Design, Pharmacology, and NMR Structure of a Minimized Cystine Knot with Agouti-Related Protein Activity†

Pilgrim J. Jackson,‡ Joseph C. McNulty,‡ Ying-Kui Yang,§ Darren A. Thompson,‡ Biaoxin Chai,§ Ira Gantz,§ Gregory S. Barsh,|| and Glenn L. Millhauser*,‡

Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064, Department of Surgery, University of Michigan Medical Center, Ann Arbor, Michigan 48109-0682, and Howard Hughes Medical Institute and Department of Pediatrics and Genetics, Stanford University Medical Center, Stanford, California 94305

Received October 31, 2001; Revised Manuscript Received April 15, 2002

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. The NMR structure of this active domain was recently determined and suggested that melanocortin receptor contacts were made primarily by two loops presented by a well-structured cystine knot domain within AGRP(87−132) [McNulty et al. (2001) *Biochemistry 40,* 15520−15527]. This hypothesis is tested here with NMR structure and activity studies of a 34-residue AGRP analogue designed to contain only the cystine knot domain. The designed miniprotein folds to a homogeneous product, retains the desired cystine knot architecture, functions as an antagonist, and maintains the melanocortin receptor pharmacological profile of AGRP(87−132). The AGRP-like activity of this molecule supports the hypothesis that indeed the cystine knot region possesses the melanocortin receptor contact points. Moreover, this potent AGRP analogue is synthetically accessible, may serve in the development of therapeutics for the treatment of diseases related to energy balance. and may also find use as a new reagent for probing melanocortin receptor structure and function.

The agouti-related protein (AGRP)[1] 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 arcuate nucleus of the hypothalamus and is a potent antagonist of α-MSH activity at melanocortin receptors (MC3R and MC4R) of the CNS. 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 in melanocytes (*1*, *2*). 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.

The physiological consequences of AGRP/MC4R interactions are now well documented. Overexpression of AGRP in mice results in hyperphagia, obesity, and symptoms equivalent to adult onset diabetes (*1*). Mice that do not express MC4R exhibit approximately the same phenotype as those overexpressing AGRP, demonstrating that this is an essential receptor in the regulation of energy balance (*9*). AGRP also selects for MC3R, and the importance of this interaction is now becoming clear. MC3R knockout mice do not exhibit increased feeding behavior; however, they do exhibit an approximately 50−60% increase in adipose mass, decreased energy expenditure, and evidence of decreased fat and carbohydrate oxidation (*10*, *11*).

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*, *12*). Second, six of the ten 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

† Supported by NIH Grants DK58606 (G.L.M.), DK48506 (G.S.B.), and DK54032 (I.G.) and by The University of Michigan Gastrointestinal Peptide Research Center (P30DK34933).

* Address correspondence to this author: phone, (831) 459-2176; fax, (831) 459-2935; e-mail, glennm@hydrogen.ucsc.edu.

‡ University of California, Santa Cruz.

§ University of Michigan Medical Center.

|| Stanford University Medical Center.

[1] Abbreviations: AGRP, agouti-related protein; α-MSH, α-melanocyte stimulating hormone; CSI, conformational shift index (experimental chemical shift minus random coil chemical shift); DQF-COSY, two-dimensional double-quantum-filtered correlation spectroscopy; GPCR, G-protein-coupled receptor; HX, hydrogen−deuterium exchange; ICK, inhibitor cystine knot; $^3J_{(HN-H\alpha)}$, three-bond HN−Hα scalar coupling constant; MC1R−MC5R, melanocortin receptors types 1−5; NDP-MSH, [Nle^4,D-Phe^7]-α-MSH, a superpotent melanocortin receptor agonist; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser spectroscopy; TOCSY, two-dimensional total correlation spectroscopy.

10.1021/bi012000x CCC: $22.00
Published on Web 05/23/2002

mammalian protein (*13*). 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 (*14*, *15*). AGRP and MC4 receptors may be ideal pharmacological targets for treating these disorders (*3*). 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 (*16*).

There are ongoing efforts to identify the fundamental features in AGRP that confer its unique affinity and selectivity for the MC3 and MC4 receptors. Previous work has demonstrated that the 46-residue, cysteine-rich C-terminal domain corresponding to AGRP(87−132) is sufficient for high-affinity antagonist function at these receptors (*6*). We recently determined the NMR structure of this essential domain (*13*). By evaluating the structure with regard to recent mutagenesis experiments, we proposed that the MCR contacts reside within two well-ordered loops of a cystine knot region within this domain (*13*).

The purpose of this current effort is to design a minimal protein with AGRP melanocortin antagonist activity. The goals are twofold. First, we are interested in testing the hypothesis that indeed AGRP activity resides within a well-ordered cystine knot subdomain of AGRP(87−132). Second, as found with other minimized proteins (*17*), this molecule may serve as an intermediate step in the development of MCR-targeted therapeutics for the treatment of disorders related to energy balance. Here we present the design, synthesis, pharmacological characterization, and NMR structure of a 34-residue sequence with AGRP activity. The resulting miniprotein folds to a homogeneous product, retains the cystine knot architecture, functions as an antagonist, and displaces agonist at MC3 and MC4 receptors with IC_{50} values in the nanomolar range. To within experimental error, the function of this designed miniprotein is indistinguishable from that of AGRP(87−132).

MATERIALS AND METHODS

Chemical Protein Synthesis and Folding. N^{α}-Acetyl-AGRP(87−120, C105A)-amide was synthesized using standard Fmoc solid-phase peptide synthesis. Oxidative folding was performed with peptide concentrations of 0.1 mM in pH = 7.5 aqueous solution with 0.5 M DMSO, 4.0 M guanidine, 10 mM reduced glutathione, and 2 mM oxidized glutathione. Folding was monitored by HPLC, which revealed a single peak for the product. The folded product was purified by HPLC, and its identity as the fully oxidized peptide was confirmed by mass spectrometry.

Binding Assays. The long-acting superpotent α-MSH analogue [Nle^4,D-Phe^7]-α-MSH (NDP-MSH) was obtained from Peninsula Laboratories (Belmont, CA). Human AGRP-(87−132) was obtained from Gryphon Sciences (South San Francisco, CA), and human AGRP(86−132) was purchased from Peptides International (Louisville, KY). Binding experiments were performed on whole HEK-293 cells stably transfected with the human melanocortin receptor subtypes hMC1R, hMC3R, hMC4R, or hMC5R. ^{125}I-NDP-MSH and ^{125}I-AGRP(86−132) were prepared by simple oxidative methods using chloramine-T and Na^{125}I (Amersham Pharmacia Biotech), as previously described, followed by HPLC purification over a C18 column (*6*). Briefly, 12 h prior to the experiments 0.3 million cells were plated on 24-well plates. Before the binding experiments were initiated, cells were washed twice with MEM medium. Cells were then incubated with different concentrations of unlabeled ligand containing 0.2% BSA and either 2×10^5 cpm of ^{125}I-NDP-MSH (200000 cpm = 0.166 nM) or 1×10^5 cpm of ^{125}I-AGRP(86−132) (100000 cpm = 0.82 nM). After 1 h incubation the cells were again washed twice with MEM medium, and the experiment was terminated by lysing the cells with 0.1 N NaOH and 1% Triton X-100. Radioactivity present in the lysate was quantified using an analytical γ counter. Nonspecific binding was determined by measuring the amount of ^{125}I-label remaining bound in the presence of 10^{-5} M unlabeled ligand, and specific binding was obtained by subtracting nonspecific bound radioactivity from total bound radioactivity. Data were analyzed using Graphpad Prism. Data are expressed as ±SEM with experimental $n = 3$ using duplicate wells with experiments being performed on different days.

cAMP Assays. cAMP assays were performed as previously described using a competitive binding assay (Amersham Pharmacia Biotech, cAMP assay kit TRK 432) (*6*). HEK-293 cells stably transfected with the human melanocortin receptor (hMCR) subtypes hMC3R and hMC4R were used for these studies. Data were analyzed using Graphpad Prism (Graphpad Software, San Diego, CA). Data are expressed as ±SEM with experimental $n = 3$ using duplicate wells with experiments being performed on different days.

NMR Sample Preparation. The NMR sample for NOESY, TOCSY, and DQF-COSY contained approximately 1.7 mM peptide in 20 mM acetate-*d* buffer at pH = 5.0 in 90% H_2O and 10% D_2O. The HX exchange experiment was carried out with approximately 1.6 mM peptide in 200 mM acetate-*d* buffer at pH = 4.0. (Note that higher buffer strength was used in the HX experiments, relative to the 2D experiments, to achieve and maintain the lower pH.)

NMR Spectroscopy. 800 MHz NMR experiments were performed on the Varian Unity+ spectrometer housed in the Stanford NMR Facility. Water suppression for the NOESY, TOCSY, and DQF-COSY experiments was achieved using the WET solvent suppression sequence (*18*). All 2D spectra were collected using States−TPPI quadrature detection. All spectra were collected with 8192 points in the t_2 dimension. The NOESY spectrum was obtained with 700 t_1 increments. TOCSY and DQF-COSY spectra were obtained with 500 t_1 increments. NOESY, TOCSY, and DQF-COSY had a sweep width of 8000 Hz in both directions.

The hydrogen−deuterium amide exchange (HX) experiments were performed using 500 MHz NMR on a Varian Unity+ spectrometer. The magnet was preshimmed on dilute AGRP(87−120, C105A) solution in D_2O/acetate-*d* buffer, pH = 4.0. The first spectrum was collected 90 min after reconstituting the protonated sample in D_2O. 1D spectra were collected every 30 min until 4 h postreconstitution. TOCSY spectra were taken 4, 8, and 24 h after reconstitution with 4096 points, 500 t_1 increments, and a sweep width of 6000 Hz. Spectra were processed using the MNMR package (Carlsberg Laboratory, Department of Chemistry, Denmark). 1,4-Dioxane was used as an internal chemical shift standard (3.743 ppm).

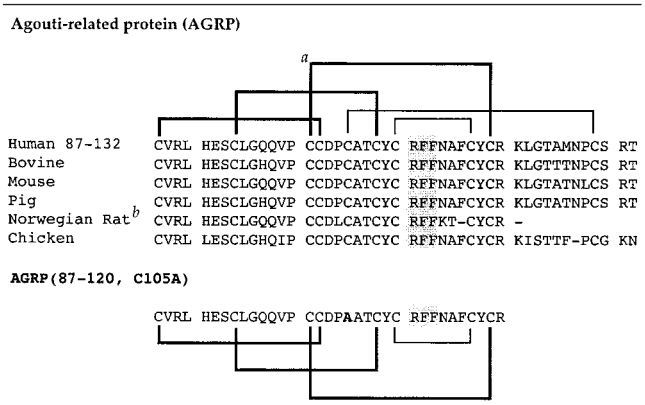
Structure Calculations. Chemical shifts were assigned, and peaks were picked manually using XEASY (*19*). Initial peak assignments were done using the NOAH module of DYANA (*20*, *21*). Remaining unassigned peaks were assigned manually in XEASY. Violations in initial structure calculations using DYANA were used to identify misassigned peaks. The final set of NOEs obtained from 150 ms 800 MHz NOESY data (362 intraresidue, 180 sequential, 83 medium range, $|j - i| < 5$, and 158 long range) along with 23 $^3J_{H\alpha-HN}$ coupling constants was processed using the CALIBA and HABAS modules in DYANA to establish distance and torsion restraints. The result was 602 upper limit restraints and 130 angle restraints. The disulfides and the assigned hydrogen bonds were incorporated as additional restraints in final structure calculations (see Results). A total of 1000 DYANA structures were calculated, and, of these, the 40 lowest energy structures make up the family reported. All 40 structures have target functions less than 2 Å^2 and have no upper limit violations greater than 0.3 Å (see Table 3). All structure representations were developed with the aid of MOLMOL (*22*).

RESULTS

Design, Synthesis, and Folding. The cysteine-rich C-terminal domain of the agouti-related protein, corresponding to AGRP(87−132), contains five disulfide bonds and exhibits MC receptor binding affinity and antagonism equivalent to that of the full-length protein (*6*). The three-dimensional structure of this domain has recently been determined by ^{1}H NMR at 800 MHz (*13*). The first 34 residues of AGRP-(87−132) are well ordered and contain a three-strand, antiparallel β-sheet where the last two strands form a β-hairpin. The turn region of the hairpin presents the RFF triplet (residues 111−113) that is essential for MCR activity (*23*−*25*). The relative spatial positioning of three of the disulfide cross-links (Cys87−Cys102, Cys94−Cys108, and Cys101−Cys119) demonstrates that the 34-residue ordered region of AGRP(87−132) adopts the inhibitor cystine knot (ICK) fold (*13*) previously identified for numerous invertebrate toxins (*26*, *27*). Interestingly, AGRP may be the first mammalian protein shown to possess a domain with this specific fold. The final 12 residues of AGRP(87−132) exhibit conformational variability, according to NMR structure calculations, and appear to form a loop that points away from the RFF triplet (*13*, *24*).

The RFF triplet is conserved across species in both AGRP and the agouti protein (*25*). We refer to the loop in AGRP closed by disulfide Cys110−Cys117, which presents this tripeptide segment, as the active loop (*24*). It has been proposed that these residues form one of the primary contacts to the melanocortin receptor, most likely at a negatively charged transmembrane pocket (*28*, *29*). Despite the importance of the active loop, it is not sufficient for full MCR activity. Cyclic peptides corresponding to AGRP's active loop bind the MC4 receptor with substantially less affinity than that of AGRP(87−132) and exhibit almost no detectable binding at the MC3 receptor (*25*). The NMR structure along with mutagenesis experiments suggests that, in addition to the active loop, the N-terminal loop [the first 16 residues of AGRP(87−132)] may also contact the MCRs and enhance affinity for both MC3R and MC4R (see discussion in ref *13*). The N-terminal loop is proximal to the RFF triplet and shows significant sequence variation to that found in the agouti protein which, in turn, selects for a different spectrum of MCR subtypes (MC1R and MC4R).

Table 1: AGRP Sequences



[a] Heavy lines indicate inhibitor cystine knot (ICK) disulfides. [b] Complete sequence not yet determined.

As noted above, the final 12 residues of AGRP(87−132) are relatively disordered. In addition, recombinant AGRP with residues 120−127 replaced by polyalanine exhibits almost the same activity at MC3R and MC4R as wild-type AGRP and AGRP(83−132) (*25*). Mutagenesis studies with the homologous agouti protein, which is also a potent antagonist of MC4R, have demonstrated that the disulfide bond anchoring the final residues of its cysteine-rich C-terminal domain is not essential for function (*30*). Taken together, it is reasonable to suggest that the final 12 residues of AGRP(87−132) do not contact the receptor.

In light of the discussion above, we propose that the ICK domain of AGRP(87−132) possesses the essential features for MCR selectivity and antagonist function. To test this hypothesis, we sought to design a minimal protein that would maintain the ICK structure but without the final unstructured residues. The sequence of this "mini-AGRP", designated AGRP(87−120, C105A), is shown in Table 1. 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 (*27*). 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 contributes stability (*13*). In addition, it has been noted that ICK proteins often fold better when there is a residue following the final ICK motif cysteine (*31*). Thus, Arg120 was included in the sequence. 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 ($-NH_2$) at the C-terminus to avoid non-native electrostatic charges at these backbone positions.

Designed human AGRP(87−120, C105A) was synthesized by stepwise addition using standard Fmoc solid-phase peptide synthesis. The resulting cleaved peptide was folded in aqueous solution under oxidizing conditions (DMSO and glutathione) over a period of several hours. HPLC revealed a single peak for the oxidized product. Oxidized AGRP(87−

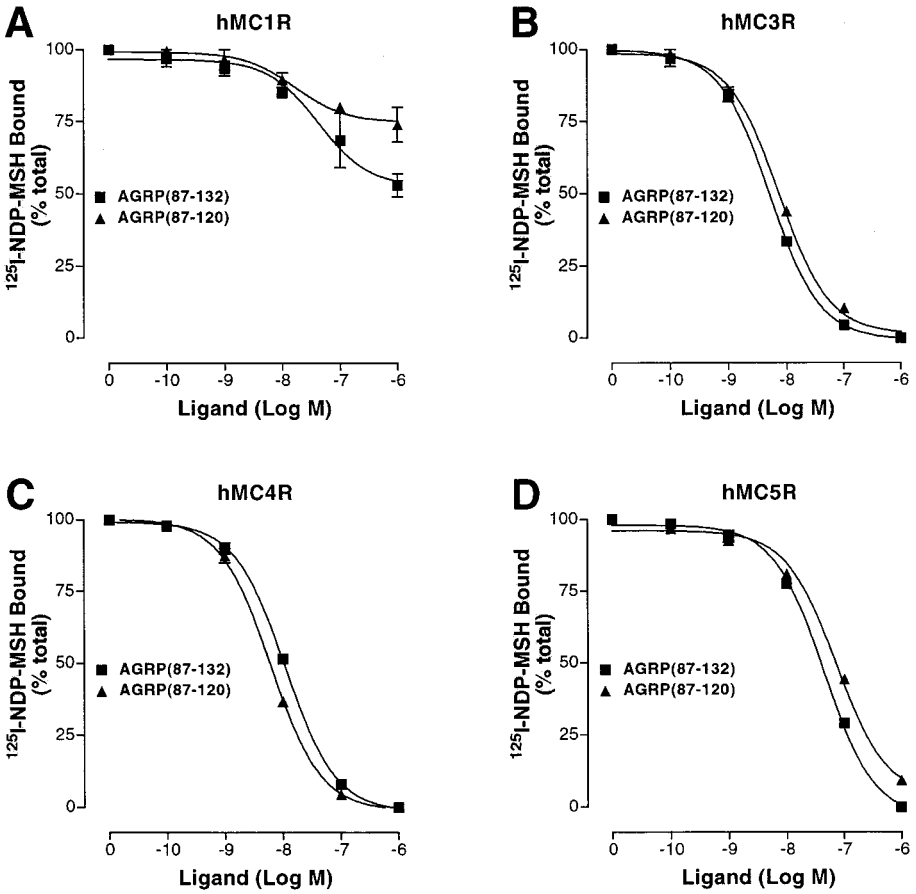


FIGURE 1: Displacement of the radioligand ^{125}I-NDP-MSH from human melanocortin receptors by AGRP(87−132) and AGRP(87−120, C105A). Receptors were stably expressed in HEK-293 cells. IC_{50} values are reported in Table 2.

120, C105A) gave a mass eight units below that of reduced form, consistent with a product containing four disulfide bonds. NMR (see below) further verified homogeneity of the final product.

Receptor Binding and Activity. To examine the receptor activity of AGRP(87−120, C105A) and compare its affinity to that of AGRP(87−132), binding displacement assays were performed on the family of human MC receptors using the agonist radioligand ^{125}I-NDP-MSH (Figure 1). (Note that since neither NDP-MSH nor AGRP(87−132) binds the MC2R, studies were not performed with this MCR subtype.) Consistent with the known low-binding affinity of AGRP for the hMC1R (*6*), both AGRP(87−132) and AGRP(87−120, C105A) were observed to only partially displace ^{125}I-NDP-MSH from the MC1R at 1.0 μM (Figure 1, panel A). AGRP(87−132) was observed to displace 47% of bound ^{125}I-NDP-MSH from the hMC1R while AGRP(87−120, C105A) displaced 26% of bound ^{125}I-NDP-MSH. As shown in Figure 1, panels B−D, AGRP(87−132) and AGRP(87−120, C105A) displayed a similar ability to displace ^{125}I-NDP-MSH at the hMC3R, hMC4R, and hMC5R. IC_{50} values are reported in Table 2. (In all assays the radioligand concentration was kept sufficiently low such that the IC_{50} values provide a good estimate of the inhibitor dissociation constant K_i.)

The MC3R and MC4R are the two MCR subtypes known to have the greatest affinity for both AGRP(87−132) and recombinant full-length AGRP (*6*). Therefore, a functional assay and additional binding studies were performed with those MCRs. AGRP(87−120, C105A) was found to exhibit potent antagonist activity at MC3 and MC4 receptors (Figure 2). Specifically, AGRP(87−120, C105A) causes a dose-dependent rightward shift in the EC_{50} of α-MSH at both MC3R and MC4R (*1*, *6*). To examine the affinity of the AGRP(87−120, C105A) at the hMC3R and hMC4R using a radiolabeled antagonist (as opposed to the radiolabeled agonist ^{125}I-NDP-MSH), we performed displacement studies using ^{125}I-AGRP(86−132) (Figure 2). Similar to the ^{125}I-NDP-MSH binding displacement data, both AGRP(86−132) and AGRP(87−120, C105A) displayed nearly identical binding affinity for the hMC3R and hMC4R. The only significant difference noted was that AGRP(87−120, C105A) gave an IC_{50} approximately 4-fold lower than that of AGRP-

Table 2: Radioligand Displacement IC_{50} Values from Human MCR Subtypes[a,b]

	displacement of ^{125}I-NDP-MSH	
	AGRP(87−132)	AGRP(87−120, C105A)
MC1R	47%[c]	26%
MC3R	5.2 ± 0.7 nM	7.5 ± 0.5 nM
MC4R	11 ± 1 nM	6.1 ± 0.5 nM
MC5R	44 ± 7 nM	71 ± 4 nM
	displacement of ^{125}I-AGRP(86−132)	
	AGRP(86−132)	AGRP(87−120, C105A)
MC3R	1.4 ± 0.1 nM	5.5 ± 0.2 nM
MC4R	1.8 ± 0.1 nM	2.8 ± 0.8 nM

[a] All measurements were performed in triplicate. [b] All IC_{50} values are ±SEM. [c] AGRP(87−132) and AGRP(87−120, C105A) incompletely displace ^{125}I-NDP-MSH from MC1R; therefore, IC_{50} values cannot be determined. Percentages represent displacement at 10^{-6} M peptide.

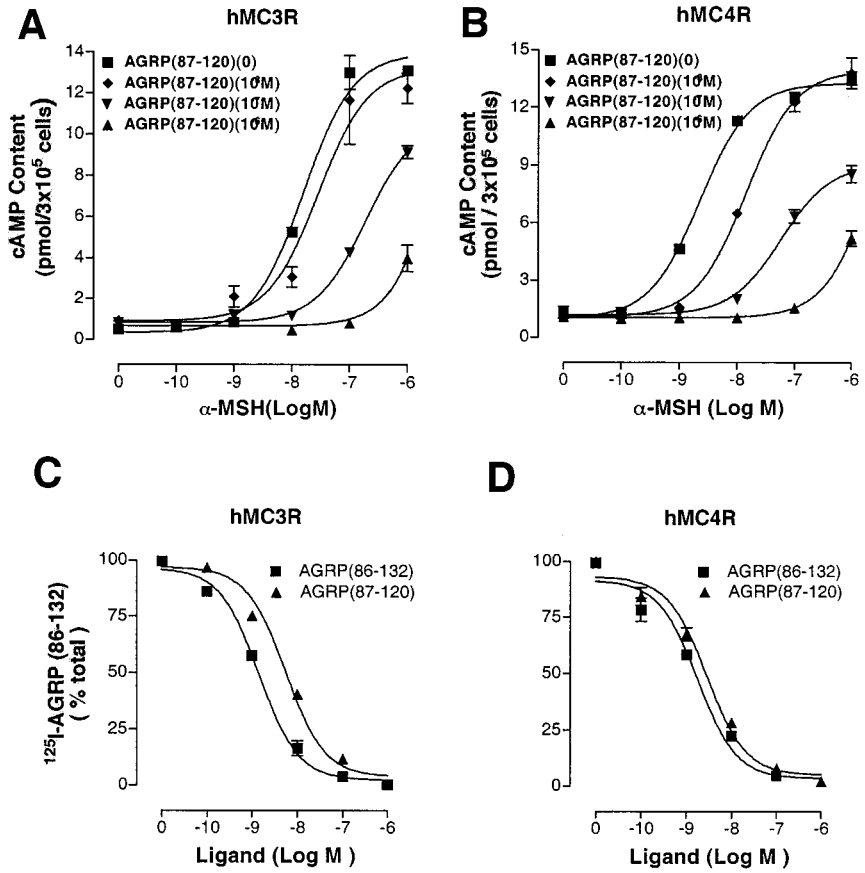


FIGURE 2: Inhibition of α-MSH-stimulated cAMP generation by AGRP(87−120, C105A) binding to human MC3 (A) and MC4 (B) receptors. The rightward shift at increasing AGRP(87−120, C105A) concentration (0 M, ■; 10^{-8} M, ◆; 10^{-7} M, ▼; 10^{-6} M, ▲) demonstrates antagonist activity. Displacement of the radioligand ^{125}I-AGRP(86−132) from human MC3 (C) and MC4 (D) receptors by AGRP(86−132) and AGRP-(87−120, C105A). IC_{50} values are reported in Table 2.

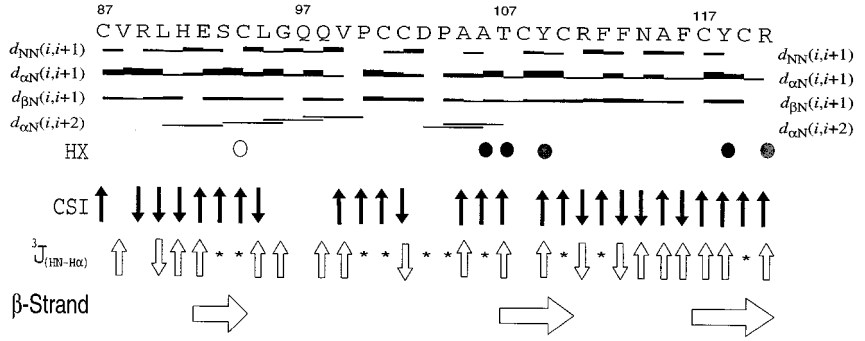


FIGURE 3: Summary of short-range backbone NOE's, HX protection, chemical shift indexes (CSI), $^3J_{(HN-H\alpha)}$ coupling constants, and β-strand assignments for AGRP(87−132). Line thickness indicates strong, medium, or weak NOEs. For HX, open circles indicate protection for >2 h, gray circles >8 h and black circles >24 h, respectively. Solid vertical arrows indicate α proton chemical shifts that vary by more than 0.10 ppm from random coil values. Hollow vertical arrows indicate $^3J_{(HN-H\alpha)}$ < 5.5 Hz or $^3J_{(HN-H\alpha)}$ > 8.5 Hz. Positive deviation of CSI and $^3J_{(HN-H\alpha)}$ for sequential residues signifies β structure whereas negative deviation signifies α structure.

(87−132) at hMC3R. Taken together, these data (Figures 1 and 2, Table 2) demonstrate that AGRP(87−120, C105A) functions as a high-affinity antagonist at MC3R and MC4R with activity that is approximately equivalent to AGRP(87−132).

NMR Structure. NOESY, TOCSY, and DQF-COSY experiments were acquired at 15 °C in 20 mM acetate-*d* buffer, pH = 5.0, at 800 MHz. NOESY spectra were acquired with a 150 ms mixing time and TOCSY with a 50 ms mixing time. The HX protection data were obtained at 15 °C in 200 mM acetate-*d* buffer at 500 MHz. Distance and torsion restraints used in the structure calculations were derived from the 150 ms NOESY and DQF-COSY spectra. The characteristics of all NMR spectra are indicative of a well-folded protein with a single predominant conformer. The presence of strong $H\alpha_i-H\delta_{i+1}$ NOE's for Val99−Pro100 and Asp103−Pro104 indicates a trans conformation for these proline residues (*32*).

Figure 3 summarizes the observed sequential NOEs, medium-range NOEs, HX protection data, α proton chemical shift indexes (CSI), and $^3J_{HN-H\alpha}$ coupling constants. No backbone−backbone or backbone−side chain $d(i,i+3)$ or $d(i,i+4)$ NOE's were observed, indicating the presence of little, if any, helical character consistent with previous find-

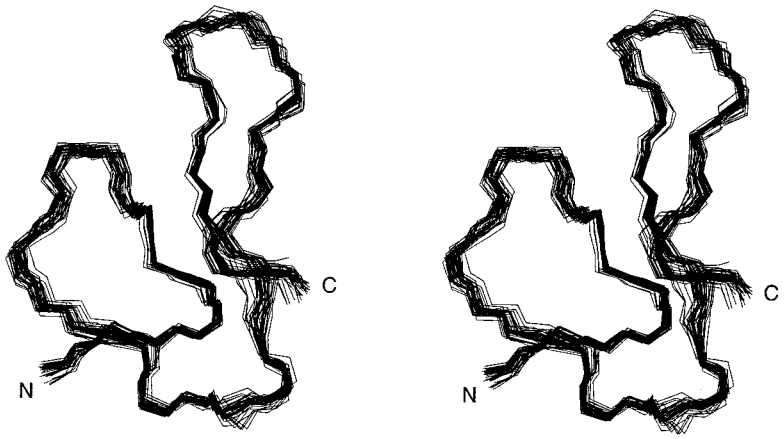


FIGURE 4: Family of 40 low-energy structures calculated in DYANA. The structures are aligned to minimize the backbone RMSD. Only backbone atoms are shown, and statistics are reported in Table 3.

ings for AGRP(87−132) (*24*, *33*). In contrast, considerable β-strand structure is suggested by the presence of medium−strong $d_{\alpha N}(i,i+1)$ NOE's, relatively weak $d_{NN}(i,i+1)$ NOEs, and a significant number of both CSI and $^3J_{HN-H\alpha}$ values greater than random coil values (*34*−*36*).

The disulfide map of AGRP(87−120, C105A) has not been determined using chemical means; however, NMR distance restraints can identify specific disulfide connectivities. The 602 NOE and 130 torsion restraints were used in DYANA to determine a family of 40 low-energy structures with RMSD = 0.47 Å. These structures were examined for pairwise placement of the cysteine sulfurs. Identification of disulfide bonds Cys87−Cys102 and Cys110−Cys117 was unambiguous and fully consistent with the disulfide connectivities previously determined for AGRP (*37*) and AGRP-(87−132) (*38*). However, covalent connectivities among the remaining four cysteines were less clear. The three possibilities (Cys94−Cys108 and Cys101−Cys119; Cys94−Cys101 and Cys108−Cys119; Cys94−Cys119 and Cys101−Cys108) were tested explicitly by performing DYANA structure calculations with each pair of specific disulfide linkages included as covalent restraints. Among the three possibilities, the structures calculated with disulfides Cys94−Cys108 and Cys101−Cys119 gave the smallest number of NOE and torsion violations and the lowest target function (by more than a factor of 2 compared to the next lowest). Strong support for the latter of these two disulfides (>95%) comes from a strong Hβ−Hβ NOE solely between the side chains of residues 101 and 119 (*39*). The overall disulfide arrangement identified for AGRP(87−120, C105A) agrees with that previously determined for AGRP and AGRP(87−132) (*37*, *38*).

A final round of DYANA calculations was performed with inclusion of the disulfide bonds and the assigned hydrogen bonds, consistent with the HX protection data in Figure 3, to produce a family of 40 low-energy structures (Figure 4). The backbone is well defined, as indicated by the overall backbone RMSD of 0.41 Å. Additional RMSD data for this family as well as their structural statistics are provided in Table 3. The final three-dimensional structure of AGRP-(87−120, C105A) is displayed as a ribbon diagram in Figure 5. The structure calculations and HX protection indicate the presence of a three-strand, antiparallel β-sheet composed of a well-defined β-hairpin across residues 106−120, with a third extended strand formed by residues 92−94. The torsion angles from the ϕ of Glu92 through the ϕ of Cys94 support the assignment of β-strand; however, this segment is somewhat irregular. The active loop of AGRP is the segment closed by the Cys110−Cys117 disulfide bond and corresponds to the turn region of the β-hairpin. Figure 5 also shows the heavy atoms of the RFF triplet side chains.

Table 3: Structural Statistics for 40 AGRP(87−120, C105A) Conformers

average target function	1.37 Å^2
NOE constraint violations	
no. >0.2 Å	5
sum of violations	6.0 ± 0.5 Å
maximum violation	0.31 ± 0.05 Å
dihedral angle constraint violations	
no. >5°	0
sum of violations	0.4 ± 0.1°
maximum violation	0.07 ± 0.02°
van der Waals violations	
sum of violations	2.3 ± 0.4 Å
maximum violation	0.17 ± 0.03 Å
NOE contacts	
long range	158
medium range	83
sequential	180
intraresidue	362
RMSD (Å) pairwise averaged over family of structures	
backbone only	0.41
heavy atoms only	0.99
all atoms	1.2
Ramachandran space statistics across entire structure family	
residues in most favorable regions	66.1%
residues in additionally favored regions	31.4%
residues in generously allowed regions	2.2%
residues in disallowed regions	0.3%

Further examination of Figure 5 shows that AGRP(87−120, C105A) adopts the ICK fold as anticipated. Specifically, disulfide Cys101−Cys119 passes from the front of the protein to the back through a ring formed by the Cys87−Cys102 and Cys94−Cys108 disulfides and their intervening peptide segments (*26*, *27*). In addition, Arg120 participates in the β-hairpin by forming complementary hydrogen bonds with Thr107.

The structures of AGRP(87−120, C105A) and AGRP-(87−132) are compared in Figure 6. The structures are aligned to minimize the RMSD between the backbones of AGRP(87−120, C105A) and the first 34 residues of AGRP-(87−132) (RMSD = 1.7 Å). The β-hairpin region from residues 106−121, which includes the active loop, shows

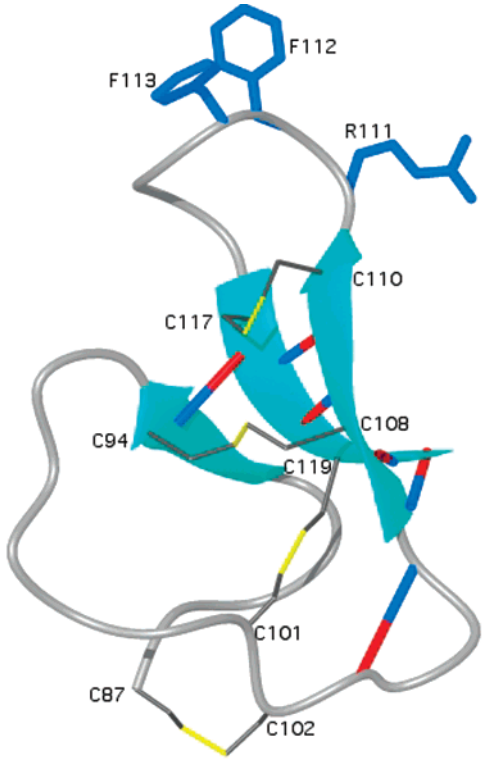


FIGURE 5: Ribbon diagram of AGRP(87−120, C105A). Backbone arrows represent assigned secondary structure identified in Figure 3. Blue and red thick-set lines represent the assigned hydrogen bonds from Figure 3. The figure also reveals the heavy atoms of the Cys side chains, with S−S bonds in yellow, and the residues of the RFF triplet.

FIGURE 6: Structural comparison of AGRP(87−120, C105A) (blue) and AGRP(87−132) (red). The ribbon diagrams are aligned to minimize the RMSD between the backbones of AGRP(87−120, C105A) and the first 34 residues of AGRP(87−132) (RMSD = 1.7 Å).

excellent agreement between the two molecules. The HX data for AGRP(87−120, C105A) and AGRP(87−132) show a similar pattern of protection from exchange for Thr107, Tyr109, Tyr118, and Arg120, further supporting similarity of their β-hairpins. The N-terminal loops, defined by the first 16 residues, also show a similar global fold. However, overlap of the structures clearly reveals differences between the proteins. 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, HX 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 (*13*).

DISCUSSION

The designed protein AGRP(87−120, C105A) folds to a homogeneous product, exhibits MCR binding and selectivity almost identical to that of AGRP(87−132), and functions as a potent antagonist at MC3R and MC4R. In addition, the mini-AGRP adopts the inhibitor cystine knot fold previously identified for the first 34 residues of AGRP(87−132). However, AGRP(87−120, C105A) seems to have more of a separation between its N-terminal loop and β-hairpin than observed for the first 34 residues of AGRP(87−132).

These findings help to identify the regions of AGRP that interact with MC receptors. Previous mutagenesis studies on MC4 receptors have suggested that a cluster of negatively charged residues emerging from two of the receptor's transmembrane helices provides the docking region for the RFF triplet (*28*, *29*). However, work with active loop cyclic peptides containing the RFF triplet demonstrated that this docking interaction alone is insufficient for full AGRP-like high-affinity binding and antagonism (*25*). On the basis of the recently reported NMR structure of AGRP(87−132), we proposed that the N-terminal loop also provides an additional interaction point where contact is most likely with one or two of the receptor's extracellular loops (*13*). The studies here further demonstrate that indeed the AGRP domain consisting of the N-terminal loop and the active loop provides the necessary contacts for AGRP- like function at MC3R and MC4R.

Despite the apparent importance of AGRP's N-terminal loop, it is possible that its influence comes not from direct contact to the receptor but instead through intramolecular contacts that help to order AGRP's β-hairpin. Specifically, the interactions between the N-terminal loop and the β-hairpin may help to structure the active loop such that it makes proper contact with the MC receptors. Interestingly, full-length AGRP with Gln97 and Gln98 both replaced by alanine exhibited binding at MC3R and MC4R equivalent to wild-type AGRP, suggesting that at least these residues are not directly involved in receptor binding (*25*). Future work with additional N-terminal loop mutants and chimeras will be needed to clearly resolve the mechanism by which this region contributes to AGRP function.

Localization of MCR activity to AGRP's residues 87−120 raises the issue as to the functions of the remaining segments of the protein. A recent paper (*40*) suggests that heparan sulfate proteoglycans may enhance the activity of AGRP in vivo, possibly by interaction with N-terminal residues outside of the cysteine-rich folded domain. In contrast, no function has yet been proposed for AGRP's final 12 residues. The sequence of this segment is well conserved across species and may contribute subtly to AGRP activity. There are a number of possibilities to consider. For instance, these residues may contribute to thermodynamic stability,

increase the rate of folding in vivo, or slow AGRP proteolysis by sterically protecting residues within the cystine knot domain. Yet another possibility is suggested by recent work demonstrating that AGRP and AGRP(83−132) function not only as antagonists of α-MSH but also independently as an inverse agonist (*41*, *42*). Perhaps the final 12 residues make a weak contact to the receptor and contribute to this subtle but important function.

The studies presented here demonstrate that the molecular features that confer AGRP's affinity for MC3R and MC4R reside largely within the 34 residues comprising the ICK domain. The receptor binding face of AGRP is likely comprised of the N-terminal loop of this domain along with the active loop that presents the RFF triplet. The potent AGRP(87−120, C105A) is synthetically accessible, may serve in the development of therapeutics for the treatment of diseases related to energy balance, and may also find use as a new reagent for probing melanocortin receptor structure and function.

ACKNOWLEDGMENT

We thank Dr. Corey Liu and the Stanford Magnetic Resonance Laboratory (J. D. Puglisi, Director) at the Stanford University School of Medicine for access to the 800 MHz spectrometer and assistance with data acquisition.

REFERENCES

1. Ollmann, M. M., Wilson, B. D., Yang, Y. K., Kerns, J. A., Chen, Y., Gantz, I., and Barsh, G. S. (1997) *Science 278*, 135−138.
2. Shutter, J. R., Graham, M., Kinsey, A. C., Scully, S., Lüthy, R., and Stark, K. L. (1997) *Genes Dev. 11*, 593−602.
3. Wilson, B. D., Ollmann, M. M., and Barsh, G. S. (1999) *Mol. Med. Today 5*, 250−256.
4. Gantz, I., Miwa, H., Konda, Y., Shimoto, Y., Tashiro, T., Watson, S. J., Delvalle, J., and Yamada, T. (1993) *J. Biol. Chem. 268*, 15174−15179.
5. Abdel-Malek, Z. A. (2001) *Cell. Mol. Life Sci. 58*, 434−441.
6. Yang, Y. K., Thompson, D. A., Dickinson, C. J., Wilken, J., Barsh, G. S., Kent, S. B., and Gantz, I. (1999) *Mol. Endocrinol. 13*, 148−155.
7. Willard, D. H., Bodnar, W., Harris, C., Kiefer, L., Nichols, J. S., Blanchard, S., Hoffman, C., Moyer, M., Burkhart, W., Weiel, J., et al. (1995) *Biochemistry 34*, 12341−12346.
8. Yang, Y. K., Ollmann, M. M., Wilson, B. D., Dickinson, C., Yamada, T., Barsh, G. S., and Gantz, I. (1997) *Mol. Endocrinol. 11*, 274−280.
9. Huszar, D., Lynch, C. A., Fairchild-Huntress, V., Dunmore, J. H., Fang, Q., Berkemeier, L. R., Gu, W., Kesterson, R. A., Boston, B. A., Cone, R. D., Smith, F. J., Campfield, L. A., Burn, P., and Lee, F. (1997) *Cell 88*, 131−141.
10. Chen, A. S., Marsh, D. J., Trumbauer, M. E., Frazier, E. G., Guan, X. M., Yu, H., Rosenblum, C. I., Vongs, A., Feng, Y., Cao, L., Metzger, J. M., Strack, A. M., Camacho, R. E., Mellin, T. N., Nunes, C. N., Min, W., Fisher, J., Gopal-Truter, S., MacIntyre, D. E., Chen, H. Y., and Van der Ploeg, L. H. (2000) *Nat. Genet. 26*, 97−102.
11. Butler, A. A., Kesterson, R. A., Khong, K., Cullen, M. J., Pelleymounter, M. A., Dekoning, J., Baetscher, M., and Cone, R. D. (2000) *Endocrinology 141*, 3518−3521.
12. Lu, D. S., Willard, D., Patel, I. R., Kadwell, S., Overton, L., Kost, T., Luther, M., Chen, W. B., Woychik, R. P., Wilkison, W. O., and Cone, R. D. (1994) *Nature 371*, 799−802.
13. McNulty, J. C., Thompson, D. A., Bolin, K. A., Wilken, J., Barsh, G. S., and Millhauser, G. L. (2001) *Biochemistry 40*, 15520−15527.
14. Friedman, J. M. (2000) *Nature 404*, 632−634.
15. Mokdad, A. H., Serdula, M. K., Dietz, W. H., Bowman, B. A., Marks, J. S., and Koplan, J. P. (1999) *J. Am. Med. Assoc. 282*, 1519−1522.
16. Marks, D. L., Ling, N., and Cone, R. D. (2001) *Cancer Res. 61*, 1432−1438.
17. Cunningham, B. C., and Wells, J. A. (1997) *Curr. Opin. Struct. Biol. 7*, 457−462.
18. Smallcombe, S. H., Patt, S. L., and Keifer, P. A. (1995) *J. Magn. Reson., Ser. A 117*, 295−303.
19. Bartels, C., Xia, T. H., Billeter, M., Guntert, P., and Wuthrich, K. (1995) *J. Biomol. NMR 6*, 1−10.
20. Guntert, P., Braun, W., and Wuthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
21. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
22. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
23. Kiefer, L. L., Veal, J. M., Mountjoy, K. G., and Wilkison, W. O. (1998) *Biochemistry 37*, 991−997.
24. Bolin, K. A., Anderson, D. J., Trulson, J. A., Thompson, D. A., Wilken, J., Kent, S. B., Gantz, I., and Millhauser, G. L. (1999) *FEBS Lett. 451*, 125−131.
25. Tota, M. R., Smith, T. S., Mao, C., MacNeil, T., Mosley, R. T., Van der Ploeg, L. H., and Fong, T. M. (1999) *Biochemistry 38*, 897−904.
26. Norton, R. S., and Pallaghy, P. K. (1998) *Toxicon 36*, 1573−1583.
27. Craik, D. J., Daly, N. L., and Waine, C. (2001) *Toxicon 39*, 43−60.
28. Yang, Y. K., Fong, T. M., Dickinson, C. J., Mao, C., Li, J. Y., Tota, M. R., Mosley, R., Van Der Ploeg, L. H., and Gantz, I. (2000) *Biochemistry 39*, 14900−14911.
29. Haskell-Luevano, C., Cone, R. D., Monck, E. K., and Wan, Y. P. (2001) *Biochemistry 40*, 6164−6179.
30. Perry, W. L., Nakamura, T., Swing, D. A., Secrest, L., Eagleson, B., Hustad, C. M., Copeland, N. G., and Jenkins, N. A. (1996) *Genetics 144*, 255−264.
31. Price-Carter, M., Gray, W. R., and Goldenberg, D. P. (1996) *Biochemistry 35*, 15547−15557.
32. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
33. Rosenfeld, R. D., Zeni, L., Welcher, A. A., Narhi, L. O., Hale, C., Marasco, J., Delaney, J., Gleason, T., Philo, J. S., Katta, V., Hui, J., Baumgartner, J., Graham, M., Stark, K. L., and Karbon, W. (1998) *Biochemistry 37*, 16041−16052.
34. Cavanagh, J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego.
35. Roberts, G. C. K. (1993) *NMR of Macromolecules: a Practical Approach*, IRL Press at Oxford University Press, Oxford and New York.
36. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR 5*, 67−81.
37. Bures, E. J., Hui, J. O., Young, Y., Chow, D. T., Katta, V., Rohde, M. F., Zeni, L., Rosenfeld, R. D., Stark, K. L., and Haniu, M. (1998) *Biochemistry 37*, 12172−12177.
38. Young, Y., Zeni, L., Rosenfeld, R. D., Stark, K. L., Rohde, M. F., and Haniu, M. (1999) *J. Pept. Res. 54*, 514−521.
39. Klaus, W., Broger, C., Gerber, P., and Senn, H. (1993) *J. Mol. Biol. 232*, 897−906.
40. Reizes, O., Lincecum, J., Wang, Z., Goldberger, O., Huang, L., Kaksonen, M., Ahima, R., Hinkes, M. T., Barsh, G. S., Rauvala, H., and Bernfield, M. (2001) *Cell 106*, 105−116.
41. Nijenhuis, W. A., Oosterom, J., and Adan, R. A. (2001) *Mol. Endocrinol. 15*, 164−171.
42. Haskell-Luevano, C., and Monck, E. K. (2001) *Regul. Pept. 99*, 1−7.

BI012000X