Slow Folding of Three-Fingered Toxins Is Associated with the Accumulation of Native Disulfide-Bonded Intermediates†

Margherita Ruoppolo,*,‡,§ Fabio Talamo,‡,§,|| Piero Pucci,⊥,#,△ Mireille Moutiez,▽ Eric Quèmèneur,○ Andrè Mènez,▽ and Gennaro Marino⊥,#,△

Dipartimento di Chimica, Università degli Studi di Salerno, Salerno, Italy, Centro Internazionale di Servizi di Spettrometria di Massa, CNR−Università di Napoli Federico II, Napoli, Italy, Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II, Napoli, Italy, CEINGE, Biotecnologie Avanzate, scrl, Napoli, Italy, CEA, Département d'Ingènierie et d'Etudes des Protèines, Saclay, France, and CEA Valrhô, SBTN, Département d'Ingénierie et d'Etudes des Protéines, Marcoule, France

Received June 11, 2001; Revised Manuscript Received September 5, 2001

ABSTRACT: Snake neurotoxins are short all-β proteins that display a complex organization of the disulfide bonds: two bonds connect consecutive cysteine residues (C43−C54, C55−C60), and two bonds intersect when bridging (C3−C24, C17−C41) to form a particular structure classified as "disulfide β-cross". We investigated the oxidative folding of a neurotoxin variant, named α62, to define the chemical nature of the three-disulfide intermediates that accumulate during the process in order to describe in detail its folding pathway. These folding intermediates were separated by reverse-phase HPLC, and their disulfide bonds were identified using a combination of tryptic hydrolysis, manual Edman degradation, and mass spectrometry. Two dominant intermediates containing three native disulfide bonds were identified, lacking the C43−C54 and C17−C41 pairing and therefore named des-[43−54] and des-[17−41], respectively. Both species were individually allowed to reoxidize under folding conditions, showing that des-[17−41] was a fast-forming nonproductive intermediate that had to interconvert into the des-[43−54] isomer before forming the native protein. Conversely, the des-[43−54] intermediate appeared to be the immediate precursor of the oxidized neurotoxin. A kinetic model for the folding of neurotoxin α62 which fits with the observed time−course accumulation of des-[17−41] and des-[43−54] is proposed. The effect of turn 2, located between residues 17 and 24, on the overall kinetics is discussed in view of this model.

Understanding protein folding requires the identification of the intermediate conformer(s) that form(s) on the pathway(s) leading the unfolded state to the native structure of a polypeptide chain. The occurrence of either a single well-defined folding pathway or an ensemble of heterogeneous conformers forming a thermodynamic funnel has been intensively debated over the past decade, although some convergent views seem now to emerge (*1*−*4*). Knowledge of the basic rules might allow the prediction of the folding properties of some proteins with a reasonable accuracy, at least in the core of small and highly cooperative ones (*5*). The question is more complex for disulfide-bonded proteins because the combination of both kinetic and thermodynamic parameters of disulfide bond formation has a great influence on the accumulation of some intermediates (*6*).

During the past 3 decades, the folding pathways of various small proteins containing disulfide bonds have been studied by identifying the nature of the intermediates that accumulate during their oxidative folding process (*6*−*13*). At present, no single folding scenario seems to emerge distinctly from these detailed studies. Therefore, the molecular mechanisms that are associated with the folding pathways of small disulfide-containing proteins still remain to be disclosed. Within this context, folding studies of novel protein models, especially those with high pharmacological potential, are very challenging.

We have been involved in the elucidation of the folding properties of snake toxins, short all-β proteins, which display a complex organization of the disulfide bonds. Two S−S bonds connect consecutive cysteine residues (C43−C54, C55−C60) and two bonds intersect when bridging (C3−C24, C17−C41) to form a particular structure classified as "disulfide β-cross" (*14*), because cysteine residues tend to make a cross symbol when viewed along the length of the β-strands. The general organization of the polypeptide chain in snake toxins generates a trefoiled structure termed "three-

† This work was supported by EC Grant BIO4-CT96-0436 to G.M. and E.Q., by Progetto Finalizzato Biotecnologie, CNR, and by Progetto di Ricerca di Interesse Nazionale, PRIN, Murst, Roma, to G.M. and P.P.

* To whom correspondence should be addressed at the Dipartimento di Chimica, Università degli Studi di Salerno, Via S. Allende, 84081 Baronissi (Salerno), Italy. Tel: 0039-089-965298. Fax: 0039-089-965296. E-mail: ruoppolo@mbox.chem.unisa.it.

‡ M.R. and F.T. equally contributed to this paper.

§ Dipartimento di Chimica, Università degli Studi di Salerno.

|| Present address: Fondazione Centro San Raffaele, DIBIT-hSR, Milano, Italy.

⊥ Centro Internazionale di Servizi di Spettrometria di Massa, CNR−Università di Napoli Federico II.

Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II.

△ CEINGE, Biotecnologie Avanzate.

▽ CEA, Département d'Ingènierie et d'Etudes des Protèines.

○ CEA Valrhô, SBTN, Département d'Ingénierie et d'Etudes des Protéines.

10.1021/bi0111956 CCC: $20.00
Published on Web 11/20/2001

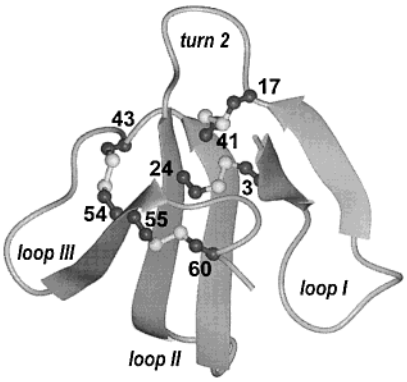

FIGURE 1: Ribbon structure of a snake neurotoxin containing 62 residues. The figure shows the general organization of the structure, particularly how the three loops fold under the control of the complex connectivity of the disulfide bonds (numbers are those of cysteine residues). Turn 2, which was described as a major determinant of the rate of folding, is pointed out. The figure was drawn with MolMol (R. Koradi, ETH Zürich) using the coordinates of erabutoxin (PDB file 1ERA).

fingers fold" (Figure 1). "Three-fingered" snake toxins act by blocking ion channels (neurotoxins), or enzymes such as acetylcholinesterase (fascicullin) or Na−K-ATPase and protein kinase C (cardiotoxins), in addition to less-defined targets. The disulfide β-cross fold is adopted by many proteins with unrelated functions, like three-fingered snake neurotoxins (*15*, *16*), skin-secreted proteins from *Xenopus* (*17*), wheat germ agglutinin (*18*), CD59, a membrane inhibitor of complement activation (*19*, *20*), each of the three domains of the extracellular portion of the urokinase-type plasminogen activator (*21*), and a ligand binding domain of type II activin receptor (*22*). The three-fingers fold seems therefore to be a common fold that can accommodate a large number of amino acid modifications in related sequences. This has recently been exemplified through the engineering of a fifth disulfide bond into loop II of neurotoxin α to shift its specificity (*23*) or the reengineering of loops I and II resulting in a fascicullin/toxin α chimera (*24*). In both cases, the ability of the proteins to fold was preserved, despite the complex disulfide bond pattern they had to acquire. Although a hydrophobic collapse might control some stages of the folding of three-fingered toxins (*25*), the major favorable component seems to be the topology of the disulfide β-cross structure which makes it a good folding nucleus (*14*), although the rationale of this assumption is still missing.

In a recent study (*26*), it was shown that a single mutation located in an appropriate site of three-fingered neurotoxin might suffice to substantially alter the rate of the sequential folding process of the protein and to deeply modify the proportion of intermediates that accumulate during its oxidative folding process. The authors used three variants of a four disulfide-containing three-fingered neurotoxin, for which the length of a large turn, turn 2 (Figure 1), was increased from 3 (variant α60, sequence CPG) to 4 (toxin α, sequence CPGE) and then to 5 (variant α62, sequence CSPGE) residues. The increase in the length of turn 2 slowed the folding process so that, after 2 h of incubation, about 90% of the native variant α60 was folded whereas it was no more than 70% for the 61-residue toxin α and 20% for the variant α62. These data are in agreement with a previous proposal (*27*) that turn 2, which connects loops I and II (Figure 1), would be directly involved in the control of the overall kinetics of the folding reaction of related neurotoxins. Interestingly, three disulfide-containing intermediates were the predominant species for all variants folding, but this population was markedly more abundant and persistent for the slowest folding neurotoxin. This finding raised the hypothesis that the rate-limiting step of the folding reaction might result from a difficult rescue of some kinetically trapped non-native intermediates (*26*).

This paper reports the characterization of the three-disulfide intermediates in the folding of neurotoxin α62 with the aim to answer the question left open in previous work and, more generally, to dissect the folding process of small proteins with complex disulfide connectivity. Using RP-HPLC,[1] two intermediates containing three disulfides were found to accumulate to detectable levels during early and late stages of neurotoxin α62 folding. Tryptic digestion and manual Edman degradation in combination with mass spectrometry led to the identification of the disulfide bonds present in these two intermediates. Both intermediates consist of chemically homogeneous species containing three of the four native disulfide bonds and lacking the C43−C54 and the C17−C41 coupling, respectively. These intermediates were individually allowed to reoxidize to the native protein. The des-[43−54] intermediate appeared to be the immediate precursor of the native species. Conversely, the des-[17−41] species was unable to form the fourth disulfide bond and had to rearrange into intermediates that can directly reach the native state. These isomerization reactions provide an explanation for the accumulation of three disulfide intermediates along the pathway, which caused the slow oxidative folding, observed for neurotoxin α62. These findings offer a new basis for understanding how a single mutation in a critical turn may affect the kinetics of folding of a three-fingered protein.

MATERIALS AND METHODS

Synthetic neurotoxin α62 (sequence: LE**C**HNQQSSQ PPTTKT**C**SPG ETN**C**YKKVWS DHRGTIIERG **C**G**C**-PTVKPGI KLN**CC**TTDK**C** NN) was prepared by solid-phase peptide synthesis on an Applied Biosystems 431A synthesizer using the Fmoc strategy. Fmoc-Asp(Rink amide MBHA resin)-OtBu (0.45 mmol/g) was from Novabiochem. Amino acid protecting groups were as follows: OtBu (E, D), trityl (C, H, N, Q), tBu (S, T, Y), Boc (K, W), and Pmc (R). Cleavage of the protein from the resin and simultaneous deprotection of the side chains were achieved by treatment with trifluoroacetic acid/triisopropylsilane/H_2O (9/0.5/0.5) for 30 min at room temperature. The crude protein was washed three times with ether and solubilized in 10% acetic acid. It was then lyophilized and purified by RP-HPLC using a C18 5 μm Bondasorb column under the following conditions:

[1] Abbreviations: CAM, carboxamidomethyl group; des-[43−54] and des-[17−41], three disulfide intermediates lacking C43−C54 and C17−C41 disulfide bonds, respectively; DTT, reduced dithiothreitol; EDTA, ethylenediaminetetraacetic acid; ESIMS, electrospray ionization mass spectrometry; GSH, reduced glutathione; GSSG, oxidized glutathione; IAM, iodoacetamide; MALDIMS, matrix-assisted laser desorption ionization mass spectrometry; MW, molecular weight; 1S, folding intermediates containing one disulfide bond; RP-HPLC, reverse-phase high-performance liquid chromatography; 3S, folding intermediates containing three disulfide bonds; 2S, folding intermediates containing two disulfide bonds; TFA, trifluoroacetic acid.

0.1% TFA (solvent A), acetonitrile (solvent B), 0−21% solvent B (5 min), 21−25% solvent B (20 min), and 25−80% solvent B (5 min).

The concentration of solutions of synthetic neurotoxin was determined using a molar extinction coefficient of 6490 M^{-1} cm^{-1} at 280 nm (*27*).

Dithiothreitol, ethylenediaminetetraacetic acid, reduced glutathione, oxidized glutathione, iodoacetamide, and trypsin were obtained from Sigma Chemical Co.; Tris was purchased from Fluka. Guanidinium chloride was acquired from Pierce. All other reagents were of the highest grade commercially available.

Folding Reaction. The neurotoxin samples were reduced and denatured as described (*26*) and used within 2 days. Reduced and oxidized glutathione stock solutions were made fresh daily in 0.1 M Tris-HCl (pH 7.5) at a concentration of 25 mM; 1 mM EDTA was added to the buffers to prevent oxidation of SH groups catalyzed by traces of heavy metals.

Lyophilized reduced and denatured neurotoxin was dissolved to a concentration of approximately 50 μM in 1% CH_3COOH and then diluted into the folding buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 7.5) to a final concentration of 15 μM. The desired amounts of GSH and GSSG stock solutions were added to initiate folding; typically, final concentrations of the glutathione species were 3 mM GSH/0.3 mM GSSG. The pH of the solution was adjusted to 7.5 with Tris base and then the reaction carried out at 25 °C under nitrogen atmosphere. The folding reaction was quenched after 90 min or 4 h by alkylation with iodoacetamide, under conditions developed to prevent reshuffling of S−S bonds (*28*, *29*). IAM was freshly dissolved in 0.1 M Tris-HCl, containing 1 mM EDTA (pH 7.5) at 65 °C, and cooled to room temperature before use. During preparation of the reagents, the solutions were protected from light to minimize photolytic production of iodine, which is a very potent oxidizing agent for thiols. The folding aliquots (1 mL) were added to an equal volume of a 2.2 M IAM solution. Alkylation was performed for 30 s, in the dark, at room temperature, under nitrogen atmosphere. After 30 s, 50 μL of 20% TFA was added, and the aliquots were quickly vortexed and fast desalted by RP-HPLC using a Vydac C18 column (5 μm, 0.46 × 25 cm). The elution system consisted of 0.1% TFA (solvent A) and 0.07% TFA in 95% acetonitrile (solvent B). The protein was fast desalted with a linear gradient of solvent B from 5% to 95% at a flow rate of 1 mL/min. The eluted protein fraction was recovered and lyophilized.

The folding was alternatively quenched by addition of hydrochloric acid to a final concentration of 3% (experimentally determined pH ∼2).

Isolation of Three-Disulfide Intermediate Population. The population of carboxamidomethylated three-disulfide intermediates was separated from the folding mixture by RP-HPLC using a Vydac C18 column (5 μm, 0.46 × 25 cm). The elution system consisted of 0.1% TFA (solvent A) and 0.07% TFA in 95% acetonitrile (solvent B). The system was equilibrated at 20% of solvent B for 5 min. Different disulfide species were separated using a linear gradient of solvent B from 20% to 27% over 18 min at flow rate of 1 mL/min. Eluted species were monitored at 220 and 280 nm, recovered, and lyophilized. Each species was characterized by ESIMS analysis.

The same procedure was also used to purify the acid-quenched three-disulfide intermediates after 90 min of folding. Eluted species were monitored at 220 and 280 nm. The HPLC chromatogram of acid-quenched intermediates (not shown) was very similar to that obtained for the carboxamidomethylated ones (shown in Figure 2A). The two major peaks were recovered and lyophilized. A small aliquot of lyophilized samples was dissolved in 1% acetic acid and added to an equal volume of 2.2 M IAM solution. The pH of the solution was adjusted to 7.5 by addition of a set aliquot of Tris base and the alkylation performed for 30 s in the dark, at room temperature, under nitrogen atmosphere. After 30 s, 50 μL of 20% TFA was added, and the samples were quickly vortexed and analyzed by RP-HPLC, exhibiting retention times shown in Figure 2A. Peaks were collected and analyzed by ESIMS, confirming that they corresponded to three-disulfide intermediates.

Peptide Mapping. Carboxamidomethylated three-disulfide intermediates contained in peaks C and D (ca. 100 μg) (Figure 2) were digested with trypsin in 50 mM ammonium acetate, pH 6.5 at 37 °C, using an enzyme:substrate ratio of 1:50 (w/w). Digestion was performed at pH 6.5 in order to avoid scrambling of the disulfide bonds. The samples were then boiled at 100 °C for 2 min and lyophilized.

Manual Edman degradation step was performed as described (*30*).

Reoxidation of Purified Des-[17−41] and Des-[43−54]. Lyophilized des-[17−41] and des-[43−54] were dissolved to a concentration of approximately 50 μM in 1% CH_3COOH and then diluted into the folding buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 7.5) to a final concentration of 15 μM. The desired amounts of GSH and GSSG stock solutions were added to give final concentrations of 3 mM GSH/0.3 mM GSSG. The pH of the solution was adjusted to 7.5 with Tris base and then the reaction carried out at 25 °C under nitrogen atmosphere. The reoxidation reaction was quenched at 0, 30, 90, and 120 min by alkylation with iodoacetamide, as described. The aliquots were fast desalted by RP-HPLC, and the eluted protein fraction was recovered and lyophilized. The reoxidation reaction was monitored by RP-HPLC using a Vydac C18 narrow-bore column (5 μm, 0.21 cm × 25 cm). The elution system consisted of 0.1% TFA (solvent A) and 0.07% TFA in 95% acetonitrile (solvent B). The system was equilibrated at 15% of solvent B for 5 min. Different disulfide species were separated using a linear gradient of solvent B from 15% to 37% over 35 min at a flow rate of 200 μL/min. Eluted species were monitored at 220 and 280 nm, recovered, and lyophilized. Each species was characterized by ESIMS analysis.

Electrospray Mass Analysis. ESIMS analyses were carried out using a BIO-Q triple quadrupole mass spectrometer equipped with an electrospray ion source (Micromass). The lyophilized samples were dissolved in H_2O containing 2% CH_3COOH and diluted 1/1 with CH_3CN, and RP-HPLC peaks were directly analyzed; 10 μL (10−20 pmol/μL) was introduced into the ion source via loop injection at a flow rate of 10 μL/min. Spectra were recorded by scanning the quadrupole at 10 s/scan. Data were acquired and elaborated by the MassLynx software. Each population of intermediates was accurately quantified by measuring the total ion current produced by each species (*8*, *9*). Mass-scale calibration was performed by means of multiply charged ions from a separate

injection of horse heart myoglobin (average molecular mass 16951.5 Da).

MALDIMS Analysis. Peptide mixtures were analyzed by MALDI-TOF mass spectrometry using a Voyager DE mass spectrometer (PerSeptive Biosystems, Boston, MA). The mass range was calibrated using bovine insulin (average molecular mass 5734.6 Da) and a matrix peak (379.1 Da) as internal standards. Samples were dissolved in 0.2% TFA at 10 pmol/μL. One microliter was applied to a sample slide before 1 μL of a solution of a cyano-4-hydroxycinnamic acid in CH_3CN/0.1% TFA (1/2) was applied (10 mg/mL). The sample and the matrix were allowed to air-dry before spectra were collected. Mass spectra were generated from the sum of 50 laser shots.

RESULTS

Isolation of Three-Disulfide Intermediate Populations. It was previously shown that 3S intermediates are abundant during most of the folding of neurotoxin α62 (*26*). The structural characterization of the 3S intermediates that accumulate during the process is now reported. The folding of denatured and reduced neurotoxin α62 was carried out as previously described (*26*). The folding intermediates, trapped in a stable form by alkylation with a large excess of iodoacetamide, were analyzed by RP-HPLC using a linear gradient of solvent B from 20% to 27% over 18 min as described in the Materials and Methods section. Two time points were selected in the initial and late phases of the folding reaction, 90 min and 4 h, respectively.

Figure 2A shows the RP-HPLC separation of intermediates trapped after 90 min of incubation. The chromatographic profile shows the presence of two main peaks, labeled C and D, respectively, and two minor peaks, named A and B, respectively. The peaks were collected and analyzed by ESIMS. The determination of their accurate molecular mass indicated that both peaks C and D contain species with three intramolecular disulfides and two carboxamidomethylated cysteines (3S2CAM) (peak C, measured MW = 6989.63 ± 0.32 Da; peak D, measured MW = 6989.48 ± 0.43 Da; calculated MW = 6989.82 Da); peak A contains the fully oxidized protein (4S) (measured MW = 6873.37 ± 0.59 Da; calculated MW = 6873.75 Da), and peak B contains species having three intramolecular disulfides, one mixed disulfide with glutathione, and one carboxamidomethylated cysteine (3S1G1CAM) (measured MW = 7238.45 ± 0.35 Da; calculated MW = 7238.14 Da).

The chromatographic analysis confirmed previous ESIMS data (*26*) in that the three-disulfide intermediates predominate from the early stages of the process (Figure 2A) when the native species is present at very low levels. The population of 3S species splits into at least two isomers, peaks C and D, suggesting that the isomeric components might have different conformations after formation of the disulfide bonds. Moreover, it should be emphasized that each peak, C and D, might in theory consist of a mixture of isomers containing three S−S bonds but involving different cysteine residues.

Figure 2B shows the RP-HPLC chromatogram of the sample withdrawn after a 4 h reaction. The RP-HPLC and ESIMS analyses revealed the occurrence of the same

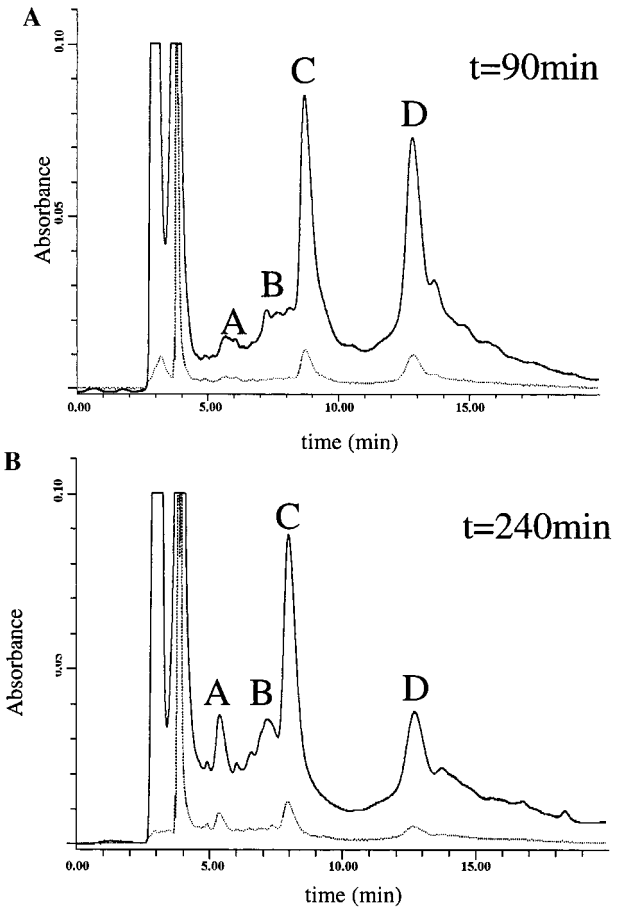

FIGURE 2: Folding of neurotoxin α62 followed by RP-HPLC at 220 nm (continuous line) and 280 nm (dotted lines). Chromatographic profiles refer to aliquots withdrawn at 90 min (A) and 4 h (B) of folding and alkylated with idoacetamide. Peaks C and D contain species with three intramolecular disulfides and two carboxamidomethylated cysteines (3S2CAM); peak A contains the fully oxidized protein (4S), and peak B contains species having three intramolecular disulfides, one mixed disulfide with glutathione, and one carboxamidomethylated cysteine (3S1G1CAM). For details see the text.

components detected after 90 min, but accumulating at different levels, with a decrease of peak D and an increase of the fully oxidized protein, peak A, whereas the amount of component C seems to remain constant. Peak B appeared increased at 240 min in comparison with its abundance at 90 min. From its chemical structure (3S1G1CAM), we assumed it derives from peak C and/or D on the route to the fully oxidized protein, peak A.

Peaks C and D were then collected for further structural and kinetic characterization. However, we could not collect enough material for peak B to achieve the resolution of its disulfide pattern as was done for peaks C and D.

Assignment of Disulfide Bonds within Three-Disulfide Intermediates. Peaks C and D corresponding to the three-disulfide intermediates produced after 90 min of folding (Figure 2A) were digested with trypsin, and the resulting peptide mixtures were directly analyzed by matrix-assisted laser desorption ionization mass spectrometry (MALDIMS). Figure 3 A shows a partial MALDI spectrum of the peptide mixture derived from the hydrolysis of peak D; mass signals were assigned to the corresponding peptides on the basis of the protein sequence and the specificity of the enzyme. The MALDI spectrum showed the presence of a signal occurring

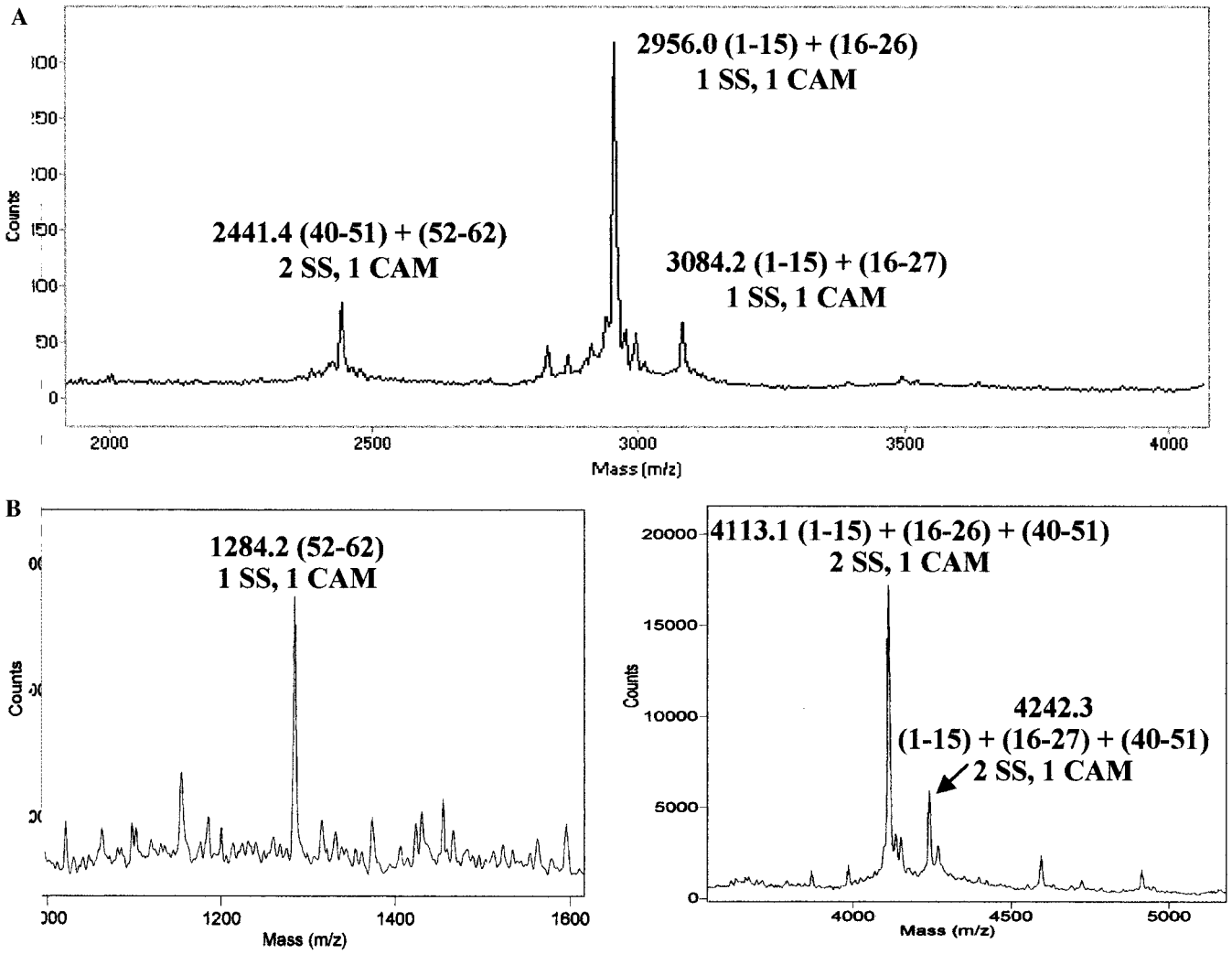

FIGURE 3: Partial MALDI spectrum of the peptide mixture derived by the tryptic hydrolysis of RP-HPLC peak D (panel A) and peak C (panel B) containing the three-disulfide species produced after 90 min of folding of neurotoxin α62 (Figure 2A). The assignment of peptides to the corresponding peaks is shown.

at *m*/*z* 2956.0 corresponding to the peptides (L1−K15) + (T16−K26) linked by one disulfide and containing one carboxamidomethylated cysteine. The peak at *m*/*z* 3084.2 represented a longer form of the same peptide pair (L1−K15) + (T16−K27), produced by a noncomplete cleavage at level of K26. The signal occurring at *m*/*z* 2441.4 is due to peptides (G40−K51) + (L52−N62) linked by two disulfides and containing one carboxamidomethylated cysteine. The presence of this cluster was due to the occurrence of a C-G-C and of a C-C sequence in neurotoxin α62 so that cysteines 41, 43, 54, and 55 cannot be separated by enzymatic hydrolysis. On the basis of the mass value alone, no unambiguous assignment of S−S bonds could be obtained since peptides (L1−K15) + (T16−K26) contain one unpaired cysteine and one disulfide, while cluster (G40−K51) + (L52−N62) is stabilized by two S−S bonds and contains an unpaired cysteine residue. These clusters could then be associated to isomeric structures with differing cysteine pairings. The structural ambiguities could be solved by submitting the peptide mixture to consecutive steps of manual Edman degradation, which selectively remove the N-terminal residues, followed by MALDIMS analysis of the truncated peptide mixture according to the strategy already described (*30*). Results are summarized in Figure 4. Two Edman degradation steps were needed to assign the disulfide bond C3−C24 occurring in the peptide pair (L1−K15) + (L16−K26). Following the first reaction cycle, the signal at *m*/*z* 2956.0 shifted back to *m*/*z* 2741.3 by the loss of L1 and T16. The second Edman step removed E2 and CAMC17, moving the signal back to *m*/*z* 2452.3 and demonstrating that C17 had been alkylated and was therefore not involved in any disulfide bond. A further cycle of Edman reaction was needed to discriminate among the possible cysteine pairings within the peptide pair (G40−K51) + (L52−N62). Following the first two degradation cycles, the signal at *m*/*z* 2441.4 moved to *m*/*z* 2270.6 (loss of G40 and L52) and then to *m*/*z* 1997.9 (removal of CAMC41 and N53), demonstrating that C41 had been alkylated and therefore not involved in any S−S bond. The third Edman step provided the chemical method to cleave the polypeptide chain between the two juxtaposed C54 and C55 residues, removing the N-terminal C54 and causing the collapse of the cluster. The corresponding MALDI spectrum in fact showed the presence of a mass signal at *m*/*z* 896.6, which was assigned to peptide (C55−N62) containing an intramolecular disulfide bond linking C55 and C60. This result was further supported by the presence of the signal at *m*/*z* 943.2 corresponding to peptide (C43−K51) containing the free C43. The second disulfide bond

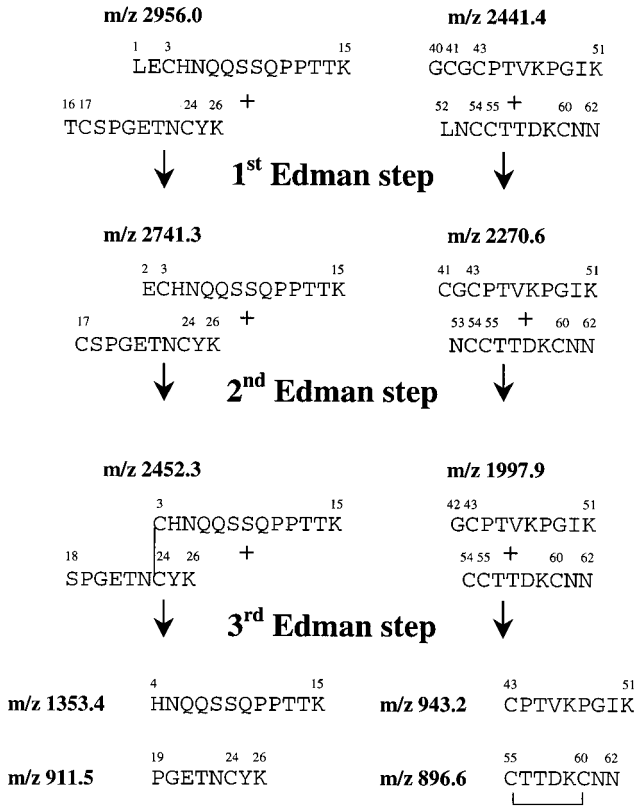

FIGURE 4: Assignment of disulfide bonds within RP-HPLC peak D containing the three-disulfide species produced after 90 min of folding of neurotoxin α62 (Figure 2A). A strategy using a combination of tryptic hydrolysis, steps of manual Edman degradation, and mass spectrometry was used.

occurring in cluster (40−51) + (52−62) was then inferred to be C43−C54 by exclusion.

The strategy illustrated above meant that it was possible to establish that peak D contains a homogeneous species characterized by the presence of three native disulfide bonds linking C3−C24, C43−C54, and C55−C60 and missing C17−C41. This intermediate was therefore named des-[17−41].

The MALDI spectrum of the peptide mixture derived from the tryptic hydrolysis of peak C (Figure 3B) showed the presence of a signal occurring at *m*/*z* 4113.1 and corresponding to the three-peptide clusters (L1−K15) + (T16−K26) + (G40−K51) linked by two disulfides and containing one carboxamidomethylated cysteine. The signal at *m*/*z* 4242.3 represented a longer form of the same peptide cluster (L1−K15) + (T16−K27) + (G40−K51), produced by a non-complete cleavage at the level of K26. In the low mass region of the spectrum, a mass peak was detected at *m*/*z* 1284.2 and assigned to peptide (L52−N62) containing one intramolecular disulfide bond and one carboxamidomethylated cysteine. As for peak D, this analysis did not lead to the assignment of the disulfide bonds pattern of the folding intermediates occurring in peak C. The peptide mixture was then submitted to three steps of Edman degradation and reanalyzed by MALDIMS as described above. The results are summarized in Figure 5. These data clearly showed that peak C consisted of a homogeneous species characterized by the presence of three native disulfides linking C3−C24, C17−C41, and C55−C60 and missing the C43−C54 coupling. This component was therefore named des-[43−54].

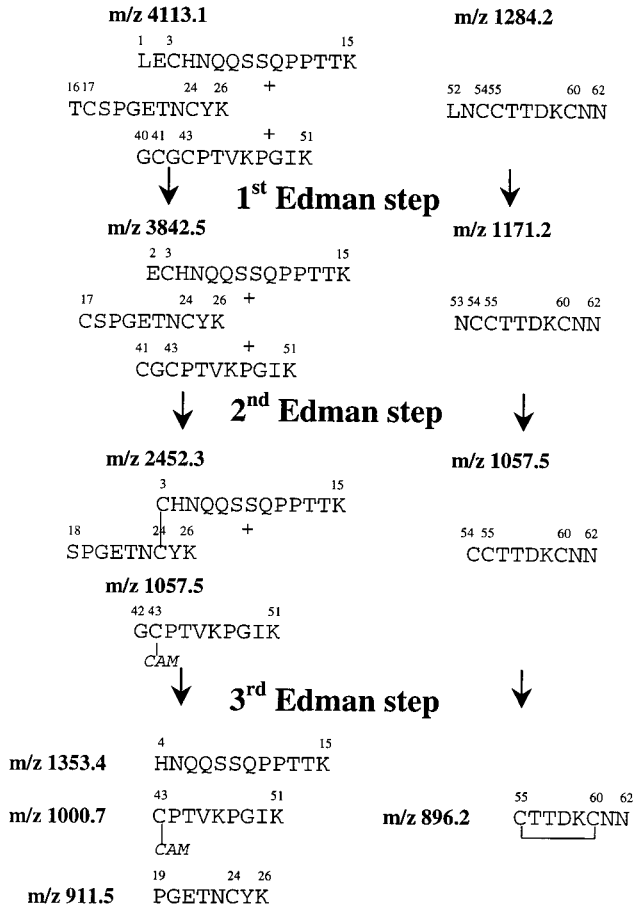

FIGURE 5: Assignment of disulfide bonds within RP-HPLC peak C containing the three-disulfide species produced after 90 min of folding of neurotoxin α62 (Figure 2A). A strategy using a combination of tryptic hydrolysis, steps of manual Edman degradation, and mass spectrometry was used.

Peaks C and D corresponding to the three-disulfide intermediates produced after 4 h of folding (Figure 2B) were also characterized using the same strategy described above. This analysis showed that they contain the same intermediates detected in the aliquot withdrawn at 90 min of the process.

Reoxidation of Purified Des-[17−41] and Des-[43−54]. A pending question was the evolution of intermediates C and D during folding in reaching the native state. This point was addressed by individually incubating each of the 3S intermediates in the appropriate folding buffer. The three-disulfide intermediates, des-[17−41] and des-[43−54], were purified from a folding experiment carried out under the conditions described above. The process was frozen by acid quenching the reaction mixture after 90 min of folding, instead of chemical quenching with iodoacetamide. The acid-quenched mixture of folding intermediates was either immediately fractionated by HPLC or stored for several hours at −50 °C prior to chromatographic analysis. Both samples yielded identical HPLC profiles, indicating that the acid quenching procedure did not produce intramolecular disulfide rearrangements.

The chromatogram of acid-quenched samples turned out to be very similar to that of carboxamidomethylated intermediates shown in Figure 2A with two major peaks eluted at 9.1 and 12.9 min, respectively, and two minor peaks. The two three-disulfide intermediates were identified by two

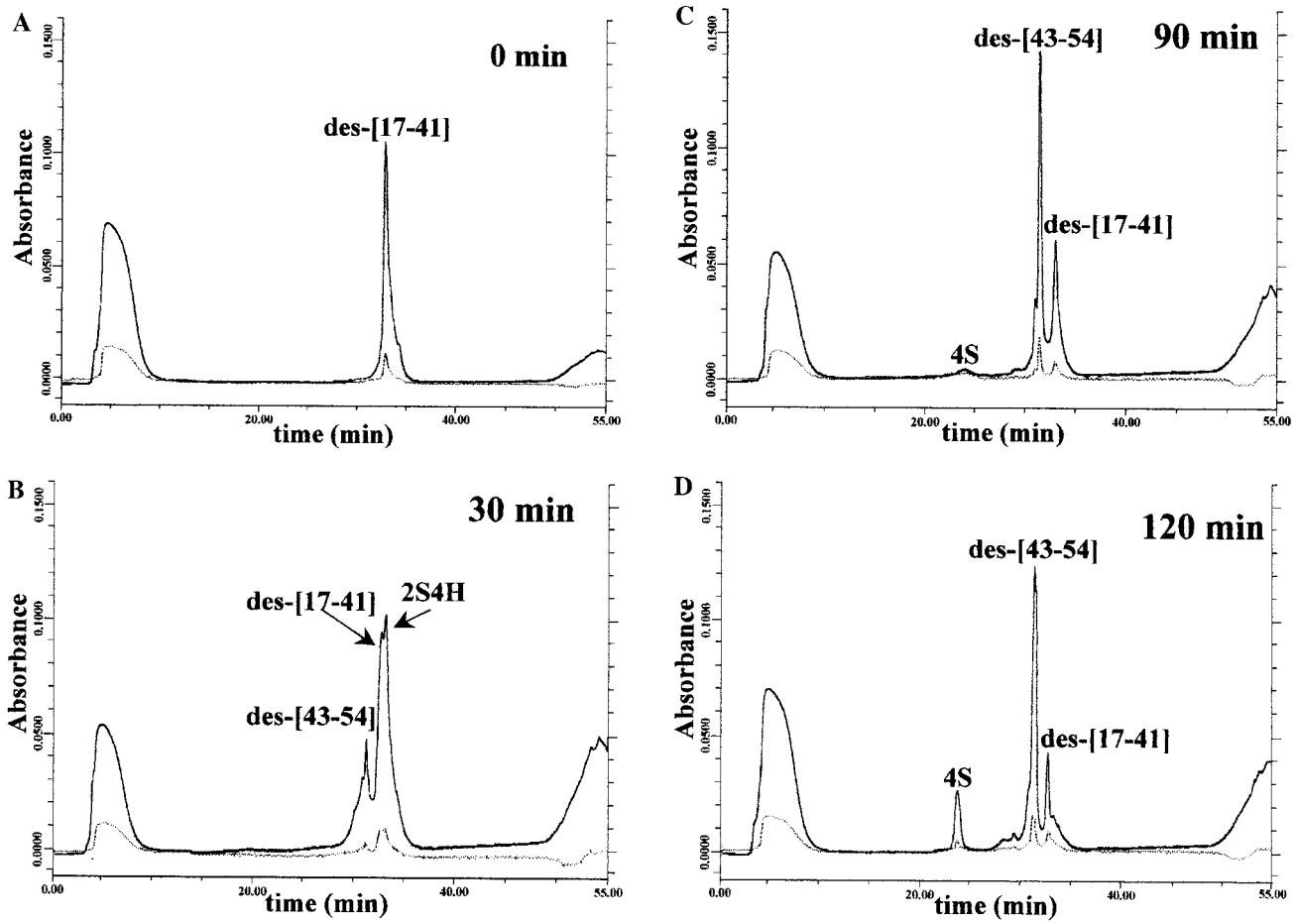

FIGURE 6: Reoxidation of the des-[17−41] intermediate followed by RP-HPLC at 220 nm (continuous line) and 280 nm (dotted lines). Chromatographic profiles refer to aliquots withdrawn at 0 (A), 30 (B), 90 (C), and 120 (D) min of the process.

independent methods (data not shown). A small aliquot of two peaks collected from the RP-HPLC separation was carboxamidomethylated as described in the Materials and Methods section and reanalyzed by RP-HPLC, revealing retention times almost coincident with those exhibited by previously characterized 3S intermediates and shown in Figure 2A. The peaks were collected and analyzed by ESIMS, showing mass values corresponding to species with three intramolecular disulfide bonds and two carboxamidomethylated cysteines. Taken together, these analyses demonstrate that the two purified acid-quenched three-disulfide intermediates were actually des-[17−41] and des-[43−54].

The reoxidation of the two intermediates, i.e., des-[17−41] and des-[43−54], was performed by reincubation in the folding buffer as described in the Materials and Methods section. Aliquots of the reoxidation mixtures were withdrawn at 0, 30, 60, 90, and 120 min of incubation, and the intermediates were trapped by alkylation with iodoacetamide. The excess blocking reagent was removed by fast desalting RP-HPLC, and the proteic peaks were collected and lyophilized. The reoxidation mixtures were then analyzed by RP-HPLC using a narrow-bore C18 column, which provides high sensitivity and resolution; the eluted peaks were characterized by ESIMS analysis. The reoxidation experiments of des-[17−41] and des-[43−54] monitored by RP-HPLC are shown in Figures 6 and 7, respectively. Figure 6A shows that a single peak corresponding to the des-[17−41] intermediate was present at time 0. After 30 min (Figure 6B) the RP-HPLC profile revealed the presence of both intermediates, des-[17−41] and des-[43−54], together with species containing two intramolecular disulfide bonds and four carboxamidomethylated cysteines (2S4H). After 90 min (Figure 6C), des-[43−54] constituted the predominant species while des-[17−41] decreased sharply and species 4S appeared in a small peak. Finally, after 120 min of incubation the chromatographic analysis showed a significant increase of species 4S. These data indicate that, during the folding process, the des-[17−41] intermediate undergoes intramolecular disulfide rearrangements to generate productive species able to evolve to the native protein.

Figure 7A shows the RP-HPLC profile of the aliquot withdrawn at time 0 of the reoxidation of the des-[43−54] intermediate, confirming the presence of a single component. The time−course analysis of the process, reported in Figure 7B−D, shows that the des-[43−54] intermediate evolves to the native species 4S, with no other components accumulating to detectable levels in the RP-HPLC chromatograms. This process, however, is quite slow with a small peak corresponding to the 4S species appearing only after 90 min.

These experimental results indicate that the appearance of the 4S occurs at about the same rate from both des-[17−41] or des-[43−54], despite the fact that des-[17−41] is a nonproductive intermediate. They can be explained on the grounds that the rate of the backward reaction of des-[17−41] to species 2S4H might be fast compared to that of the formation of the native protein from des-[43−54].

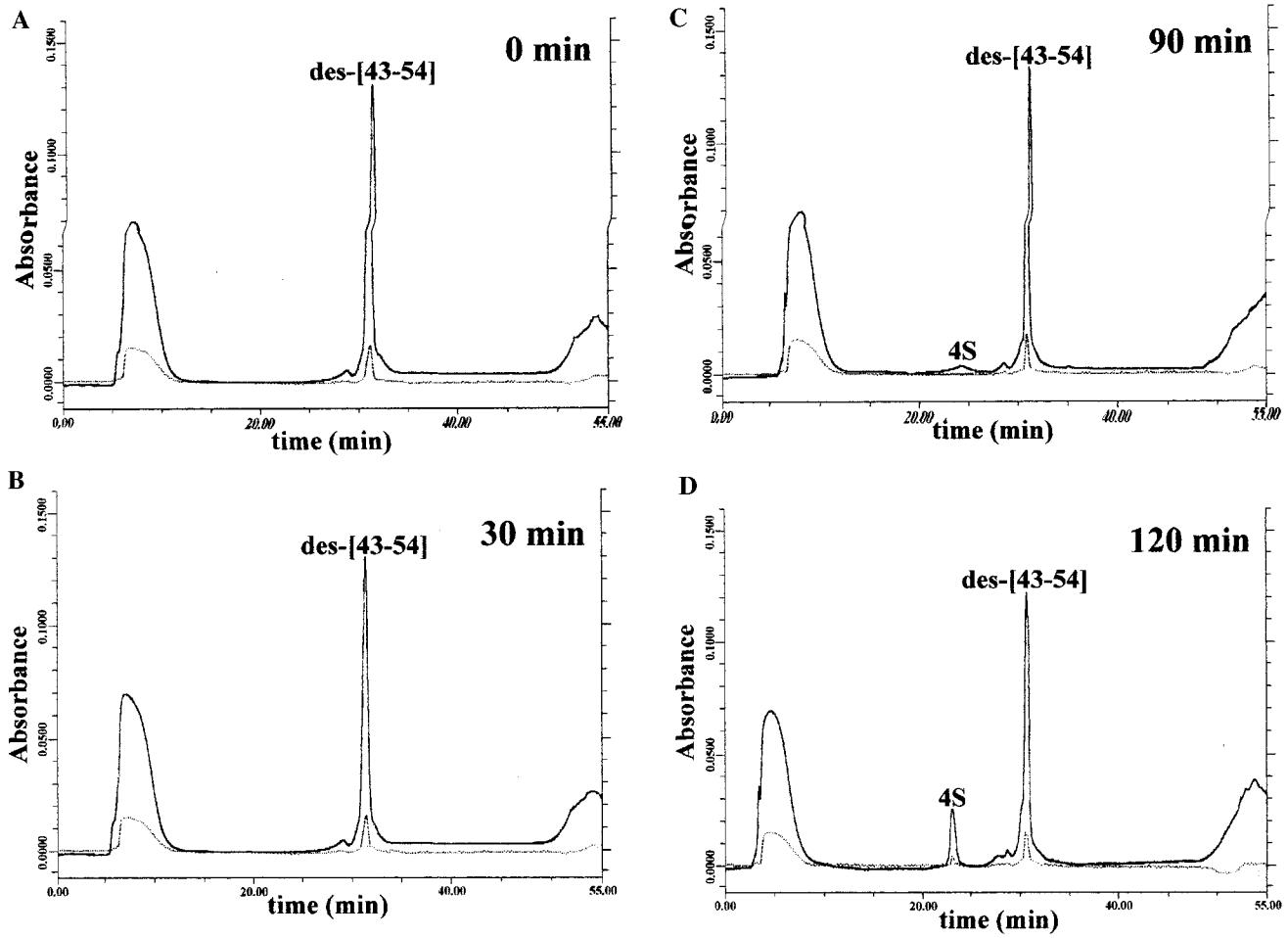

FIGURE 7: Reoxidation of the des-[43−54] intermediate followed by RP-HPLC at 220 nm (continuous line) and 280 nm (dotted lines). Chromatographic profiles refer to aliquots withdrawn at 0 (A), 30 (B), 90 (C), and 120 (D) min of the process.

DISCUSSION

It was previously demonstrated that three-fingered snake neurotoxins fold according to a sequential mechanism via reiteration of two steps: (i) formation of a mixed disulfide with glutathione and (ii) internal attack of a free SH group to form an intramolecular disulfide bond (*26*). Turn 2 (Figure 1) controlled the overall kinetics of the folding reaction: increasing the length of the turn resulted in a slowing down of the folding process with the accumulation of three-disulfide-containing intermediates (*26*). It was then suggested that the isomerization reactions occurring within this population of intermediates were the rate-limiting step of the entire process. The occurrence of isomerization reaction of the 3S population was supported by the analysis of the folding pathway of toxin α62 in which the intermediates 2S4H and the glutathione mixed disulfide species 3S1G1H accumulate and last longer throughout the folding process. These data suggested the occurrence of forward and backward reactions to rescue kinetically trapped off-pathway intermediates. The key element in elucidating the toxin folding pathway would therefore consist of the structural characterization of three-disulfide intermediates accumulating at different stages of the process.

This paper describes the characterization of the three-disulfide intermediates formed in the folding pathway of neurotoxin α62, which exhibited the slowest reoxidation rate (*26*). Reverse-phase HPLC was used to isolate two major 3S intermediates, which predominate at early and late stages of folding (peaks C and D in Figure 2). The key point to define the actual chemical nature of these species consisted of the identification of the disulfide bonds present in this population of intermediates. The occurrence of a C41-G42-C43 sequence and the presence of two juxtaposed cysteines at positions 54 and 55 exacerbated the problem. The peptide bonds between these cysteine residues had in fact to be cleaved in order to isolate each cysteine residue into a single peptide and hereby unambiguously assign the disulfide bonds. A combination of tryptic hydrolysis, manual Edman degradation steps, and mass spectrometry was found to be highly effective for the purpose. The Edman reaction provided the chemical method needed to cleave between cysteine residues where the polypeptide chain had proved to be resistant to proteases, leading to the assignment of cysteine couplings within the two purified intermediates. This combination of methodology constitutes a fine example of how the exploitation of state-of-the-art technology in conjunction with classical biochemical methods can lead to a very detailed structural characterization of disulfide bonding patterns in protein folding intermediates.

Both three-disulfide intermediates were shown to contain a different set of native disulfide bonds, missing C17−C41 (peak D) and C43−C54 (peak C) couplings, respectively. No differences were observed in the disulfide bond patterns between the intermediates formed at early and late stages of

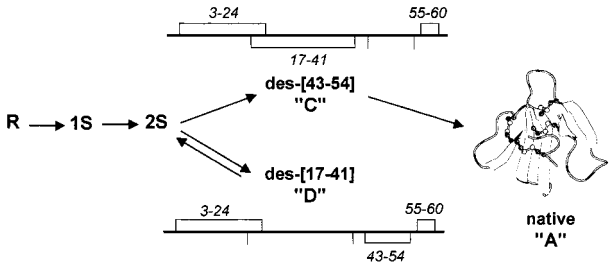

FIGURE 8: Schematic view of the main phases in the oxidative folding of neurotoxin α62. R is the reduced protein; 1S and 2S represent the population of intermediates containing one and two disulfide bonds, respectively. The two major species, which can be detected in the 3S population, are depicted in addition to the native state. A, C, and D recall the identification of these species in the chromatograms shown in Figure 2. Both 3S species contain native disulfide bonds but only des-[43−54] successfully evolves toward the native structure. Worth noting is the fact that turn 2 is located between residues 14 and 23 and might thus be critical for the formation of the disulfide β-cross involving C3−C24 and C17−C41 disulfide bonds.

the process. The rate-limiting step of the whole process should then consist of the conversion of these intermediates into the native neurotoxin. Since both species contain only native couplings and lack a single disulfide bond, this scenario might be depicted by two alternative hypotheses: (i) the two intermediates could evolve into the native structure forming the last disulfide bond independently, following different pathways, or (ii) a rearrangement of disulfide bonds should occur in order to generate the effective 3S precursor which can eventually form the native neurotoxin. Des-[17−41] and des-[43−54] intermediates were then purified and individually allowed to reoxidize in the folding buffer. The time−course analysis of these experiments showed that the des-[43−54] intermediate can evolve into the native species, whereas the des-[17−41] intermediate is not able to form the last disulfide. This species in fact had to undergo intramolecular rearrangements back to 2S intermediates and then form a productive species.

In light of these results, a model of the folding pathway of neurotoxin α62 can be proposed (Figure 8). The main features of this model can be summarized as follows. The initial step of the process is the fast formation of the three-disulfide intermediates, des-[17−41] ("D") and des-[43−54] ("C"), containing only native disulfide bonds. These intermediates were shown to be chemically homogeneous although they might be constituted by different conformers. Des-[17−41] is clearly nonproductive, at least under the present experimental conditions. It is not able to directly form the native species, but it has to isomerize into des-[43−54] intermediates before the native protein occurs. The kinetic trapping of des-[17−41] might also fit with the accumulation of species 3S1G1H at longer folding times. On the other hand, the des-[43−54] intermediate seems to constitute the effective precursor of the native protein. However, it is not possible to rule out that des-[43−54] can undergo slow intramolecular rearrangements to generate a productive intermediate which rapidly forms the native species. In this case, the final intermediate might not accumulate to detectable levels.

This model can explain the predominance of three disulfide intermediates observed along the folding of neurotoxin α62 (*26*), confirming that intramolecular rearrangements at the level of the 3S species constitute the rate-limiting step of the entire process. Moreover, the model proposed here suggests the occurrence of a temporal hierarchy in the formation of disulfide bonds in neurotoxin folding, in which the sequential order of cysteine couplings greatly affects the total process rate.

The different behavior of the two three-disulfide containing intermediates might give rise to some hypothesis on their conformation. Examination of the 3D structure of toxin α suggests that the disulfide bond C17−C41 is located at the top of both loop I (strands β1 and β2) and loop II (strands β3 and β4). It brings together strands β2 and β4, and it is critical in maintaining the integrity of loops I and II. On the other hand, the disulfide bond C43−C54 largely maintains or depends on the correct structure of loop III. The absence of pairing C17−C41 in the des-[17−41] intermediate suggests, therefore, that neither loop I nor loop II is formed, resulting in a poorly ordered polypeptidic chain between residues 25 and 42. Consequently, the polypeptide chain has to be completely twisted to form loops I and II and hence the last disulfide bond. In the des-[43−54] intermediate, both loops I and II are probably formed; thus, the occurrence of the last disulfide essentially depends on the formation of loop III. It is tempting to suggest that the concomitant presence of disulfides C3−C24 and C17−C41 might favor the formation of a more nativelike conformation of the polypeptide chain. A possible explanation is that a higher number of local constraints tends to favor native contacts which are critical in β-proteins because they contribute to the correct positioning of amino acids in early folding intermediates (*31*, *32*). Although the formation of a small hydrophobic core has been reported to control some stages of folding for a structurally related protein, cardiotoxin (*25*), the present experiments do not provide evidence for a similar situation in the case of neurotoxin α62. It should be emphasized again that species des-[17−41] and des-[43−54] are probably not conformationally homogeneous. The conformational properties of the two intermediates should therefore be further investigated to draw any definitive conclusions.

A striking observation is that turn 2, which is specifically associated with the slow rate of folding of variant α62 compared to shorter variants α60 and wild-type α61 (*26*), is located within this complex disulfide β-cross motif formed by the two disulfide bonds, C3−C24 and C17−C41. According to the proposed folding scheme (Figure 8), turn 2 might control the overall folding kinetics by (a) acting on the rate of formation of des-[43−54] and, thus, directly controlling the relative ratio of the two 3S intermediates or (b) controlling the rate of conversion of des-[17−41] into the native protein. Thus the longer turn 2 in variant α62 might impair the formation of loops I and II, slowing down the kinetics of formation of the disulfide bond C17−C41. This could decrease the overall rate of appearance of des-[43−54] with the concomitant accumulation of the dead-end intermediate des-[17−41]. Shortening this turn, as in variant α60, would increase the rate of formation of the productive intermediate, disfavoring the route to des-[17−41]. Alternatively, the long turn of variant α62 would decrease the kinetics of a putative direct conversion of des-[17−41] into the native protein to almost undetectable levels, making it impossible to observe this route under the present conditions. If this is true, a productive pathway via des-[17−

41] might be anticipated for variant α60. The current set of data does not permit a firm rejection or confirmation of these hypotheses. The characterization of 3S intermediates in the folding of variant α60 would greatly help in sorting these ambiguities.

The particular pathway adopted by neurotoxin α62 shares a number of common features with that followed by another well-studied small protein, BPTI, although the two proteins differ both in the number of disulfides and in their secondary structures. The main similarity is that no scrambled fully oxidized species occurs as the folding intermediate. In addition, all but one-disulfide-containing intermediates (respectively 3S and 2S) exclusively possess native disulfide bonds (*33*, *34*). Finally, the native form for both the variant α62 and BPTI results from an intramolecular rearrangement within their respective all but one native disulfide populations, with one predominantly productive species, des-[43−54] for the variant α62 and [Cys^{30}−Cys^{51}, Cys^{5}−Cys^{55}] for BPTI (*35*). However, formation of the native form from the productive species seems much faster in BPTI (*35*) than in neurotoxin α62. It is noteworthy that the folding pathway of variant α62 is much more comparable to that of BPTI than to those of other all-β disulfide-containing proteins, such as epidermal growth factor, EGF (*36*). In this case disulfide-scrambled isomers accumulate along the pathway, and they have to rearrange before forming the native structure.

On a more general ground, the results presented here show the complexity of the folding process of small all-β proteins with a high disulfide bond content. Species with varying degrees of structures form rapidly and interconvert each other at different rates. The observation that only intermediates with native disulfide bonds accumulate strongly suggests that native disulfide bonds are dominant early in the folding, resulting in funneling the conformations toward the native state. However, even the order of formation of the native disulfide bonds seems to be important to prevent the accumulation of nonproductive intermediates.

ACKNOWLEDGMENT

We are very grateful to Dr. Bernard Gilquin and Dr. Gilles Mourier for helpful discussions.

REFERENCES

1. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) *Proteins: Struct., Funct., Genet. 21*, 167−195.
2. Chan, H. S., and Dill, K. A. (1998) *Proteins: Struct., Funct., Genet. 30*, 2−23.
3. Pande, V. S., Grosberg, A. Y., Tanaka, T., and Roshkar, D. (1998) *Curr. Opin. Struct. Biol. 8*, 68−79.
4. Dobson, C. M., and Karplus, M. (1999) *Curr. Opin. Struct. Biol. 9*, 92−101.
5. Baker, D. (2000) *Nature 405*, 39−42.
6. Creighton, T. E., Darby, N. J., and Kemmink, J. (1996) *FASEB J. 10*, 110−118.
7. Isaacs, N. W. (1995) *Curr. Opin. Struct. Biol. 5*, 391−395.
8. Ruoppolo, M., Freedman, R. B., Pucci, P., and Marino, G. (1996) *Biochemistry 35*, 13636−13646.
9. Ruoppolo, M., Lundström-Ljung, L., Talamo, F., Pucci, P., and Marino, G. (1997) *Biochemistry 36*, 12259−12266.
10. Norton, R. S., and Pallaghy, P. K. (1998) *Toxicon 36*, 1573−1583.
11. De Young, L. R., Schmelzer, C. H., and Burton, L. E. (1999) *Protein Sci. 8*, 2513−2518.
12. Darling, R. J., Ruddon, R. W., Perini, F., and Bedow, E. (2000) *J. Biol. Chem. 275*, 15413−15421.
13. Rattenholl, A., Ruoppolo, M., Flagiello, A., Monti, M., Vinci, F., Marino, G., Lilie, H., Schwarz, E., and Rudolph, R. (2001) *J. Mol. Biol. 305*, 523−533.
14. Harrison, P. M., and Sternberg, M. J. E. (1996) *J. Mol. Biol. 264*, 603−623.
15. Ménez, A., Bontems, F., Roumestand C., Gilquin B., and Toma, F. (1992) *Proc. R. Soc. Edinburgh 99B*, 83−103.
16. Servent, D., and Ménéz, A. (2001) in *Neurotoxicology handbook: natural toxins from animal origin*, Humana Press (in press).
17. Kolbe, H. V. J., Huber, A., Cordier, P., Rasmussen, U. B., Bouchon, B., Vlasak, R., Délot, E. C., and Kreil, G. (1993) *J. Biol. Chem. 268*, 16458−16464.
18. Drenth, J., Low, B., Richardson, J. S., and Wright, C. S. (1980) *J. Biol. Chem. 255*, 2652−2655.
19. Fletcher, C. M., Harrison, R. A., Lachmann, P. J., and Neuhaus, D. (1994) *Structure 2*, 185−199.
20. Kieffer, B., Driscoll, P. C., Campbell, I. D., Willis, A. C., van der Merve, A. P., and Davis, S. J. (1994) *Biochemistry 33*, 4471−4482.
21. Ploug, M., and Ellis, V. (1994) *FEBS Lett. 349*, 163−168.
22. Greenwald, J., Fischer, W. H., Vale, W. W., and Choe, S. (1999) *Nat. Struct. Biol. 6*, 18−22.
23. Mourier, G., Servent, D., Zinn-Justin, S., and Ménez, A. (2000) *Protein Eng. 13*, 217−225.
24. Ricciardi, A., le Du, M. H., Khayati, M., Dajas, F., Boulain, J. C., Ménez, A., and Ducancel, F. (2000) *J. Biol. Chem. 275*, 18302−18310.
25. Sivaraman, T., Kumar, T. K. S., Chang, D. K., Lin, W. Y., and Yu, C. (1998) *J. Biol. Chem. 273*, 10181−10189.
26. Ruoppolo, M., Moutiez, M., Mazzeo, M. F., Pucci, P., Menez, A., Marino, G., and Quéméneur, E. (1998) *Biochemistry 37*, 16060−16068.
27. Ménez, A., Bouet, F., Guschlbauer, W., and Fromageot, P. (1980) *Biochemistry 19*, 4166−4172.
28. Torella, C., Ruoppolo, M., Marino, G., and Pucci, P. (1994) *FEBS Lett. 352*, 301−306.
29. Gray, W. R. (1993) *Protein Sci. 2*, 1732−1748.
30. Morris, H. R., and Pucci, P. (1985) *Biochem. Biophys. Res. Commun. 126*, 1122−1128.
31. Capaldi, A. P., and Radford, S. E. (1998) *Curr. Opin. Struct. Biol. 8*, 86−92.
32. Baldwin, R. L., and Rose, G. D. (1999) *Trends Biochem. Sci. 24*, 77−83.
33. van Mierlo, C. P., Darby, N. J., Neuhaus, D., and Creighton, T. E. (1991) *J. Mol. Biol. 222*, 373−390.
34. Eigenbrot, C., Randal, M., and Kossiakoff, A. A. (1990) *Protein Eng. 3*, 591−598.
35. Weissman, J. S., and Kim, P. S. (1991) *Science 253*, 1386−1393.
36. Chang, J. Y., Li, L., and Bulychev, A. (2000) *J. Biol. Chem. 275*, 8287−8289.

BI0111956