ELSEVIER

Peptides, Vol. 18, No. 3, pp. 337–346, 1997

0196-9781/97 $17.00 + .00

PII S0196-9781(96)00323-3

MALDI-MS as a Monitor of the Purification and Folding of Synthetic Eclosion Hormone

YAJUN J. WANG,* LALE YURTTAS,† BRUCE E. DALE,†[1] DAVID H. RUSSELL,‡ GARY KINSEL,‡ LISA M. PRESTON–SCHAFFTER,‡ VIRGINIA JOHNSON* AND TIMOTHY K. HAYES*[2]

Department of Entomology, †Department of Chemical Engineering, and ‡Department of Chemistry, Texas A&M University, College Station, TX 77843

Received 24 January 1996; Accepted 5 November 1996

WANG, Y. J., L. YURTTAS, B. E. DALE, D. H. RUSSELL, G. KINSEL, L. M. PRESTON–SCHAFFTER, V. JOHNSON AND T. K. HAYES. *MALDI-MS as a monitor of the purification and folding of synthetic eclosion hormone.* PEPTIDES **18**(3) 337–346, 1997.—Analogues of the small protein *Manduca sexta* eclosion hormone (62 amino acids) were synthesized by Fmoc solid-phase methodology. Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) was used to analyze the products of the syntheses and this information was used to design an efficient purification scheme. MALDI-MS was used to monitor the target products through purification and it was also used to monitor folding of the purified materials. The folded EH analogues were shown to be biologically active proteins with an in vivo bioassay using pharate adult moths, *Heliothis virescens.* © 1997 Elsevier Science Inc.

Eclosion hormone Fmoc solid-phase peptide synthesis
Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) Protein folding

IN recent years, matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) has become a powerful and flexible technique for the analysis of peptides and proteins (10). In MALDI-MS, samples are cocrystallized with a matrix compound on the probe tip of the mass spectrometer. A laser impulse excites the matrix molecules, which transfer this energy to ionize analyte molecules. Ionization occurs in a time-of-flight vacuum tube in an electric field. Analytes are accelerated in the field at a rate that is related to the inverse of their mass/charge ratios. Some of the properties (10,13) of MALDI-MS that make it a superb analytical tool for peptides and proteins include: subpicomole sensitivity, analysis of complex mixtures, high mass resolution, simplicity of operation, rapid analysis, and rugged instrument performance.

Eclosion hormone (EH) is a small 62-amino acid insect neuroprotein (12,20) that is small enough to be the subject of solid-phase peptide synthesis. EH is released from the nervous system of insects and stimulates ecdysis (i.e., the complex behavior required for an insect to emerge from its old exoskeleton during a molt). Recent studies (22) have demonstrated that EH acts through endocrine glands (i.e., the epitracheal glands) to release an ecdysis-triggering hormone (ETH), which has direct action on the nervous system to stimulate ecdysis behavior. The study of interactions of EH with its target cell receptors could benefit from the synthesis of structural analogues. However, EH is a small protein that has the conformational restriction imposed by three disulfide bonds. The proper folding of EH is likely important for it to interact with its target cell receptors. This article demonstrates how MALDI-MS can be rapidly used to monitor purification and folding of the three EH analogues synthesized.

METHOD

EH Analogue Synthesis

Three EH analogues, synthetic *Manduca sexta* EH (designated as EHQ), [Asn^{20}]EH (EHN), and a chemically stabilized EH analogue, [Nle^{11}, Asn^{20}, Nle^{24}, NpA^{28}]EH (EHS) were synthesized using 9-fluorenylmethyloxycarbonyl (Fmoc) SPPS (2,6) on a MilliGen 9050 synthesizer. The resin used was Fmoc-L-Leu-PepSyn KA (PerSeptive Biosystems/Biosearch, Framingham, MA). The general methods were similar to those previously described for the synthesis of the Tat protein (7). Coupling activation was achieved with benzotriazolyl-oxy-tris(dijmethylamino)-phosphonium-hexafluorophosphate/1-hydroxy-benzotriazole (BOP/HOBt). Side-chain protection for the synthesis of EHQ and EHN was specific for the following amino acids. Asn and Gln: 2,4,6-trimethoxybenzyl (Tmob); Asp and Glu: *O*-butyl (OtBu); Cys: trityl; Lys: *t*-butyloxycarbonyl (Boc); Ser, Thr, and Tyr: butyl (tBu). For EHS synthesis, Trityl instead of Tmob was used for Asn protection.

Coupling times for individual amino acids varied depending on the amino acid, side-chain blocking group, and its position in the sequence. Coupling times for the first 20 residues from the *C*-ter-

[1] Current address: Department of Chemical Engineering, Michigan State University, East Lansing, MI 48824.
[2] Requests for reprints should be addressed to Dr. Timothy K. Hayes at his current address: Biological Products, Bayer Corporation, 1017 Main Campus Drive, Suite 3800, Raleigh, NC 27606.

minus were 1 h, for the next 20 residues 2 h, and so forth. Coupling of and to beta-branched amino acids, Tmob-protected amino acids, and Trityl-protected amino acids was lengthened by 1 h to help counteract the reduced coupling kinetics produced by steric hindrance. An additional 1 h was added to the coupling of amino acids to Gly and Ala. Coupling times are summarized in Fig. 1.

After each coupling was accomplished, a 10-min capping procedure was performed with a solution of 6.4% (w/v) 1-hydroxybenzotriazole (HOBt) and 5% (v/v) acetic anhydride in dimethylformamide (DMF). The Fmoc group was removed by washing with 20% piperidine in DMF prior to the next coupling cycle.

After the complete protein was constructed, side-chain blocking groups were removed while the protein was simultaneously cleaved from the resin support at room temperature (details follow).

Standard scale cleavage. Peptide resin (200 mg) was cleaved with 15 ml of the cleavage cocktail for 2 h. After being separated from the cleavage cocktail, the resin was recleaved for an additional 2 h. Vacuum filtration was employed to separate the resin from the cocktail. The filtrates from both cleavages were combined, rotary evaporated to remove most of the reagents, and then cooled to 4°C. The protein was precipitated in ice-cold anhydrous diethyl ether containing 1% β-mercaptoethanol and washed thoroughly with the same. The precipitate was vacuum filtered, air dried, and subjected to amino acid analysis, MALDI-MS, and purification.

Micro scale cleavage. Peptide resin (15 mg) was cleaved twice with 3 ml of the cleavage cocktails, each for 2 h. Then the resin was filtered in a Pasteur pipette filled with glass fiber. The volatile reagent in the filtrate was removed using a SpeedVac concentrator (Savant Instruments, Inc., Hicksville, NY). The protein was precipitated and washed with ice-cold anhydrous diethyl ether containing 1% β-mercaptoethanol. A centrifuge step was followed to separate the protein from the ether wash.

To optimize cleavage conditions, aliquots of the crude protein resin were cleaved in the micro scale with different cleavage cocktails: reagent R (trifluoroacetic acid/thioanisole/1,2-ethanedithiol/anisole = 90:5:3:2) (1), dimethyl sulfide (DMS) cocktail (trifluoroacetic acid/DMS/1, 2-ethanedithiol/thioanisole = 80:12:3:5), and trialkylsilanes [triethylsilane (TES) or triisopropylsilane (TIPS)] cocktail [trifluoroacetic acid/trialkylsilanes = 95:5) (15)]. A dual cleavage using either reagent R or DMS cocktail followed by a TES cocktail was also performed, in which the second cocktail was introduced after the first one was removed.

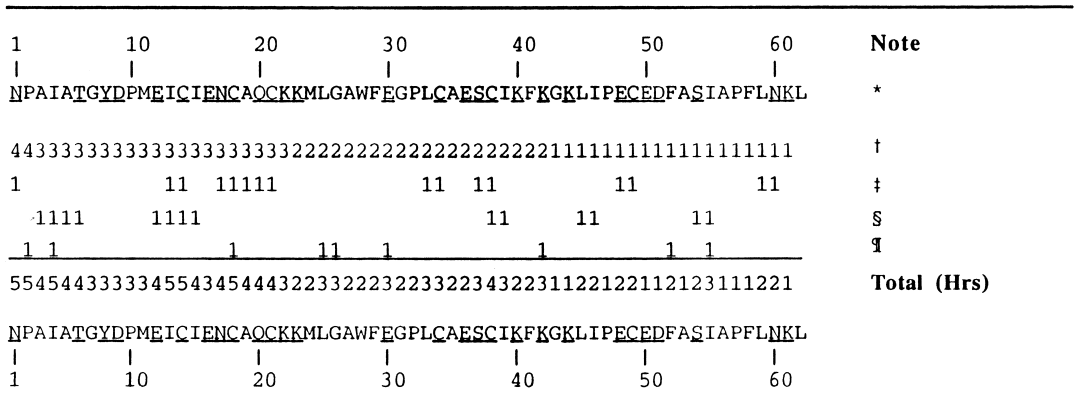
The bulk protein was cleaved with TES/TIPS cocktail at standard scale.

Purification

Gel filtration chromatography (step 1). The separation of crude protein was carried out on a 10 × 1000 mm Sephadex G-25 column. The protein was eluted with 0.02 *M* Tris, 10% acetonitrile (pH 8.0) buffer at a flow rate of 200 μl/min. cytochrome C (60 nmol) and 10 μl of 0.5% bromophenol blue were added into the sample as indicators. Fractions were collected at a rate of 10 min/fraction. The fractions coeluted with the visibly red cytochrome C were pooled and desalted with several C18 Sep-Pak® cartridges (Waters, Milford, MA) for MALDI-MS analysis and further purification.

Anion exchange chromatography (step 2). A 10 × 200 mm glass column was packed with Accell™ plus quaternary methylamine (QMA) anion exchange media (Waters, Millipore) for protein purification. The proteins were eluted with a linear gradient of 0–500 m*M* sodium acetate in 20 m*M* Tris buffer (pH 8.0) for 50 min at a flow rate of 0.8 ml/min. The absorbance at 262 nm was monitored with time. EH analogue was recovered in the 35–50-min fractions as determined by MALDI-MS analysis.

Reverse-phase HPLC (step 3). The fractions containing EH analogue from step 2 were combined and adjusted to pH 2.0 by adding 0.1% TFA. This mixture was applied to a C18 semipreparative column (Vydac, 10 μm, 10 × 250 mm) and eluted with a linear gradient of 0–90% acetonitrile in 0.1% TFA over 49 min



* Side chain protected amino acid are underlined. These protections were: N, Q (Tmob); D, E (OtBu); C (Trt); K (Boc); and S, T, Y (tBu).

The numbers in the body of the figure under each residue are the coupling time in hours due to the potential coupling problems foreseen in the sequence as specified below:

† Chain length.
‡ Steric hindrance (Tmob, Trt protections).
§ Steric hindrance (ß-branched residues).
¶ Coupled to Gly or Ala resides.

FIG. 1. EHQ sequence and the coupling time during synthesis.

at a flow rate of 2.0 ml/min. The absorbance at 280 nm was recorded. The 40–55-min fractions were collected at a rate of 0.5 min/fraction and subjected to MALDI-MS.

Characterization and Folding

Analytical HPLC. An aliquot of the EH analogue-containing fraction was applied to a C18 analytical column (Vydac, 5 μm, 4.6 $\times$ 250 mm). The elution was carried out with a linear gradient of 0–90% acetonitrile in 0.1% TFA over 49 min at a flow rate of 1.0 ml/min. The absorbance at 214 nm and/or 280 nm was monitored.

MALDI mass spectrometry. MALDI-MS was performed on a linear time-of-flight mass spectrometer [built in-house, for details refer to Preston et al. (16)]. Sample was prepared by the addition of 2 μl of the protein solution (50 pmol) to 2 μl of matrix solution (0.1 *M* α-cyano-4-hydroxycinnamic acid in methanol) prior to the analysis. A drop of 2 μl of matrix/analytes was applied to the probe tips, which were coated with nitrocellulose. The N_2 laser (337 nm) was triggered at 4 Hz. The voltage for the source was 18 kV. The acquired spectra were a summation of ion intensity for 100 laser pulses and then were transferred to an Intel® 80386-based microcomputer for analysis. Grams/386™ software (Galactic Industries Corporation, Salem, NH) was used for data processing.

Amino acid analysis. Crude and/or purified samples were hydrolyzed for 1 h at 150°C in 6 N HCl (gas phase). Analysis was performed on a Waters PicoTag™ amino acid analysis system. Methods were those supplied by the manufacturer (Waters, Milford, MA).

N-terminal sequence analysis. Sequencing analysis was accomplished on an Applied Biosystems 470A protein sequencer with an on-line Applied Biosystems 120A PTH analyzer. Samples were sequenced by using standard manufacturer cycles on a polybrene glass fiber membrane (9).

Cys reduction and alkylation. Proteins were denatured in 6 *M* guanidine hydrochloride (Gdn-HCl) and reduced with 0.1 *M* dithiothreitol (DTT) in 0.2 *M* Tris buffer (pH 8.7) for 2 h at room temperature. The reduced proteins were alkylated with either iodoacetamide or iodoacetic acid at a total concentration of 0.1 *M* at pH 8.7 for 2 min. The reaction was immediately quenched by the addition of 50% acetic acid and the reagents were removed using a C18 Sep-Pak® cartridge followed by vacuum drying.

Protein folding. Purified reduced protein (10 nmol) was dissolved in 100 μl of 6 *M* Gdn-HCl, 0.1 *M* Tris buffer (pH 8.7), then dialyzed against the following redox buffers containing 4, 2, or 0 *M* Gdn-HCl each for 12 h consecutively. Besides Gdn-HCl, the redox buffers (pH 8.7) were comprised of 0.1 *M* Tris, 0.1 *M* sodium phosphate, 0.1 *M* sodium chloride, 1 m*M* oxidized glutathione (GSSG), and 1 m*M* reduced glutathione (GSH) (18). An aliquot of approximately 100 pmol was taken at each Gdn-HCl concentration for MALDI-MS analysis. These samples were alkylated to quench the Cys oxidation as described above. The desalted samples were brought up with 20 μl of the 5 m*M* TFA and acetonitrile mixture (1:1, v/v). Only about 1 μl of the dissolved protein (5 pmol) was used in MALDI-MS analysis.

For bioassay, the folded protein in the redox buffer underwent a 24-h buffer exchange with physiological saline, which was prepared in accordance with *H. virescens* hemolymph composition analysis (3). The dialysis was carried out on a microdialyzer system 100 (Pierce, Rockford, IL) with a 5,000 molecular weight cutoff membrane. For monitoring protein folding process, a continuous flow mode at a constant flowrate was also used.

Bioassay. EH activity of the folded protein was detected using a pharate adult of *Heliothis virescens* in vivo assay, similar to those described by Kataoka et al. (11). The *Heliothis virescens* pupae (5–7 days old) were purchased from USDA Agricultural Research and Services in Stonville, MS. Upon arrival, insects were kept in an incubator at 29°C in an 8-h dark/16-h light period for 4–6 days prior to assay. Insects were staged daily by observing their appearance either under the microscope or with the naked eye. Those whose molting fluid was mostly reabsorbed, had red tarsi, and dark, soft bodies were staged as day 12. Only these pupae were selected for the assay. Insects were injected in the dorsal thorax with 5 μl of folded EH analogue at 1200 h, approximately 7 h before normal eclosion, which occurs at sunset. Moths that eclosed within 3 h of the injection were scored as a positive EH response. Refolded bovine pancreatic trypsin inhibitor (BPTI) was used as a negative control.

Circular dichroism. CD experiments were performed on a Jasco J-600 Spectropolarimeter with a 0.02-cm flow cell. About 10 nmol refolded EH analogue was dissolved in 500 μl trifluoroethanol (TFE). CD spectra were recorded at room temperature in the far UV region (185–250 nm) with step resolution of 0.2 nm. An average of eight traces was taken. The profile was compared with other small proteins (8).

RESULTS

Characterization of Crude Products

The 62-amino acid sequence of each EH analogue was assembled as described using Fmoc solid-phase methodology. The rationale for the use of Fmoc chemistry over Boc chemistry was primarily because the polar side chains of the amino acids Asn and Gln could be blocked. The alternative Boc chemistry did not have blocking groups available for this purpose. Hydrophobic blocking of these primary amide side chains reduced the potential of resin-bound protein aggregation. This aggregation could result in reduced coupling kinetics that would work against efforts to lengthen the protein further through solid-phase methodology (17). Tmob protection of Asn and Gln was readily available for the synthesis of EHQ and EHN. However, Trityl protection was reported superior and was used when EHS was synthesized. The Tmob deprotection adduct was reported to have a high propensity to alkylate Trp (6), a residue found in both the earlier EH analogues. Recently, a Fmoc-Trp (Boc) derivative (21) has been discovered that helps protect against this type of alkylation during deprotection. However, this was not used for the EHS synthesis because Trp was replaced with naphthyl alanine (NpA).

Aliquots of the protein resin were cleaved using several TFA cleavage protocols: reagent R, a DMS cocktail, trialkylsilane (i.e., TES/TIPS) cocktails, and a dual cleavage method (reagent R or DMS cocktail followed by TES cocktail). The crude product mixture was characterized using analytical reverse-phase HPLC and MALDI-MS.

Analyses of the crude products by MALDI-MS. The molecular weight of the target product varied with the different cleavage protocols employed (Table 1). Reagent R or the DMS cocktail resulted in products with 240–390 Da higher molecular weight than the theoretical mass. TIPS cocktail-cleaved proteins were about 180 to 200 daltons higher. The masses of TES cocktail-cleaved proteins and the dual cleaved proteins agreed with the expected values.

To investigate the causes for the different masses of target products between different cleavage protocols, we compared the masses of crude EHN product cleaved just by reagent R with the masses of those cleaved with reagent R followed by a second TES cocktail cleavage (Table 2). The mass spectral profiles for the crude sample were more or less the same for both samples

TABLE 1

MALDI-MS ANALYSIS OF SYNTHETIC EH ANALOGUE

Protein	Cleavage Reagent	Theoretical Mass (D) A	Observed Mass* (D) B	Difference† (D) C = B − A	Predicted Adduct‡
EHQ	Reagent R	6813.13	7203	390	Trt + adduct from EDT + TFA
	TIPS		6995	182	Tmob
	TES		6817	4	
	Dual§		6871	58	Na^+, K^+
EHN	DMS	6798.92	7058	259	Trt
	Reagent R		7071	272	Trt + salt adduct
	TIPS		7007	208	Tmob + acetyl
	TES		6840	41	K^+
	Dual§		6827	28	Na^+
EHS	Reagent R	6773.81	7010	236	Trt
	TES		6777	3	
	Dual§		6772	−2	

* Observed MALDI mass of the target product in crude proteins.
† Defined as difference between the observed mass and the theoretical mass.
‡ The probable adducts are those deemed possible from the synthesis whose mass closely matches the mass difference.
§ Cleaved first with reagent R followed by a second cleavage with TES cocktail.

[refer to Fig. 3(A) for reagent R-cleaved product]. However, the masses of the dual cleaved products were very close to the calculated values, whereas the masses of the reagent R-cleaved products, except for product No. 3, were about 200 units higher than the calculated values.

Monitoring Purification with MALDI-MS

The first attempts to purify EH analogues were done by semipreparative C18 RP-HPLC alone (Fig. 2). The shadowed fractions were found to contain full-length EH analogues by MALDI-MS and amino acid composition analysis. The EHS fraction at 48.5 min was relatively pure and a single peak of about 6800 Da was observed in the MALDI-MS analysis. This EHS fraction accounted for about 7% of the crude product mixture.

TABLE 2

MALDI-MS ANALYSIS OF THE CRUDE EHN CLEAVED WITH DIFFERENT CLEAVAGE REAGENTS*

Product No.	Calculated Mass† (D) A	(Reagent R/TES) Observed B	(Reagent R/TES) Difference‡ C = B − A	(Reagent R) Observed D	(Reagent R) Difference E = D − A
1	802	§		814	12
2	1221	§		1272	51
3	1907	1891	−16	1879	−28
4	2911	2920	9	3104	193¶
5	3040	3066	26	3270	230¶
6	4789	4804	15	5005	216¶
7	6799	6827	28	7071	272¶

* The cleavage reagents are indicated in the parentheses.
† See Table 3 for the assumed sequence from which the mass was calculated.
‡ These differences are likely due to small ion adducts (Na^+, K^+, etc.) and mass assignment error due to broad peaks in this crude sample.
§ These peaks were not seen in the MALDI mass spectrum of the dual cleaved EHN.
¶ These differences are suggested due to a large moiety (i.e., Trityl group), which is covalently attached to the cleavage products.

Even though the EHQ fraction at 45 min was not resolved from other components by the HPLC separation alone, only one peak of about 6830 Da was observed by MALDI-MS analysis. The EHN analogue was primarily contained in the 44-min fraction. There are substantially more peaks observed in the EHN MALDI-MS analysis [Fig. 3(A)] than were found in either of the EHS or EHQ analogs.

The EH analogues containing fractions for EHQ and EHS were pooled and repurified once on the same semipreparative RP-HPLC system. The second HPLC purification yielded clean products as confirmed by both analytical HPLC (Fig. 4) and MALDI-MS (Fig. 5). The yields of the purified unfolded EHQ and EHS were 13% and 3% from the crude mass, respectively. About 360 mg of the crude mass of EHS was purified in multiple batches and yielded 10 mg of purified unfolded protein. Note that crude masses were not pure protein material alone as each included water, salts, and byproducts of the deprotection and cleavage reactions.

In the case of EHN purification, no fraction from RP-HPLC of the crude synthetic product contained a single protein. A combined purification scheme including gel filtration, anion-exchange, and RP-HPLC was designed from MALDI-MS data of the crude product [Fig. 3(A)]. A capping step was included in the synthetic procedure to avoid the deletion product formation and the lower masses in the MALDI-MS analysis correspond to prematurely truncated analogues. Based on this information, the protein sequence was deduced from the molecular weight. The isoelectric points (pI) were calculated for each presumed sequence (Table 3). The resultant purification scheme included: (a) Sephadex G-25 gel filtration chromatography to remove peaks 1–3 with lower molecular weight [Fig. 3(B)]; (b) anion exchange chromatography followed by RP-HPLC to remove peaks 4–6 with higher pIs than the EH analogues [Fig. 3(C)]. In the final MALDI mass spectrum, only the EH analogue and its doubly charged signal were detected. The yield of the final purified unfolded EHN was approximate 1% from the crude mass.

Characterization of the Purified EH Analogues

Purified EH analogues were characterized by HPLC (Fig. 4), MALDI-MS (Fig. 5) analyses, Edman degradation, and amino

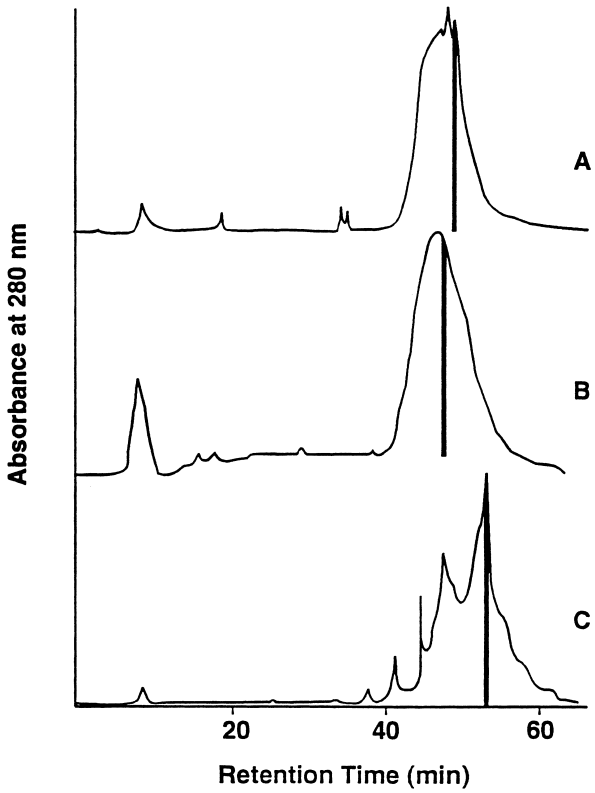


FIG. 2. Semipreparative purification of EH analogues. (A) EHQ; (B) EHN; and (C) EHS. A 10 μm, 10 $\times$ 250 mm C18 column (Vydac) was used. Proteins were eluted at 2.0 ml/min with a gradient of 0–70% solvent B over 49 min, where solvent A = 0.1% TFA; solvent B = 90% acetonitrile (MeCN) in 0.1% TFA. Absorbance at 280 nm was monitored with elution time. Fractions were collected at 0.5 min/fraction and subjected to MALDI-MS analysis. The shadowed fractions contained EH analogues at the highest apparent purity and were pooled for further purification.

acid analysis (Table 4). The analogues at this point were present at various stages of disulfide bond oxidation. Despite this fact, the HPLC analyses illustrated relatively clean products. The observed molecular weight of the products were: EHQ 6816 (calculated: 6813); EHN + K^+ 6839 (calculated: 6838); and EHS 6776 (calculated: 6774). The expected sequences for the first 22 residues from the *N*-terminus for all three proteins were observed during analysis by Edman degradation. The amino acid analysis for 6 N HCl hydrolysis of EH analogues agreed with the theoretical composition within the accuracy limits of the analysis regardless the cleavage protocol used (Table 4).

Monitoring Folding With RP-HPLC and MALDI-MS

The proper folding of EH analogues results in the formation of three disulfide bonds from six free thiols. The shift from free thiols to disulfide bonds can be monitored with the aid of an alkylation reaction because only free thiols and not disulfide bonds will react with reagents like iodoacetic acid and iodoacetamide. Analytical methods like C18 reversed-phase HPLC and MALDI were used to monitor the shift from free thiols to disulfide bonds as each EH analogue folded. EH analogues were prepared for folding by complete reduction to the free thiol state in 6 M Gdn-HCl. The presence of both reduced and oxidized glutathione at pH 8.7 allowed for the free exchange of disulfide bonds to allow folding to proceed as the Gdn-HCl was removed by dialysis. EHS folding is the example used to demonstrate the monitoring of the folding process.

Analyses of EHS folded states by RP-HPLC (Fig. 6) and MALDI-MS (Fig. 7) are distinctive for reduced EHS [Figs. 6(A) and 7(A)], fully carboxymethylated EHS [Figs. 6(B) and 7(B), unfolded], and refolded EHS [Figs. 6(C) and 7(C), carboxymethylation resistant]. The fully folded EHS had the shortest retention time on HPLC. This species was followed closely by the fully carboxymethylated protein. The reduced, unalkylated protein had the longest retention time. The fully reduced unalkylated EHS had an observed mass of 6778 [Fig. 7(A)]. The fully reduced alkylated EHS had an observed mass that increased by 354 D over the unalkylated material [Fig. 7(B)]. This mass observation is very close to the 348-Da increase expected for six carboxymethylations.

Snapshots of the folding process were observed by MALDI analysis (Fig. 8) of samples taken from the folding reaction as EHS equilibrated with 4, 2, 1.5, 0.5, and 0.2 *M* Gdn-HCl in 1 m*M* reduced/oxidized glutathione, 0.1 *M* Tris at pH 8. The first indication of molecular weight shift and lowered alkylation of free thiols was not noted until EHS had equilibrated with 1.5 *M* Gdn-HCl. After EHS had equilibrated with 0.5 and 0.2 *M* Gdn-HCl, the observed molecular weight of the sample had shifted to lower values. This indicated that very little alkylation had occurred to the molecules in these samples. The MALDI peaks representing EHS through the folding process were broad.

The secondary structure of the folded EHQ was examined using circular dichroism in trifluoroethanol (TFE) (Fig. 9). There were two minima at 208 and 222 nm and a maximum at 192 nm. The profile is similar to other small proteins that are dominated by α-helical secondary structure (8).

Biological Activity

The appearance of the native conformation in folded EH analogues was tested by the precocious eclosion bioassay. All three

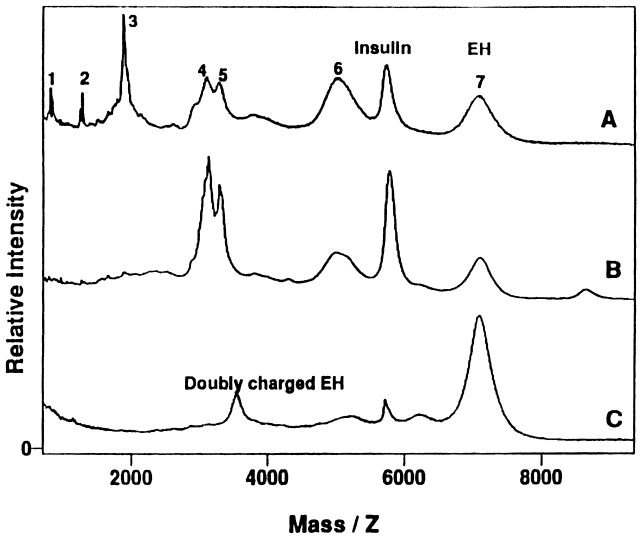


FIG. 3. MALDI mass spectra of EHN at each purification step. (A) Crude protein, where peaks 1–6 were impurities, peak 7 was EH analogue, and insulin was used as internal standard; (B) after gel filtration step, where peaks 1–3 were removed; (C) after anion exchange chromatography and RP-HPLC, where peaks 4–6 were further removed. Only EH analogue was detectable at the last step.

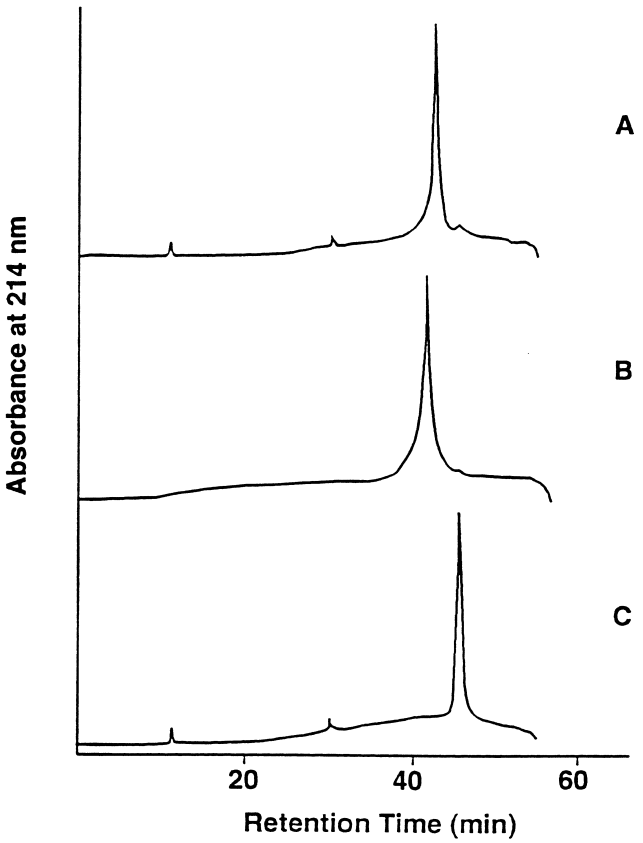


FIG. 4. Analytical HPLC analysis of the purified EH analogues. (A) EHQ; (B) EHN; and (C) EHS. A 5 μm, 4.6 $\times$ 250 mm C18 column (Vydac) was used. Samples were eluted at 1.0 ml/min with a gradient of 0–70% solvent B over 49 min, where solvent A = 0.1% TFA; solvent B = 90% acetonitrile (MeCN) in 0.1% TFA. Absorbance at 214 nm was monitored with elution time.

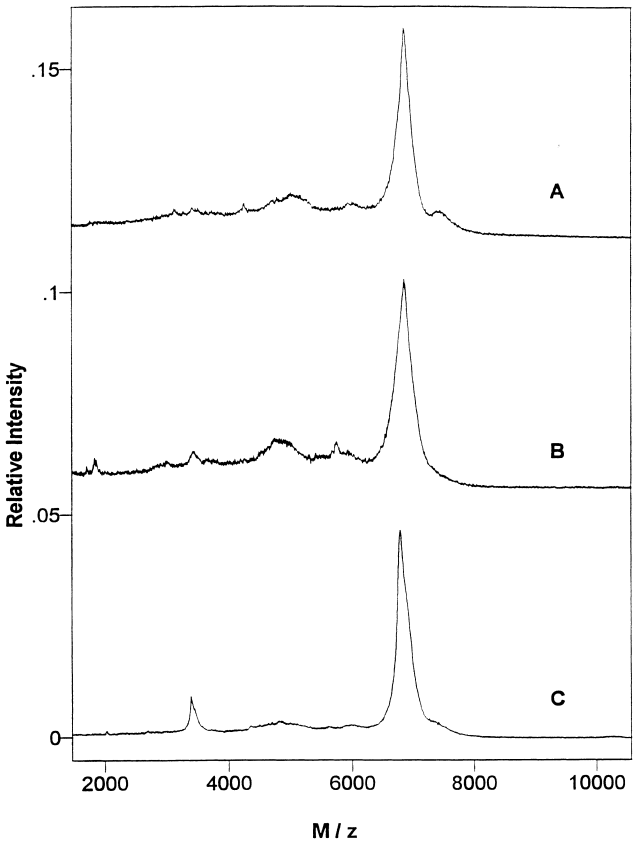


FIG. 5. MALDI mass spectra of the purified EH analogues. The molecular weights were found to be: EHQ 6816 (A); EHN 6836 (B); and EHS 6777 (C).

proteins at pharmacological doses (1 nmol/animal) triggered about 70–80% of the injected pharate adults to emerge earlier than normal (Fig. 10). The response to EHS occurred over a 3-h period whereas the other two EH analogues stimulated a complete response by 1 h postinjection. No animals emerged early after the control of 1 nmol BPTI (bovine pancreatic trypsin inhibitor) was injected. Folded EHS stimulated eclosion in a dose-responsive manner with an apparent ED_{50} of approximately 2 pmol. Reduced and alkylated EHS exhibited a much higher ED_{50} of around 500 pmol (Fig. 11).

DISCUSSION

EH Analogue Synthesis

A better assembly was achieved for EHS than for EHQ and EHN. The reason(s) for this are unknown. However, several changes in the EHS synthesis are worth noting: (a) Trityl rather than Tmob-protected Asn was used; (b) labile Trp was replaced with L-3-(2-naphthyl)-alanine (NpA); and (c) Met residues

TABLE 3

PREDICTED BYPRODUCTS OF EHN SYNTHESIS

Product No.	Assumed Protein Sequence*	pI
1	APFLNKL	9.67
2	FASIAPFLNKL	9.67
3	IPECEDFASIAPFLNKL	4.00
4	SCIKFKGKLIPECEDFASIAPFLNKL	7.74
5	ESCIKFKGKLIPECEDFASIAPFLNKL	6.35
6	NCKKMLGAWFEGPLCAESCIKFKGKLIPECEDFASIAPFLNKL	7.74
7	NPAIATGYDPMEICIENCANCKKMLGAWFEGPLCAESCIKFKGKLIPECEDFASIAPFLNKL	4.68

* Cys^{38} is outlined to indicate the possible underprotected site; see Discussion section for details.

TABLE 4

AMINO ACID ANALYSIS OF EH ANALOGUES*

Amino Acid	Expected	EHQ (TES)	EHQ (DMS)	EHN (TIPS)	EHN (DMS)	EHS (TES)	ESH (Dual†)
Asx	5/6‡	5.1	4.9	6.6	6.1	6.6	6.3
Glx	7/6‡	7.0	7.1	6.3	6.3	6.6	6.5
Ser	2	1.6	2.1	2.4	2.2	1.5	1.9
Thr	1	0.8	0.7	0.6	0.7	0.8	1.3
Gly	4	4.0	4.0	4.0	4.2	4.4	4.5
Ala	7	6.6	6.8	7.0	7.0	6.7	6.8
Pro	5	4.8	5.1	5.0	5.3	5.0	5.4
Tyr	1	1.0	0.7	0.7	0.7	0.8	1.0
Met	2	2.3	1.8	1.8	1.8	—	—
norLeu§	(2)	—	—	—	—	2.2	¶
Ile	6	5.3	5.2	5.1	5.4	5.3	4.7
Leu	5	5.3	5.7	5.5	5.9	5.3	5.1
Phe	4	4.4	4.6	4.4	4.8	4.1	4.1
Lys	6	7.0	6.3	5.8	6.5	6.0	4.8

* the abbreviations in the parentheses indicate the cleavage protocols (see text for details).
† Cleaved first with reagent R followed by a second cleavage with TES cocktail.
‡ Slash is used to distinguish the composition in EHQ (left), EHN and EHS (right).
§ Norleucine was used to replace Met in EHS only.
¶ Norleucine content was not quantified for this assay.

were substituted with nor-Leu (Nle). Each of these replacements made the product protein less susceptible to alkylation. The second and third changes also made the product more resistant to oxidation.

Cleavage of EH analogues with reagent R resulted in products that exceeded the theoretical mass by 180–400 mass units. In most cases, the additional mass (i.e., observed by MALDI analysis in Tables 1 and 2) approached 242 Da, which is consistent with a Trityl group attached to the protein (6). The mass accuracy for values obtained from the spectra of the crude products is poor because the peaks were broad [Fig. 3(A)] when compared to purified products (Fig. 5). This peak broadening resulted from sample heterogeneity and poor resolution. These unresolved product ions, mainly due to loss of H_2O and/or NH_3, shift the centroid of an unresolved multicomponent peak (19). However, mass accuracy is still sufficient to allow the detection of an EH analogue that retained a blocking group or an analogue that was alkylated by a moiety the size of a trityl (MW 240) or Tmob group (MW 180). The DMS cocktail resulted in similar cleavage products with a mass higher than was calculated for any of the EH analogues tested.

In contrast to the carbonium ion scavengers that are components of classical Fmoc cleavage reagents (i.e., reagent R), trialkylsilanes have been reported to irreversibly prevent Trityl groups from reattaching to Cys residues (15). Unfortunately, the slower reacting triisopropylsilane (TIPS) failed to prevent products of the different EH analogue syntheses from having higher than expected molecular weights. The more rapid reacting triethylsilane (TES) kept the molecular weights of the products very close to the calculated values. Additionally, EH analogues that were first cleaved with reagent R resulted in the predicted molecular weight only after being recleaved with the TES cocktail. Both of these last points support the hypothesis that EH analogues cleaved with reagent R contain at least one Cys that remains blocked with a Trityl or covalently linked to a Tmob residue during cleavage.

The Cys residue that retains a Trityl residue can be inferred from an examination of the molecular weights of the protein mixture in the crude product of an EH analogue synthesis. The validity of this inference depends on the assumption that the major byproducts of synthesis assembly are *N*-terminally acetylated truncations of the target EH analogue (Table 3). These byproducts would be expected because a capping regimen was used in the assembly process. When EHN was cleaved by reagent R, byproducts 4, 5, and 6 (Table 2) had higher masses than when they were subsequently cleaved with the TES cocktail. In comparison, the mass of byproduct 3 did not change between the two cleavage protocols. Therefore, a Trityl or Tmob group was likely either uncleaved or reattached to Cys^{38} during the reagent R cleavage and a second TES cleavage removed the extra mass from the product.

As a note in added proof, other noncysteine peptides synthesized and cleaved with reagent R in our laboratory failed to pick up this extra mass and failed to benefit (i.e., weight reduction) by a second cleavage with the TES cocktail.

Purification

A proper bioassay is the classical method to follow the purification of a biologically active protein. However, the assay for EH required multiple nanomoles of the crude denatured protein (i.e., for folding, purification, analysis, and injection), in at least eight replicates. Additionally, even though the assay itself takes several hours, at least 7 days are required to accurately stage each animal so that they will be in the correct physiological condition to detect active EH analogues.

Alternatively, matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) is a breakthrough in providing the information for developing and monitoring purification schemes for small denatured proteins. MALDI-MS has been demonstrated to promote the ionization and subsequent precise mass analysis of proteins to over 200 kDa at low picomole levels and with a accuracy of better than 0.1% (10). MALDI-MS analysis takes less than 10 min/sample. The technique also offers great reproducibility. Another advantage of MALDI-MS is that mixtures of protein byproducts are amenable to analysis, unlike

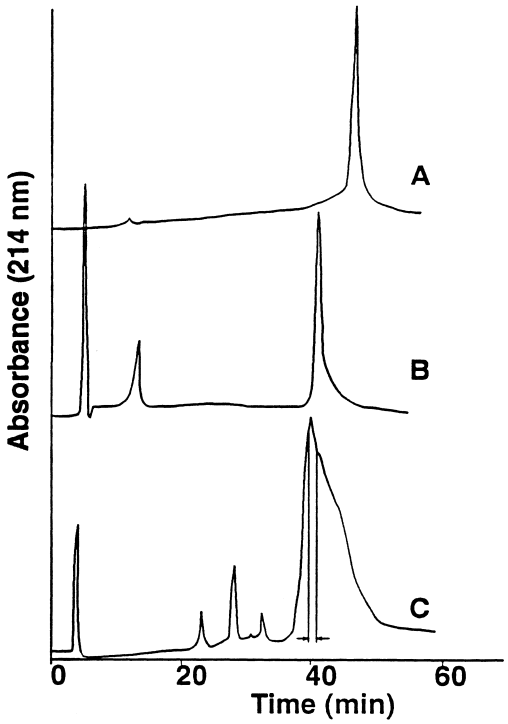


FIG. 6. Analytical RP-HPLC chromatograms of different forms of EHS. (A) Linear protein with reduced Cys residues, (B) EHS with carboxymethylated Cys, and (C) EHS after folding and alkylation of the free thiols. The fraction between arrows is fully folded with three disulfide bonds. Conditions were the same as those described in Fig. 4.

some other types of MS analysis. Thus, the synthetic mixture, without purification, can be directly applied to the sample probe for analysis and detection. In this last case, peaks can be broad as in the case of EH analogue synthesis. This results in a reduction of mass accuracy for components that are part of a crude synthetic mixture. However, despite the reduced mass accuracy, the method is sufficient to predict the general nature of protein byproducts and accurately predict some chromatographic behavior for the design of purification schemes.

MALDI-MS and the Development of a Purification Scheme

Automated solid-phase peptide synthesis instruments often produce the intended amino acid sequence in high yield. However, as peptides become longer the yield of the final target peptide drops rapidly. For a peptide with 50–70 residues, a yield of about 1–20% can be expected if a catastrophic failure does not occur (6). Because a capping protocol was used in the synthesis of each EH analogue, the major protein impurities were likely truncated species that are *N*-acetylated. This added step in the synthesis cycle limits the number of protein impurities found with the target product and simplifies the analysis and the assignment of crude synthetic protein mixtures. MALDI-MS analysis of the crude synthetic product of EH analogue synthesis revealed the approximate number of impurities and the molecular weight of each. The molecular weight was used to predict a *C*-terminal protein sequence that corresponds to each impurity. The sequence information was used to calculate a pI and relative hydrophobicity value for each protein impurity. These four pieces of information formed the basis for successful design of more complex purification schemes as were needed for the purification of EHN.

EH Analogue Folding and Biological Activity

Biological activity depends on the folded structure of EH analogue. Folding can be monitored by following disulfide bond formation (4). Cys oxidation to Cys-Cys can be quenched by alkylation with iodoacetate or iodoacetamide, which irreversibly blocks sulfhydryl groups. Alkylation with iodoacetate gives rise to one negative charge and 58 added mass units for each free sulfhydryl group. These properties can be used with appropriate analytical tools to determine the number of Cys residues in a protein and to determine the number of free thiols during the folding process. Creighton (5) first demonstrated this concept using electrophoresis to count the number of Cys residues per polypeptide chain of bovine pancreatic trypsin inhibitor.

MALDI-MS can be used in a similar manner to monitor folding. Each carboxymethylation increases the mass of EH analogue by 58. This change in mass can be observed with the mass resolution of MALDI-MS. Additionally, only about 5 pmol of EH analogue was needed for a MALDI-MS analysis. However, much more sample is needed for standard electrophoretic and HPLC detection of EH analogues.

Even though RP-HPLC requires more material than that needed for a MALDI-MS analysis, it provides complementary information. Different species were discovered by HPLC anal-

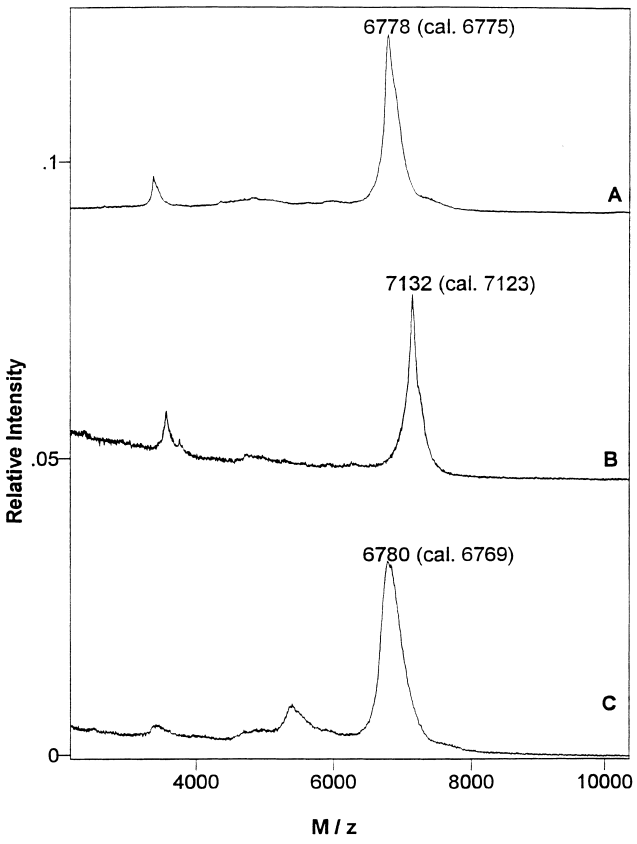


FIG. 7. MALDI mass spectra of different forms of EHS. (A) Linear protein with reduced Cys residues, (B) linear protein with reduced and carboxymethylated Cys residues, and (C) the fully folded protein.

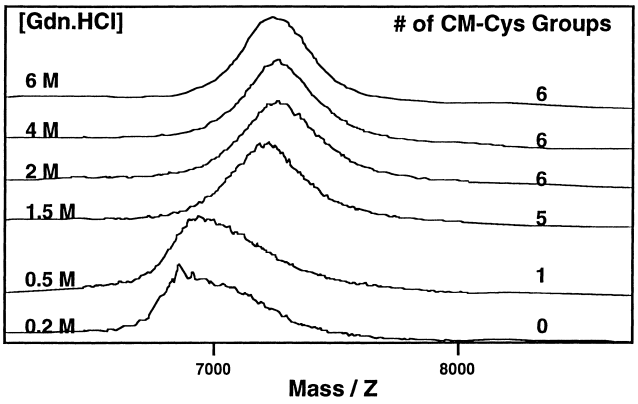


FIG. 8. MALDI mass spectra of EHS folding intermediates. Complete denatured and reduced EHS was equilibrated with 6, 4, 2, 1.5, and 0.2 *M* Gdn-HCl in 1 m*M* reduced and oxidized glutathione, 0.1 *M* Tris buffer (pH 8.7). Aliquots were taken at each Gdn-HCl concentration and were alkylated with iodoacetate. The number of CM-Cys groups in each intermediate was calculated from the difference between the observed mass and the mass for the oxidized molecule.

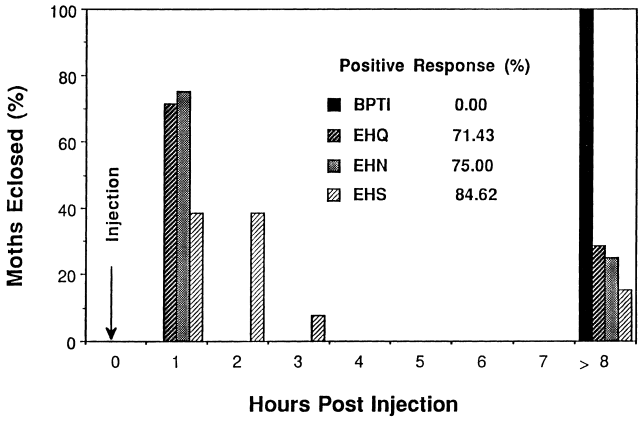


FIG. 10. EH activity of the refolded EH analogues was detected in vivo using pharate adult *Heliothis virescens.* Insects were injected with 5 μl of protein 7 h before normal eclosion. Those moths eclosed within 3 h postinjection were scored as positive responses. Bovine pancreatic trypsin inhibitor (BPTI) was used as control.

ysis as shoulders trailing the main peak [Fig. 6(C)] for the EHS sample that had been exposed to folding conditions. HPLC fractions taken from these regions of the chromatogram contain compounds with very similar molecular weights (i.e., determined by MALDI-MS analysis). When EHS folds, it is possible and likely that different disulfide bonds form and that these bonds form in different orders. All of these different molecules would then have different conformations that would expose different amounts of hydrophobicity to the stationary phase in RP-HPLC. The molecules would also have very similar molecular weights as observed for the EHS sample that had undergone the folding process. Thus, RP-HPLC is a good tool to detect a sample that contains different folded conformers. RP-HPLC could also be used to monitor the folding process to help develop folding conditions that favor a single conformer.

A unique feature of using MALDI-MS to monitor a folding process as shown in Fig. 8 is that one can infer useful stability information from these mass spectra. The formation of all three disulfide bonds took place in a cooperative manner. They formed within 1 Gdn-HCl molar concentration unit. This narrow window corresponds to the transition region in a protein

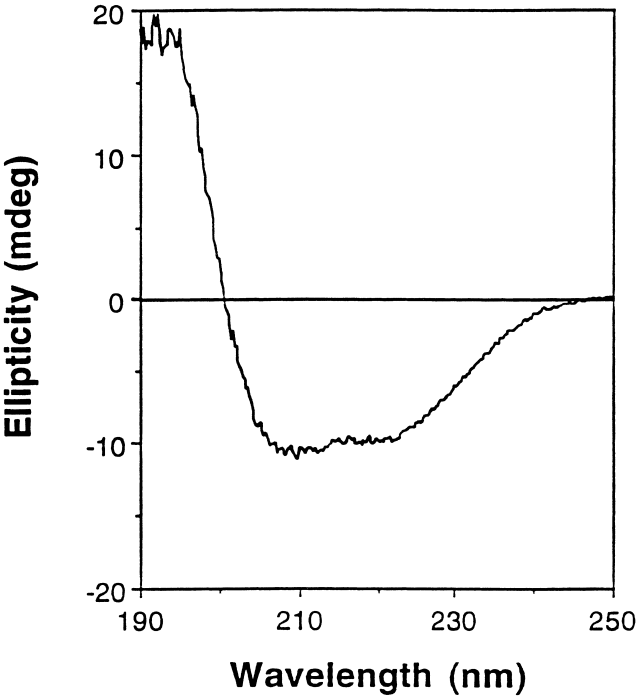


FIG. 9. Circular Dichroism spectrum of refolded EHQ. About 10 nmol refolded EHQ was dissolved in 500 μl trifluoroethanol (TFE). The CD spectrum of an average of eight traces was recorded on a JASCO J-600 spectropolarimeter with a 0.02-cm path length cell.

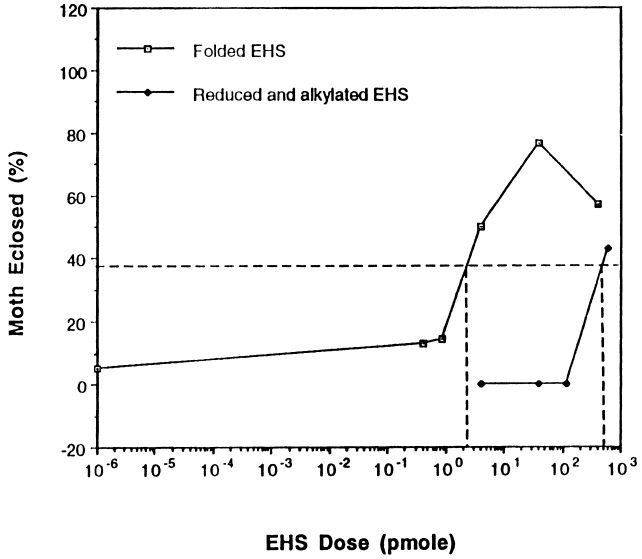


FIG. 11. Dose–response of the synthetic EHS to EH activity. Reduced and alkylated EHS [sample from Fig. 6(B)] or folded EHS [sample from the fraction between arrows in Fig. 6(C)] were injected to the animals 7 h before normal eclosion.

unfolding curve (14). This particular Gdn-HCl concentration reflects thermodynamic stability of the folded protein (i.e., a protein like EHS with a transition region between 1.5 and 0.5 *M* is less stable than a protein with a transition region at about 4–3 *M*).

EHN analogues were developed to give options for future studies. The sequence change between synthetic EHQ and EHN is very small (i.e., Asn has one less -CH_2- in its side chain than Gln) and only occurs at position 20. Both EHQ and EHN were active in the bioassay. EHN-based analogues have the potential for enzymatic cleavage (i.e., Asn C in conjunction with V8) between each Cys residue to help define each disulfide bond pair. Tools are now available with these analogues to develop studies that explore protein folding pathways.

EHS was designed to be an analogue with increased resistance toward oxidation and alkylation. However, the structural changes also made EHS a more hydrophobic analogue. Either the increased hydrophobicity or the resistance to chemical deterioration could be the reason for the increased duration of action (Fig. 9). The folded EHS sample was active in pharate adult at about 300 fmol, which is about 20-fold less potent than that reported for the native EH as isolated from *M. sexta* (11). This difference is partly due to the heterogeneity in conformation. As expected, the conformation of EH analogues is likely important for interaction with EH receptors. The 250-fold decrease in potency for the reduced alkylated EHS clearly conforms with this hypothesis.

The MALDI mass spectrometric approach offers several advantages over the traditional analytical tool for supporting the synthesis, purification, and folding of small bioactive proteins. It allows one to measure protein masses with a typical accuracy of 0.1–0.01% in complex protein mixtures, such as obtained from solid-phase peptide synthesis or in the refolding medium. The sensitivity of the method is high and lies in the subpicomole level. The time needed for an analysis is about 10 min. It helps bring the advantages of solid-phase peptide synthesis to small proteins.

ACKNOWLEDGEMENTS

We are grateful to the Texas Agricultural Experiment Station, the Institute of Biosciences and Technology, Biotechnology Instrumentation Laboratory, and the Engineering Biosciences Research Center at Texas A&M University for their support of this research. We also thank Dr. Mark Wright for his assistance with sequence analysis and Freddie Nails, Jr. for help with the amino acid analysis. The CD instrument was kindly provided by Dr. Thomas O. Baldwin at the Center for Macromolecular Design and the Department of Biochemistry and Biophysics at Texas A&M University.

REFERENCES

1. Albericio, F.; Kneib–Cordonier, N.; Biancalana, S.; Gera, L.; Masada, R. I.; Hudson, D.; Barany, G. Preparation and application of the 5-(4-(9-fluorenylmethyloxycarbonyl)aminomethyl-3,5-dimethoxyphenoxy)-valeric acid (PAL) handle for the solid-phase synthesis of *C*-terminal peptide amides under mild conditions. J. Org. Chem. 55:3730–3743; 1990.
2. Barany, G.; Kneib–Cordonier, N.; Mullen, D. G. Solid-phase peptide synthesis: A silver anniversary report. Int. J. Pept. Protein Res. 30:705–739; 1987.
3. Bindokas, V. P.; Adams, M. E. Hemolymph composition of the tobacco budworm, *Heliothis virescens* F (Lepidoptera: Noctuidae). Comp. Biochem. Physiol. [A] 90(1):151–155; 1988.
4. Creighton, T. E. Experimental studies of protein folding and unfolding. Prog. Biophys. Molec. Biol. 33:231–297; 1978.
5. Creighton, T. E. Counting integral numbers of amino acid residues per polypeptide chain. Nature. 284:487–489; 1980.
6. Fields, G. B.; Noble, R. L. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. Int. J. Pept. Protein Res. 35:161–214; 1990.
7. Frankel, A. D.; Biancalana, S.; Hudson, D. Activity of synthetic peptides from the Tat protein of the immunodificiency virus type I. Proc. Natl. Acad. Sci. USA 86:7397–7401; 1989.
8. Greenfield, N.; Fasman, G. D. Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry 8(10):4108–4116; 1969.
9. Hunkapiller, M. W. Gas phase sequence analysis of proteins/peptides. In: Bhown, A. S., Ed. Protein/peptide sequence analysis: Current methodologies. Boca Raton, FL: CRC Press, Inc.; 1988:87–117.
10. Karas, M.; Bahr, U. Laser desorption ionization mass spectrometry of large biomolecules. Trends Anal. Chem. 9:321–325; 1990.
11. Kataoka, H.; Troetschler, R. G.; Kramer, S. J.; Cesarin, B. J.; Schooley, D. A. Isolation and primary structure of the eclosion hormone of the tobacco hornworm, *Manduca sexta.* Biochem. Biophys. Res. Commun. 146(2):746–750; 1987.
12. Marti, T.; Takio, K.; Walsh, K. A.; Terzi, G.; Truman, J. W. Microanalysis of the amino acid sequence of the eclosion hormone from the tobacco hornworm *Manduca sexta.* FEBS Lett. 219:415–418; 1987.
13. Mock, K. K.; Davey, M.; Stevenson, M. P.; Cottrell, J. S. The integration of mass spectrometry into the biochemistry laboratory. Biol. Mass Spectrom. 19:948–953; 1991.
14. Pace, C. N.; Shirley, B. A.; Thomson, J. A. Measuring the conformational stability of a protein. In: Creighton, T. E., Ed. Protein structure: A practical approach. New York: IRL Press; 1989:311–329.
15. Pearson, D. A.; Blanchette, M.; Baker, M. L.; Guindon, C. A. Trialkylsilanes as scavengers for the trifluoroacetic acid deblocking of protecting groups in peptide synthesis. Tetrahedron Lett. 30(21):2739–2742; 1989.
16. Preston, L. M.; Murray, K. K.; Russell, D. H. Reproducibility and quantitation of matrix assisted laser desorption ionization mass spectrometry: Effects of nitrocellulose on peptides ion yields. J. Biol. Mass Spectrom. 22:544–550; 1993.
17. Rovero, P.; Quartara, L.; Fabbri, G. Solid-phase synthesis of neurokinin A antagonists: Comparison of the Boc and Fmoc methods. Int. J. Pept. Protein Res. 37:140–144; 1991.
18. Sabatier, J. M.; Darbon, H.; Fourquet, P.; Rochat, H.; Rietschoten, J. V. Reduction and reoxidation of the neurotoxin II from the scorpion *Androctonus australis Hector*. Int. J. Pept. Protein Res. 30:125–134; 1987.
19. Solouki, T.; Gillig, K. J.; Russell, D. Mass measurement accuracy of matrix-assisted laser desorbed biomolecules: A Fourier-transform ion cyclotron resonance mass spectrometry study. Rapid Commun. Mass Spectrom. 8:26–31; 1994.
20. Truman, J. W. The eclosion hormone systems of insects. Prog. Brain Res. 92:361; 1992.
21. White, P. Fmoc-Trp(Boc)-OH: A new derivative for the synthesis of peptides containing tryptophan. In: Smith, J. A.; Rivier, J. E., Eds. Peptides: Chemistry, structure & biology: Proceedings of the 12th APS. Leiden: ESCOM Science Publisher B. V.; 1992:537–538.
22. Zitnan, D.; Kingan, T. G.; Hermesman, J. L.; Adams, M. E. Identification of ecdysis-triggering hormone from an epitracheal endocrine system. Science 271:88–91; 1996.