NMR structure of a minimized human agouti related protein prepared by total chemical synthesis

Kimberly A. Bolin[a], D. Joe Anderson[a], Julie A. Trulson[a], Darren A. Thompson[c], Jill Wilken[c], Stephen B.H. Kent[c], Ira Gantz[b], Glenn L. Millhauser[a,*]

[a] *Department of Chemistry and Biochemistry, University of California, Santa Cruz, CA 95064, USA*
[b] *Department of Surgery, University of Michigan Medical Center, 6504 MSRB I 1150 W. Medical Center Drive, Ann Arbor, MI 48109-0682, USA*
[c] *Gryphon Sciences, 250 East Grand Avenue, Suite 90, San Francisco, CA 94080-4824, USA*

Received 18 March 1999; received in revised form 8 April 1999

Abstract The structure of the chemically synthesized C-terminal region of the human agouti related protein (AGRP) was determined by 2D ^{1}H NMR. Referred to as minimized agouti related protein, MARP is a 46 residue polypeptide containing 10 Cys residues involved in five disulfide bonds that retains the biological activity of full length AGRP. AGRP is a mammalian signaling molecule, involved in weight homeostasis, that causes adult onset obesity when overexpressed in mice. AGRP was originally identified by homology to the agouti protein, another potent signaling molecule involved in obesity disorders in mice. While AGRP's exact mechanism of action is unknown, it has been identified as a competitive antagonist of melanocortin receptors 3 and 4 (MC3r, MC4r), and MC4r in particular is implicated in the hypothalamic control of feeding behavior. Full length agouti and AGRP are only 25% homologous, however, their active C-terminal regions are ~40% homologous, with nine out of the 10 Cys residues spatially conserved. Until now, 3D structures have not been available for either agouti, AGRP or their C-terminal regions. The NMR structure of MARP reported here can be characterized as three major loops, with four of the five disulfide bridges at the base of the structure. Though its fold is well defined, no canonical secondary structure is identified. While previously reported structural models of the C-terminal region of AGRP were attempted based on Cys homology between AGRP and certain toxin proteins, we find that Cys spacing is not sufficient to correctly determine the 3D fold of the molecule.

Key words: Agouti related protein; Minimized human agouti related protein; Nuclear magnetic resonance structure; Total chemical synthesis

1. Introduction

Recent biochemical investigations have identified agouti related protein (AGRP) as playing a major role in the regulation of mammalian feeding behavior. AGRP is a naturally occurring competitive antagonist of melanocortin receptors 3 and 4 (MC3r and MC4r), the overexpression of which results in adult onset obesity and diabetes in mice [1–3]. AGRP binding to MC4r in particular is the subject of intense interest since knockout mice that do not express MC4r exhibit the same phenotype as caused by overexpression of AGRP [2]. There is also evidence for the parallel expression of AGRP and neuropeptide Y in the arcuate nucleus of the hypothalamus, with neuropeptide Y known to stimulate feeding [3]. This region of the brain also expresses MC4r and is involved in energy homeostasis.

The growing body of evidence linking AGRP to weight control has yet to elucidate its exact mechanism of action. However, studies on AGRP do benefit from analogy to the much more widely studied agouti protein, as AGRP was originally identified through the homology of its C-terminal region with the same region of the agouti protein [1]. The agouti protein has been a focal point in obesity research for a number of years, since ectopic expression of the wild-type protein in mice leads to obesity and related disorders, i.e. the same symptoms as the overexpression of the more recently identified AGRP [4,5]. However, unlike AGRP, agouti has distinct expression patterns in mice and humans [6], making in vivo work with mice less applicable to human obesity disorders. AGRP, like agouti, is selective for MC3r and MC4r but has approximately 100-fold greater binding affinity than agouti at these receptors [7].

While full length agouti and AGRP are only 25% homologous, in their 46 residue Cys-rich C-terminal regions nine of the 10 Cys residues are spatially conserved and there are a further 10 identical residues giving ~40% sequence identity. Three consecutive, conserved residues RFF (111–113 in human AGRP) were determined to be essential to the biological activity of both agouti [8,9] and AGRP [10]. Two recent investigations have shown that the chemically synthesized C-terminal region of AGRP competitively antagonizes α-melanocyte stimulating hormone (α-MSH) at melanocortin receptors with equal or greater potency than the full proteins [11,12], consistent with similar findings for agouti [13]. Thus the C-terminal region of AGRP (MARP, Fig. 1) is a prime candidate for structural studies.

Despite the important biological activities of AGRP, no experimental 3D structure is available for this protein. The inhibitor cystine knot (ICK) family of proteins are also disulfide-rich and the structures of these invertebrate toxins have been used to suggest possible structures for the agouti and AGRP C-terminal regions [9,10]. Indeed, the recently reported

*Corresponding author. Fax: (1) (831) 459-2935.
E-mail: glennm@hydrogen.ucsc.edu

Abbreviations: $^3J_{HN\alpha}$, three bond αH-NH scalar coupling constant; α-MSH, α-melanocyte stimulating hormone; AGRP/ART, agouti related protein/agouti related transcript; CD, circular dichroism; Conformational shift, experimental chemical shift−random coil chemical shift; DQF-COSY, two-dimensional double-quantum filtered correlation spectroscopy; HX, hydrogen-deuterium exchange; ICK, inhibitor cystine knot; MARP, minimized human agouti related protein, residues 87–132 of human AGRP; MC3r/MC4r, melanocortin receptor 3/4; MRE, mean residue ellipticity; NDP-MSH, [Nle^4,D-Phe^7]α-MSH, a superpotent melanocortin agonist; nOe, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser spectroscopy; TOCSY, two-dimensional total correlation spectroscopy

PII: S0014-5793(99)00553-0

disulfide map for AGRP and a construct containing the C-terminal domain demonstrates ICK-like pairings [14] of the 10 Cys residues: 1-16, 8-22, 15-33, 19-43, 24-31 (using MARP numbering, see below) [15]. Beyond such modeling, the only structural data published on either agouti or AGRP are circular dichroism (CD) spectra which suggest that both proteins have little regular secondary structure, although there may be some β-sheet, consistent with ICK structural characteristics [13,16].

In this paper we report the 3D structure in solution of the human AGRP Cys-rich, C-terminal region as determined by ^{1}H NMR using protein prepared by total chemical synthesis. Because biochemical investigations demonstrate that this minimal region retains full biological activity, we refer to this protein as minimized agouti related protein, MARP (human AGRP residues 87–132). The protein's topology is characterized by three large loops and an absence of canonical secondary structure such as helices or sheets. Two of the three loops are structurally well characterized by the NMR data as indicated by low RMSDs. The region of MARP containing the RFF triplet [10] (residues 25–27 in MARP) necessary for function is located in one of the best defined regions of the protein. Despite the conservation of Cys spacing and the disulfide map between MARP and other small disulfide-rich proteins from the ICK family, it is apparent from the structure reported here that MARP does not adopt an ICK-like fold.

2. Materials and methods

2.1. Chemical protein synthesis

N^{α}-Acetyl-MARP and N^{α}-acetyl-MARP(Arg25Ala) were synthesized, folded and purified to each give a protein containing five disulfide bonds, as reported previously [12]

2.2. CD experiments

CD spectra were recorded at 25°C on an Aviv 60DS spectropolarimeter in a rectangular 1 mm path length cuvet for concentrations up to 60 μM, for concentrations higher than this a round cell with a 0.1 mm path length was used. All CD samples were 50 mM potassium phosphate, pH 4.25. Concentration dependence was ruled out in the range 20 μM–1 mM. Temperature dependence was determined for 5–85°C. The spectra are superimposable from 5 to 45°C.

2.3. NMR sample preparation

The activity of MARP used for the NMR sample and that of a single mutant were assayed by measuring the inhibition of cAMP production in the presence of [Nle^4,D-Phe^7]α-MSH (NDP-MSH) [12] in HEK-293 cells transfected with human MC4r. Control experiments were performed with no MARP. The NMR samples were found to be native-like with complete activity. NMR samples contained approximately 1.9 mM MARP at pH 4.2 in 50 mM KH_2PO_4 buffer in 90% H_2O/10% D_2O. Additional samples for hydrogen-deuterium exchange (HX) experiments were prepared by lyophilization of protonated samples followed by reconstitution in 700 μl D_2O.

2.4. NMR experiments

^{1}H 2D NMR spectra were principally acquired at 15°C on a Varian 500 Unity Plus spectrometer using inverse probes. NMR data were routinely acquired with a 6000 Hz spectral width, 4096 complex points in t_2 and 512 (TOCSY/DQF-COSY) or 700 (NOESY) increments in t_1. All spectra were processed using the MNMR package (Carlsberg Laboratory, Department of Chemistry, Denmark) and analyzed using XEASY [17], with chemical shifts referenced to 1,4-dioxane at 3.743 ppm. Sequential assignments of all backbone and $>90\%$ of side chain protons were accomplished using standard methods [18,19] for 50 ms TOCSY, 150 ms NOESY and DQF-COSY data. Additional data sets were acquired at 25°C and 30°C to resolve ambiguities. Examination of the three Pro residues identified nOes consistent only with trans-Pro. Four additional peaks were identified in the αN region of the TOCSY spectrum, however, associated spin systems could not be identified and neither could nOes to the peaks in question.

NOESY data for distance restraints were collected at 15°C using the WET sequence [20] for water suppression, 1.6 s recycle delay and a mixing time of 80 ms in both H_2O and D_2O. $^3J_{HN\alpha}$ coupling constants at 25°C were determined by both linear least squares fitting of the antiphase doublets in a DQF-COSY and also using the INFIT [21] module of XEASY with 150 ms NOESY data. These methods agreed to within ± 0.5 Hz for all of the measured coupling constants. At 15°C larger intrinsic linewidths precluded accurate measurement of $^3J_{HN\alpha}$.

For amide exchange experiments, the magnet was preshimmed on a 21 residue peptide sample at pH 4 in D_2O/phosphate buffer. The first TOCSY experiment was begun 23 min after reconstituting the protonated sample in D_2O. Four TOCSY experiments identical to those described above, except for the number of t_1 increments, were acquired back to back at 15°C over a period of 24 h. The first three consisted of 150 t_1 increments and the final experiment 300 increments. NOESY and DQF-COSY spectra were also acquired as described above.

2.5. Structure calculations

Final structure calculations included the covalent connectivity of the published disulfide map [15], and were based on a total of 414 interproton distance constraints derived from the 80 ms 2D NOESY spectra and 34 backbone Φ dihedral angle constraints derived from coupling constant measurements, giving a total of 448 total restraints, or 9.7 restraints per residue. The distance restraints can be broken down into 228 intraresidue (backbone to side chain only), 129 sequential, 20 medium range ($1 < |i-j| \leq 5$) and 37 long range ($|i-j| > 5$) restraints. These restraints were assigned as strong, medium or weak. The total numbers of restraints in each category were 95 strong, 246 medium, and 77 weak. All categories had a lower limit of 1.6 Å, with upper limits of 2.8, 3.5 and 5 Å for the strong, medium and weak categories, respectively. Trial structures were generated using the simulated annealing protocol from CNS version 0.4a (anneal.inp) with SUM averaging for the nOe distances [22−26]. Twenty structures with no bond or nOe angle violations were used to represent the solution structure of MARP (see Table 1 for RMSDs). Structures were displayed using MOLMOL [27].

3. Results

3.1. Chemical synthesis and characterization of MARP

The synthesis and biochemical characterization of MARP were previously reported [12]. The N-terminal residue (Cys-1) of MARP corresponds to the first Cys (Cys-87) of the Cys-rich region (Fig. 1) in full length, 132 residue human AGRP. The material used for MARP NMR sample shows native-like activity as measured by its ability to competitively inhibit NDP-MSH at MC4r (Fig. 2a), as has been shown in previous studies [12,28]. Previous mutational studies of agouti and AGRP showed residues Arg-25, Phe-26, Phe-27 (the RFF triplet [10]) and Asp-17 (in agouti) to be determinants of receptor binding [8–10]. In the work reported here, the replacement of Arg-25 by Ala results in a complete loss of inhibitory activity (Fig. 2a).

The far-UV CD spectrum of MARP in Fig. 2b is similar to that reported for a similar C-terminal fragment (85–132) of

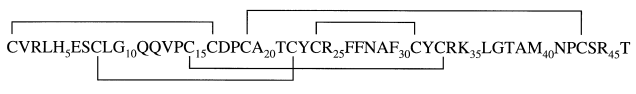

Fig. 1. Covalent structure of MARP, corresponding to the Cys-rich C-terminus of the human agouti related protein. (Human AGRP numbering is obtained by adding 86 to MARP numbering.) Disulfide bonds are indicated by lines joining Cys residues [15].

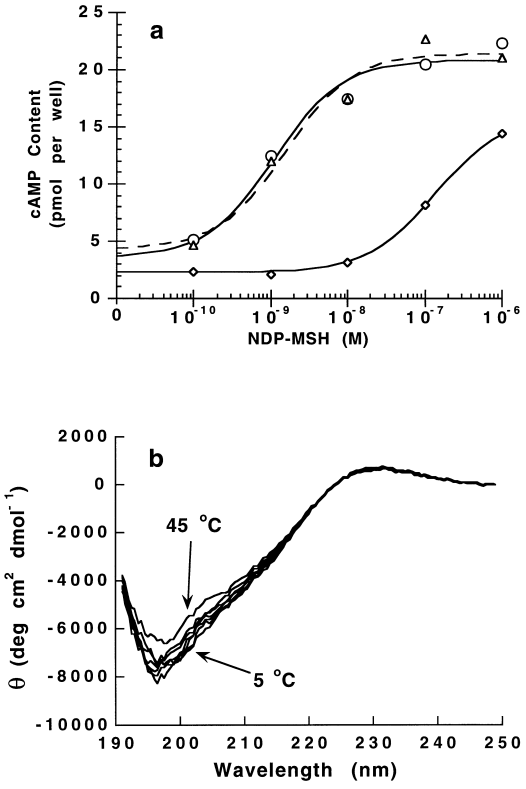

Fig. 2. a: Inhibition of NDP-MSH stimulated cAMP generation in cells transfected with MC4r. Control experiment with no MARP (○); MARP added (5.0×10^{-6} M) demonstrating competitive inhibition of NDP-MSH (◇); MARP with Arg-25 to Ala substitution added (5.0×10^{-6} M) demonstrating loss of inhibition due to a mutation in the active loop (△). b: Far-UV CD spectra of MARP in 50 mM phosphate buffer at pH 4.25 as a function of temperature with spectra shown every 5°C from 5°C to 30°C and also at 45°C. ←

AGRP reported elsewhere [16] and is characterized by a negative maximum at approximately 198 nm and a slight negative maximum at 245 nm. There is little indication of canonical α-helix, however the spectrum does perhaps suggest some β character or turns. The near-UV CD spectrum (not shown), indicative of tertiary structure, shows a weak minimum at approximately 275 nm which can be attributed to the disulfide bonds and possibly restricted orientations of the side chains of the Tyr-23 and Tyr-32.

At 25°C between 20 μM and 1.0 mM there is no observable concentration dependence as measured by CD. Between 1 mM and 1.9 mM there are no concentration dependent changes in the NMR spectra (i.e. linewidths, chemical shifts, etc. all remain constant). The far-UV CD spectrum remains constant between 5°C and 45°C (Fig. 2b), a temperature range well beyond that of the present NMR experiments. Thus, by all indications, MARP exists as a monomer and does not exhibit temperature dependence under the conditions of the NMR experiments. The characteristics of the NMR spectra are indicative of a well folded protein with a single predominant conformer. $^{3}J_{NH\alpha}$ coupling constants and temperature coeffi-

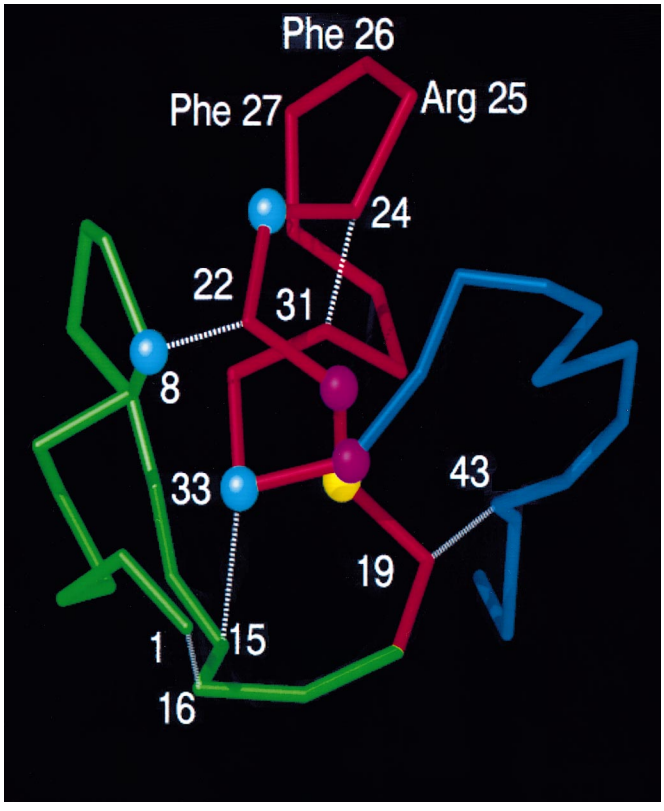

Fig. 3. C_α backbone of the MARP minimized average structure. The N-terminal loop is in green, central loop in red, C-terminal loop in blue. Disulfide bonds are represented by dashed lines. Spheres represent residues with amides protected from HX, cyan >12 h, yellow >24 h and magenta >8 days.

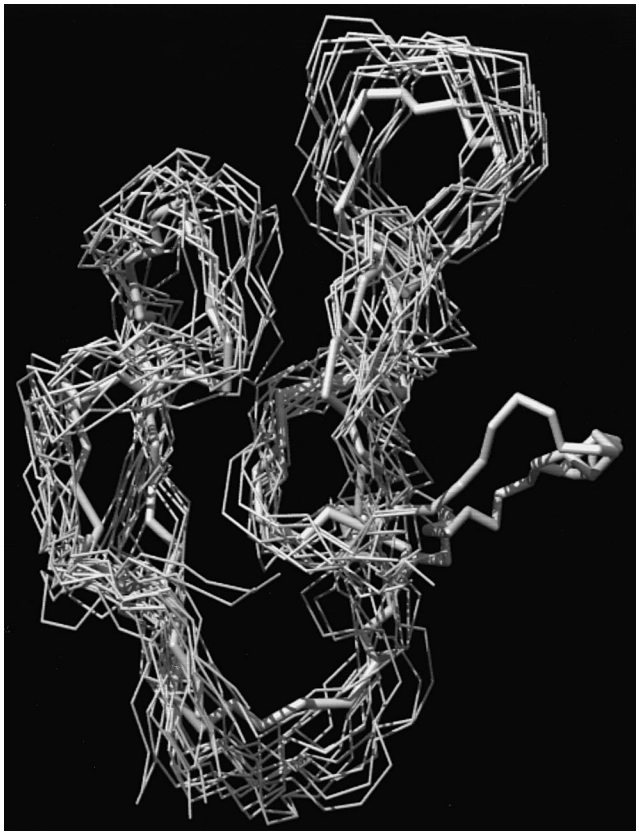

Fig. 4. All atom backbone representation of MARP for 14 NMR structures with residues 1–34 fit to the minimized average structure (RMSD 1.49 Å). Only the minimized average structure (thick cylinder) is shown for the more disordered C-terminal loop.

cient also indicate a fully folded, non-random coil conformer (see Appendix).

3.2. Structural description of MARP

The minimized average NMR structure of MARP is shown in Fig. 3. Consistent with the far-UV CD spectrum, MARP shows little evidence of helix or sheet secondary structure. The disulfide bonds (1-16, 8-22, 15-33, 19-43 and 24-31) appear to form a scaffold upon which the structure is apportioned into three major loops, which we refer to as the N-terminal loop (residues 1–18), the central loop (residues 19–34) and the C-terminal loop (residues 35–46) which are indicated by different colors in Fig. 3. RMSDs for the individual loops are reported in Table 1. The N-terminal and central loops are much better defined both within the loops and with respect to each other than the C-terminal loop. The backbone RMSD for the entire protein (2.54 Å) is of the same order as that of the C-terminal loop (2.36 Å), while the backbone RMSD for residues 1–34 (1.66 Å) is of the order of the individual N-terminal and central loops. To demonstrate limited backbone structure variability of the N-terminal and central loops, a superposition of 14 structures (selected for clarity) for residues 1–34 and the MARP minimized average structure (residues 1–46) is shown in Fig. 4.

Four of the five disulfide bonds are located at the base of the structure where they appear to pinch together the bottoms of the loops to form the 'core' of the protein (Fig. 3). The exception is disulfide bond 24-31 which stabilizes the central loop. The central loop, residues 19–34, contains the RFF triplet determined to be critical for activity. This motif is situated within an even smaller, well defined loop bound by Cys-24 and Cys-31 which we refer to as the 'active' loop. The side chain atoms of the RFF triplet residues are located at the surface of the protein as depicted in Fig. 5. Recent experiments further highlight the importance of this active loop. These studies demonstrate that short cyclic peptides corresponding to residues 24–31 of human AGRP do in fact antagonize MC3r and MC4r [10].

Inspection of the family of NMR structures and consideration of the observed HX (Fig. 6) reveal a structure for the central loop that can be best described as an irregular hairpin with a well defined loop from Cys-24 to Cys-31 (RMSD 0.6 Å Fig. 5) and a stem region which is both twisted around and curved along its *z*-axis (Fig. 3). This characterization is supported by critical examination of the nOe, $^{3}J_{HN\alpha}$ and chemical shift data (Fig. 6). As shown in Fig. 5, the active loop is highly constrained with the RFF triplet side chains exposed to solvent. Arg-25 and Phe-27 point out into the solvent, while one face of the Phe-26 aromatic ring rests parallel against the surface of the protein. Though the active loop satisfies several of the determinants for an Ω-loop [29], the side chain orientation of Arg-25 and Phe-27 precludes its definition as such since Ω-loop side chains generally pack within the loop of backbone atoms.

HX experiments demonstrate that the amide protons of residues Cys-8, Ala-20, Thr-1, Tyr-23, Tyr-2, Cys-3 and Arg-34 are protected from exchange with solvent. To explore whether these results are consistent with the average structure in Fig. 3, the program DSSP [30] was used to identify potential hydrogen bonds. DSSP identified the backbone amides of Ala-20, Thr-21, Tyr-23, and Arg-34 as potential hydrogen bond donors. In addition, solvent accessible surface area calculations showed that residues Cys-8 and Cys-33 are completely buried from solvent, though in the D_2O spectrum the αN crosspeaks of these two residues overlap, thus their individual protection from exchange is uncertain. Tyr-32 has only 8% solvent accessible surface area at the C_δ protons.

The NMR structure gives a well resolved fold, however, as

Table 1
Summary of MARP backbone and heavy atom RMSDs

Region (residues)	Backbone RMSD[a] (Å)	Heavy atom RMSD[a] (Å)
Global (1-46)	2.54	3.26
N- and active loops (1–34)	1.66	2.38
N-terminal loop (1–18)	1.31	2.03
Central loop (19–34)	1.51	2.22
Active loop (24–31)	0.69	1.53
C-terminal loop (35–46)	2.36	3.43

[a]Determined by fitting the family of 20 NMR structures to the minimized average structure.

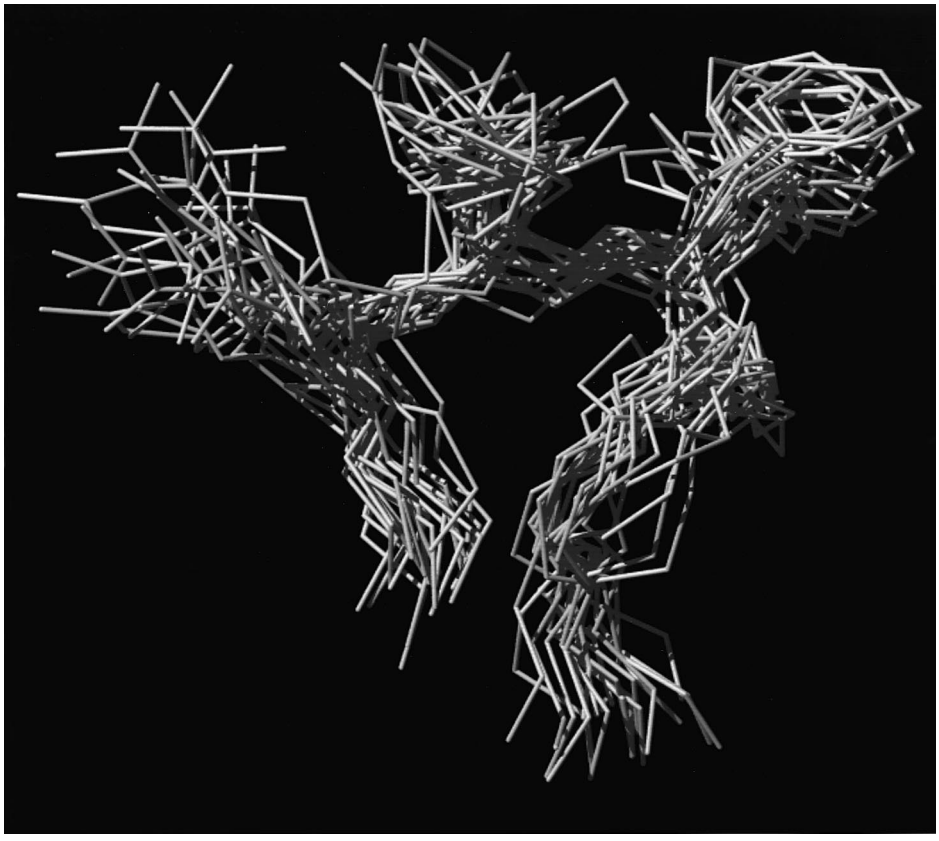

Fig. 5. Backbone atoms for residues 24–31 of the family of 20 structures with residues 24–31 fit to the minimized average structure. The side chain heavy atoms of residues 25, 26 and 27, essential for activity, are shown.

mentioned previously, canonical helices and β-sheets were not identified on the basis of nOes or other protocols including the chemical shift index [31] or $^3J_{NH\alpha}$ coupling constants [18].

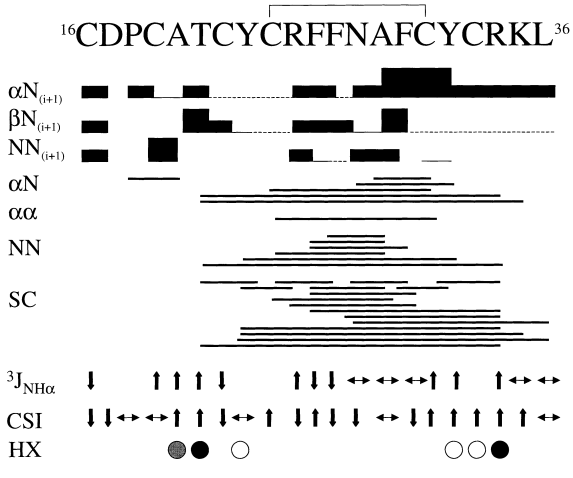

Fig. 6. nOe diagram for the central loop and flanking residues from the 80 ms NOESY data. For sequential nOes the height of the bar indicates the strength of the nOe, dashed lines indicate overlapping nOes. For all other nOes, only non-sequential non-overlapping nOes are reported, intensity is not indicated. For $^3J_{NH\alpha}$ up arrows indicate $J > 8$ Hz, down arrows $J < 6$ Hz and horizontal arrows $6 < J < 8$ Hz. For chemical shift index (CSI) up arrows indicate $\Delta\delta_{H\alpha} > 0.1$ ppm, down arrows $\Delta\delta_{H\alpha} > |-0.1|$ ppm, horizontal arrows 0.1 ppm $< \Delta\delta_{H\alpha} < |-0.1|$ ppm. For HX, open circles represent HX protection for >12 h, gray circles >24 h and black circles >8 days.

The guidelines for these protocols assign secondary structure on the basis of four or more consecutive residues with similar conformational shifts or $^3J_{HN\alpha}$. Helical structure is characterized by $^3J_{HN\alpha} < 6$ Hz and negative conformational shifts and β-sheet by $^3J_{HN\alpha} > 8$ Hz and positive conformational shifts. As shown in Fig. 6, even in the active loop and stem region of the central loop, no regular secondary structure is identified by these criteria. However, the chemical shift index points towards a possible extended strand from residue 31 to residue 35.

4. Discussion

The 3D structure of MARP is characterized by three loops held together at the base by an apparent scaffold of four disulfide bonds 1-16, 8-22, 15-33 and 19-43. The fifth disulfide bond, 24-31, further stabilizes the base of the active loop which presents the RFF triplet on the protein surface. There is no identifiable canonical helical or sheet structure. It is clear from biochemical data that the RFF triplet is critical for the activity of MARP as a competitive antagonist of α-MSH stimulated activation of MC4r signaling. The structure presented here shows that MARP is structured to present the side chains of the RFF triplet on the surface of the protein and to the surrounding solvent. Recent work demonstrates that MARP is much more active than smaller AGRP derived peptides containing the RFF triplet [10]. Thus, the detailed fold of the central loop and perhaps the presence of the N- and C-terminal loops are critical for AGRP function. In addition, based on work with chimeras of melanocortin receptors (I.

Gantz, unpublished data), it is possible that the N- and C-terminal loops may confer receptor subtype specificity. Further work is needed to support this hypothesis.

The previous absence of structural data on both AGRP and agouti encouraged the modeling of the C-terminal regions of these proteins onto the ICK family [14] which is characterized by homologous Cys spacing [9,10]. The ICK family of proteins primarily consists of small (<60 residues) disulfide-rich (three or four disulfides) toxin proteins from the venom of spiders and cone snails, which function as ion channel antagonists [14]. The coincidence between the function of the majority of these toxins and the recent description of part of the agouti protein's mechanism of action being calcium dependent [32–34] further encouraged these homology modeling efforts. The ICK motif in particular is characterized by the topology of the three disulfide bonds corresponding to 1-16, 8-22 and 15-33 in MARP. In the ICK motif the first two disulfide bonds with their intervening main chain atoms form a topological circle through which the third disulfide bond passes, forming the *cystine knot* [14]. The motif is further characterized by the identification of an irregular triple stranded antiparallel β-sheet, roughly corresponding to residues 6–8, 20–24 and 31–34 in MARP. The remaining two disulfide bonds in MARP each occur in individual ICK proteins as separate examples of potential 'non-motif' disulfide bonds (though it should be noted that so far there are no examples of ICK motif proteins with five disulfide bonds).

Despite these apparent similarities, the experimental structure of MARP shows that this protein does not satisfy the criteria required for inclusion in the ICK family. While the first two disulfide bonds in MARP, 1-16 and 8-22, together with the polypeptide backbone form a topological circle, none of the remaining disulfides passes through the circle to form a cystine knot. Instead, disulfide bond 15-33 is positioned adjacent to the circle with all of the fold on one side of this circle. In addition, MARP lacks the β-sheet found in ICK family proteins. The experimental determination of the distinctive 3D structure of MARP reported here suggests that Cys spacing and even the disulfide map of small Cys-rich proteins may not always be sufficient to accurately predict protein folds. These results speak to the potential limitations of 'homology modeling' of protein structures, and may have important implications for the emerging field of genomic structural biology.

Acknowledgements: This work was supported by grants from the NIH (GM46870), the UC Biotechnology Program (97-18) and the Petroleum Research Fund (32164-AC4) administered by the ACS. K.A.B. and G.L.M. thank the donors of the PRF for their support. NMR instrumentation was supported in part by the Elsa U. Pardee Foundation and a grant from the W.M. Keck Foundation. The authors wish to thank Mr. Jim Loo for technical assistance in the acquisition of NMR data and Mr. Eliah Aronoff-Spencer for technical assistance.

Appendix A. Supporting information available

Tables of the ^{1}H chemical shifts at 15°C, $^3J_{NH\alpha}$ at 25°C, and NH temperature coefficients. Ordering information is given on any current masthead page.

Supporting information for microfilm edition

MARP ^{1}H chemical shifts at 15°C in 50 mm phosphate buffer at pH 4.2

Residue	NH (ppm)	H_α (ppm)	H_β (ppm)	Other (ppm)
Cys-1	8.00	5.00	3.12, 2.81	
Val-2	8.92	4.18	1.94, 2.08	CH_3^γ 0.91, 0.80
Arg-3	8.91	4.01	1.67, 1.95	H^γ1.69, 1.82, H^δ 3.24 3.26, NH 7.31
Leu-4	8.13	3.72	1.34, 1.54	H^γ1.03, CH_3^δ 0.79
His-5	8.99	4.25	3.70 3.42	H^δ 7.22, H^ε8.51
Glu-6	8.04	4.66	2.16	H^γ2.24
Ser-7	8.63	4.68	3.84	
Cys-8	8.08	4.96	3.57 3.13	
Leu-9	7.92	4.08	1.57, 1.44	H^γ1.44, CH_3^δ0.82
Gly-10	8.78	4.07, 3.69		
Gln-11	8.06	4.24	1.95	H^γ2.28, $H^{\varepsilon 2}$ 6.86, 7.49
Gln-12	8.68	4.27	2.20, 1.95	H^γ2.30, 2.34, $H^{\varepsilon 2}$ 6.82, 7.45
Val-13	7.48	4.47	2.05	CH_3^γ0.79, 0.90
Pro-14	8.81	4.55	2.25, 1.99	H^γ1.84, 1.99, H^δ3.65, 3.77
Cys-15		4.92	3.35, 2.74	
Cys-16	9.58	4.22	2.62, 3.17	
Asp-17	8.19	4.78	2.63 2.41	
Pro-18	8.88	4.50	2.35	H^γ1.98, H^δ3.89, 4.04
Cys-19		4.72	2.94, 3.47	
Ala-20	8.01	4.94	1.26	
Thr-21	8.76	4.58	4.00	CH_3^γ1.16
Cys-22	8.94	4.58	2.94, 3.02	
Tyr-23	8.71	4.62	2.79	H^δ6.93 H^ε 6.74
Cys-24	8.30	4.91	3.23, 2.60	
Arg-25	8.29	3.84	1.83, 1.55	H^γ1.56, 1.23, H^δ3.06
Phe-26	7.91	4.74	2.80, 3.32	H^δ7.26 H^ε 7.42
Phe-27	8.54	4.19	3.12, 3.02	H^δ7.17, H^ε7.33
Asn-28	8.48	4.20	2.39, 2.77	$H^{\delta 2}$ 6.66, 7.30
Ala-29	7.75	4.38	1.25	
Phe-30	8.38	4.21	3.32, 3.36	
Cys-31	8.34	5.63	2.59, 3.03	
Tyr-32	8.88	5.20	2.59, 2.81	H^δ6.92, H^ε 6.66
Cys-33	8.19	4.96	3.21, 2.66	
Arg-34	9.43	4.75	1.83, 1.70	H^γ1.57, 1.69, H^δ2.62, 2.88, NH 7.08
Lys-35	9.09	4.47	1.82, 1.68	H^γ1.25, 1.43, H^δ1.64, H^ε2.89
Leu-36	8.77	4.3640	1.63	H^γ1.52, CH_3^δ0.82, 0.70
Gly-37	8.48	4.09, 4.01		
Thr-38	7.82	4.42	4.42	CH_3^γ1.21
Ala-39	8.53	4.16	1.42	
Met-40	8.06	4.40	1.95, 2.09	H^γ2.51, 2.62
Asn-41	7.74	5.00	2.62, 2.81	$H^{\delta 2}$ 7.67, 6.95

MARP ^{1}H chemical shifts at 15°C in 50 mm phosphate buffer at pH 4.2

Residue	NH (ppm)	H_α (ppm)	H_β (ppm)	Other (ppm)
Pro-42	8.68	4.47	2.29	H^γ1.94, H^δ3.61, 3.67
Cys-43		4.59	3.22, 3.13	
Ser-44	8.37	4.46	3.86	
Arg-45	8.26	4.43	1.93, 1.75	H^γ1.66, H^δ3.21, NH 7.25
Thr-46	7.86	4.15	4.23	CH_3^γ1.15

MARP coupling constants at 25°C and NH temperature coefficients in 50 mm phosphate buffer at pH 4.2

Residue	$^3J_{NH\alpha}$ coupling constant (Hz)	NH temperature coefficient (ppb/°C)
Cys-1		6.93
Val-2	8.79	3.63
Arg-3	7.00	6.37
Leu-4	5.64	3.47
His-5	7.50	2.58
Glu-6	9.34	2.90
Ser-7		5.86
Cys-8	9.92	3.84
Leu-9	4.40	2.43
Gly-10		7.06
Gln-11	8.14	4.94
Gln-12	8.00	4.38
Val-13	9.03	5.56
Pro-14		0.00
Cys-15	4.33	6.74
Cys-16	5.42	3.10
Asp-17		4.81
Pro-18		0.00
Cys-19	8.81	3.84
Ala-20	9.64	3.23
Thr-21	9.36	4.04
Cys-22	5.76	7.56
Tyr-23		2.01
Cys-24		3.14
Arg-25	5.90	2.93
Phe-26	8.70	3.67
Phe-27	3.68	6.21
Asn-28	7.82	5.66
Ala-29	7.56	2.10
Phe-30	7.59	4.84
Cys-31	9.17	5.84
Tyr-32	9.49	4.43
Cys-33		1.13
Arg-34	9.34	2.74
Lys-35	7.25	8.80
Leu-36	7.74	6.43
Gly-37		7.00
Thr-38	8.07	2.07
Ala-39	4.48	7.67
Met-40	7.50	2.98
Asn-41	8.19	1.08
Pro-42		0.00
Cys-43	7.24	7.86
Ser-44	7.26	8.38
Arg-45	7.54	4.47
Thr-46		4.61

References

[1] Shutter, J.R., Graham, M., Kinsey, A.C., Scully, S., Luthy, R. and Stark, K.L. (1997) Genes Dev. 11, 593–602.
[2] 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.
[3] Hahn, T.M., Breininger, J.F., Baskin, D.G. and Schwartz, M.W. (1997) Nature Neurosci. 1, 271–272.
[4] Klebig, M.L., Wilkinson, J.E., Geisler, J.G. and Woychik, R.P. (1995) Proc. Natl. Acad. Sci. USA 92, 4728–4732.
[5] Michaud, E.J., Mynatt, R.L., Miltenberger, R.J., Klebig, M.L., Wilkinson, J.E., Zemel, M.B., Wilkison, W.O. and Woychik, R.P. (1997) J. Endocrinol. 155, 207–209.
[6] Wilson, B.D., Ollmann, M.M., Kang, L., Stoffel, M., Bell, G.I. and Barsh, G.S. (1995) Hum. Mol. Genet. 4, 223–230.
[7] Fong, T.M., Mao, C., MacNeil, T., Kalyani, R., Smith, T., Weinberg, D., Tota, M.R. and Van der Ploeg, L.H. (1997) Biochem. Biophys. Res. Commun. 237, 629–631.
[8] Kiefer, L.L., Ittoop, O.R., Bunce, K., Truesdale, A.T., Willard, D.H., Nichols, J.S., Blanchard, S.G., Mountjoy, K., Chen, W.J. and Wilkison, W.O. (1997) Biochemistry 36, 2084–2090.
[9] Kiefer, L.L., Veal, J.M., Mountjoy, K.G. and Wilkison, W.O. (1998) Biochemistry 37, 991–997.
[10] Tota, M.R., Smith, T.S., Mao, C., MacNeil, T., Mosley, R.T., Van der Ploeg, L.H.T. and Fong, T.M. (1999) Biochemistry 38, 897–904.
[11] Quillan, J.M., Sadee, W., Wei, E.T., Jimenez, C., Ji, L. and Chang, J.K. (1998) FEBS Lett. 428, 59–62.
[12] Yang, Y.-K., Thompson, D.A., Dickinson, C.J., Wilken, J., Barsh, G.S., Kent, S.B.H. and Gantz, I. (1999) Mol. Endocrinol. 13, 148–155.
[13] Willard, D.H., Bodnar, W., Harris, C., Kiefer, L., Nichols, J.S., Blanchard, S., Hoffman, C., Moyer, M., Burkhart, W. and Weiel, J. et al. (1995) Biochemistry 34, 12341–12346.
[14] Norton, R.S. and Pallaghy, P.K. (1998) Toxicon 36, 1573–1583.
[15] 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.
[16] 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.
[17] Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) J. Biomol. NMR 5, 1–10.
[18] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley and Sons, New York.
[19] Redfield, C. (1993) in: NMR of Macromolecules: A Practical Approach (Roberts, G.K.C., Ed.), pp. 71–99, IRL Press at Oxford University Press, Oxford.
[20] Smallcombe, S.H., Patt, S.L. and Keifer, P.A. (1995) J. Magn. Reson. Ser. A 117, 295–303.
[21] Szyperski, T., Güntert, P., Otting, G. and Wüthrich, K. (1992) J. Magn. Reson. 99, 552–560.
[22] Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse Kunstleve, 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 Biol. Crystallogr. 54, 905–921.
[23] Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) FEBS Lett. 239, 129–136.
[24] Nilges, M., Gronenborn, A.M., Brunger, A.T. and Clore, G.M. (1988) Protein Eng. 2, 27–38.
[25] Kuszewski, J., Gronenborn, A.M. and Clore, G.M. (1996) J. Magn. Reson. Ser. B 112, 79–81.
[26] Stein, E.G., Rice, L.M. and Brunger, A.T. (1997) J. Magn. Reson. 124, 154–164.
[27] Koradi, R., Billeter, M. and Wüthrich, K. (1996) J. Mol. Graphics 14, 51–55.
[28] 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.
[29] Leszczynski, J.F. and Rose, G.D. (1986) Science 234, 849–855.
[30] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577–2637.
[31] Wishart, D.S., Sykes, B.D. and Richards, F.M. (1992) Biochemistry 31, 1647–1651.
[32] Kim, J.H., Mynatt, R.L., Moore, J.W., Woychik, R.P., Moustaid, N. and Zemel, M.B. (1996) FASEB J. 10, 1646–1652.
[33] Kim, J.H., Kiefer, L.L., Woychik, R.P., Wilkison, W.O., Truesdale, A., Ittoop, O., Willard, D., Nichols, J. and Zemel, M.B. (1997) Am. J. Physiol. 272, E379–384.
[34] Jones, B.H., Kim, J.H., Zemel, M.B., Woychik, R.P., Michaud, E.J., Wilkison, W.O. and Moustaid, N. (1996) Am. J. Physiol. 270, E192–196.