J|A|C|S
ARTICLES
Published on Web 01/20/2004

Total Chemical Synthesis of Crambin

Duhee Bang,[†,§] Neeraj Chopra,[†,‡] and Stephen B. H. Kent*[,†,‡,§]

Contribution from the Institute for Biophysical Dynamics, Department of Biochemistry and Molecular Biology, and Department of Chemistry, The University of Chicago, 920 East 58th Street, Chicago, Illinois 60637

Received September 13, 2003; E-mail: skent@uchicago.edu

Abstract: Crambin is a small (46 amino acids) protein isolated from the seeds of the plant *Crambe abyssinica.* Crambin has been extensively used as a model protein for the development of advanced crystallography and NMR techniques and for computational folding studies. We set out to establish synthetic access to crambin. Initially, we synthesized the 46 amino acid polypeptide by native chemical ligation of two distinct sets of peptide segments (15 + 31 and 31 + 15 residues). The synthetic polypeptide chain folded in good yield to give native crambin containing three disulfide bonds. The chemically synthesized crambin was characterized by LC−MS and by 2D-NMR. However, the 31-residue peptide segments were difficult to purify, and this caused an overall low yield for the synthesis. To overcome this problem, we synthesized crambin by the native chemical ligation of three segments (15 + 16 + 15 residues). Total synthesis using the ligation of three segments gave more than a 10-fold increase in yield and a protein product of exceptionally high purity. This work demonstrates the efficacy of chemical protein synthesis by the native chemical ligation of three segments and establishes efficient synthetic access to the important model protein crambin for experimental studies of protein folding and stability.

Introduction

Our aim is to apply chemistry to the study of the principles governing protein folding. Facile total chemical synthesis of a model protein will enable a wide range of experimental folding studies by providing chemical control of the protein structure. For our model protein, we chose crambin (Figure 1).

Crambin is a small (46 amino acid) protein isolated from the seeds of the plant *Crambe abyssinica*.[1,2] Although no biological function has been ascribed to crambin, it is homologous with membrane-active plant toxins.[3] The crambin molecule contains three disulfide bonds and displays β-strand, β-turn, and helical elements of protein secondary structure. Crambin has been widely used for the development of advanced NMR methods[4] and for computational studies of protein folding.[5,6] Crambin has also been used for ultrahigh-resolution X-ray crystallography[7−9] because crambin crystals are unusually well-ordered and diffract to remarkably high resolution. Because it is stable in organic media, crambin has been a useful model protein for designing enzymes to work in nonaqueous solvents.[10] Despite the importance of this protein, the expression of recombinant crambin has been problematic and results in low yields.[11] These difficulties have been attributed to the hydrophobic nature of the crambin molecule.

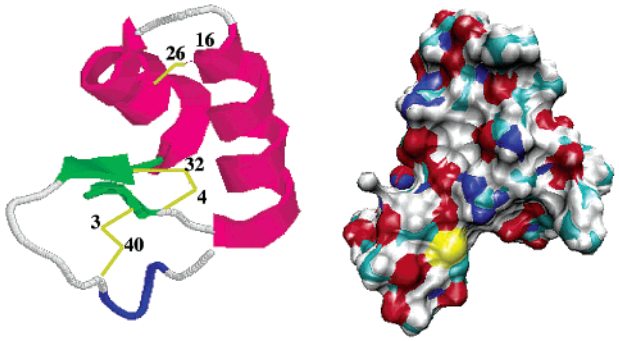


Figure 1. Molecular structure of crambin isolated from seeds of the plant *Crambe abyssinica.* Positions of the disulfides are shown in yellow. Coordinates were obtained from the Protein Data Bank (PDB accession no. 1AB1).

As the first step to understanding the molecular basis of crambin folding, we set out to establish an efficient total chemical synthesis of crambin. Using modern synthetic methods, we prepared the 46 amino acid residue polypeptide chain by the native chemical ligation of unprotected peptide segments.[12,13] Here we report a systematic exploration of synthetic routes to

[†] Institute for Biophysical Dynamics.
[‡] Department of Biochemistry and Molecular Biology.
[§] Department of Chemistry.

(1) VanEtten, C. H.; Nielsen, H. C.; Peters, J. E. *Phytochemistry* **1965**, *4*, 467−473.
(2) Teeter, M. M.; Mazer, J. A.; L'Italien, J. J. *Biochemistry* **1981**, *20*, 5437−5443.
(3) Rao, U.; Stec, B.; Teeter, M. M. *Acta Crystallogr., Sect. D* **1995**, *51*, 914−924.
(4) Xu, Y.; Wu, J.; Gorensteine, D.; Braun, W. *J. Magn. Reson.* **1999**, *136*, 76−85.
(5) Shimada, J.; Kussell, E.; Shakhnovich, E. *J. Mol. Biol.* **2001**, *308*, 79−95.
(6) Caves, L. S. D.; Evanseck, J. D.; Karplus, M. *Protein Sci.* **1998**, *7*, 649−666.
(7) Jelsch, C.; Teeter, M. M.; Lamzin, V.; Pichon-Pesme, V.; Blessing, R.; Lecomte, C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3171−3176.
(8) Teeter, M. M.; Roe, S. M.; Heo, N. H. *J. Mol. Biol.* **1993**, *230*, 292−311.
(9) Hendrickson, W. A.; Teeter, M. M. *Nature* **1981**, *290*, 109−113.
(10) Arnold, F. H. *Protein Eng.* **1988**, *2*, 21−25.
(11) Lobb, L.; Stec, B.; Kantrowitz, E. K.; Yamano, A.; Stojanoff, V.; Markman, O.; Teeter, M. M. *Protein Eng.* **1996**, *9*, 1233−1239.
(12) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−779.
(13) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923−960.

SI-form Crambin

Thr-Thr-Cys-Cys-Pro-Ser-Ile-Val-Ala-Arg-Ser-Asn-Phe-Asn-Val-Cys-Arg-Leu-Pro-Gly-Thr-Ser-Glu-Ala-Ile-Cys-Ala-Thr-Tyr-Thr-Gly-Cys-Ile-Ile-Ile-Pro-Gly-Ala-Thr-Cys-Pro-Gly-Asp-Tyr-Ala-Asn[46]

PL-form [V15A]Crambin

Thr-Thr-Cys-Cys-Pro-Ser-Ile-Val-Ala-Arg-Ser-Asn-Phe-Asn-Ala-Cys-Arg-Leu-Pro-Gly-Thr-Pro-Glu-Ala-Leu-Cys-Ala-Thr-Tyr-Thr-Gly-Cys-Ile-Ile-Ile-Pro-Gly-Ala-Thr-Cys-Pro-Gly-Asp-Tyr-Ala-Asn[46]

Figure 2. Target crambin amino acid sequences: the "SI" form has Ser[22] and Ile[25]; the "PL" form has Pro[22] and Leu[25]. In the latter case, a Val[15]-Ala substitution was made in the target sequence at the Val[15]-Cys[16] ligation site to preempt anticipated slow reaction arising from the β-branched nature of the C-terminal residue in the Cram[1−15]$^{\alpha}$thioester segment.

crambin. First, we describe synthetic schemes involving the ligation of two segments that resulted in low overall yields. Then, we describe the efficient total chemical synthesis of crambin by means of the native chemical ligation of three unprotected peptide segments. For both synthetic routes, efficient folding of the 46-residue polypeptide to form native crambin is described. The covalent and folded molecular structures of the synthetic protein products were rigorously characterized. The optimized synthetic route can provide several hundred milligrams of synthetic crambin protein from a single lab scale synthesis.

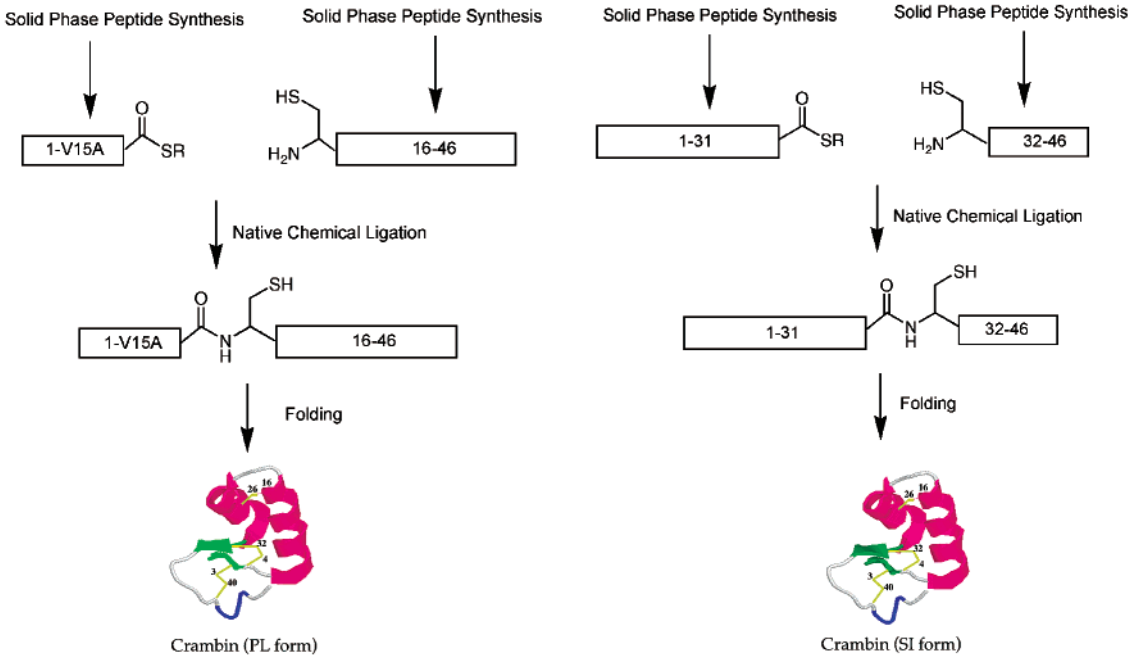
Results

Synthetic Design. The amino acid sequences of the target 46-residue polypeptide chain are shown in Figure 2. Crambin is isolated from nature in two forms: one has Ser[22] and Ile[25] ("SI" form) and the other has Pro[22] and Leu[25] ("PL" form).[2] Therefore, we explored the synthesis of both the SI form and the PL form; the SI form was explored by ligation of Cram[1−31]$^{\alpha}$thioester and Cram[Cys[32]-46], and the PL form was explored by ligation of Cram[1-V15A]$^{\alpha}$thioester and Cram[Cys[16]-46].

Initial Syntheses by Ligation of Two Segments (Scheme 1). Native chemical ligation of two peptide segments went essentially to completion in 24 h for both the SI and PL forms. In both cases, the synthetic polypeptide chain folded in good yield to give native crambin containing three disulfide bonds. The folded crambin proteins were characterized by reversed-phase high-performance liquid chromatography−mass spectrometry (LC−MS) and by two-dimensional ^{1}H NMR. However, overall yields of purified folded protein were low (~10 mg per synthesis). This was not because of the ligation reactions or the folding of the linear polypeptide chain, which proceeded in good yield. The most difficult part of the synthesis of crambin by this two-segment approach was synthesis of the purified 31-residue peptides. The crude products resulting from synthesis of Cram[1−31]$^{\alpha}$thioester and Cram[Cys[16]-46] contained substantial amounts of impurities eluting very close to the main peak in analytical reversed-phase high-performance liquid chromatography (HPLC). The presence of these impurities caused purification by preparative HPLC to be very difficult, even with repeated purification steps. For example, from several consecutive purifications of the crude peptide Cram[Cys[16]-46] (977 mg of crude), only 46 mg of purified peptide of acceptable purity was recovered (yield = 4.7%). Because of the low yields of this peptide, the overall yield for synthesis of the folded protein was very low. Therefore, we decided to study the nature of the impurity peak, with the goal of avoiding it.

Optimization. LC−MS analysis of the synthetic Cram[Cys[16]-46] showed that a des-Thr peptide was the major impurity present in the crude peptide (Figure 3A). Because of the presence of four Thr residues in the sequence of this relatively long (for MS/MS analysis) peptide, it was not possible to identify where the des-Thr occurred. Thus, we synthesized Cram[Cys(Acm)[16]-31]$^{\alpha}$thioester to systematically investigate the nature of the des-Thr byproduct and to test the feasibility of a

Scheme 1. Synthetic Strategy for the Preparation of Crambin by Ligation of Two Peptide Segments



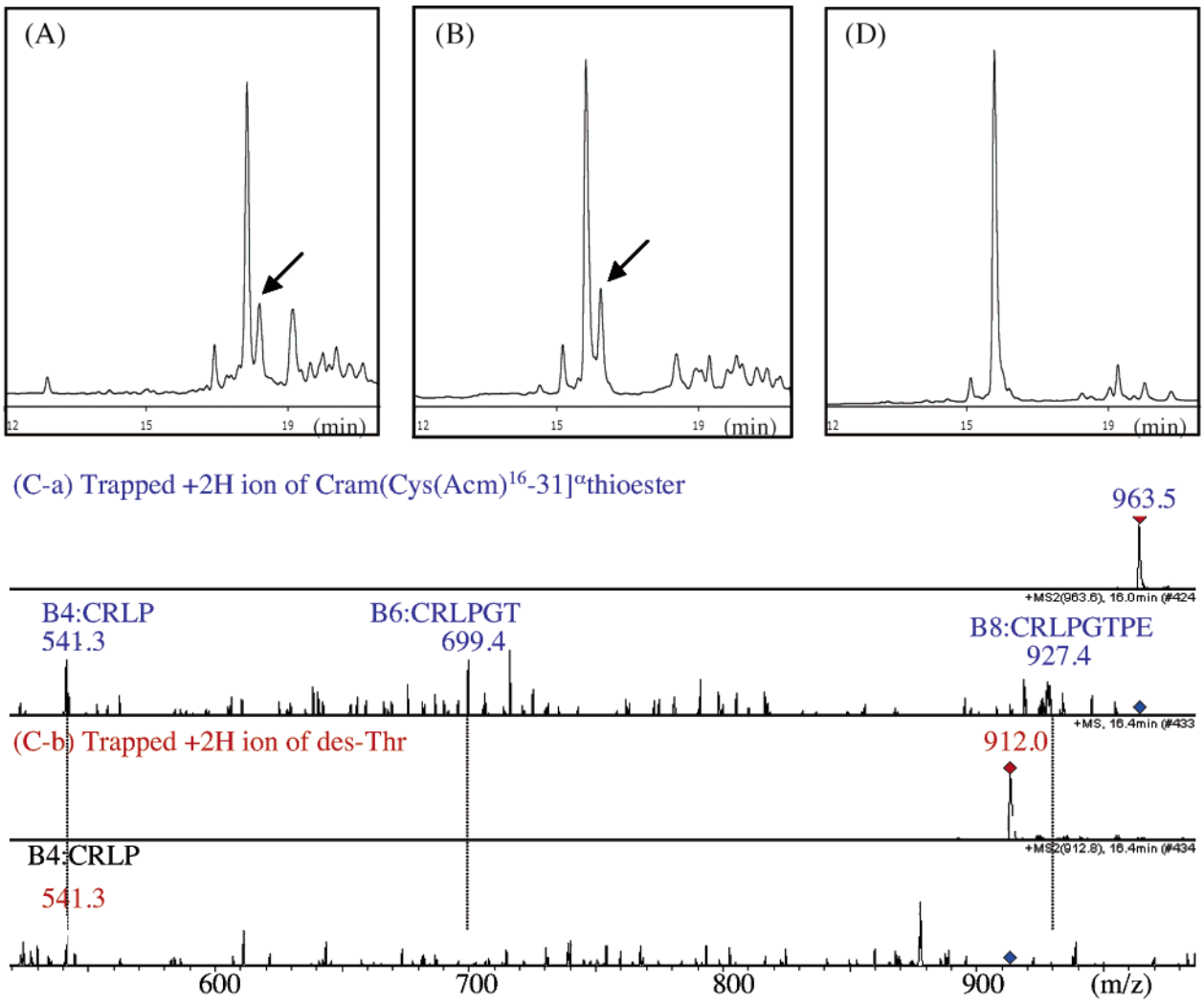


Figure 3. Optimization of peptide segment synthesis. Analytical HPLC profiles of crude peptides, monitored at 214 nm. (A) 31-residue crude peptide, Cram[Cys[16]-46] showed a large des-Thr byproduct peak (arrow). (B) 16-Residue crude peptide, Cram[Cys(Acm)[16]-31]$^{\alpha}$thioester also showed a large des-Thr peak. (C) Mass spectrometric daughter ion analyses. The main peak and the des-Thr peak from Cram[Cys(Acm)[16]-31]$^{\alpha}$thioester were analyzed by LC−MS/MS. The data showed that of the three Thr residues present in the target sequence, Thr [21] was missing. (C−a) shows trapping of $+2H^{+}$ ion from the main peak and its MS/MS singly charged daughter ions (blue). (C−b) shows $+2H^{+}$ ion of the des-Thr peak and its MS/MS data (red). B6 and B8 ions did not occur in the des-Thr byproduct. (D) Crude peptide Cram[Cys(Acm)[16]-31]$^{\alpha}$thioester obtained after triple coupling of the Thr[21] showed no des-Thr peak.

three-segment ligation. Interestingly, we also found a des-Thr impurity as a major byproduct in synthetic Cram[Cys(Acm)[16]-31]$^{\alpha}$thioester (Figure 3B). We applied LC−MS/MS (i.e., daughter ion analysis of the byproduct peak in the LC−MS run) to find which of the three Thr residues (Thr[21], Thr[28], or Thr[30]) was missing in the des-Thr byproduct. LC−MS/MS showed that the culprit was Thr[21] (Figure 3C). Subsequent systematic studies showed that triple coupling (!) of Thr[21] was necessary to remedy this problem. An optimized synthesis of Cram-[Cys(Acm)[16]-31]$^{\alpha}$thioester is shown (Figure 3D). By contrast, we found that synthesis of Cram[Cys[16]-46] gave a product that still contained substantial des-Thr byproduct even with triple coupling of the Thr[21]. Consequently, we decided to apply a three-segment ligation strategy to the synthesis of crambin.

Crambin Synthesis by Ligation of Three Segments (Scheme 2). Since the middle segment has both an N-terminal cysteine and a C-terminal thioester moiety, either the α-amino or the side-chain thiol of the N-terminal cysteine must be protected until after the first ligation to prevent undesired cyclization.[14] The acetamidomethyl (Acm) group was used to protect the side-chain thiol of the N-terminal cysteine of the middle peptide segment.[15,16] Three peptide segments, Cram[1-V15A]$^{\alpha}$thioester, Cram[Cys(Acm)[16]-31]$^{\alpha}$thioester, and Cram[Cys[32]-46] were synthesized. Purified synthetic peptides were recovered in excellent yield (∼50% for each peptide) (Table 1); in each case only a single preparative HPLC run was needed to obtain high-purity peptide.

Ligation of Cram[Cys(Acm)[16]-31]$^{\alpha}$thioester and Cram-[Cys[32]-46], Followed by Acm Group Removal. The ligation reaction of Cram[Cys(Acm)[16]-31]$^{\alpha}$thioester and Cram[Cys[32]-46] was essentially complete after 6 h (Figure 4), and the ligation product was readily purified. Initially, the Acm group was removed by the widely used deprotection protocol using mercury

(14) Tam, J. P.; Lu, Y. I. *Protein Sci.* **1998**, *7*, 1583−1592.
(15) Canne, L. E.; Botti, P.; Simon, R. J.; Chen, Y.; Dennis, E. D.; Kent, S. B. H. *J. Am. Chem. Soc.* **1999**. *121*, 8720−8727.
(16) Kochendoerfer, G. G.; Chen, S. Y.; Mao, F.; Cressman, S.; Traviglia, S.; Shao, H.; Hunter, C. L.; Low, D. W.; Cagle, E. N.; Carnevali, M.; Gueriguian, V.; Keogh, P. J.; Porter, H.; Stratton, S. M.; Wiedeke, M. C.; Wilken, J.; Tang, J.; Levy, J. J.; Miranda, L. P.; Crnogorac, M. M.; Kalbag, S.; Botti, P.; Schindler-Horvat, J.; Savatski, L.; Adamson, J. W.; Kung, A.; Kent, S. B. H.; Bradburne, J. A. *Science* **2003**, *299*, 884−887.

Scheme 2. Synthetic Strategy for the Preparation of Crambin by Ligation of Three Peptide Segments

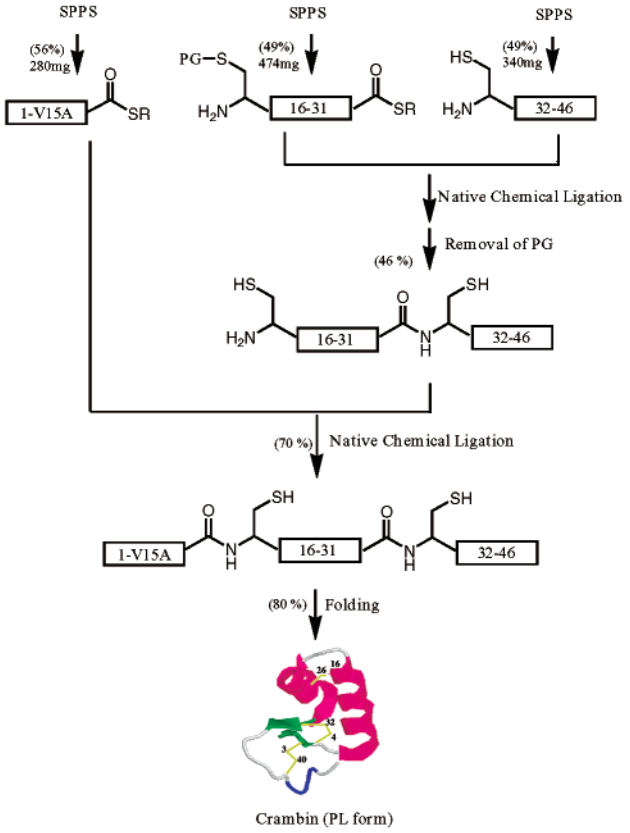


Table 1. Yields of Synthetic Peptide Segments

peptide segment	resin scale (mmol)	crude peptide amount (mg)	purified peptides (mg)	yield (%)
Cram[16−46]	0.4	977	46	4.7
Cram[Cys(Acm)^{16}-31]$^{\alpha}$ thioester	0.5	*a*	474 (0.246 mmol)	49.2
Cram[32−46]	0.4	*a*	340 (0.226 mmol)	56.4
Cram[1-V15A]$^{\alpha}$ thioester	0.4	572	280	49.0

[a] Crude peptides from HF cleavage mixture were directly injected onto the preparative HPLC column, and therefore no crude peptide weight is available.

acetate ($Hg(OAc)_2$), and the reaction was quenched with β-mercaptoethanol.[15,16] However, the quenching reaction took several hours even with 20% β-mercaptoethanol (v/v) and resulted in significant loss of deprotected peptide as Cys(SH)−mercury adducts. To optimize the removal of the Acm group, systematic studies were performed. The deprotected peptide was prepared with minimal problems with a shorter overall reaction time by the use of $Ag(OAc)_2$ with DTT quenching.[17] The recovered yield after the first ligation and Acm-removal steps was improved to 46% (details in the Experimental Section).

Ligation of Cram[1-V15A]$^{\alpha}$thioester and Cram[Cys^{16}-46], Followed by Folding of the Synthetic Polypeptide Chain. Ligation of Cram[1-V15A]$^{\alpha}$thioester and Cram[Cys^{16}-46] was essentially complete after 24 h (Figure 5). Yield of the purified

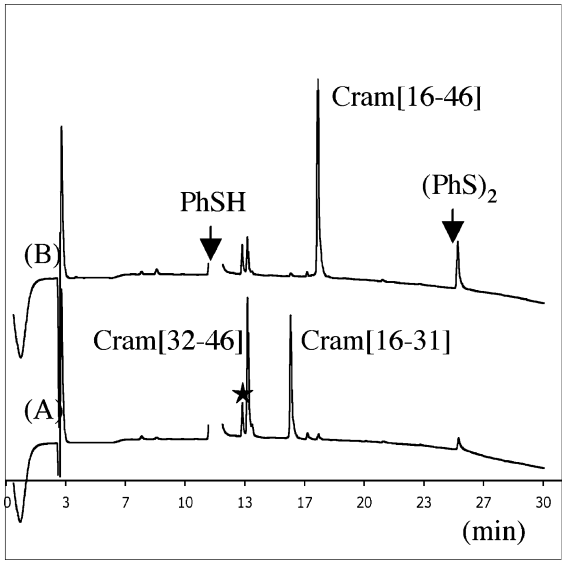


Figure 4. Ligation reaction of Cram[Cys(Acm)^{16}-31]$^{\alpha}$thioester with Cram-[Cys^{32}-46] was monitored by analytical HPLC (UV profiles at 214 nm are shown). (A) At $t = 5$ min, a small amount of a cyclized (with the internal thiol of Cys^{26}) thioester form (★) of Cram[Cys(Acm)^{16}-31]$^{\alpha}$thioester was observed. (B) At $t = 6$ h, formation of the product Cram[Cys(Acm)^{16}-46] was essentially complete.

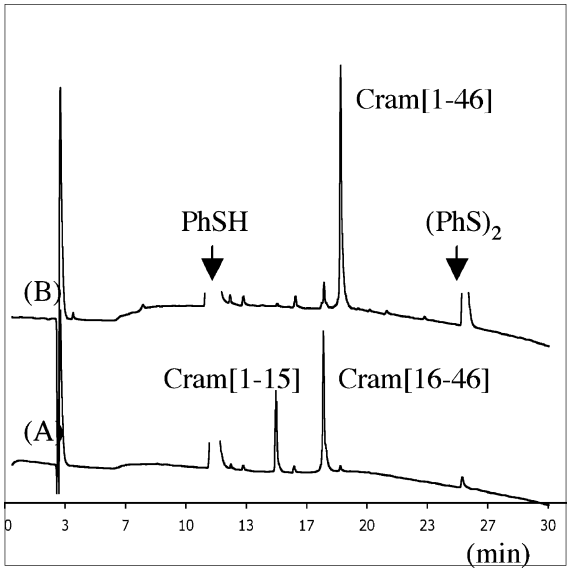


Figure 5. Ligation reaction of Cram[1-V15A]$^{\alpha}$thioester with Cram[Cys^{16}-46] was monitored by analytical HPLC (UV profiles at 214 nm are shown). (A) Reaction mixture at $t = 0$. (B) At $t = 24$ h, formation of the product Cram[1−46] was essentially complete.

ligation product Cram[1−46] was ∼70%. Folding of the polypeptide using a customized variant of standard folding/disulfide formation conditions in the presence of 2 M guanidine−HCl was complete in less than 1 h and was very efficient giving a near-quantitative yield of folded protein (Figure 6). The synthetic protein was readily purified by preparative HPLC to give highly purified crambin in good yield (∼80% based on linear polypeptide).

Characterization. Electrospray MS of the linear polypeptide gave an observed mass (obsd: 4708.6 ± 0.2 Da; calcd: 4708.4 Da) consistent with the covalent structure of the desired 46-residue target sequence. The measured mass of folded crambin was consistent with formation of three disulfide bonds (obsd:

(17) Kawakami, T.; Toda, C.; Akaji, K.; Nishimura, T.; Nakatsuji, T.; Ueno, K.; Sonobe, M.; Sonobe, H.; Aimoto, S. *J. Biochem. (Tokyo)* **2000**, *128*, 455−461.

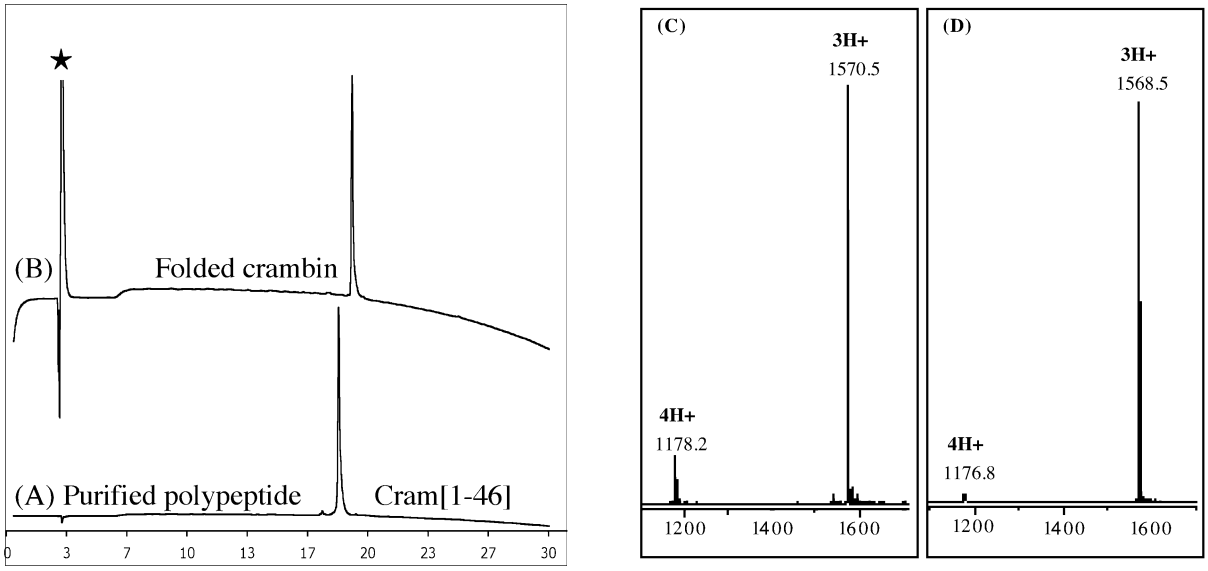


Figure 6. Folding of the polypeptide Cram[1−46] was monitored by LC−MS; (A and B) UV profiles at 214 nm. (A) Purified polypeptide, Cram[1−46] prior to the folding reaction. (B) Folding mixture at $t = 1$ h. Folded crambin eluted *later* than the linear polypeptide. (C) Mass data corresponding to the peak in (A). (D) Mass data corresponding to the peak in (B). On folding, the mass decreased by 6 Da, reflecting formation of three disulfide bonds. (★ = cysteine and cystine peak).

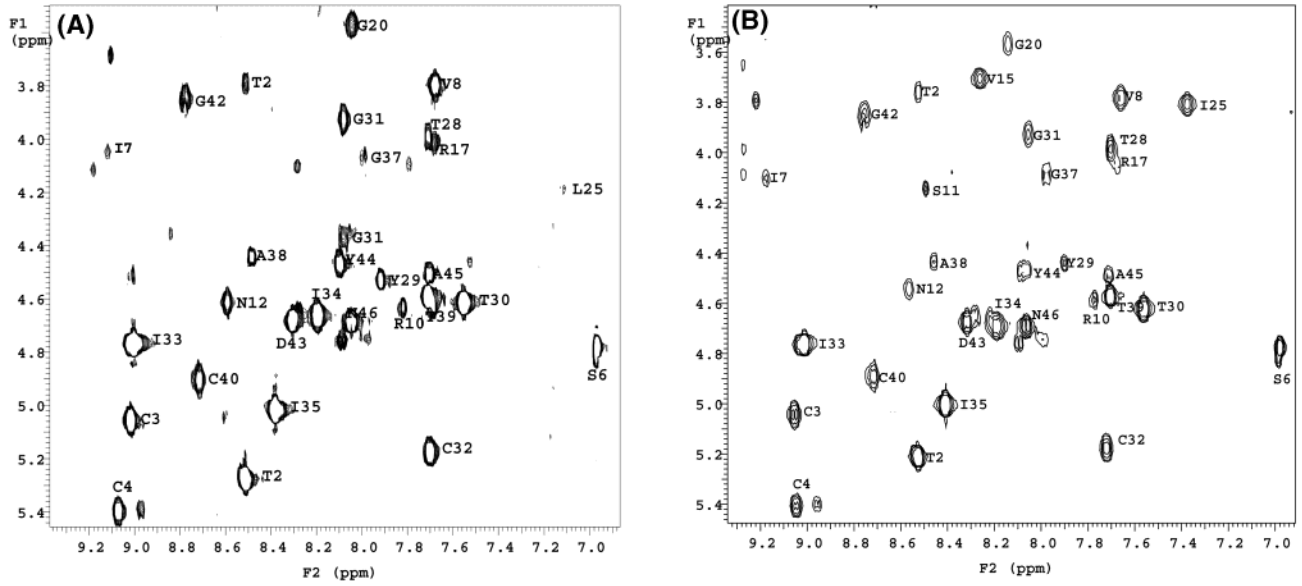


Figure 7. Two-dimensional TOCSY ^{1}H NMR spectra showing the fingerprint region of folded synthetic crambin. (A) PL-[V15A]crambin. (B) SI-crambin. Assignments were by comparison with previously published data.[4,18]

4702.8 ± 0.4 Da; calcd: 4702.4 Da; i.e., 5.8 ± 0.4 Da decrease, corresponding to the loss of six protons).

The shape of the circular dichroism spectrum of synthetic crambin was identical to that of the natural material. The folded (tertiary) structure of crambin was characterized in detail by two-dimensional ^{1}H NMR (Figure 7). Each amino acid peak from the TOCSY fingerprint region was assigned by comparison from reported data,[4,18] and the observed pattern of cross-peaks and the chemical shifts were identical with those previously reported. Thus, chemically synthesized crambin was shown to be structurally identical to crambin isolated from nature.

Discussion

Past attempts to produce the protein crambin by recombinant means have not met with success. The only published description of the recombinant expression of crambin reported a yield of only ~600 μg/L of purified folded crambin, derived from a fusion protein construct.[11] Attempts to directly express crambin in multiple expression systems led to unacceptably low yields. To the present day, an efficient production of crambin by expression in recombinant DNA engineered cells has not been achieved (M. Teeter and C. Rienstra, private communications).

In the work reported here, high-purity crambin has been prepared in good yield by optimized total chemical synthesis using modern ligation methods. Although crambin is a small protein, a synthetic approach to the 46-amino acid polypeptide chain by ligation of two unprotected peptide segments gave poor yields because of difficulties associated with the stepwise solid-

(18) Lamerichs, R. M.; Berliner, L. J.; Boelens, R.; De Marco, A.; Llinas, M.; Kaptein, R. *Eur. J. Biochem.* **1988**, *171*, 307−312.

phase synthesis of the longer peptide segments. Systematic studies showed that a three-segment ligation strategy enabled the optimization of the synthesis of all peptide segments and that it was a much more efficient synthetic approach. Systematic optimization of the chain assembly and peptide purification by a single preparative HPLC step gave good recovered yields of the three short peptide segments in high purity (Table 1). About 300 mg of each purified peptide segment was obtained from a 0.4 mmol scale solid-phase peptide synthesis.

The three-segment ligation strategy, comprising two ligation reactions and the associated deprotection, purification, and lyophilization steps, required several days to be carried out, starting from the three purified peptide segments. Several times, we conducted chemical syntheses of crambin at a multi-tens-of-milligrams-of-peptide scale. From this systematic approach, we were able to improve to 46% the overall yield for the first ligation, removal of the Acm-protecting group, and preparative HPLC. Effective removal of the Acm group from the ligation product was key to this optimization. Deprotection by mercury acetate took several hours for complete reaction and required the addition of a large amount (20% v/v) of β-mercaptoethanol to quench the reaction and disrupt the Cys(SH)−mercury complexes. Even then, some mercuric adducts of the peptide thiols still remained. The silver acetate method was easier to handle, and a moderate amount (3% v/v) of β-mercaptoethanol or dithiothreitol (DTT) could be used to quench the reaction and completely disrupt the {peptide-thiol}−metal adducts because of the formation of a metal−thiol (e.g., DTT) precipitate. This made it possible to avoid a lyophilization step after preparative HPLC of the first ligation reaction. That is, after the first ligation and subsequent preparative HPLC step, silver acetate could be added directly to the pooled fractions in HPLC solvent, and after a brief (1 h) reaction, the addition of β-mercaptoethanol or DTT was enough to immediately quench the reaction. Furthermore, simple centrifugation to remove the precipitate was followed by preparative HPLC of the supernatant solution to give a significantly improved yield of deprotected ligation product compared to the use of mercury acetate.

For the second ligation reaction, in repeated trials the average recovered yield after preparative HPLC purification of full-length 46-residue polypeptide was around 70%.

The folding reaction was very efficient. Beginning with purified polypeptide, folding gave a single protein product as a sharp and symmetrical peak by analytical HPLC, with only trace amounts of coproducts (Figure 6). Interestingly, the folded crambin protein molecule eluted *later* than the linear polypeptide in the analytical HPLC profile (Figure 6). For typical globular proteins, folding of the polypeptide chain results in *earlier* elution on HPLC analysis because of the burial of hydrophobic side chains in the interior of the protein molecule. The contrary behavior observed for crambin on folding is consistent with the hydrophobic nature of the folded protein. Average recovered yield from the folding reaction and subsequent preparative HPLC step was 80%. Thus, if all of the peptide was used for ligation and folding, it would be possible to make ~250 mg (~0.05 mmol) of high-purity crambin protein from a single lab scale synthesis (see Scheme 2). Overall, the three-segment strategy gave substantially increased yield of high-purity synthetic protein.

Despite the intense interest in crambin, there have been no reports of the study of this molecule by protein engineering. The inability to use protein engineering to systematically study this protein has led to a significant gap in our understanding of this protein model system and is the consequence of the failure to produce crambin by recombinant means.[11] Efficient chemical synthetic access to crambin will enable systematic studies of the molecular basis of a range of protein properties. To our knowledge, there have been no experimental folding studies done on crambin, in contrast to the numerous *theoretical* folding studies of crambin as a model system. Because of the facile synthetic access we report here, the molecular basis of crambin folding and the role of the disulfide bonds in its high stability can now be addressed by systematic variation of the covalent structure of the protein. In addition, synthesis of crambin analogues will be greatly facilitated by this three-segment approach: preparation of each analogue protein molecule will involve the straightforward (re)synthesis of only the variant small segment(s). Chemically synthesized crambin will also enable structure determinations using methods such as protein solid-state NMR (in conjunction with site-specific isotope labeling) and will provide for the further exploration of racemic protein crystallography.[19,20]

Conclusion

We have developed an efficient three-segment ligation approach to the total chemical synthesis of high-purity crambin in good yield. Hundred-milligram amounts of the hydrophobic protein crambin can be readily prepared in a single lab scale synthesis. For the first time, this provides complete experimental control over the protein molecule so that any desired chemical modifications of crambin can be readily achieved. This control over the covalent structure of the molecule will enable the application of chemistry to the experimental elucidation of the principles of protein folding in this important model protein that has been the focus of numerous structural and theoretical studies.

Experimental Section

Abbreviations for Peptides. Cram[1-V15A]$^{\alpha}$thioester: TTCCPSIVARSNFNA-SCH_2CH_2CO-Leu; Cram[Cys(Acm)^{16}-31]$^{\alpha}$thioester: C(Acm)RLPGTPEALCATYTG-SCH_2CH_2CO-Leu; Cram[Cys^{32}-46]: CIIIPGATCPGDYAN.

Materials. Boc-amino acids, S-trityl-β-mercaptopropionic acid, and 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were obtained from Peptide Institute, Inc. (Osaka, Japan). Boc-Asn-OCH_2-Pam-resin, Boc-Leu-OCH_2-Pam-resin, and *N*,*N*-diisopropylethylamine (DIEA) were obtained from Applied Biosystems (Foster City, CA). *N*,*N*-Dimethylformamide (DMF), dichloromethane (DCM), and acetonitrile were purchased from Fisher (Chicago, IL). Thiophenol and *p*-cresol were purchased from Sigma-Aldrich (St. Louis, MO). Trifluoroacetic acid (TFA) was from Halocarbon (New Jersey).

HPLC and LC−MS. Analytical reversed-phase HPLC and LC−MS were performed on an Agilent 1100 Series chromatography instrument equipped with an MSD ion trap, using Vydac C4 columns (5 μm, 0.46 × 25 cm). Chromatographic separations were performed using a linear gradient (10−60%) of buffer B in buffer A over 25 min at a flow rate of 1 mL/min. Buffer A = 0.1% TFA in water; buffer B = 0.08% TFA in acetonitrile. For MS data, observed masses were

(19) Mackay, A. L. *Nature* **1989**, *342*, 133.
(20) Hung, L. W.; Kohmura, M.; Ariyoshi, Y.; Kim, S. H. *J. Mol. Biol*. **1999**, *285*, 311−321.

derived from the *m*/*z* values for all protonation states of a molecular species. Calculation of theoretical masses was based on average isotope composition. Preparative HPLC was performed on a Waters Prep LC 4000 system using Vydac C4 column (12 μm, 2.2 × 25 cm) at a flow rate of 10 mL/min, with a gradient of 20−50% buffer B in buffer A over 60 min. Fractions were pooled based on LC−MS analysis.

Peptide Segment Synthesis. Peptides and peptide-$^{\alpha}$thioesters were made manually by "in situ neutralization" Boc chemistry stepwise solid-phase peptide synthesis,[21] on $-OCH_2$-Pam-resins (free $^{\alpha}$carboxyl peptides) or on $HSCH_2CH_2CO$-Leu-OCH_2-Pam-resin ($^{\alpha}$thioester peptides).[22] Side-chain protection for amino acids was as follows: Arg(Tos), Asn(Xan), Asp(OcHex), Cys(4-CH_3Bzl) or Cys(Acm), Glu(OcHex), Lys(2-Cl-Z), Ser(Bzl), Thr(Bzl), Tyr(Br-Z). For triple coupling of Thr[21], two additional coupling steps were performed with only DMF washes intervening. After chain assembly was complete, peptides were deprotected and simultaneously cleaved from the resin support by treatment with anhydrous HF containing *p*-cresol (90:10, v/v) for 1 h at 0 °C. After evaporation of the HF under reduced pressure, crude products were precipitated and triturated with chilled diethyl ether, and the peptide products were dissolved in 50% aqueous acetonitrile containing 0.1% TFA. Optimally, the peptide solution was diluted with water containing 0.1% TFA and injected into the preparative HPLC system without prior lyophilization. Peptide compositions were confirmed by LC−MS. Cram[1-V15A]$^{\alpha}$thioester (obsd mass 1785.3 ± 0.4 Da, calcd average mass 1785.9 Da), Cram[Cys(Acm)16−31]$^{\alpha}$thioester (obsd, 1925.0 ± 0.3 Da; calcd, 1925.3 Da), Cram(Cys32−46) (obsd, 1507.3 ± 0.5 Da; calcd, 1507.7 Da).

Native Chemical Ligation of Cram[Cys(Acm)^{16}-31]$^{\alpha}$thioester and Cram[Cys^{32}-46]. Native chemical ligation reactions[12] were performed in aqueous solution containing 6 M GnHCl, 100 mM phosphate, pH 7.5, 1% thiophenol (v/v), at concentration of ∼2 mg/mL for each peptide. The ligation reaction was complete in 6 h. The product was characterized by LC−MS analysis: Cram[Cys(Acm)^{16}-46] (obsd, 3213.3 ± 0.3 Da; calcd, 3213.7 Da).

Removal of the Acm-Protecting Group from Cram[Cys(Acm)^{16}-46]. a. Acm Group Removal by Hg(OAc)$_2$ and β-Mercaptoethanol. The ligated peptide Cram[Cys(Acm)^{16}-46] was purified and lyophilized. Then, 78 mg (0.024 mmol) of purified peptide Cram[Cys(Acm)^{16}-46] was dissolved in 39 mL of 20% aqueous acetonitrile containing 0.1% TFA. A 6-fold excess of $Hg(OAc)_2$ (0.146 mmol) was added to the solution. Reaction was complete in 30 min by LC−MS, and Cram[Cys^{16}-46]−mercury complex was formed. For the quenching reaction, 20% β-mercaptoethanol (v/v) was added, and after 3 h, no further reaction occurred. The solution was injected onto preparative HPLC and fractions were identified by LC−MS, and the purified peptide was lyophilized (26 mg, 33% yield).

b. Acm Group Removal by Ag(OAc)$_2$ and DTT or β-Mercaptoethanol. Acm group removal by $Ag(OAc)_2$ was performed directly in solution containing HPLC gradient buffer (no lyophilization step) after purification of the peptide from the ligation reaction of 74 mg (0.038 mmol) of Cram[Cys(Acm)^{16}-31]$^{\alpha}$thioester and 57 mg (0.0378 mmol) of Cram[Cys^{32}-46]. $Ag(OAc)_2$ (2 mmol) was added to ∼100 mL of the solution from preparative HPLC. Reaction was complete (LC−MS) in 1 h, and Cram[Cys^{16}-46]−silver complex was formed. DTT (2.4 mmol) was added to the solution for quenching, and a precipitate formed immediately. After centrifugation, the supernatant was filtered and used for preparative HPLC. Fractions containing purified peptide were identified by LC−MS, combined, and lyophilized (55 mg, 46.3% yield). Another Acm group removal reaction was carried out on a similar scale under identical conditions except that 3% β-mercaptoethanol (v/v) was used in the quenching reaction. Precipitation occurred, and the suspension was centrifuged. After filtration, the solution was used directly for preparative HPLC purification. Purified peptides were lyophilized (40 mg, 41.5% yield).

Native Chemical Ligation of Cram[1-V15A]$^{\alpha}$thioester and Cram[Cys^{16}-46]. Native chemical ligation reactions were performed in 6 M GnHCl, 100 mM sodium phosphate, pH 7.5, and 1% thiophenol (v/v) at a concentration of 2 mg/mL for each peptide. Ligation reactions were complete in 24 h. The reactions were performed at 0.013−0.017 mmol scale, and purified yields obtained varied from 64 to 73%. Cram[1−46] (obsd, 4708.6 ± 0.2 Da; calcd, 4708.4 Da).

Folding. The purified polypeptide chain Cram(1−46) was folded in 2 M GnHCl, 100 mM Tris, 8 mM cysteine, 1 mM cystine, pH 8.0, at a concentration of ∼0.3 mg/mL with exclusion of air. During the folding reaction, no stirring was performed. The reactions were performed at 0.003−0.008 mmol scales, and purified yields varied from 69 to 88%. Folded crambin was characterized by LC−MS (obsd: 4702.8 ± 0.4 Da; calcd: 4702.4 Da).

2D ^{1}H NMR. 2D-TOCSY spectra were taken on a Varian Inova 600 spectrometer by a previously reported protocol.[4] Crambin (3.5 mg) was dissolved in 500 μL of 75% d_6-acetone/20% H_2O/5% D_2O. The spectra (45-ms mixing time) were acquired with 256 complex points in the t_1 dimension and 4096 complex points in the t_2 dimension using a sweep width of 7654 Hz. Water suppression was achieved by presaturation. Zero-filling (×2) was applied in the t_1 dimension. Acquired data were processed with a 90° shifted sine-bell function.

Acknowledgment. Duhee Bang was supported by funding from the NSF MRSEC program at the University of Chicago (Grant DMR 0213745). We thank Josh Kurutz and Shohei Koide for instruction in advanced protein NMR techniques. We thank Wendy Gordon for helpful discussions on the manuscript.

JA0385078

(21) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180−193.
(22) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068−10073.