

Loops and Links: Structural Insights into the Remarkable Function of the Agouti-Related Protein

GLENN L. MILLHAUSER,^a JOE C. McNULTY,^a PILGRIM J. JACKSON,^a
DARREN A. THOMPSON,^a GREGORY S. BARSH,^b AND IRA GANTZ^c

^a*Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064, USA*

^b*Howard Hughes Medical Institute and the Department of Pediatrics and Genetics, Stanford University Medical Center, Stanford, California 94305, USA*

^c*Department of Surgery, University of Michigan Medical Center, Ann Arbor, Michigan 48109, USA*

ABSTRACT: The agouti-related protein (AGRP) is an endogenous antagonist of the melanocortin receptors MC3R and MC4R found in the hypothalamus and exhibits potent orexigenic activity. The cysteine-rich C-terminal domain of this protein, corresponding to AGRP(87-132), exhibits receptor binding affinity and antagonism equivalent to that of the full-length protein. We recently determined the NMR structure of AGRP(87-132) and demonstrated that a portion of the domain adopts the inhibitor cystine-knot fold. Remarkably, this is the first identification of a mammalian protein with this specific architecture. Further analysis of the structure suggests that melanocortin receptor contacts are made primarily by two loops presented within the cystine knot.¹⁰ To test this hypothesis we designed a 34-residue AGRP analogue corresponding to only the cystine knot. We found that this designed miniprotein folds to a homogeneous product, retains the desired cystine-knot architecture, functions as a potent antagonist, and maintains the melanocortin receptor pharmacological profile of AGRP(87-132).²⁶ The AGRP-like activity of this molecule supports the hypothesis that indeed the cystine-knot region possesses the melanocortin receptor contacts. Based on these design and structure studies, we propose that the N-terminal loop of AGRP(87-132) makes contact with a receptor exoloop and helps confer AGRP's selectivity for the central MCRs.

KEYWORDS: melanocortin receptor; agouti-related protein; nuclear magnetic resonance

INTRODUCTION

The agouti-related protein (AGRP) plays a key role in the regulation of feeding behavior and energy homeostasis in mammals.¹⁻³ This orexigenic, paracrine signaling molecule is produced in the hypothalamus and is a potent antagonist of α -MSH

Address for correspondence: Glenn L. Millhauser, Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064. Voice: 831-459-2176; fax: 831-459-2935. glennm@hydrogen.ucsc.edu

Ann. N.Y. Acad. Sci. 994: 27-35 (2003). © 2003 New York Academy of Sciences.

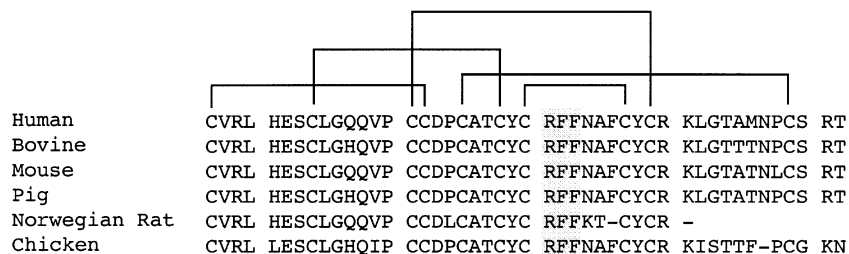
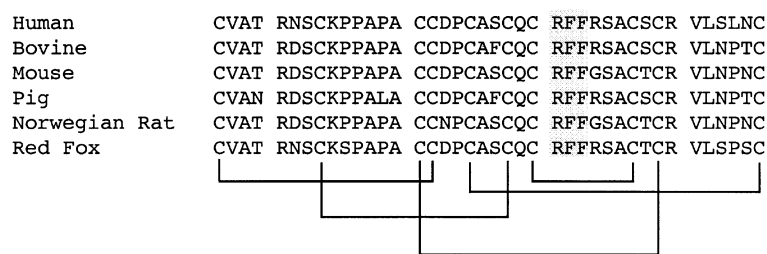
Agouti-related protein (AGRP)**Agouti protein**

FIGURE 1. Sequences of the agouti and AGRP cysteine-rich, C-terminal domains for various species. Disulfide bonds between cysteines are indicated. Residues of the RFF triplet, which dock at a receptor transmembrane pocket, are shaded.

activity at melanocortin receptors (MC3R and MC4R) of the central nervous system. The melanocortin receptor family consists of five subtypes (MC1R–MC5R) and belongs to the superfamily of G-protein coupled receptors (GPCRs) which activate the adenylate cyclase signal transduction pathway.^{4,5} AGRP was discovered by its homology with the agouti protein, which participates in the control of coat pigmentation in rodents by antagonizing MC1 receptors.^{1,2} As shown in FIGURE 1, both AGRP and the agouti protein possess a cysteine-rich C-terminal domain, and these domains alone are sufficient for high-affinity MCR binding, selectivity, and potent antagonism.^{6,7} Between AGRP and the agouti protein, the C-terminal domains are approximately 40% identical, yet they bind to distinct sets of MCR subtypes; AGRP binds with high affinity to MC3R and MC4R whereas the agouti protein binds to MC1R and MC4R.^{1,8} The molecular basis for this receptor subtype selectivity is currently unknown.

AGRP is the focus of current attention for several reasons. First, AGRP and the agouti protein exhibit unique biochemical function, as they are the only known endogenous competitive antagonists of GPCRs.^{1,9} Second, 6 of the 10 cysteines in AGRP's C-terminal domain participate in a network of disulfide cross-links with spatial positioning that is more reminiscent of an invertebrate toxin than a mammalian protein.¹⁰ Third, controlling the interaction between AGRP and its receptors in

the brain may open up new avenues for treating consumptive disorders. Indeed, obesity and other diseases of energy balance such as diabetes are approaching epidemic proportions in the United States.¹¹ MC3R and MC4R, the two receptors for which AGRP exhibits high affinity, may be ideal pharmacological targets for treating these disorders.^{3,12} In addition, the wasting condition known as cachexia is one of the major contributors to declining health in AIDS and cancer. It has been suggested that molecules with AGRP-like activity may be useful in treating this condition.¹³

STRUCTURE OF AGRP(87-132) AND DESIGN OF A MINI-AGRP

Shortly after the discovery of AGRP, we sought to determine the three-dimensional structure of this novel protein. Given the unique function of AGRP as an endogenous antagonist and its emerging importance in diabetes and obesity research, structural studies were clearly warranted. Disulfide connectivities had been determined for both agouti⁷ and AGRP.¹⁴ The Cys spacing within the C-terminal domains and the disulfide maps are similar to that found in a number of invertebrate toxins. Homology models indeed suggested that agouti and AGRP would adopt a specific toxin-like fold called an inhibitor cystine-knot (ICK) motif.^{15,16} Although such a prediction seems reasonable by the criteria of homology models, suggesting an ICK fold for AGRP or agouti is remarkable because this type of structure had not been previously observed in mammalian proteins. In addition, members of the ICK motif class differ widely in their detailed structures, and the pseudoknot nature of this fold allows for non-ICK arrangements of the polypeptide.¹⁷⁻¹⁹ Thus, we concluded that structural modeling alone would not provide a robust structural characterization.

Active AGRP(87-132) was prepared by total chemical synthesis.²⁰ Nuclear magnetic resonance (NMR) data were obtained at 800 MHz from a sample at 15°C in pH 5.0 phosphate buffer.¹⁰ A suite of homonuclear experiments yielded 799 distance restraints and 92 torsion restraints. Using these restraints along with known disulfide connectivities, we calculated a family of 40 low-energy structures as shown in FIGURE 2. The backbone fold is well defined as indicated by the overall backbone RMSD of 0.535 Å. The first 34 residues corresponding to AGRP(87-120) show an even greater degree of order with a backbone RMSD of 0.360 Å. In contrast, the final 12 residues show significant conformational variability that, in turn, reflects the near absence of long-range NOEs linking this region of the protein to the first 34 residues.

A ribbon diagram of the three-dimensional structure of AGRP(87-132) is shown in FIGURE 2.¹⁰ The protein is characterized by a three-strand anti-parallel β -sheet stabilized by a network of disulfide bonds. We refer to the short loop presenting residues 111-113 as the active loop, and the flanking loops as the N-terminal and C-terminal loops, respectively. A defining feature of the ICK motif is that the disulfide emerging from the third cysteine threads through a topological circle formed by two other disulfides and their intervening peptide segments. Examination of AGRP(87-132) indeed demonstrates that the disulfide involving the third Cys passes from the front of the protein to the back through the appropriate circle. Thus, AGRP(87-132) adopts the ICK fold and is the first mammalian protein assigned to this structure class.

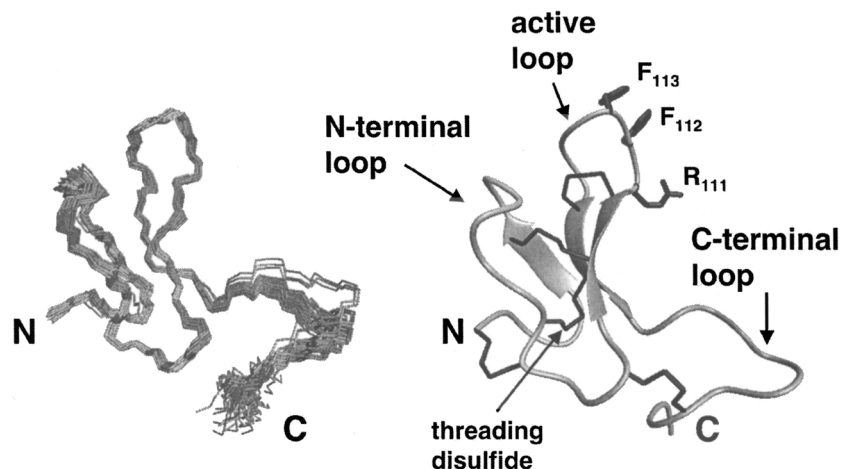


FIGURE 2. NMR structure of the AGRP C-terminal domain, AGRP(87-132). (*Left*) Family of 40 low-energy structures. The structures are aligned from residues 87–120 to highlight the observed ordering of this region and the relative disorder of residues 121–132. Only backbone atoms are shown. (*Right*) Ribbon diagram of AGRP(87-132). Backbone *arrows* identify a three-strand β -sheet; the Cys side chains involved in disulfide bonds are in *black*. The threading disulfide assigns this protein to the ICK family. Relevant loops and residues of the RFF triplet are indicated.

The triplet of residues at positions 111–113, highlighted for both agouti and AGRP sequences in FIGURE 1, is essential for MC4R antagonism. An alanine scanning experiment on agouti showed that replacement of any residue within the RFF triplet caused at least a 40-fold loss of affinity for MC4R binding.¹⁵ We have shown that replacement of R with A in the triplet leads to a complete loss of AGRP antagonism.²¹ FIGURE 2 shows that the RFF triplet resides within a short octapeptide loop—the active loop—closed by a single disulfide bond.

Tota *et al.* reasoned that the active loop forms a primary contact point for the ligand–receptor interaction.¹⁶ They tested their hypothesis by screening short cyclic peptides derived from the RFF-containing region of agouti and AGRP for MC3R and MC4R antagonism. Indeed, they demonstrated inhibition of NDP-MSH binding and inhibition of α -MSH–stimulated cAMP production. However, the AGRP-derived cyclic peptide exhibited a greater than 10-fold lower affinity than that of the full AGRP C-terminal domain at MC4R. Also, their cyclic peptides showed almost no affinity for MC3R regardless of whether the peptide sequences were derived from agouti or AGRP. Thus, other structural determinants must play a role in selectivity and tight binding at central MCRs.

Mutagenesis studies suggest that the RFF triplet interacts with its cognate receptors by docking to a cluster of negatively charged residues presented by two adjacent transmembrane helices.^{22,23} Both agonists and antagonists appear to share this common transmembrane MCR docking location. However, there are additional regions in the receptors that influence only the binding of AGRP and the agouti protein. Cassette mutagenesis studies have shown that the second and third extracellular loops

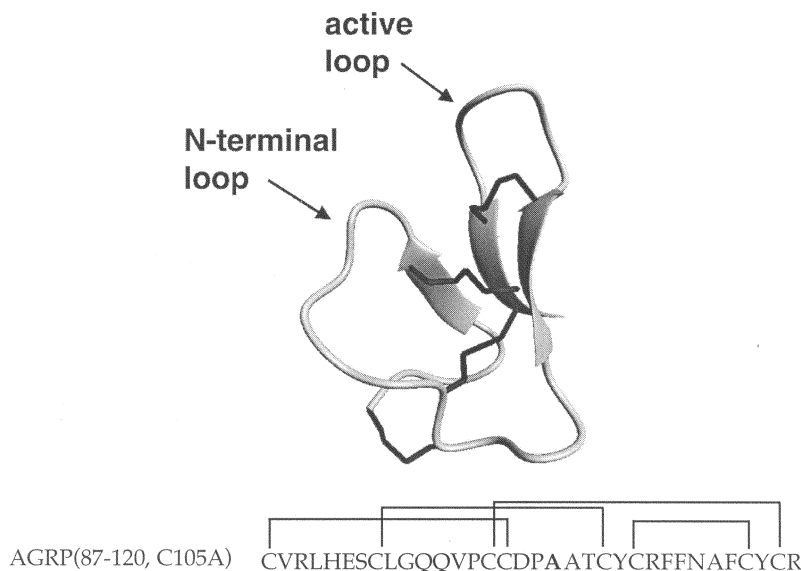


FIGURE 3. Structure and sequence of AGRP(87-120, C105A), mini-AGRP. Structural features follow FIGURE 2 legend.

(exoloops) connecting transmembrane helices are particularly important.^{24,25} Although the MC1R does not bind AGRP, a chimeric MC1R containing the second and third exoloops from MC4R results in a receptor that binds AGRP with high affinity.²⁴ In addition, chimeric MC4R with MC1R exoloops loses its high affinity for AGRP.

Based on these mutagenesis studies and the AGRP(87-132) structure, we hypothesized that the ordered region of AGRP's C-terminal domain possessed the determinants that control its MCR function. To test this hypothesis, we developed a miniprotein designed to present only the ordered region.²⁶ The sequence of this "mini-AGRP," designated AGRP(87-120, C105A), is shown in FIGURE 3. The rationale of this sequence is as follows. First, the ICK motif is found for numerous toxins with widely varying sequences and thus should fold as an independent domain.¹⁷ Second, while the 33-residue segment corresponding to AGRP(87-119) contains all of the disulfide bonds involved in the ICK structure, the NMR structure demonstrates that Arg120 participates directly in the hydrogen bonded β -sheet and thus was included in the sequence.¹⁰ To avoid having a free cysteine that in turn might lead to non-native disulfides, we eliminated Cys105 by replacing it with alanine. Finally, AGRP(87-120, C105A) was synthesized with an acetyl group at the N-terminus and an amide ($-\text{NH}_2$) at the C-terminus to avoid non-native electrostatic charges at these backbone positions.

Side-by-side receptor activity assays were performed on AGRP(87-120, C105A) and AGRP(87-132).²⁶ AGRP(87-120, C105A) functions as an antagonist and displaces NDP-MSH with IC_{50} values of 7.5 nM at MC3R and 6.1 nM at MC4R. These values are within experimental error of those determined for AGRP(87-132) (5.2 nM

and 11.0 nM). Considering the full characterization over the family of MCRs, the function of mini-AGRP is indistinguishable from that of AGRP(87-132). We also determined the NMR structure of this designed protein following the protocols used for AGRP(87-132).²⁶ The structure family revealed a low RMSD of 0.41 Å and thus a well-defined backbone fold. The ribbon diagram, shown in FIGURE 3, shows that the β-hairpin region from residues 106–120, which includes the active loop, is structured the same as in AGRP(87-132). On the other hand, it appears that the N-terminal loop and the β-hairpin are further separated from each other in AGRP(87-120, C105A) than in AGRP(87-132). Interestingly, hydrogen/deuterium exchange experiments identify only one marginally protected hydrogen bond between the β-strand from residues 92–94 and the β-hairpin in AGRP(87-120, C105A), whereas in AGRP(87-132) two hydrogen bonds are identified with each showing significant protection from exchange.²⁶

INSIGHTS INTO POSSIBLE RECEPTOR CONTACTS

The above findings identify clearly the core region of AGRP responsible for MCR selectivity and antagonism. We may now begin to identify regions of AGRP that interact with MC receptors. There are no structures available for any member of the melanocortin receptor family. However, the MCRs are homologous to rhodopsin, another member of the GPCR family for which a high-quality X-ray structure has been determined.²⁷ Homology models of MC1R^{28,29} and MC4R²³ have been developed and locate the transmembrane region responsible for AGRP active loop recognition to helices 2 and 3 near the receptor's extracellular surface.^{22,23} FIGURE 4 shows the rhodopsin structure, representative of MC4R, along with AGRP's C-terminal domain. AGRP(87-132) is oriented such that its active loop presenting the RFF triplet points towards the transmembrane binding region, and its N-terminal loop is directed to the receptor's third exoloop (between helices 6 and 7). The C-terminal loop, which our miniprotein shows is not needed for MCR recognition, is pointed away from the receptor surface. The bars in the figure are drawn to establish the distance between the relevant interaction sites. In the receptor, the bar reaches from the transmembrane site (left) to the third exoloop (right), whereas in AGRP the bar reaches from the active loop to a distal point on the N-terminal loop. In each case, the bar is approximately 24 Å long and thus suggests that the relevant regions of AGRP can indeed span the portions of the receptor that influence AGRP binding. Moreover, the specific interaction between the AGRP(87-132) N-terminal loop and the receptor exoloop may be responsible for AGRP's ability to bind selectively to central MCRs.

Also highlighted in FIGURE 4 is a collection of basic residues in AGRP(87-132) that contribute to a patch of positive charge on the protein surface. One of these residues is from the RFF triplet. However, the remaining four are all within, or near, the C-terminal loop. Although no function has yet been ascribed to this loop, its conserved sequence and participation in surface charge is suggestive. Recent studies suggest that syndecans may play a role in potentiating the effects of AGRP. Syndecans are a class of heparin sulfate proteoglycans (HSPGs), which are found ubiquitously on cell surfaces. Syndecan-3 is found on neurons in the hypothalamic feeding centers; feeding state regulates its level of expression, and syndecan-3 knockout

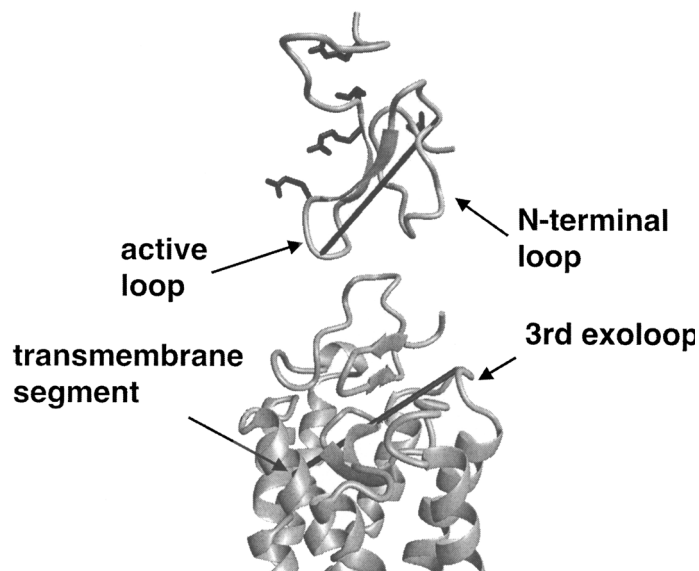


FIGURE 4. Model of AGRP(87-132) oriented for receptor docking. AGRP is positioned above the extracellular surface of rhodopsin, a GPCR with homology to MCRs. The bars are 24 Å and show that the distance between AGRP's active loop and N-terminal loop is approximately matched to the distance between the receptor transmembrane segment and the third exoloop. The Arg and Lys residues on AGRP(87-132) shown in black (positions 89, 111, 120, 121 and 131, FIG. 1) contribute to a patch of positive charge on the AGRP surface. Arg111, in the active loop, docks to the transmembrane segment. The remaining positively charged residues point away from the receptor surface and may facilitate interactions with syndecans.

mice show reduced reflex hyperphagia in response to food deprivation. Pharmacological studies by Reizes *et al.* suggest that syndecan-3 potentiates AGRP function at MC4R.³⁰ Although it is currently thought that the AGRP N-terminus mediates this interaction, the finding of a positive patch on a surface region of AGRP(87-132) that points away from the receptor suggests that the C-terminal domain may participate as well.

By combining protein design with structure determination, we have identified the core region responsible for AGRP function at MCRs. In light of our studies, and those of others, it seems very plausible that the active loop and flanking region of the N-terminal loop comprise the receptor-binding interface, with the C-terminal loop perhaps a contributor to AGRP's interaction with syndecans. The next step is to directly test these hypotheses. An unexpected benefit of our investigations is the finding that the synthesis of mini-AGRP is straightforward. Thus, mini-AGRP may serve as a scaffold into which one may engineer biophysical probes for exploring receptor interactions. Such studies, now ongoing in our laboratory, should greatly advance the understanding of how AGRP uses its unique fold to exert its antagonist function at melanocortin receptors.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants DK58606 (G.L.M), DK64265 (G.L.M.), DK48506 (G.S.B.), and DK54032 (I.G.) and by the University of Michigan Gastrointestinal Peptide Research Center (P30DK34933).

REFERENCES

1. OLLMANN, M.M. *et al.* 1997. Antagonism of central melanocortin receptors in vitro and in vivo by agouti-related protein. *Science* **278**: 135–138.
2. SHUTTER, J.R. *et al.* 1997. Hypothalamic expression of ART, a novel gene related to agouti, is up-regulated in obese and diabetic mutant mice. *Genes Dev.* **11**: 593–602.
3. WILSON, B.D., M.M. OLLMANN & G.S. BARSH. 1999. The role of agouti-related protein in regulating body weight. *Mol. Med. Today* **5**: 250–256.
4. GANTZ, I. *et al.* 1993. Molecular cloning, expression, and gene localization of a 4th melanocortin receptor. *J. Biol. Chem.* **268**: 15174–15179.
5. ABDEL-MALEK, Z.A. 2001. Melanocortin receptors: their functions and regulation by physiological agonists and antagonists. *Cell. Mol. Life Sci.* **58**: 434–441.
6. YANG, Y.K. *et al.* 1999. Characterization of agouti-related protein binding to melanocortin receptors. *Mol. Endocrinol.* **13**: 148–155.
7. WILLARD, D.H. *et al.* 1995. Agouti structure and function: characterization of a potent alpha-melanocyte stimulating hormone receptor antagonist. *Biochemistry* **34**: 12341–12346.
8. YANG, Y.K. *et al.* 1997. Effects of recombinant agouti-signaling protein on melanocortin action. *Mol. Endocrinol.* **11**: 274–280.
9. LU, D.S. *et al.* 1994. Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. *Nature* **371**: 799–802.
10. McNULTY, J.C. *et al.* 2001. High resolution NMR structure of the chemically-synthesized melanocortin receptor binding domain AGRP(87-132) of the agouti-related protein. *Biochemistry* **40**: 15520–15527.
11. FRIEDMAN, J.M. 2000. Obesity in the new millennium. *Nature* **404**: 632–634.
12. CHEN, A.S. *et al.* 2000. Inactivation of the mouse melanocortin-3 receptor results in increased fat mass and reduced lean body mass. *Nat. Genet.* **26**: 97–102.
13. MARKS, D.L., N. LING & R.D. CONE. 2001. Role of the central melanocortin system in cachexia. *Cancer Res.* **61**: 1432–1438.
14. BURES, E.J. *et al.* 1998. Determination of disulfide structure in agouti-related protein (AGRP) by stepwise reduction and alkylation. *Biochemistry* **37**: 12172–12177.
15. KIEFER, L.L. *et al.* 1998. Melanocortin receptor binding determinants in the agouti protein. *Biochemistry* **37**: 991–997.
16. TOTA, M.R. *et al.* 1999. Molecular interaction of agouti protein and agouti-related protein with human melanocortin receptors. *Biochemistry* **38**: 897–904.
17. CRAIK, D.J., N.L. DALY & C. WAINE. 2001. The cystine knot motif in toxins and implications for drug design. *Toxicon* **39**: 43–60.
18. PALLAGHY, P.K. *et al.* 1994. A common structural motif incorporating a cystine knot and a triple-stranded beta-sheet in toxic and inhibitory polypeptides. *Protein Sci.* **3**: 1833–1839.
19. NORTON, R.S. & P.K. PALLAGHY. 1998. The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon* **36**: 1573–1583.
20. YANG, Y.K. *et al.* 1999. Characterization of agouti-related protein binding to melanocortin receptors. *Mol. Endocrinol.* **13**: 148–155.
21. BOLIN, K.A. *et al.* 1999. NMR structure of a minimized human agouti related protein prepared by total chemical synthesis. *FEBS Lett.* **451**: 125–131.
22. YANG, Y.K. *et al.* 2000. Molecular determinants of ligand binding to the human melanocortin-4 receptor. *Biochemistry* **39**: 14900–14911.
23. HASKELL-LUEVANO, C. *et al.* 2001. Structure activity studies of the melanocortin-4 receptor by in vitro mutagenesis: identification of agouti-related protein (AGRP),

- melanocortin agonist and synthetic peptide antagonist interaction determinants. *Biochemistry* **40**: 6164–6179.
24. YANG, Y.K. *et al.* 1999. Contribution of melanocortin receptor exoloops to agouti-related protein binding. *J. Biol. Chem.* **274**: 14100–14106.
 25. OOSTEROM, J. *et al.* 2001. Common requirements for melanocortin-4 receptor selectivity of structurally unrelated melanocortin agonist and endogenous antagonist, agouti protein. *J. Biol. Chem.* **276**: 931–936.
 26. JACKSON, P.J. *et al.* 2002. Design, pharmacology and NMR structure of a minimized cystine knot with agouti-related protein activity. *Biochemistry* **41**: 7565–7572.
 27. PALCZEWSKI, K. *et al.* 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**: 739–745.
 28. PRUSIS, P. *et al.* 2001. PLS modeling of chimeric MS04/MSH-peptide and MC1/MC3-receptor interactions reveals a novel method for the analysis of ligand-receptor interactions. *Biochim. Biophys. Acta* **1544**: 350–357.
 29. PRUSIS, P. *et al.* 1997. Modeling of the three-dimensional structure of the human melanocortin 1 receptor, using an automated method and docking of a rigid cyclic melanocyte-stimulating hormone core peptide. *J. Mol. Graphics Modelling* **15**: 307–317.
 30. REIZES, O. *et al.* 2001. Transgenic expression of syndecan-1 uncovers a physiological control of feeding behavior by syndecan-3. *Cell* **106**: 105–116.