High-Resolution NMR Structure of the Chemically-Synthesized Melanocortin Receptor Binding Domain AGRP(87−132) of the Agouti-Related Protein[†,‡]

Joseph C. McNulty,[§] Darren A. Thompson,[§,||] Kimberly A. Bolin,[§] Jill Wilken,[||] Gregory S. Barsh,[⊥] and Glenn L. Millhauser*,[§]

Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064, Howard Hughes Medical Institute and the Department of Pediatrics and Genetics, Stanford University Medical Center, Stanford, California 94305, and Gryphon Sciences, 250 East Grand Avenue, Suite 90, South San Francisco, California 94080-4824

Received August 23, 2001; Revised Manuscript Received October 17, 2001

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 (appetite-stimulating) activity. The cysteine-rich C-terminal domain of this protein, corresponding to AGRP(87−132), contains five disulfide bonds and exhibits receptor binding affinity and antagonism equivalent to that of the full-length protein. The three-dimensional structure of this domain has been determined by ^{1}H NMR at 800 MHz. The first 34 residues of AGRP(87−132) are well-ordered and contain a three-stranded antiparallel β sheet, where the last two strands form a β hairpin. The relative spatial positioning of the disulfide cross-links demonstrates that the ordered region of AGRP(87−132) adopts the inhibitor cystine knot (ICK) fold previously identified for numerous invertebrate toxins. Interestingly, this may be the first example of a mammalian protein assigned to the ICK superfamily. The hairpin's turn region presents a triplet of residues (Arg-Phe-Phe) known to be essential for melanocortin receptor binding. The structure also suggests that AGRP possesses an additional melanocortin-receptor contact region within a loop formed by the first 16 residues of its C-terminal domain. This specific region shows little sequence homology to the corresponding region of the agouti protein, which is an MC1R antagonist involved in pigmentation. Consideration of these sequence differences, along with recent experiments on mutant and chimeric melanocortin receptors, allows us to postulate that this loop in the first 16 residues of its C-terminal domain confers AGRP's distinct selectivity for MC3R and MC4R.

Obesity and associated disorders such as diabetes are now at epidemic levels in the United States and other developed countries (*1, 2*). An understanding at the molecular level of the normal processes governing energy homeostasis and weight regulation is therefore essential for developing a comprehensive understanding of these disorders and producing effective pharmaceutical treatments. Recent studies have demonstrated the key role played by brain melanocortin receptors (MCRs) in the regulation of energy homeostasis (*3−6*). Although MCRs are found in various tissues, these studies indicate that the specific receptors MC3R and MC4R are directly implicated in energy balance. All MCRs are G-protein-coupled receptors (GPCRs) that up-regulate the production of cAMP in the presence of small endogenous peptide agonists (*7*). These agonists are derived in vivo from the single pro-opiomelanocortin peptide (POMC) that is cleaved post-translationally to form the melanocortin-stimulating hormones (MSHs) and the adrenocorticotropin hormone (ACTH). MCR function is also modulated by potent endogenous antagonists known as the agouti protein and the agouti-related protein (AGRP) that indeed may exert even greater control over MCR signaling than the peptide agonists (*8−11*). Relative to the small peptide agonists, these antagonists are larger in size and possess a five-disulfide bond network within domains that are presumably well structured in solution (*12−14*). AGRP selectively antagonizes MC3R and MC4R, and consequently, knowledge of its solution structure is essential for understanding the molecular basis of energy regulation and feeding behavior.

Each MCR exhibits unique affinities for both the agonist and antagonist peptides. The first MCR discovered (MC1R) is responsible for transducing the signal to modulate skin and coat color in mammals (*15*). This signal is provided by periodic changes of the agouti protein (ASIP) in the skin (*9*). The agouti lethal yellow mutation (A^{y}/a) in the mouse results from the ectopic expression of the agouti protein in tissues outside of the skin, including the brain (*16*). The resulting phenotype is characterized by a single nonmodulated coat color and obesity. Investigations into the cause of the obesity phenotype identified AGRP by a match of its C-terminal cysteine spacing to that found in the agouti protein (*8, 17*). Additional studies have shown that while AGRP is

[†] Supported by NIH Grants DK58606 (G.L.M.) and DK48506 (G.S.B.).

[‡] Protein Data Bank accession code: the coordinates for the family of 40 structures included and described in this paper have been deposited in the RCSB−PDB under accession code 1HYK for immediate availability.

* Corresponding author. Phone: (831) 459-2176. Fax: (831) 459-2935. E-mail: glennm@hydrogen.ucsc.edu.

[§] University of California.

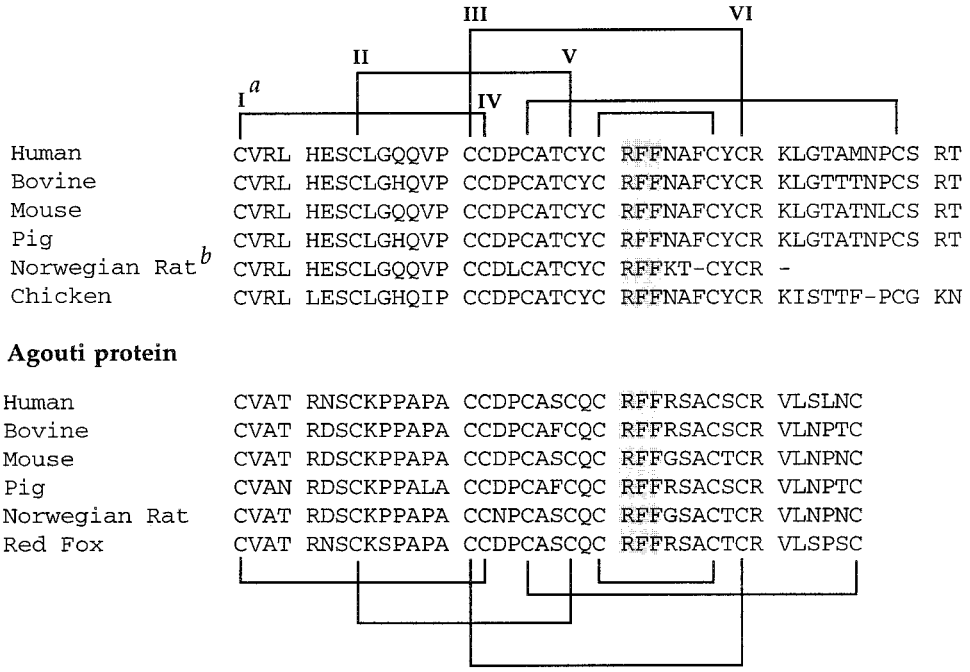
[⊥] Stanford University Medical Center.

[||] Gryphon Sciences.

10.1021/bi0117192 CCC: $20.00 © 2001 American Chemical Society
Published on Web 11/29/2001

Table 1: AGRP and Agouti Protein Sequences

Agouti-related protein (AGRP)



Agouti protein

[a] Roman numerals identify consensus cystine knot cysteine residues. [b] Complete sequence not yet determined.

an antagonist of MC3R and MC4R, agouti is an antagonist of MC1R and MC4R (*18*). The overlap of agouti protein and AGRP activity at the MC4R is responsible for the obese phenotype in the (A^y/a) mouse (*8, 19*). AGRP's function at MC3R as affecting energy metabolism rather than feeding behavior is just now becoming clear. Thus, knockout mice lacking MC3R do not exhibit increased feeding behavior; however, they do exhibit an approximately 50−60% increase in adipose mass, a decreased energy expenditure, and evidence of decreased fat and carbohydrate oxidation (*4, 6*).

Mature human AGRP is an 112-residue protein (the 20-residue signal sequence is cleaved off of the N terminus) consisting of two domains. The 46 residue cysteine-rich C-terminal domain, AGRP(87−132) (Table 1), possesses the same MCR activity and specificity in vitro as the full protein. Identical disulfide maps have been determined for both the 112-residue protein and AGRP(85−132) (*12, 14*). A detailed solution structure of the AGRP C-terminal domain is essential for interpretation of AGRP/MCR interactions and for understanding AGRP's unique receptor subtype specificity for MC3R and MC4R. Here, we present the first high-resolution NMR structure for human AGRP(87−132). In previous work, we carried out a 500 MHz NMR structural investigation of the AGRP C-terminal domain (*20*). Although this previous study correctly identified the multiloop structure for this biologically active domain, the high cysteine content and lack of regular secondary structure hampered the acquisition of the long-range structural restraints needed to clearly assign AGRP(87−132) to a particular fold motif. Sequence homology studies have suggested that AGRP may adopt the inhibitor cystine knot (ICK) fold found for certain invertebrate toxins. The current structure allows us to directly address this issue. These studies performed primarily at 800 MHz provide a family of low rmsd structures that allow us to clearly identify the fold motif of AGRP's C-terminal domain and the positions in this three-dimensional (3D) structure of residues known to be critical for the protein's function. Finally, torsion angles from ECOSY experiments reveal the extent of side-chain conformational freedom.

MATERIALS AND METHODS

Sample Preparation. N^{α}-Acetyl human AGRP(87−132) was synthesized using native chemical ligation (*21*), and its activity was tested as previously described (*10, 20*). The NMR sample contained approximately 1.8 mM of peptide in a 50 mM phosphate buffer at a pH of 5.0 and 90% H_2O/10% D_2O. The ECOSY spectrum was obtained using the same buffer system in 100% D_2O (pH was not corrected for D_2O).

NMR Spectroscopy. The 800 MHz NMR experiments were performed on the Varian Unity+ spectrometer housed in the Stanford NMR Facility. The 500 MHz NMR experiments were performed on a Varian Unity+ spectrometer at UCSC. The collection of coupling constants, HX protection data, and variable-temperature experiments were described previously (*20*). Water signal suppression for NOESY and TOCSY was achieved using the WET sequence (*22*). All of the two-dimensional spectra were collected using States-TPPI quadrature detection. All of the spectra were collected with 4096 points in the t_2 dimension. The ECOSY spectrum was obtained with 512 t_1 increments, the 800 MHz NOESY spectrum with 500 t_1 increments, and the TOCSY spectrum with 350 t_1 increments. All of the spectra had a 6000 Hz sweep width in both dimensions. Spectra were processed using the MNMR package (Department of Chemistry, Carlsberg Laboratory, Copenhagen, Denmark) and assigned using XEASY (*23*). 1,4-Dioxane was used as an internal chemical-shift standard (3.743 ppm).

Structure Calculations. The unique NOEs obtained from 150 millisecond 800 MHz NOESY data (445 intraresidual,

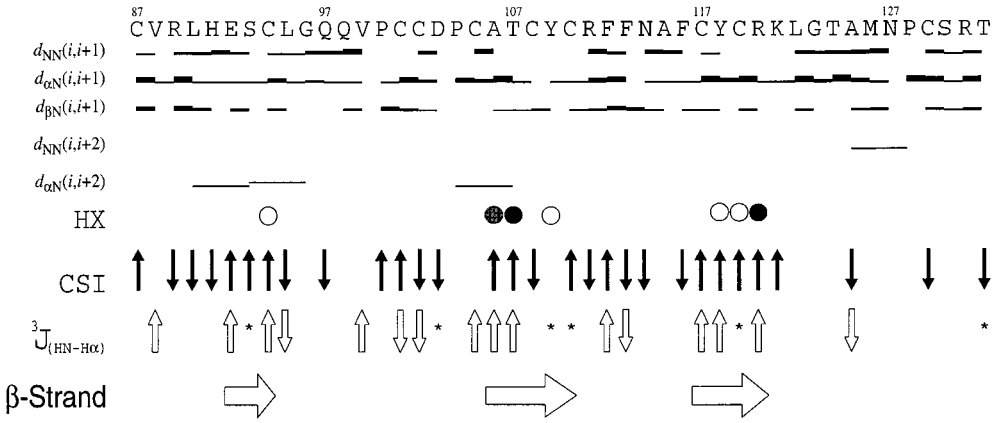


FIGURE 1: Summary of short-range backbone NOEs, HX protection, α-proton chemical shift indexes (CSI), $^3J_{HN-H\alpha}$ coupling constants, and β-strand assignments for AGRP(87−132). The line thickness indicates strong, medium, or weak NOEs. For HX, open circles indicate protection for >8 h, gray circles for >20 h, and black circles for >8 days. Solid vertical arrows indicate α-proton chemical shifts that vary by more than 0.10 ppm from the random coil values. Hollow vertical arrows indicate $^3J_{HN-H\alpha}$ < 5.5 Hz or $^3J_{HN-H\alpha}$ > 8.5 Hz. A positive deviation of CSI and $^3J_{HN-H\alpha}$ for sequential residues signifies a β structure, whereas a negative deviation signifies an α structure.

187 sequential, 57 $|j - i| < 5$ medium range, and 110 long range), along with 35 $^3J_{H\alpha-HN}$ and 44 $^3J_{H\alpha-H\beta}$ coupling constants,[11] were initially processed using the CALIBA and HABAS modules in DYANA (*24, 25*). The disulfides and assigned hydrogen bonds were incorporated as pseudo NOE restraints (*26, 27*). The best 20 out of 5000 DYANA structures, all with a target function of less than 3.0 Å^2, were used in a round of automated stereospecific assignment using GLOMSA (*24*). The seven resulting side-chain stereospecific assignments obtained were then used in a final round of DYANA structure calculations.

To obtain a refined family of structures, the lowest target function structure from the DYANA calculations was entered as the prefolded starting structure for simulated annealing in CNS (*28*). In preparation for CNS calculations, the distance restraints calculated by DYANA were converted to the XPLOR format using AQUA 2.0 (*29*). These converted restraints were modified to give a lower bound of 1.6 Å and an upper bound for each NOE as determined by the CALIBA module of DYANA. The CNS calculations included experimental NOE and *J* coupling restraints, the covalent disulfide connectivities, assigned hydrogen bonds consistent with observed HX protection, and $^3J_{H\alpha-HN}$ values. A family of 40 structures was calculated in this manner. No NOE violations of greater than 0.3 Å resulted from this method of calculation. All of the structure representations were developed with the aid of MOLMOL (*30*).

RESULTS

All of the NMR spectra were recorded at 15 °C in a pH 5.0, 50 mM phosphate buffer. The 800 MHz NOESY spectra were acquired with 80 and 150 ms mixing times, and the TOCSY spectra were acquired with a 50 ms mixing time. The ECOSY spectrum, DQF−COSY spectrum, and HX protection data were obtained at 500 MHz (*20*). The distance and torsion restraints used in the structure calculations were derived from the 150 ms NOESY, DQF−COSY, and ECOSY spectra. The presence of strong $H\alpha_i \rightarrow H\delta_{i+1}$ NOEs for Val99-Pro100, Asp103-Pro104, and Asn127-Pro128 indicates a trans conformation for all three of the proline residues (*31*).

Figure 1 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 $i \rightarrow i + 3$ or $i \rightarrow i + 4$ NOEs were observed, as was consistent with the previously published circular dichroism data that indicated the presence of little, if any, helical character (*13, 20*). In contrast, considerable β-strand structure is suggested by the presence of medium-strong $d_{\alpha N(i,i+1)}$ NOEs, a near absence of $d_{NN(i,i+1)}$ NOEs, and a significant number of both CSI and $^3J_{HN-H\alpha}$ values greater than random coil values (*32−34*). The consecutive HX protection for residues 118−120 indicates their possible participation as an interior strand of a β sheet (*33*).

The data of Figure 1, along with specific distance restraints, identified a secondary structure within AGRP(87−132) consisting of a three-stranded antiparallel β sheet as shown in Figure 2. (Note also the location of the strands in Figure 1.) All of the long-range NOEs involving backbone protons are shown, as are the short- and medium-range NOEs observed among residues 103−106. Amide protons exhibiting HX protection (Figure 1) are indicated in the context of the proposed secondary structure, and these suggest the presence of a well-defined β hairpin across residues 106−121, with a third extended strand in near proximity formed by residues 92−94. In addition to the β sheet, a "tight turn" is identified for residues 103−106. All of the NOEs (solid arrows) are directly supportive of the assignment of a tight turn (*31*). In addition, the assignment of trans to the Pro104 amide bond places the proposed acceptor carbonyl into the proper position to account for HX protection of the Ala106 amide proton and is consistent with all of the observed NOEs.

[1] Abbreviations: $^3J_{HN-H\alpha}$, three-bond HN−Hα scalar coupling constant; α-MSH, α-melanocyte stimulating hormone; AGRP, agouti-related protein; CSI, conformational shift index (experimental chemical shift-random coil chemical shift); DQF−COSY, two-dimensional double-quantum filtered correlation spectroscopy; ECOSY, two-dimensional exclusive correlation spectroscopy; GPCR, G-protein-coupled receptor; HX, hydrogen−deuterium exchange; ICK, inhibitor cystine knot; MC3R/MC4R, melanocortin receptor 3/4; 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.

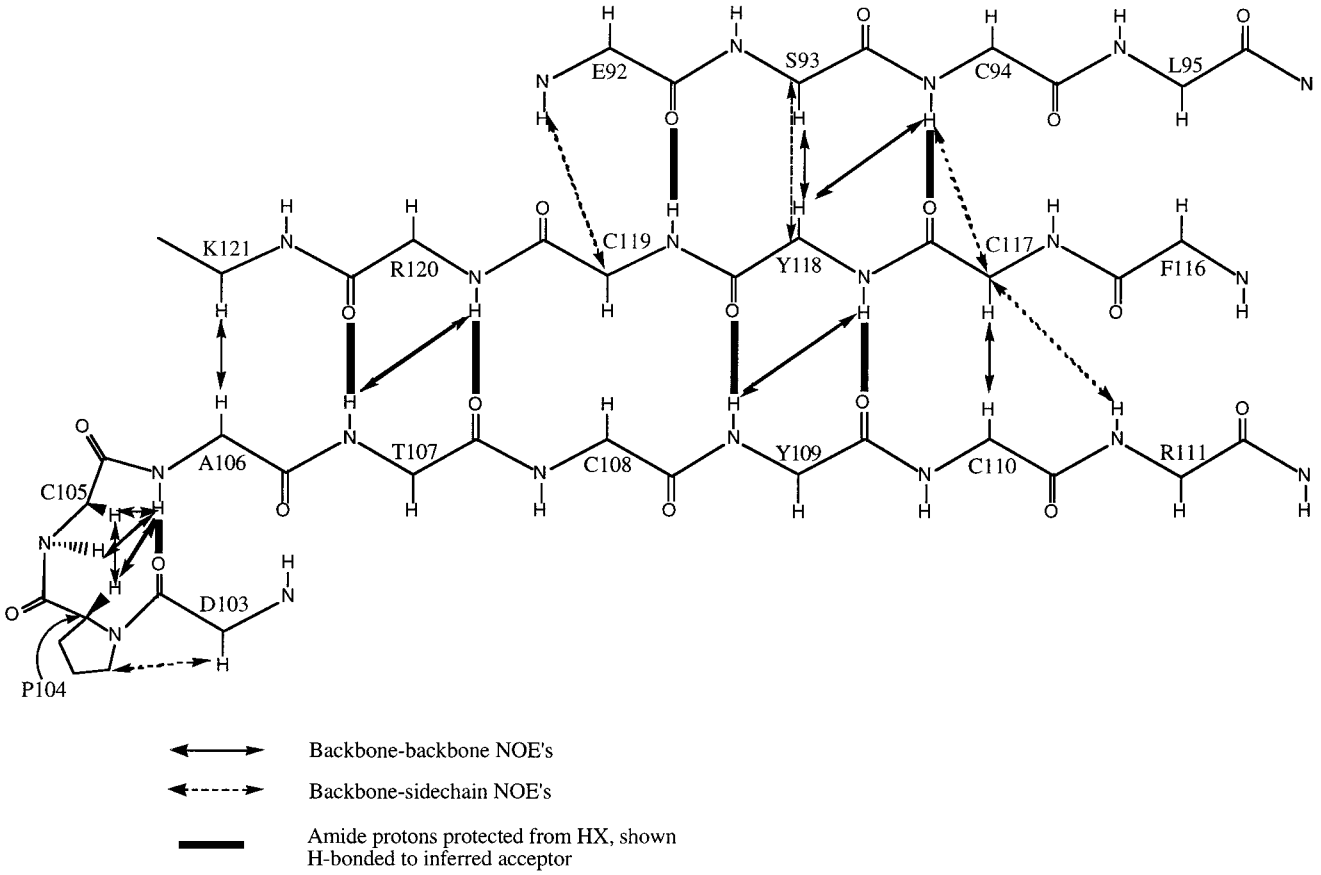


FIGURE 2: Secondary structure assignment showing hydrogen bonds determined from the NOE, HX, and coupling data in Figure 1 and from the indicated long-range backbone NOEs. The P104 side chain has been drawn to show the presence of $H\delta_i \rightarrow H\alpha_{i+1}$ NOEs.

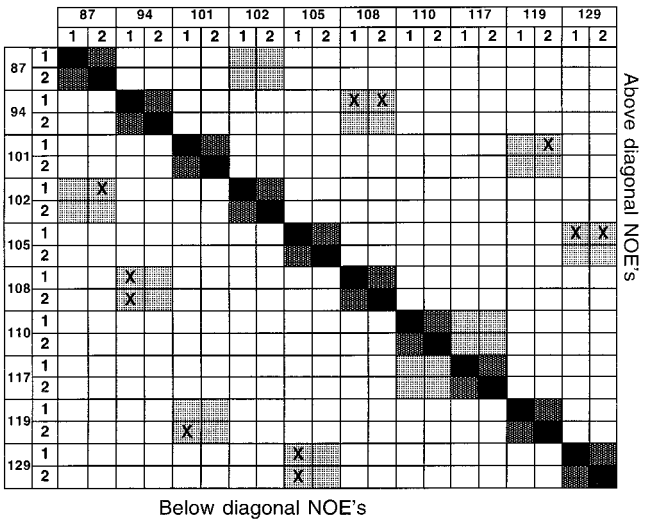


FIGURE 3: Matrix showing possible and observed NOEs between cysteine β protons. Cysteines are numbered according to sequence, and indices 1 and 2 refer to the two resolved β protons. The light gray squares correspond to those consistent with the chemically determined disulfide map for recombinant human AGRP, and the observed NOEs are indicated by an ×.

The disulfide-bond network has been chemically determined for recombinant AGRP(85−132) (*12, 14*). NMR evidence demonstrating that the chemically synthesized peptide used in this study has the same disulfide map is shown in Figure 3. Among all of the possible Cys $H\beta_i$−$H\beta_j$ NOEs, only those consistent with the chemically determined disulfide map are observed. In addition, of the five possible disulfide bonds, the NOEs from four are unambiguously observed in the NOESY spectra.

The family of 40 low-energy structures calculated in CNS is shown in Figure 4. In addition to the 799 distance restraints and 92 backbone and side-chain torsion restraints, these calculated structures also incorporated restraints because of the disulfides (Figure 3) and assigned hydrogen bonds (Figure 2). The complete rmsd data for this final family as well as their structural statistics is provided in Table 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 which, in turn, is reflected in the near absence of long-range NOEs linking this region of the protein to the first 34 residues.

To examine how well the NMR data alone define the fold of AGRP(87−132), structure calculations were repeated but without the disulfide and hydrogen-bond restraints. The resulting structures did give larger backbone rmsd values of approximately 0.84 Å for the first 34 residues and 1.91 Å for the entire molecule. However, the overall fold remained unchanged. In addition, all of the cysteine sulfurs were placed in pairwise positions consistent with the chemically determined disulfide map. These results indicate the consistency and redundancy of the available restraints on the relative positions of the backbone atoms.

The final 3D fold of the molecule is shown as a ribbon diagram in Figure 5. The left-hand-most strand of the β sheet is drawn to include residues 92−94, as assigned by the

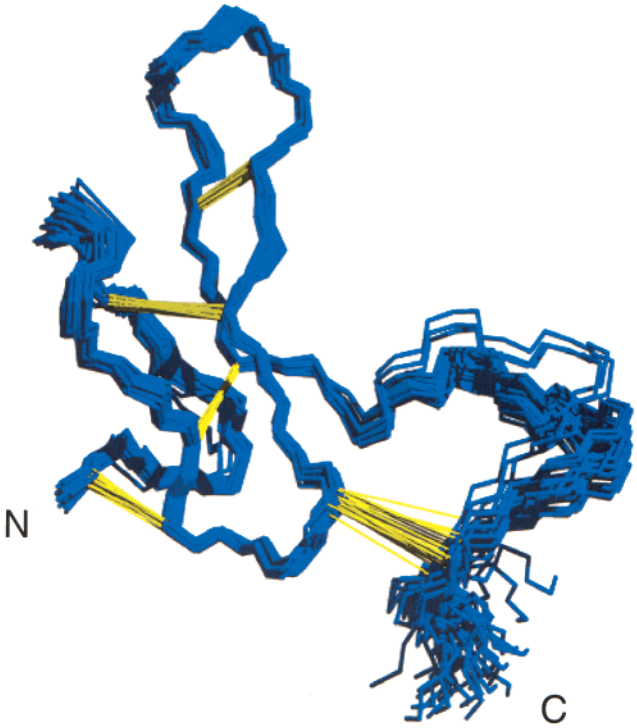


FIGURE 4: Family of 40 low-energy structures calculated in CNS at 15 °C. The structures are aligned from residues 87−120 to highlight the observed ordering of this region and the comparative disorder of residues 121−132 (Table 2). Only backbone atoms are shown. The bonds between backbone atoms are shown in blue; yellow lines represent the disulfides.

SWISS-MODEL program (*35*). Consistent with the well-ordered region of the protein noted previously, all of the significant secondary structural elements and hydrogen bonds are present in the first 34 residues. The prominent loop closed by the Cys110-Cys117 disulfide bond (Figure 5) corresponds to the turn region of the β hairpin and contains the RFF triplet (Table 1, residues 111−113) that has been established as essential for AGRP's antagonist activity (*36*). We refer to this loop as the active loop. To further emphasize the structure of this important region, Figure 5 also shows the surface-exposed region of the active loop and the heavy atoms of the RFF triplet side chains. The prominence of this triplet at the protein surface is consistent with the suggestion that these residues directly contact the MCR in the bound configuration (see the following discussion).

The ECOSY-derived side-chain Hα−Hβ coupling constants (Figure 6) were used to assess the side-chain rotameric conformational freedom and to aid in the stereospecific assignment of β-side-chain protons. Coupling constants between 6 and 8 Hz are not consistent with one stable rotamer of the χ_1 side-chain torsion angle (*32, 33*). Of the 49 measured coupling constants, 26 fall within this range, suggesting either multiple stable conformations or rotational averaging of the side chains on the NMR timescale for these particular residues. Side-chain rotational freedom is expected for those residues exposed on the protein surface. In support of the solvent exposure of the β hairpin containing the RFF triplet (Figure 5), most of the couplings from residues 111−114 fall within the disordered region. For those residues with coupling constants that do suggest a single side-chain rotameric state, most are from cysteine residues, consistent with their involvement in disulfide bonds.

DISCUSSION

Pharmacological studies carried out on heterologous cells that express specific melanocortin receptors have demonstrated that AGRP's cysteine-rich C-terminal domain possesses the necessary features for full MCR selectivity and antagonist function. A recent paper (*37*) suggests that heparan sulfate proteoglycans may enhance the activity of AGRP in vivo, possibly by an interaction with the amino-terminal fragment. Physiological studies, nevertheless, demonstrate that the AGRP C-terminal domain alone is capable of exerting potent effects on energy balance. The work presented here provides a detailed 3D structure of this essential domain. Within AGRP(87−132), the residues 87−120 are well-ordered, whereas the final 12 residues exhibit conformational heterogeneity. A well-defined fold for residues 87−120 is readily understood given that four of the five disulfide bonds are within this region, as are all of the observed hydrogen-bond contacts. This region contains a three-stranded antiparallel β sheet with two of the strands forming a β hairpin. The turn of this hairpin contains the RFF triplet within the active loop that is essential for AGRP's MCR affinity.

Previous modeling efforts have suggested that AGRP(87−132) shares homology with ω-agatoxin IVB, which adopts the ICK fold (*36*). First identified for nerve growth factors, the cystine knot is now recognized as a general scaffold for a wide range of proteins (*38, 39*). The hallmark of this motif is a loop formed by two disulfide bonds, and their connecting backbone segments, threaded by a third disulfide. Currently, there are two known cystine knot classes: the growth factor cystine knot and ICK. Several mammalian protein structures fall within the growth factor superfamily, including the nerve growth factor, platelet-derived growth factor, and transforming growth factor β. The ICK family includes numerous invertebrate toxins, such as the ω-conotoxins and the ω- and μ-agatoxins, and plant inhibitor peptides, including gurmarin and potato carboxypeptidase inhibitor. Both cystine knot classes are defined by a three-disulfide-bond topology; however, they differ in which specific disulfide threads the ring that is formed by the remaining two. Labeling the participating cysteine residues I−VI (Table 1), each class possesses a I−IV, II−V, III−VI disulfide bond pattern. In the growth factor class, cystine I−IV is the threading disulfide, while in the inhibitor class, III−VI is the threading disulfide. In addition, a triple-stranded antiparallel β sheet and a β hairpin between cysteines V and VI are common components in the ICK motif.

On the basis of the cysteine spacing found for the consensus cystine knot sequence, the motif disulfides in AGRP(87−132) are Cys87-Cys102, Cys94-Cys108, and Cys101-Cys119 (Table 1). Analysis of the 800 MHz NMR structure (Figures 4 and 5) shows that disulfide 101−119 (III−VI) passes from the front of the protein to the back through a loop formed by two remaining motif disulfides and their intervening peptide segments. Thus, AGRP(87−132) adopts the ICK 3D structure. Further support for this structural assignment is in the identification of the characteristic three-stranded β sheet. Within this β sheet, the segment forming the β hairpin (residues 106−121) makes regular antiparallel β-strand contacts through the four hydrogen bonds (Figure 2). However, the first strand of the β sheet (residues 92−94) is twisted with respect to the plane

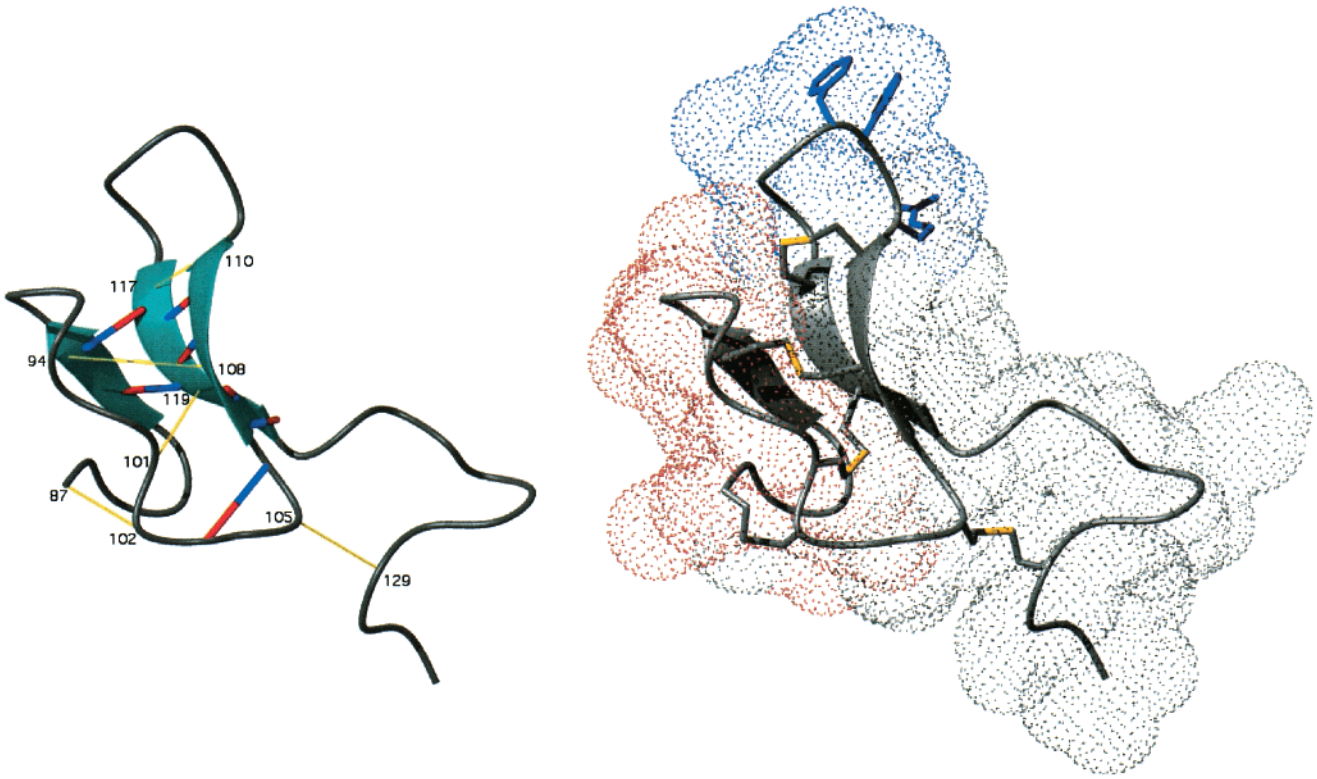


FIGURE 5: Ribbon diagrams of AGRP(87−132). Backbone arrows represent the assigned secondary structure identified in Figure 1. For the figure on the left, blue and red thickset lines represent the assigned hydrogen bonds from Figure 2, and the thin yellow lines are the numbered disulfides. The figure on the right reveals the heavy atoms of the cysteine side chains and the residues of the RFF triplet. The solvent-exposed surface area is also shown with the active loop shaded blue and the N-terminal loop, implicated in MCR selectivity, shaded red.

Table 2: NMR Structural Statistics from CNS Family of 40 Structures

Energies (kcal/mol)	
$E_{overall} = 84 \pm 3$	$E_{vdw} = 25 \pm 2$
$E_{bond} = 4.2 \pm 0.2$	$E_{noe} = 17.5 \pm 0.5$
$E_{angle} = 27 \pm 1$	$E_{coup} = 6.3 \pm 0.8$
$E_{improp} = 1.4 \pm 0.2$	

RMS Deviations from Idealized Geometry

bond lengths (Å): 0.00250 ± 0.00005
angles (deg): 0.380 ± 0.007
impropers (deg): 0.15 ± 0.01
rms deviations across all of the distance restraints: 0.020 ± 0.001 Å

RMSD from Mean (Å)	residues included in RMSD		
	87−120	121−132	87−132
backbone only	0.360	0.419	0.535
heavy atoms only	0.925	1.017	1.041
all atoms	1.146	1.264	1.261

Ramanchandran Space Statistics Across Entire Structure Family
residues in most favorable regions: 56.9%
residues in additionally allowed regions: 31.9%
residues in generously allowed regions: 10.2%
residues in disallowed regions: 1.0%

defined by the remaining two strands. Such disposition of the N-terminal strand of the β sheet is, nevertheless, a common feature of the ICK motif (*39*, *40*). To our knowledge, this is the first classification of a mammalian protein into the ICK superfamily and the first identification of an ICK protein as a ligand for a GPCR.

Residues of the RFF triplet, within the active loop and conserved in both agouti protein and AGRP sequences (Table 1), are essential for MC4R antagonism. Figure 5 shows that these residues are present in a solvent-exposed region of the β hairpin. An alanine-scanning experiment on the agouti protein showed that the replacement of any residue within the RFF triplet caused a 40-fold loss of affinity for MC4R binding (*41*). In AGRP(87−132), the replacement of triplet residue Arg111 with Ala eliminates the protein's ability to inhibit NDP-MSH (a potent analogue of α-MSH) stimulated cAMP production in MC4R transfected cells (*20*). Tota et al. (*36*) reasoned that the RFF triplet forms a primary contact point for the ligand−receptor interaction. Their hypothesis was tested by screening short cyclic peptides (corresponding to the active loop defined here) derived from the RFF-containing region of agouti and AGRP for MC3R and MC4R antagonism. Indeed, they demonstrated the inhibition of NDP-MSH binding and inhibition of α-MSH stimulated cAMP production. However, the best cyclic peptide antagonist exhibited a more than 10-fold lower affinity for MC4R than that of the full AGRP C-terminal domain. Also, their cyclic peptides showed a substantially greater affinity for MC4R than MC3R, regardless of whether the peptide sequences were derived from the agouti or AGRP. This result is in contrast to the distinct MC3R/4R selectivity behaviors of agouti and AGRP (and their C-terminal domains) discussed in the introduction (*8*, *9*). Thus, other structural determinants must play a role in receptor selectivity and high-affinity binding.

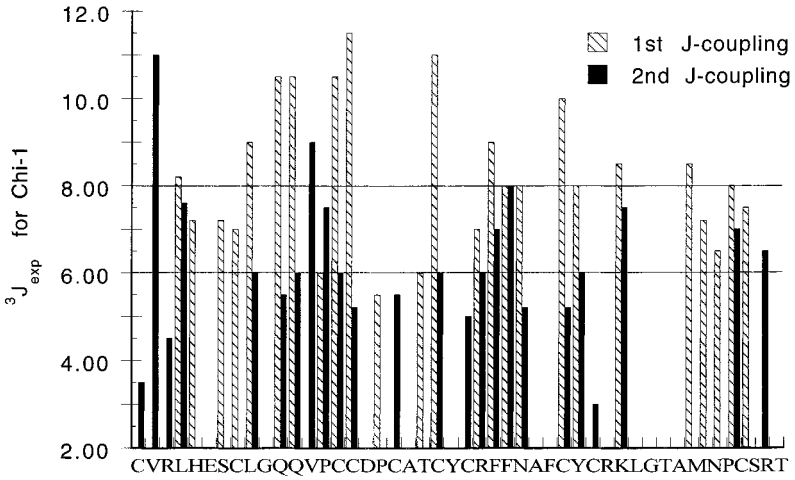


FIGURE 6: Bar graph showing ECOSY-derived $^3J_{H\alpha-H\beta}$ coupling constants. Horizontal lines indicate the region (6−8 Hz) consistent with averaging between multiple rotameric conformations.

The noted lack of structure for the final 12 residues of AGRP(87−132) may provide insight into its MCR selectivity. Studies of Cys → Ser mutations in the Cys-rich C-terminal region of the homologous agouti protein showed that 8 of the 10 Cys residues are essential for function (*19*). Interestingly, the two mutations that result in only a partial loss of antagonist activity correspond to the pair of cysteine residues that anchor the final seven residues of the agouti protein (see Table 1). Thus, structural order of these seven residues is not essential for agouti function. This result parallels our own finding of large rmsd values for the loop formed by the final 12 residues of AGRP(87−132). Because these proteins share common receptor binding sites at MC4R, our data showing conformational variability of this C-terminal loop suggests that this loop may not comprise part of the protein's binding surface. In turn, this suggests that antagonist activity and receptor selectivity resides mainly within the ICK domain of AGRP(87−132), defined by the first 34 residues. Within this domain, the residues between the first and third cysteine show little homology to the corresponding region of the agouti protein (e.g., VRLHESCLGQQVP in AGRP versus VATRNSCKPPAPA in agouti; Table 1). This polypeptide region forms a loop closed by the Cys87-Cys102 disulfide bond. We refer to this segment as the N-terminal loop. Interestingly, the structure in Figure 5 shows that the crest of the N-terminal loop of AGRP (residues 94−98) is spatially proximal to the RFF triplet. Thus, if the agouti protein C terminus adopts approximately the same fold as the AGRP C terminus, this highly variable N-terminal loop may confer the differential MCR selectivity of agouti and AGRP.

Mutagenesis studies on MC4 receptors have identified a cluster of negatively charged residues, emerging from two adjacent transmembrane helices, that is essential for peptide-agonist and AGRP binding. Substantial evidence suggests that this cluster serves as a docking point for arginine found in both the HFRW message sequence of the agonist peptides and the RFF triplet of AGRP and the agouti protein (*42*, *43*). Thus, the protrusion of AGRP's active loop (Figure 5) may allow for the RFF triplet to extend into the receptor and reach this essential negatively charged cluster.

Although agonists and antagonists appear to share a common transmembrane MCR docking location, 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 (exoloops) connecting transmembrane helices are particularly important (*44*, *45*). As discussed previously, AGRP exhibits high affinity for MC4R but little affinity for MC1R. However, chimeric MC1R containing the second and third exoloops from MC4R results in a receptor that now binds AGRP with high affinity. In addition, chimeric MC4R with MC1R exoloops loses its high affinity for AGRP.

Taken together, these mutagenesis experiments suggest that AGRP makes at least two contacts with MC receptors. One of these contacts is between the agouti/AGRP RFF triplet and the MCR transmembrane charge cluster. The other contact involves the second and third MCR exoloops. Given that the ordered N-terminal loop of AGRP(87−132) is a putative docking interface, we propose that this loop directly interacts with the second and third exoloops of MC3R and MC4R. In this scenario, AGRP's selectivity for MC3R and MC4R is therefore determined by complementary interactions between its N-terminal loop and the second and third receptor exoloops.

CONCLUSION

This study has determined the structure of AGRP(87−132) and shows that this important domain contains a region assignable to the ICK fold. A turn within a β hairpin presents the RFF triplet established to be essential for MCR binding, most likely in a transmembrane region. It is proposed that an additional MCR binding site involves the N-terminal loop of AGRP(87−132). This loop may interact with MCR exoloops and confer AGRP's unique selectivity for MC3 and MC4 receptors.

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 NMR spectrometer and assistance with data acquisition. We also thank Dr. Sean Cahill for technical assistance with structure calculations and Dr. Stephen Kent for critical reading of the manuscript.

REFERENCES

1. Friedman, J. M. (2000) *Nature 404*, 632−4.
2. Mokdad, A. H., Serdula, M. K., Dietz, W. H., Bowman, B. A., Marks, J. S., and Koplan, J. P. (1999) *JAMA, J. Am. Med. Assoc. 282*, 1519−22.
3. 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*, 131−41.
4. 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*, 3518−21.
5. Vaisse, C., Clement, K., Durand, E., Hercberg, S., Guy-Grand, B., and Froguel, P. (2000) *J. Clin. Invest. 106*, 253−62.
6. 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.
7. 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.
8. Ollmann, M. M., Wilson, B. D., Yang, Y. K., Kerns, J. A., Chen, Y., Gantz, I., and Barsh, G. S. (1997) *Science 278*, 135−8.
9. Ollmann, M. M., Lamoreux, M. L., Wilson, B. D., and Barsh, G. S. (1998) *Genes Dev. 12*, 316−30.
10. Yang, Y. K., Thompson, D. A., Dickinson, C. J., Wilken, J., Barsh, G. S., Kent, S. B., and Gantz, I. (1999) *Mol. Endocrinol. 13*, 148−55.
11. Nijenhuis, W. A., Oosterom, J., and Adan, R. A. (2001) *Mol. Endocrinol. 15*, 164−71.
12. 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*, 12172−7.
13. 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*, 16041−52.
14. Young, Y., Zeni, L., Rosenfeld, R. D., Stark, K. L., Rohde, M. F., and Haniu, M. (1999) *J. Pep. Res. 54*, 514−21.
15. Robbins, L. S., Nadeau, J. H., Johnson, K. R., Kelly, M. A., Rosellirehfuss, L., Baack, E., Mountjoy, K. G., and Cone, R. D. (1993) *Cell 72*, 827−834.
16. Yen, T. T., Gill, A. M., Frigeri, L. G., Barsh, G. S., and Wolff, G. L. (1994) *FASEB J. 8*, 479−488.
17. Shutter, J. R., Graham, M., Kinsey, A. C., Scully, S., Lüthy, R., and Stark, K. L. (1997) *Genes Dev. 11*, 593−602.
18. Yang, Y. K., Ollmann, M. M., Wilson, B. D., Dickinson, C., Yamada, T., Barsh, G. S., and Gantz, I. (1997) *Mol. Endocrinol. 11*, 274−80.
19. 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*, 255−64.
20. 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*, 125−31.
21. Dawson, P. E., Muir, T. W., Clarklewis, I., and Kent, S. B. H. (1994) *Science 266*, 776−779.
22. Smallcombe, S. H., Patt, S. L., and Keifer, P. A. (1995) *J. Magn. Reson., Ser. A 117*, 295−303.
23. Bartels, C., Xia, T. H., Billeter, M., Guntert, P., and Wuthrich, K. (1995) *J. Biomol. NMR 6*, 1−10.
24. Guntert, P., Braun, W., and Wuthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
25. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−98.
26. Güntert, P. (1998) *Q. Rev. Biophys. 31*, 145−237.
27. Williamson, M. P., Havel, T. F., and Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295−315.
28. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., GrosseKunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr., Sect. D 54*, 905−921.
29. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−86.
30. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
31. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley, New York.
32. Cavanagh, J. (1996) *Protein NMR spectroscopy: principles and practice*, Academic Press, San Diego, CA.
33. Roberts, G. C. K. (1993) *NMR of macromolecules: a practical approach*, IRL Press at Oxford University Press, Oxford, NY.
34. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR 5*, 67−81.
35. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis 18*, 2714−2723.
36. 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.
37. 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*, 105−16.
38. Isaacs, N. W. (1995) *Curr. Opin. Struct. Biol. 5*, 391−5.
39. Craik, D. J., Daly, N. L., and Waine, C. (2001) *Toxicon 39*, 43−60.
40. Norton, R. S., and Pallaghy, P. K. (1998) *Toxicon 36*, 1573−83.
41. Kiefer, L. L., Veal, J. M., Mountjoy, K. G., and Wilkison, W. O. (1998) *Biochemistry 37*, 991−7.
42. 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*, 14900−11.
43. Haskell-Luevano, C., Cone, R. D., Monck, E. K., and Wan, Y. P. (2001) *Biochemistry 40*, 6164−79.
44. Yang, Y. K., Dickinson, C. J., Zeng, Q., Li, J. Y., Thompson, D. A., and Gantz, I. (1999) *J. Biol. Chem. 274*, 14100−6.
45. Oosterom, J., Garner, K. M., den Dekker, W. K., Nijenhuis, W. A., Gispen, W. H., Burbach, J. P., Barsh, G. S., and Adan, R. A. (2001) *J. Biol. Chem. 276*, 931−6.

BI0117192