helix consisting of D-amino acids can function satisfactorily as a structural unit in a  $\beta$ -endorphin-like peptide. What this says is that for these properties of  $\beta$ -endorphin it is the surface features of the C-terminal region rather than the backbone which is important. This observation suggests that it may be possible to build systems in which the region corresponding to the residues 13-31 of  $\beta$ -endorphin can be constructed out of non-peptide-like molecules.

#### Active site design

Finally, in all of the discussion that has been presented in this article the focus has been on designing regions other than "active" sites. In another kind of approach we have taken a given tertiary structure, as in the hydrolytic enzyme papain, and by appropriate chemical modification, we have introduced covalently a flavin analog (Ref. 27, 28 and 29). In the optimal system which has been developed to date, resulting from the alkylation of the active site sulfhydryl group of the Cys-25 residue of papain with 8-bromoacetyl-10-methylisoalloxazine, we have found that the flavopapain generated is capable of oxidizing alkyl dihydro-nicotinamides with remarkable efficiency, comparable to some of the more effective naturally occurring flavoenzymes (Ref. 29). For example, the oxidation of the N-hexyl-1,4-dihydro-nicotinamide by the flavopapain proceeds with a  $k_{cat}/K_m$  value of approximately  $600,000 \text{ M}^{-1} \text{ s}^{-1}$  at 25° and pH 7.5.

In conclusion, considerable progress has been made in designing peptide systems ranging from apolipoproteins through toxins to hormones based upon an understanding of the amphiphilic secondary structures present in these molecules. We have focused on regions other than the "active" sites in these systems. In contrast, in the case of enzymes we have concentrated on designing new active sites taking the existing tertiary structure as a starting point. It is our eventual goal to put together the two approaches which have been developed in our laboratory so that it will be feasible to design tertiary structures containing active sites from the constituent amino acids.

**Acknowledgement** - The work presented here was supported by Public Health Service Program Project HL-18577, a grant from the Dow Chemical Company Foundation and National Science Foundation grants DAR 7910245 and CHE 8218637.

#### REFERENCES

1. D.J. Kroon, J.P. Kupferberg, E.T. Kaiser and F.J. Kézdy, *J. Am. Chem. Soc.* **100**, 5975-5977 (1978).
2. W.M. Fitch, *Genetics* **86**, 623-644 (1977).
3. A.D. McLachlan, *Nature (London)* **267**, 465-466 (1977).
4. D. Fukushima, E.T. Kaiser, F.J. Kézdy, D.J. Kroon, J.P. Kupferberg and S. Yokoyama, *Ann. N.Y. Acad. Sci.* **348**, 365-377 (1980).
5. D. Fukushima, J.P. Kupferberg, S. Yokoyama, D.J. Kroon, E.T. Kaiser and F.J. Kézdy, *J. Amer. Chem. Soc.* **101**, 3703-3704 (1979).
6. S. Yokoyama, D. Fukushima, F.J. Kézdy and E.T. Kaiser, *J. Biol. Chem.* **255**, 7333-7339 (1980).
7. E.T. Kaiser and F.J. Kézdy, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1137-1143 (1983).
8. S.H. Nakagawa and E.T. Kaiser, unpublished results.
9. S.H. Nakagawa and E.T. Kaiser, *J. Org. Chem.* **48**, 678-685 (1983).
10. W.F. DeGrado and E.T. Kaiser, *J. Org. Chem.* **47**, 3258-3261 (1982).
11. W.F. DeGrado and E.T. Kaiser, *J. Org. Chem.* **45**, 1295-1300 (1980).
12. S.H. Lau, unpublished results.



13. W.F. DeGrado, F.J. Kézdy and E.T. Kaiser, J. Am. Chem. Soc. 103, 679-681 (1981).
14. W.F. DeGrado, G.F. Musso, M. Lieber, E.T. Kaiser and F.J. Kézdy, Biophys. J. 37, 329-338 (1982).
15. E. Habermann, Science 177, 314-322 (1972).
16. E. Schröder, K. Lübke, M. Lehmann and I. Beitz, Experientia 27, 764-765 (1971).
17. E.T. Kaiser and F.J. Kézdy, Proc. Natl. Acad. Sci. U.S.A. 80, 1137-1143 (1983).
18. E.T. Kaiser and F.J. Kézdy, Science 223, 249-255 (1984).
19. G.R. Moe, R.J. Miller and E.T. Kaiser, J. Am. Chem. Soc. 105, 4100 (1983).
20. G.R. Moe and E.T. Kaiser, manuscript in preparation.
21. S.H. Lau, J. Rivier, W. Vale, E.T. Kaiser and F.J. Kézdy, Proc. Natl. Acad. Sci. U.S.A. 80, 7070-7074 (1983).
22. J.W. Taylor, R.J. Miller and E.T. Kaiser, J. Am. Chem. Soc. 103, 6965-6966 (1981).
23. J.W. Taylor, R.J. Miller and E.T. Kaiser, Mol. Pharmacol. 22, 657-666 (1983).
24. J.W. Taylor, R.J. Miller and E.T. Kaiser, J. Biol. Chem. 258, 4464-4471 (1983).
25. J. Blanc, R.J. Miller and E.T. Kaiser, J. Biol. Chem. 258, 8277-8284 (1983).
26. J. Blanc and E.T. Kaiser, submitted for publication.
27. H.L. Levine and E.T. Kaiser, J. Am. Chem. Soc. 100, 7670-7677 (1978).
28. E.T. Kaiser, H.L. Levine, T. Otsuki, H.E. Fried and R.-M. Dupeyre, Adv. Chem. Ser. 191, 35-48 (1980).
29. J.T. Slama, S.R. Oruganti and E.T. Kaiser, J. Am. Chem. Soc. 103, 6211-6213 (1981).