Determination of Disulfide Structure in Agouti-Related Protein (AGRP) by Stepwise Reduction and Alkylation

Edward J. Bures, John O. Hui, Yunjen Young, David T. Chow, Vishwanatham Katta, Michael F. Rohde, Lisa Zeni, Robert D. Rosenfeld, Kevin L. Stark, and Mitsuru Haniu*

Amgen, Inc., One Amgen Drive, Thousand Oaks, California 91320

Received May 11, 1998; Revised Manuscript Received July 9, 1998

ABSTRACT: The agouti-related protein gene (*Agrp*) plays an important role in body weight regulation. The mature human protein is a single polypeptide chain of 112 amino acid residues, consisting of an N-terminal acidic region and a unique C-terminal cysteine-rich domain. The disulfide structure of recombinant human AGRP was determined by chemical methods using partial reduction with tris(2-carboxyethyl)phosphine under acidic conditions, followed by direct alkylation with *N*-ethylmaleimide or fluorescein-5-maleimide. Partial reduction and alkylation provided several forms of AGRP that were modified in a stepwise fashion. The resulting proteins were characterized by peptide mapping, sequence analysis, and mass spectrometry, showing that AGRP contained a highly reducible disulfide bond, C85−C109, followed by less reactive ones, C90−C97, C74−C88, C67−C82, and C81−C99, respectively. The chemically defined disulfide connectivity of the recombinant human AGRP was homologous to that of ω-agatoxin IVB except for an additional disulfide bond, C85−C109.

Agouti-related protein gene (*Agrp*)[1] was first identified during a search for genes related to agouti in the public EST database (*1*). The human *Agrp* gene predicts a 132 amino acid protein which is 25% identical to human agouti. Both mature agouti and AGRP contain 10 cysteines, forming 5 disulfide linkages. *Agrp* is expressed primarily in the adrenal gland, subthalamic nucleus, and hypothalamus. The hypothalamic expression of *Agrp* was elevated approximately 10-fold in *ob*/*ob* and *db*/*db* mice (*1*, *2*). Overexpression of *Agrp* in transgenic mice leads to obesity, hyperinsulinemia, and hyperglycemia (*2*, *3*). The expression pattern and transcriptional regulation of *Agrp* suggest a role for the protein in the regulation of melanocortin receptors (MCRs) within the hypothalamus and implicate the *Agrp* gene product in the control of feeding (*2*). Agouti and AGRP are potent antagonists of central MCRs (*2*, *4*, *5*). The location of cysteine residues in agouti and AGRP is partially homologous to those of ω-conotoxin, spider toxin, or ω-agatoxin IVB despite the latters lacking one or two disulfide bonds (*1*, *5*−*9*). Based on these observations and site-directed mutagenesis, the tentative disulfide structure of agouti has been postulated to be a cystine knot similar to that of ω-agatoxin IVB or spider toxin plectoxin VIII (*5*, *7*). However, no experimental evidence has been reported for disulfide connectivity of agouti or AGRP.

As a first step toward the elucidation of the structure and function relationship of this protein, we attempted to determine its disulfide linkages. However, assignment of disulfide linkages, a critical aspect in protein structural characterization, is often a formidable task for the protein biochemist. The protein of interest is usually subjected to extensive proteolytic or chemical degradation; the peptides that are linked by a disulfide bond are to be purified and analyzed by sequencing or mass spectrometry (*10*, *11*). In the case of proteins that are highly cross-linked by disulfide bridges, well-known examples being the conotoxins from sea snails, Gray (*12*, *13*) introduced a selective reduction method using the chemical reagent tris(2-carboxyethyl)-phosphine (TCEP) under acidic conditions. The objective of his study was to reduce one or two of the disulfide bridges present in the protein of interest while leaving the remaining unmodified. The initially reduced disulfide was then determined by a combination of protein chemistry techniques. Because of the reactivity of free sulfhydryl groups, the partially reduced protein is usually quenched rapidly in concentrated iodoacetamide (2.2 M) at pH 8.0. Even under such conditions, disulfide scrambling has been reported (*12*). Therefore, a reagent that would alkylate free sulfhydryl groups under acidic conditions so that disulfide exchange would be minimized is of interest.

Based on such observations, Volkmann and colleagues (*14*) have employed *N*-(hydroxymethyl)benzamide, a reagent that would alkylate cysteines under acidic conditions, to modify the selectively reduced disulfide bridges found in ω-agatoxins IVB and IVC after TCEP treatment. However, this reagent is not readily available. In a recent study, Wu and Watson (*15*) reported a novel method for determining disulfide linkages, where they cyanylated the cysteinyl

* To whom correspondence should be addressed. Telephone: (805)-447-3117. FAX: (805)499-7464. E-mail: mhaniu@amgen.com.

[1] Abbreviations: *Agrp*, agouti-related protein gene; AGRP, agouti-related protein; TCEP, tris(2-carboxyethyl)phosphine; EST, expressed sequence tag; FM, fluorescein-5-maleimide; NEM, *N*-ethylmaleimide; IAA, iodoacetate; 4-HCCA, 4-cyanohydroxycinnamic acid; MCR, melanocortin receptor; DMSO, dimethyl sulfoxide; MALDI, matrix-assisted laser desorption ionization; HPLC, high-performance liquid chromatography.

S0006-2960(98)01082-4 CCC: $15.00
Published on Web 08/14/1998

residues of partially reduced proteins with the reagent 1-cyano-4-(dimethylamino)pyridinium tetrafluoroborate (CDAP) at pH 3.0. Subsequent cleavage of the peptide bonds at the N-terminal side of the cyanylated cysteines followed by complete reduction of the remaining disulfide bridges and mass spectrometric analysis allowed the identification of the initially reduced disulfide bonds. However, upon cleavage, the series of 2-iminothiazolidine-4-carboxylyl peptides are N-terminally blocked and thus refractory to Edman sequencing. In certain situations, as in the complete characterization of a recombinant protein for therapeutic use, it may be necessary to sequence through the modified cysteinyl residues. In this study, we show that the sulfhydryl-modifying agent *N*-ethylmaleimide (NEM) or fluorescein-5-maleimide (FM) (*16*−*18*) is able to alkylate free cysteines under acidic conditions. We demonstrate its use in identifying the highly reducible disulfide bond(s) as well as other disulfides in recombinant human AGRP.

MATERIALS AND METHODS

Materials. *N*-Ethylmaleimide and TCEP were purchased from Sigma Chemical Co. (St. Louis, MO). Fluorescein-5-maleimide (FM) was obtained from Molecular Probes (Eugene, OR). Thermolysin and pepsin were products made by Boehringer/Mannheim (Indianapolis, IN). 4-HCCA was purchased from Aldrich (Milwaukee, WI).

Preparation and Purification of Agouti-Related Protein. Preparation and purification of recombinant human AGRP are to be described elsewhere (*19*). The final product showed a single peak on reversed-phase HPLC and a single band on SDS−PAGE. Mass spectrometry and sequence analysis showed that the protein contained 109 amino acid residues initiated by M-K-A-P-M; therefore, it was named MK-des 5 AGRP.

Partial Reduction with TCEP and Alkylation with NEM or FM. The purified AGRP sample (90 μg) was dissolved in a denaturing buffer, consisting of a 3 M sodium acetate (pH 4.6)−8 M GdnHCl mixture (1:3 v/v), and was reduced in 25 mM TCEP at 37 °C for 10 min. Excess NEM was added at this point. After 60 min of reaction, the sample was directly subjected to reversed-phase (RP) HPLC using a Vydac C18 column (2.1 × 150 mm). The protein was eluted with a linear gradient from 2% to 65% solvent B over 60 min using solvent A (0.1% TFA) and solvent B (0.1% TFA−90% acetonitrile−9.9% water). At this point, the protein peaks were submitted to mass spectrometry. The protein peaks were further reduced with 20 mM DTT in 0.2 M Tris−6 M GdnHCl, pH 8.5, and carboxymethylated with excess iodoacetate. The protein was purified by reversed-phase HPLC as above.

Modification with fluorescein-5-maleimide (FM) was performed in a similar way except for the reducing conditions and the following procedures. The AGRP (100 μg) was dissolved in 0.1% TFA (100 μL) and reduced with 10 mM TCEP at 45 °C for 10−30 min. To the reduced sample was added 20 mM FM (dissolved in DMSO) together with an equal volume of 8 M GdnHCl. The solution was adjusted to pH 6.0 with 1 M Tris base, and the alkylation reaction was carried out at 25 °C for 60 min. The modified protein was passed through a PD-10 column (10 × 50 mm, Pharmacia) and eluted with 0.1 M Tris-HCl buffer, pH 7.5. The protein fractions were further subjected to HPLC purification using a Vydac C18 column (2.1 × 150 mm) and were eluted with a linear gradient from 35% to 60% solvent B. The protein was detected by the UV absorbance at 215 nm. The flow rate during the chromatography was 0.25 mL/min.

Complete Reduction and S-Alkylation of AGRP. The remaining disulfide bonds in partially reduced and alkylated AGRP samples were fully reduced with 10 mM DTT in 6 M Gdn HCl−0.2 M Tris-HCl, pH 8.0, at 45 °C for 60 min, followed by alkylation with 20 mM iodoacetate or 2 μL of neat 4-vinylpyridine at 25 °C for 30 min. The resulting proteins were purified by a reversed-phase HPLC using a Vydac C18 column (2.1 × 150 mm). The protein elution and the detection were performed as above.

Proteolytic Digestions of Partially Reduced and Alkylated AGRP. For tryptic digestion, the procedures of Stone and Williams (*19*) were followed. The intact protein or partially reduced and alkylated proteins were fragmented by proteolytic digestions using thermolysin or pepsin (2% w/w). Thermolytic digestion was performed in 0.1 M triethylamine hydrochloride buffer, pH 6.0, or 0.1 M Tris-HCl buffer, pH 7.5, at 37 °C for 12 h. Peptic digestion was performed in 0.02 N HCl (pH 2) at 37 °C for 12 h. The digested materials were directly subjected to HPLC using a Vydac C18 column (2.1 × 150 mm) as described later.

Peptide Separation by HPLC. Peptides were chromatographed through a Vydac C18 column (2.1 × 150 mm), using an HP1090 liquid chromatograph. Elution of peptides was performed by linear gradient system from 2% to 40% solvent B over 30 min, followed by second gradient from 40% to 60% solvent B over 10 min. Eluted peptides were detected by the absorbance at 215 nm, and fluorescein-labeled peptides were detected by the absorbance at 440 nm.

Sequence Analysis and Mass Spectrometry. Determination of the molecular weight of protein or peptide was by MALDI-mass spectrometry using either a Kratos MALDI IV for peptides or a PerSeptive Biosystems Voyager for protein samples. 4-HCCA was used as a matrix on which the protein sample was crystallized. Sequence analysis of the peptide was performed by an Applied Biosystems gas-phase sequencer, Model 477A or 494.

RESULTS AND DISCUSSION

We have attempted to establish techniques for determining the disulfide linkages of the AGRP. The primary structure of AGRP shows that the 10 cysteinyl residues are concentrated at the C-terminal region (*1*). Ellman's reagent failed to indicate any free sulfhydryl group, and because the protein migrated as a monomer under nonreducing conditions on SDS−PAGE, the 10 cysteinyl residues were therefore expected to exist in the forms of 5 intramolecular disulfide bridges. As expected from the disulfide structure, the AGRP was resistant to extensive proteolysis. Hence, another approach was necessary to identify the disulfide connectivity in AGRP.

Gray (*12, 13*) and others (*14, 20, 21*) have consistently demonstrated the power of the chemical reagent tris(2-carboxyethyl)phosphine in establishing complex disulfide linkages in proteins and peptides after partial reduction in acidic conditions. Hence, this approach was chosen in our

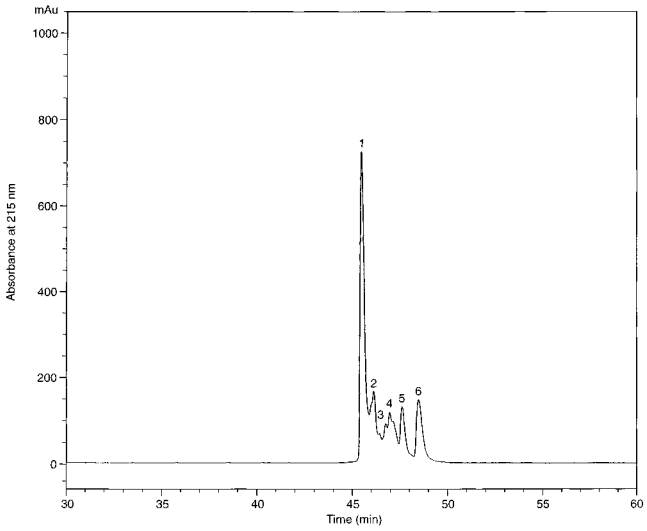

FIGURE 1: HPLC separation of partially reduced and NEM-modified AGRP. The AGRP sample was reduced with 10 mM TCEP and alkylated with NEM as described under Materials and Methods. The modified proteins were purified by reversed-phase HPLC using a Vydac C18 column (2.1 × 150 mm). The protein was eluted with a linear gradient from 2% to 65% solvent B over 60 min. Peaks 1−6 are denoted as NEM-AGRP1−6, respectively.

study on the disulfide connectivity in AGRP. However, because of the reactivity of the free sulfhydryl group, disulfide scrambling is always a concern to the protein chemist. During our literature search, we became aware of the early studies (*22*), on the alkylation of free sulfhydryl groups by the reagent NEM, where the experiments were actually performed under acidic conditions. Therefore, this is the approach that we have undertaken. The AGRP protein was partially reduced with TCEP in 1.0 M sodium acetate at pH 4.6, and to minimize disulfide exchange, the disulfide bonds initially reduced were directly alkylated with NEM in the same buffer. The remaining cystine bridges were completely reduced and differentially labeled with either 4-vinylpyridine or iodoacetate. In a previous communication, we have reported the chromatographic behavior of NEM-modified cysteines from PTH analysis (*23*), where the PTH derivative of NEM-alkylated Cys was eluted after Pro. The maleimide ring was susceptible to hydrolysis at alkaline pH, such as that employed for tryptic digestion; the PTH derivative of NEM-modified Cys was often eluted as multiple peaks. Hence, such residues derived from the partial reduction experiment of AGRP can be readily identified by peptide mapping and a combination of mass spectrometry and N-terminal sequence analysis. In this study, we demonstrate the feasibility of this approach, and we also show that the chromophoric reagent fluoresein-5-maleimide (FM) can be employed in a fashion similar to identify the reducible disulfides.

Purification of Partially Reduced and Alkylated AGRP. Native AGRP was reduced with TCEP under acidic conditions (pH 4.6) and alkylated directly with NEM as described under Materials and Methods. Examination of the products using reversed-phase HPLC shows that six peaks were obtained (Figure 1). To determine the number of cysteinyl residues modified by NEM (hence, the number of cystine bridges reduced), each purified peak was subjected to mass spectral analysis. Figure 2 shows mass spectral results of selected proteins. As summarized in Table 1, the values

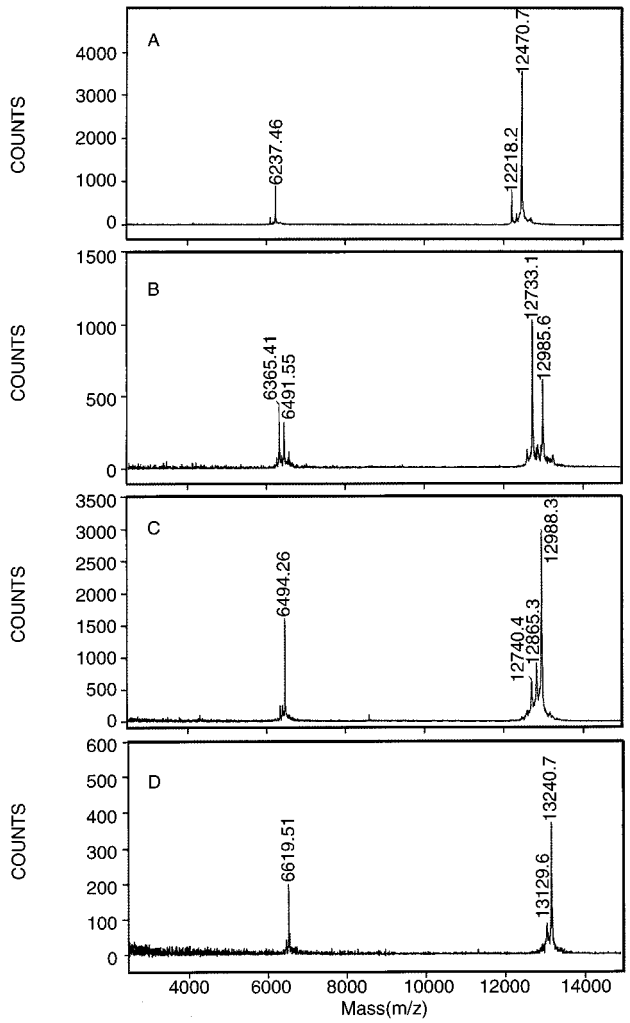

FIGURE 2: MALDI-mass spectrometry of NEM-modified AGRP proteins. Mass numbers were determined using a PerSeptive Biosystems Voyager with a matrix 4-HCCA. (A) NEM-AGRP1; (B) NEM-AGRP3; (C) NEM-AGRP4; and (D) NEM-AGRP5.

Table 1: Mass Spectrometry of NEM-Modified Proteins[a]

peak no.	obsd mass (amu)	no. of disulfide reduced/alkylated
1	(12 218.2)	intact protein
	12 470.7	one disulfide
2	12 729.0	two disulfides
3	12 733.1	two disulfides
	(12 985.6)	three disulfides
4	12 988.3	three disulfides
5	13 240.7	four disulfides
6	13 493.3	five disulfides

[a] Parentheses show minor components. The theoretical mass of MK-des 5 AGRP is 12 218.2.

ranged from one cystine bridge reduced (such as peak 1) to a completely reduced and NEM modified protein (peak 6). For instance, peak 1 (NEM-AGRP1) showed double signals of mass 12 218.2 and 12 470.7, corresponding to the intact protein and the partially alkylated protein, respectively.

Identification of the First Disulfide Bridge C85−C109. Results from mass spectrometry indicated that NEM-AGRP1 (Figure 2) contained two coeluted proteins from the HPLC column. The molecular mass of 12 218 Da was the unmodified AGRP (theoretical mass =12 218.2 Da), and the molecular mass of 12 470.7 Da was due to the protein after

Table 2: Sequence Analyses of NEM-Modified Peptides[a]

no.	peptide name	sequence	modified cysteine
1	AGRP-1B-T1, 2	LGTAMNP**C**SRT (102−112)	C109
2	AGRP-1B-T3, 4	LHESCLGQQVPCCDP**C**ATCY (70−89)	C85
3	AGRP-3-Pep 3a	**C**RF (90−92)	C90
4	AGRP-3-Pep 3b	**C**YCRKLGTAM (97−106)	C97

[a] Underlined boldfaced letters show the residue detected as an *N*-ethylsuccinimidocysteine. Other cysteines were detected as pyridylethylcysteine or intact cystine. Parentheses show the residue number.

Table 3: Sequence Results of Specific FM-Peptides[a]

no.	peak	sequence	FM-Cys
1	2a−d	A-M-N-P-**C**-S-R-T (105−112)	C109
		M-N-P-**C**-S-R-T (106−112)	C109
2	2e−h	L-G-Q-Q-V-P-Cm-Cm-D-P-**C**-A-T-Cm-Y-Cm-R (75−91)	C85
3	3a	F-**C** (96−97)	C97
4	3b−c	L-G-Q-Q-V-P-Cm-Cm-D-P-**C**-A-T-Cm-Y-**C**-R (75−91)	C85, C90
5	5a−b	L-H-E-S-**C** (70−74)	C74
6	5c−d	L-G-Q-Q-V-P-Cm-Cm-D-P **C**-A-T-**C**-Y-**C**-R (75−91)	C85, C88, C90
7	6a−b	S-R-R-**C** (64−67)	C67
8	6c	F-**C**-Y-**C**-R-K (96−101)	C97, C99[b]
9	6d−e	L-G-Q-Q-V-P-Cm-**C**-D-P-**C** A-T-**C**-Y-**C**-R (75−91)	C82, C85, C88, C90
10	7a	Y-**C**-R-K (98−101)	C99
11	7b	F-**C**-Y-**C**-R-K (96−101)	C97, C99
12	7c−d	L-G-Q-Q-V-P-**CC**-D-P-**C**-A-T-**C**-Y-**C**-R (75−91)	C81, C82, C85, C88, C90

[a] **C** shows FM-modified Cys, and Cm is the residue detected as a carboxymethylcysteine. The cysteines newly modified are denoted as C67, C74, etc. Parentheses denote the residue number. The table does not list all FM-peptides, but key peptides only. [b] C99 was unexpectedly modified in protein peak 6, probably because peak 6 might be contaminated with other components.

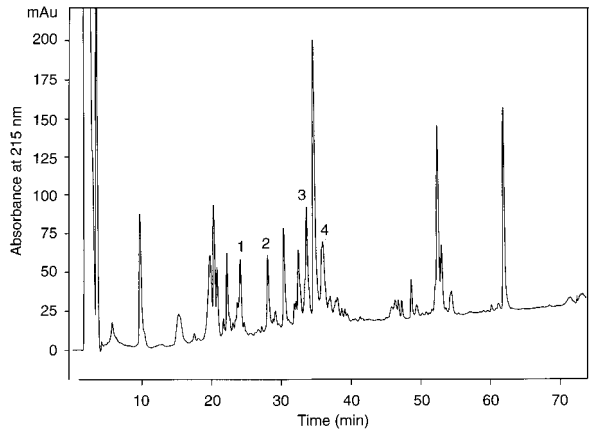

FIGURE 3: Tryptic peptide map of NEM-AGRP1B. Reduced and pyridylethylated NEM-AGRP1B was digested with trypsin and the peptides were separated by reversed phase HPLC as described under Materials and Methods. Numbered peaks were further characterized by N-terminal sequence analysis. Peaks 1−4 are denoted as NEM-AGRP1B T1−T4, respectively.

reduction of one disulfide bridge followed by NEM modification of the two free sulfhydryl groups. Thus, to identify the NEM-alkylated cysteinyl residues, NEM-AGRP1 was completely reduced and pyridylethylated. Analysis by RP-HPLC allowed the separation of the two protein peaks NEM-AGRP1A and -1B, where NEM-AGRP1A was the completely reduced and pyridylethylated protein and NEM-AGRP1B was the NEM-modified protein after complete reduction and pyridylethylation. NEM-AGRP1B was treated with trypsin, and the peptides were purified by HPLC as shown in Figure 3. Of special interest are the four peptides labeled 1−4. Analyses of peptide 1 (NEM-AGRP1B-T1) gave the sequence L-G-T-A-M-N-P-C(NEM)-S-R-T (residues 102−112), where the residue in parentheses shows the presence of an NEM-modified cysteinyl residue after hydrolysis of the maleimide ring. It gave a molecular mass of 1180.1 Da (theoretical mass = 1180.2 Da). Peptide 2 (NEM-AGRP1B-T2) was essentially the same peptide L-G-T-A-M-N-P-C(NEM)-S-R-T (residues 102−112) except the maleimide ring was intact; PTH derivatives of *N*-ethylsuccinimidocysteine were detected at cycle 8 of the Edman reaction. Its molecular mass was 1162.3 Da (theoretical mass = 1162.0 Da). Peptide 3 (1B-T3) gave the sequence L-H-E-S-C(PE)-L-G-Q-Q-V-P-C(PE)-C(PE)-D-P-C(NEM)-A-T-C(PE)-Y (residues 70−91), where C(PE) was used to denote a pyridylethylated Cys. Its molecular mass was 2622.0 Da (theoretical mass = 2620.9 Da). Peptide 4 (1B-T4) has the sequence L-H-E-S-C(PE)-L-G-Q-Q-V-P-C(PE)-C(PE)-D-P-C(NEM)-A-T-C(PE)-Y(residues 70−91) with a molecular mass of 2603.5 Da (theoretical mass = 2602.9 Da). Because C85 and C109 were the only two residues modified by NEM, this shows that they were originally disulfide-linked. The sequence results are summarized in Table 2.

Determination of the Second Disulfide Linkage C90−C97. The NEM-modified protein NEM-AGRP-3 (see Figure 1) was directly digested with pepsin to find the next disulfide linkage. The pepsin-generated map of NEM-AGRP3 was prepared by RP-HPLC (data not shown). Sequence analysis indicated that two peptides, Pep-3a and -3b, contained NEM-modified cysteines, C90 and C97, respectively. As shown in Table 2, Pep-3a has the sequence C(NEM)-R-F (residues 90−92), while Pep-3b has the sequence C(NEM)-Y-C-R-K-L-G-T-A-M (residues 97−106), where C(NEM) indicates *N*-ethylsuccinimidocysteine and C shows undetectable cysteine. These results together with mass spectral results lead us to assign the second disulfide bond, C90−C97.

Analysis of Fluorescein-Labeled Peptides for Determination of Disulfide Bonds C67−C82, C74−C88, C85−C109, and C90−C97. To confirm the above assignment and to determine the remaining disulfide linkages, the protein AGRP

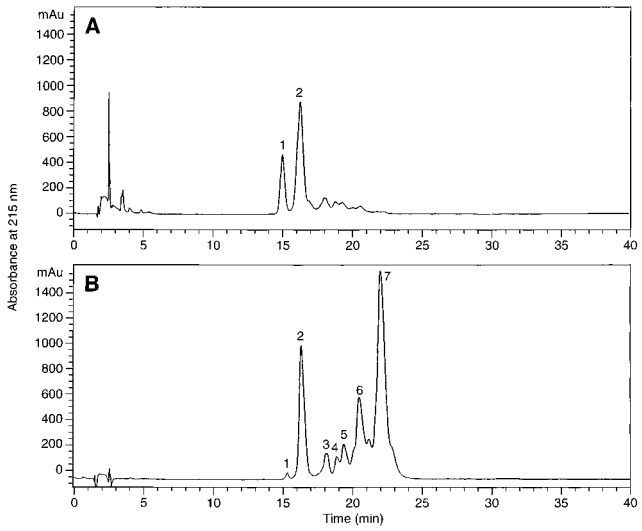

FIGURE 4: HPLC separation of FM-labeled AGRP. The AGRP sample was partially reduced with 10 mM TCEP and modified with fluorescein-5-maleimide as described under Materials and Methods. (A) 10 min of TCEP reduction and (B) 30 min of TCEP reduction. The modified protein was purified by reversed-phase HPLC using a Vydac C18 column (2.1 × 150 mm) as described under Materials and Methods. Peaks 1−7 are denoted as P1−P7, respectively.

was partially reduced with 10 mM TCEP and alkylated with fluorescein-5-maleimide (FM). Reversed-phased HPLC of FM-labeled protein is shown in Figure 4, indicating that multiple peaks were readily produced by stepwise reduction and alkylation. The FM-protein peaks P1−P7 were completely reduced and carboxymethylated prior to the thermolysin digestion. The thermolytic peptide maps of the major peaks P2, P3, P5, P6, and P7 are shown in Figure 5, indicating that the FM-labeled peptide peaks increased in the differentially modified proteins. For example, although the thermolytic digest of intact protein P1 did not show any fluorescent peptide peak (data not shown), protein P2 indicated several peptide peaks detected at 440 nm. These peptide peaks, P2a−2d, were determined to have similar sequences, A-M-N-P-X-S-R-T (residues 105−112) or M-N-P-X-S-R-T (residues 106−112), where X was the undetermined residue by sequence analysis because of FM-labeled Cys (Table 3). Multiplicity of the labeled peptides is probably due to diastereoisomers or hydrolysis of the succinimide ring since these peptides gave the same molecular weight by mass spectrometry. This conversion of *N*-ethylsuccinimido cysteine has been reported earlier (*16*, *17*). These results indicate that C109 was selectively reduced and alkylated in the protein P2. Other cysteine-containing peptides were found in peptide peaks P2e−h, showing the sequence L-G-Q-Q-V-P-C(CM)-C(CM)-D-P-X-A-T-C(CM)-Y-C(CM)-R (residues 75−91), where X was an unidentified residue and C(CM) corresponds to carboxymethylcysteine. Mass spectrometry suggests that X was FM-Cys, C85, hence demonstrating that the first disulfide bond, C85−C109, was preferentially reduced with TCEP and alkylated, while the remaining disulfide bonds were still intact.

Similarly, FM-modified protein P3 contained additional FM-labeled peptides, as compared with the peptide map of protein P2. Peptide P3a had the sequence F-X (residues 96−97), and peptides P3b and P3c have the same sequence, L-G-Q-Q-V-P-C(CM)-C(CM)-D-P-X-A-T-C(CM)-Y-X-R (residues 75−91) (Table 3), in which an additional cysteine, C90,

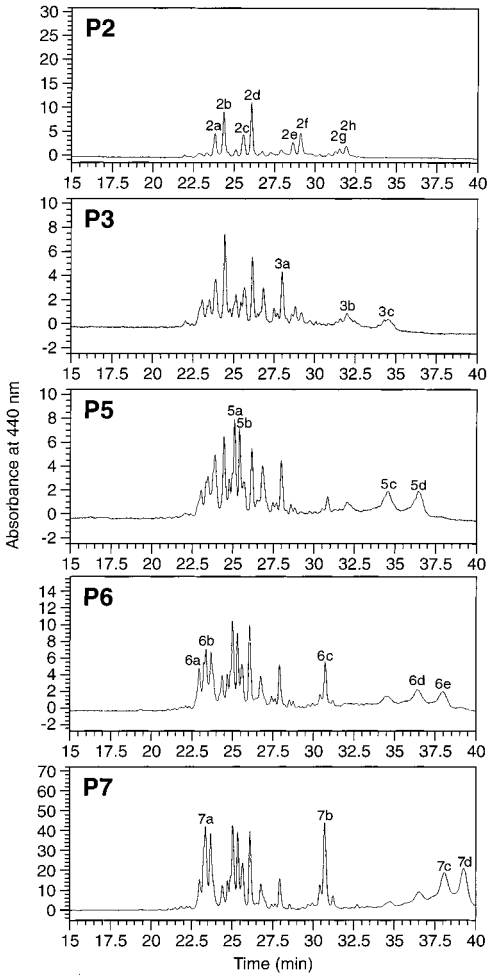

FIGURE 5: Thermolytic peptide maps of FM-proteins P2−P7. FM-labeled proteins were fully reduced and carboxymethylated as described under Materials and Methods. The proteins were digested with thermolysin in 0.1 M triethylamine buffer, pH 6.0. The FM-labeled peptides were purified by reversed-phase HPLC using a Vydac C18 column (2.1 × 150 mm), detecting with absorbance at 440 nm. Peptide maps of proteins P1 and P4 are omitted because no positive peak appeared at 440 nm or poor recovery of the peptides.

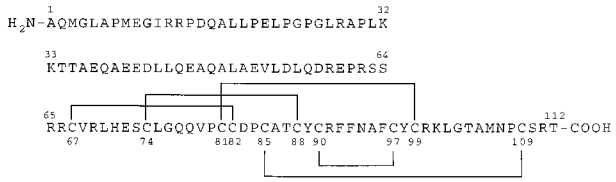

FIGURE 6: Disulfide connectivity of AGRP. Four disulfide bonds are highly homologous to that of ω-agatoxins (6), except for one extra disulfide bond (C85−C109).

was modified with FM. Based on these results, C90 is assigned to be linked to C97.

Meanwhile, one of the major peaks, P5, gave several additional peaks, P5a−5d. However, peptides P5a and P5b showed the same sequence L-H-E-S-X (residues 70−74), where X was FM-Cys. Meanwhile, peptides P5c and P5d gave the sequence L-G-Q-Q-V-P-C(CM)-C(CM)-D-P-X-A-T-X-Y-X-R (residues 75−91), in which C88 was further modified with FM. These results demonstrate that C74 linked to C88.

The peptide map of protein P6 was more complicated than the others in determining the disulfide linkage, whereas the

disulfide bond C67−C82 was suggested to be present since peptides P6a−6b had the sequence S-R-R-X (residues 64−67), and peptide P6c, F-X-Y-X-R-K (residues 96−101), where X was an FM-labeled cysteine. Peptides P6d−6e had the sequence L-G-Q-Q-V-P-C(CM)-X-D-P-X-A-T-X-Y-X-R (residues 75−91), where only C81 was detected as carboxymethylcysteine and other cysteines denoted by X were undetectable because of FM-modified cysteine. Mass spectrometry supported these modifications. Thus, it is concluded that C67 linked to C82 although peptide P6c containing C97 and C99 was slightly detected. The detection of FM-labeled peptide P6c might be derived from other modified proteins.

Finally, protein fraction P7 was assigned to be the fully reduced and alkylated protein according to the mass spectral results. Peptides P7a−7b exhibited similar peptides, Y-X-R-K (residues 98−101) and F-X-Y-X-R-K (residues 96−101), where X was FM-modified cysteine. Peptides P7c−7d had the same sequence, L-G-Q-Q-V-P-X-X-D-P-X-A-T-X-Y-X-R (residues 75−91), in which all cysteines were modified with FM. Determination of the remaining disulfide in protein P6 was slightly difficult because of the high backgrounds of the FM-labeled peaks. Thus, assignment of disulfide linkage between C81 and C99 was confirmed later by an alternative method.

Assignment of the Less Reactive Disulfide Linkage, C81−C99. FM-modified protein P6 should contain one intact disulfide bond according to the mass spectral analyses (Table 1). To identify the unreactive disulfide, protein P6 was directly digested with thermolysin as described under Materials and Methods. The peptide map of P6 indicated many FM-modified peaks, whereas peptide P6-TH-A gave two sequences, F-X-Y-X-R-K (residues 96−101) and L-G-Q-Q-V-P-X-X-D-P-X (resiudes 75−85), where X was an undetermined residue. Although mass spectral data were not available for this peptide because of the small amount, it was confirmed by isolation of two individual peptides from peptide P6-TH-A after reduction and carboxymethylation (data not shown). Since the sequence analysis of one of them (named P6-TH-A-CM2) indicated that only C81 was carboxymethylated, C81 might originally link to one of two cysteine residues, C97 or C99. Another partner peptide, P6-TH-A-CM1, F-C-Y-C-R-K (residues 96−101), contained C97 and C99. However, since residue C97 has been already assigned to link C90 as shown above, it was concluded that C81 was linked to C99. Based upon these experimental findings, the overall disulfide structure of AGRP was assigned as shown in Figure 6. However, since we have analyzed the major products derived from partial reduction/alkylation of AGRP, it cannot be denied that minor components at low level would have some disulfide rearrangements.

A Comparison of Disulfide Structures of AGRP and ω-Agatoxin IVB. The determined disulfide structure is significantly homologous to that of ω-agatoxin, a spider toxin, except for one disulfide, C85−C109 (*7*). The toxicity of AGRP in the animal body remains to be investigated; also it is necessary to know whether disulfide bonds and amino acid side chains are critical for biological activity. Site-directed mutagenesis of specific cysteines of agouti revealed that alteration of C85 and C109 partially reduced the biological activity (*7*). However, mutation of the other eight cysteines completely lost the activity. Based on these observations, a disulfide bond of C85−C109 may not be strongly essential for its neurotoxicity. Although the disulfide structures of AGRP and ω-agatoxin IVB were highly homologous, the neurotoxicity of the spider toxin may depend on the amino acid side chains, since sequence homology between these proteins is not significant. Recently, Kiefer et al. (*9*) reported the results from site-directed mutagenesis of agouti, in which MC4-R-binding determinants were postulated to be the highly conserved region including R-F-F (residues 116−118). However, for understanding the functional difference between agouti protein and ω-agatoxin, further structure−function study including the tertiary structural analysis and the receptor specificity would be necessary.

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. Ollmann, M., Wilson, B. D., Yang, Y.-K., Kerns, J. A., Chen, Y., Gantz, I., and Barsh, G. S. (1997) *Science 278*, 135−138.
3. Graham, M., Shutter, J. R., Sarmiento, U., Sarosi, I., and Stark, K. L. (1997) *Nat. Genet. 17*, 273−274.
4. Fong, T. M., Mao, C., MacNeil, T., Kalyani, R., Smith, T., Weinberg, D., Tota, M. R., and Van der Ploeg, L. H. T. (1997) *Biochem. Biophys. Res. Commun. 237*, 629−631.
5. Willard, D. H., Bodnar, W., Harris, C., Kiefer, L., Nichols, J. S., Blanchard, S., Hoffman, C., Moyer, M., Burkhart, W., Weiel, J., Luther, M. A., Wilkison, W. O., and Rocque, W. J. (1995) *Biochemistry 34*, 12341−12346.
6. Olivera, B., Miljanich, G. P., Ramachandran, J., and Adams, M. E. (1994) *Annu. Rev. Biochem. 63*, 823−867.
7. 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.
8. Figueiredo, S. G., Garcia, M., Valentim, A., Cordeiro, M., Diniz, C., and Richardson, M. (1995) *Toxicon 33*, 83−93.
9. Kiefer, L. L., Veal, J. M., Mountjoy, K. G., and Wilkison, W. O. (1998) *Biochemistry 37*, 991−997.
10. Sun, Y., and Smith, D. L. (1988) *Anal. Biochem. 172*, 130−138.
11. Jones, M. D., Hunt, J., Liu, J. L., Patterson, S. D., Kohno, T., and Lu, H. S. (1997) *Biochemistry 36*, 14914−14923.
12. Gray, W. R. (1993) *Protein Sci. 2*, 1732−1748.
13. Gray, W. R. (1993) *Protein Sci. 2*, 1749−1755.
14. Heck, S. D., Kelbaugh, P. R., Kelly, M. E., Thadeio, P. F., Saccomano, N. A., Stroh, J. G., and Volkmann, R. A. (1994) *J. Am. Chem. Soc. 116*, 10426−10436.
15. Wu, J., and Watson, J. T. (1997) *Protein Sci. 6*, 391−398.
16. Wu, C.-W., Yarbrough, L. R., and Wu, F. Y-H. (1976) *Biochemistry 15*, 2863−2868.
17. Ishii, Y., and Lehrer, S. S. (1986) *Biophys. J. 50*, 75−80.
18. Palmer, M., Buchkremer, M., Valeva, A., and Bhakdi, S. (1997) *Anal. Biochem. 253*, 175−179.
19. Rosenfeld, R. D., and Stark, K. L. (1998) *Biochemistry* (submitted for publication).
20. Stone, K. L., Elliott, J. L., Peterson, G., McMurray, W., and Williams, K. R. (1990) *Methods Enzymol. 193*, 389−412.
21. Schutte, C. G., Lemm, T., Glombitza, G. J., and Sandhoff, K. (1998) *Protein Sci. 7*, 1039−1045.
22. Gorin, G., Martic, P. A., and Doughty, G. (1966) *Arch. Biochem. Biophys. 115*, 593−597.
23. Hui, J. O., Le, J., Katta, V., Rohde, M. F., and Haniu, M. (1997) *Techniques in Protein Chemistry VIII*, pp 277−287, Academic Press, San Diego.

BI981082V