J|A|C|S
ARTICLES
Published on Web 09/20/2003

Structure of Microcin J25, a Peptide Inhibitor of Bacterial RNA Polymerase, is a Lassoed Tail

Kelly-Anne Wilson,[†] Markus Kalkum,[‡] Jennifer Ottesen,[§] Julia Yuzenkova,[||,#] Brian T. Chait,[‡] Robert Landick,[⊥] Tom Muir,[§] Konstantin Severinov,[||] and Seth A. Darst*,[†]

Contribution from the Laboratory of Molecular Biophysics, Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, and Laboratory of Synthetic Protein Chemistry, The Rockefeller University, New York, New York 10021, Department of Genetics, Waksman Institute, Piscataway, New Jersey 08854, and Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706

Received June 18, 2003; E-mail: darst@rockefeller.edu

Abstract: Microcin J25 (MccJ25) is a 21-amino acid peptide inhibitor active against the DNA-dependent RNA polymerase of Gram negative bacteria. Previously, the structure of MccJ25 was reported to be a head-to-tail circle, *cyclo*(-G^{1}GAGHVPEYF^{10}VGIGTPISFY^{20}G-). On the basis of biochemical studies, mass spectrometry, and NMR, we show that this structure is incorrect, and that the peptide has an extraordinary structural fold. MccJ25 contains an internal lactam linkage between the α-amino group of Gly1 and the γ-carboxyl of Glu8. The tail (Tyr9-Gly21) passes through the ring (Gly1-Glu8), with Phe19 and Tyr20 straddling each side of the ring, sterically trapping the tail in a noncovalent interaction we call a lassoed tail.

Introduction

DNA-dependent RNA polymerase (RNAP), the central enzyme of bacterial gene expression, is a complex molecular machine.[1,2] The catalytically competent core RNAP has subunit composition $\alpha_2\beta\beta'\omega$ and a total molecular weight near 400 kDa. In contrast to the ribosome, surprisingly few small-molecule inhibitors specific for bacterial RNAPs have been described, despite their potential as tools to investigate RNAP mechanism as well as for antibacterial drugs. Only one inhibitor, rifampicin (or its derivatives), has reached medical use,[3] where it is a key component of anti-tuberculosis therapy. Thus, recent findings that an antibacterial peptide, Microcin J25 (MccJ25), functions through inhibition of the bacterial RNAP[4,5] has generated wide interest.

MccJ25 is a 21-residue, ribosomally synthesized peptide [6,7] produced by strains of *Escherichia coli* harboring a plasmid-borne synthesis, maturation, and export system.[8] MccJ25 exhibits bacteriocidal activity against a range of Gram-negative bacterial species.[6] Genetic and biochemical results suggest that MccJ25 inhibits transcription by binding within the RNAP secondary channel,[5] a 12−15 Å diameter channel that branches off from the main active site channel, through which the nucleotide substrates can diffuse to the enzyme active site.[9,10] Binding of MccJ25 within this channel could block nucleotide substrates from entering the enzyme active site. An important step toward further understanding and manipulating the activity of MccJ25 is the characterization of its chemical and three-dimensional structure.

On the basis of biochemical and NMR studies, the primary structure of MccJ25 was reported to be a 21-residue, head-to-tail cyclic peptide, [*cyclo*(-G^{1}GAGHVPEYF^{10}VGIGTPISFY^{20}G-), or G^{1}GAGHVPEYF^{10}VGIGTPISFY^{20}G, using a numbering scheme where G^1 is the N-terminally coded amino acid of the mature peptide] and the three-dimensional solution structure was reported.[7,11] Here, we start with the key observation that synthetic peptides based on this reported structure are biologically inactive. We show, using a combination of biochemical studies, mass spectrometry, and NMR, that the reported

[†] Laboratory of Molecular Biophysics, The Rockefeller University.
[‡] Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University.
[§] Laboratory of Synthetic Protein Chemistry, The Rockefeller University.
[||] Department of Genetics, Waksman Institute.
[⊥] Department of Bacteriology, University of Wisconsin, Madison.
[#] On leave from the Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, Russia.

(1) Murakami, K.; Darst, S. A. *Curr. Opin. Struct. Biol.* **2003**, *13*, 31−39.
(2) Darst, S. A. *Curr. Opin. Struct. Biol.* **2001**, *11*, 155−162.
(3) Sensi, P. *Rev. Infect. Dis.* **1983**, *5*, Supp. 3, 402−406.
(4) Delgado, M. A.; Rintoul, M. R.; Farias, R. N.; Salomon, R. A. *J. Bacteriol.* **2001**, *183*, 4543−4550.
(5) Yuzenkova, J.; Delgado, M.; Nechaev, S.; Savalia, D.; Epshtein, V.; Artsimovitch, I.; Mooney, R. A.; Landick, R.; Farias, R. N.; Salomon, R.; Severinov, K. *J. Biol. Chem.* **2002**, *277*, 50867−50875.
(6) Salomon, R. A.; Farias, R. N. *J. Bacteriol.* **1992**, *174*, 7428−7435.
(7) Blond, A.; Peduzzi, J.; Goulard, C.; Chiuchiolo, M. J.; Barthelemy, M.; Prigent, Y.; Salomon, R. A.; Farias, R. N.; Moreno, F.; Rebuffat, S. *Eur. J. Biochem.* **1999**, *259*, 747−755.
(8) Solbiati, J. O.; Ciaccio, M.; Farias, R. N.; Gonzalez-Pastor, J. E.; Moreno, F.; Salomon, R. A. *J. Bacteriol.* **1999**, *181*, 2659−2662.
(9) Zhang, G.; Campbell, E. A.; Minakhin, L.; Richter, C.; Severinov, K.; Darst, S. A. *Cell* **1999**, *98*, 811−824.
(10) Korzheva, N.; Mustaev, A.; Kozlov, M.; Malhotra, A.; Nikiforov, V.; Goldfarb, A.; Darst, S. A. *Science* **2000**, *289*, 619−625.
(11) Blond, A.; Cheminant, M.; Segalas-Milazzo, I.; Peduzzi, J.; Barthelemy, M.; Goulard, C.; Salomon, R.; Moreno, F.; Farias, R.; Rebuffat, S. *Eur. J. Biochem.* **2001**, *268*, 2124−2133.

cyclic structure is incorrect, and that the peptide contains an extraordinary structural fold. The structure contains a lactam linkage between the α-amino group of Gly1 and the γ-carboxyl of Glu8, forming an 8-residue ring (Gly1-Glu8) that we term the lariat ring. The "tail" (Tyr9-Gly21) passes through the lariat ring, with Phe19 and Tyr20 straddling each side of the lariat ring, sterically trapping the tail within the ring in a noncovalent interaction that is resistant to strongly denaturing conditions. We call this structure a lassoed tail.

Experimental Section

General Materials and Methods. All amino acid derivatives and resins were purchased from Novabiochem (San Diego, CA) except Boc-L-Glu(OAllyl)-OH and Boc-Gly-PAM resin, which were purchased from Neosystems (Strasbourg, France). All other chemicals were purchased from Sigma-Aldrich Chemical Co (St. Louis, MO). Thermolysin was purchased from Calbiochem (Darmstadt, Germany). Analytical gradient HPLC was performed on a Hewlett-Packard 1100 series instrument with detection at 214 and 280 nm. Analytical HPLC was performed on a Vydac C18 column (5 micron, 4.6 × 150 mm) at a flow rate of 1 mL/min. Preparative HPLC was routinely performed on a Waters DeltaPrep 4000 C18 column (15−20 micron, 50 × 250 mm) at a flow rate of 5 mL/min. Two solvent systems were used on linear gradients; (1) 0.1% aqueous TFA (solvent A) vs 90% Acetonitrile plus 0.1% TFA (solvent B); (2) 0.05% aqueous TFA (solvent C) vs methanol plus 0.05% TFA (solvent D).

Bacterial Culture and Microcin Purification. MccJ25 was purified from minimal M9 media plus 1% glucose and 1 μg/mL thiamine as previously described [7] except that the *E. coli* strain, SBG231 (AB259 *sjmA1*; spontaneous MccJ25 resistant mutant, *β′*T931I; [4]) was transformed with the pTUC200 plasmid; a pB*R*322 backbone with a 13 Kb insert encompassing *McJ* genes *A*, *B*, *C*, and *D*.

Activity Assays. For the in vitro abortive initiation transcription assay, *E. coli* RNAP holoenzyme was combined with the T7 A1 promoter-containing DNA fragment, CpA primer, and [α^{32}P]UTP in the presence and absence of the natural and synthetic microcin variants at indicated concentrations.[5] Reactions were incubated at 37 °C for 10 min, and the products were resolved by denaturing PAGE and visualized by autoradiography.

For the in vivo assay, an agar overlay assay test was used.[5] Double dilutions of MccJ25 variants (10μL) were applied to a solid LB plate, allowed to dry, and overlaid with 2 mL top agar (0.7%) containing a 200 μL stationary phase innoculate of sensitive (AB259) or resistant (SBG231) *E. coli*. The plates were incubated overnight at 37 °C and examined for zones of inhibition.

Peptide Synthesis. General. All peptides were synthesized according to the in-situ neutralization/HBTU activation protocol for Boc SPPS.[12] Following chain assembly, global de-protection and cleavage from the support was achieved by treatment with HF containing 4% v/v *p*-cresol, for 1 h at 0 °C. Following removal of the HF, the crude peptide products were precipitated and washed with anhydrous cold Et_2O before being dissolved in aqueous acetonitrile (50% B) and lyophilized.

Peptide 1: Linear Microcin = H−GGAGHVPEYFV-GIGTPISFYG−OH. The peptide was synthesized (0.5 mmol scale) on a PAM resin. Following chain assembly and global deprotection/cleavage, the crude material was purified by RP-HPLC (50−60% D over 60 min.) to give the desired material, peptide **1**, which was characterized by MS; found *m*/*z* 2125.71 ± 0.05, expected 2125.04, and MS^2; fragmented with appropriate b and y ions series for this sequence.

Peptide 2: Cyclic Cysteine Microcin = GGAGHVPEY-FVGIGTPICFYG. The synthesis of circular peptide **2** by native chemical ligation[13] required the preparation of a linear peptide α-thioester precursor corresponding to the permuted and mutated microcin sequence; *C*FYGGGAGHVPEYFVGIGTPI, where the italicized residue indicates a Ser-Cys mutation. This sequence was synthesized (0.25 mmol scale) on a 3-mercaptopropioamide derivatized MBHA resin.[13] Following chain assembly and global deprotection/cleavage, the 3-mercaptopropionamide-(DNP-His)-peptide (found *m*/*z* 2393.16 ± 0.74; expected 2394) was cyclized by dissolving the crude lyophilized material (50 mg, 20.9 μmole) in ligation buffer (6 M guanidine-HCl, 100 mM Tris, pH 8.5, 1 mM EDTA, 0.1 M mercaptoethanesulfoninc acid) to a final concentration of 400 μM. The reaction was deemed complete after stirring at room temperature for 16 h, at which point the reaction mixture was purified directly by preparative RP-HPLC (27−43% B over 60 min.). This yielded the desired cyclic, fully deprotected peptide (8 mg, 3.8 μmole) with 95% purity and 16% yield (from the 3-mercaptopropioamide-[DNP-His]-peptide), which was characterized by MS; found *m*/*z* 2123.03 ± 0.03, expected 2123.01.

Peptide 4: Lariat Microcin = GGAGHVPEYFVGIGT-PISFYG. The peptide was synthesized (0.5 mmol scale) on a PAM resin. To allow on-resin cyclization between the α-amino and Glu8 γ-carboxyl groups, the side-chain of Glu8 was orthogonally protected as an allyl ester. Following chain assembly and N^{α} deprotection, the allyl group was removed by treatment with $Pd(PPh_3)_4$. The desired lactam linkage was then formed on the solid phase using PyBOP (5 eq.) as the activating agent and either DMF or DMSO as the solvent. Cyclization was monitored using the quantitiative ninhydrin test, which indicated that the coupling efficiency reached a maximum of ∼50% after three 15 min coupling procedures. Following global deprotection/cleavage, the crude material was purified by RP-HPLC (50−60% D over 60 min.) to give the desired material, peptide **4**, (4.5 mg, 2.1 μmole; 2% yield, based on original loading of the resin), which was characterized by MS; found *m*/*z* 2107.07 ± 0.03, expected 2107.03, and MS^2; fragmented with appropriate b and y ions series for this sequence (Figure 2C). The purified material contained a small amount of a peptidic contaminant with *m*/*z* 2244.

Thermolysin Cleavage of MccJ25. Extensive Digestion. MccJ25 (10 ng, 4.7 pmole) in 5 μL water, was digested with 1000 U thermolysin (suspended in 10 mM Tris-HCl pH 8.0) for 2 h at room temperature. An aliquot of this digestion mixture (1 μL) was diluted 10-fold with aqueous 0.1% TFA, absorbed to a ZipTip microcolumn (Millipore), washed with 20 μL of 0.1% aqueous TFA, and eluted with 1 μL half-saturated 2,5-

(12) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. *Int. J. Pept. Protein Res.* **1992**, *40*, 180−193.

(13) Camaerero, J. A.; Cotton, G. J.; Adeva, A.; Muir, T. W. *J. Peptide Res.* **1998**, *51*, 303−316.

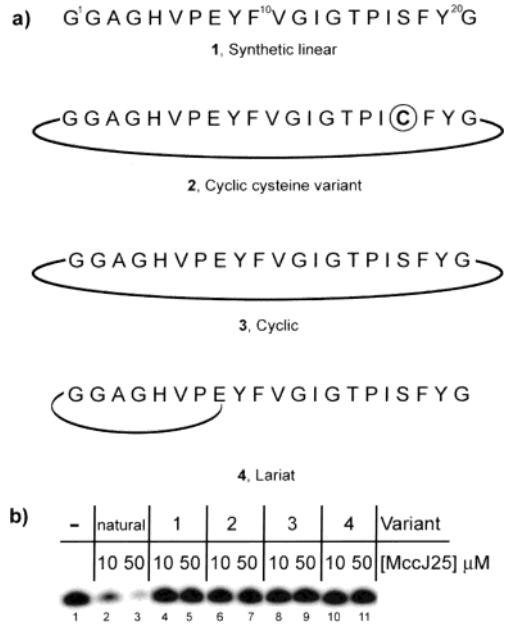

Figure 1. Structure and activity of four microcin variants. (a) Primary structures of four microcin variants, the synthetic linear peptide (1) is numbered according to the encoded sequence of the mature MccJ25. (b) MccJ25 inhibits E. coli RNAP in an in vitro abortive initiation transcription assay, rendering the polymerase unable to extend a CpA primer with [α^{32}P]-UTP (compare lanes 2 and 3 with lane 1). The four MccJ25 variants have no effect on RNAP activity (lanes 4−11).

dihydroxy benzoic acid solution (see below) directly onto a MALDI target for analysis.

Minimal Digestion. MccJ25 (7.4 μg, 3.5 nmole) in 20 μL 18.75 mM ammonium bicarbonate, pH 8.0, was digested with 1 U thermolysin (suspended in 10 mM Tris-HCl pH 8.0) for 1 h at room temperature. The digestion was stopped by addition of 25 μL 100 mM EDTA. The digestion mixture was diluted with methanol to a final concentration of 45% methanol, centrifuged at 12 000 rpm for 2 min, and the supernatant warmed at 45 °C for 5 min. The singly cleaved peptide was purified by analytical RP-HPLC on a 30 min, 45−60% methanol gradient at 45 °C to give the desired material at ~83% purity, which was characterized by MS; found *m*/*z* 2125.01 ± 0.03, expected 2125.04, and MS^2; found appropriate y and b series for a single cleavage between between Phe10 and Val 11. This peptide was sequenced by Edman degradation at the Protein Chemistry Laboratory of The University of Texas Medical Branch, Galveston, TX.

Treatment of Digested MccJ25 with Denaturing Agents. The remaining extensive digestion mix (from above) was diluted 10-fold with 0.1% aqueous TFA, and 4 μL aliquots were incubated with each of the following treatments at both room temperature and 95 °C for 1 h; 1 M NaOH, 5 M guanidine-HCl, 50% TFA, or 1% SDS (all final concentrations). Each sample was then diluted 5-fold with 0.1% aqueous TFA, except the 50% TFA condition, which was diluted with H_2O. A portion of each sample (1 μL) was then analyzed by MALDI−MS.

Mass Spectrometry. Aliquots (1 μL) of peptide solution (0.5−0.1 pmol/μL) were applied onto the MALDI compact disk-sample plate [14] and mixed with equal volumes of half-saturated 2,5-dihydroxy benzoic acid (DHB) in 50% methanol/20% acetonitrile/0.1% TFA and allowed to dry. Accurate masses of the peptides **1**-**4**, and of proteolytic cleavage products, were measured using a prototype MALDI-QqTOF instrument.[15] The MS^2 and MS^3 experiments were performed as desribed [16] using a prototype MALDI-quadrupole ion trap.[14] Ion injection durations were 200−400 ms. Resonant excitation was performed using empirically optimized instrument specific parameters (isolation width: 2.5 Da, normalized collision energy: 35−38%, activation Q: 0.240, activation time: 300 ms). Spectral acquisition times varied from 30 s to 1 min in order to achieve acceptable signal/noise ratios.

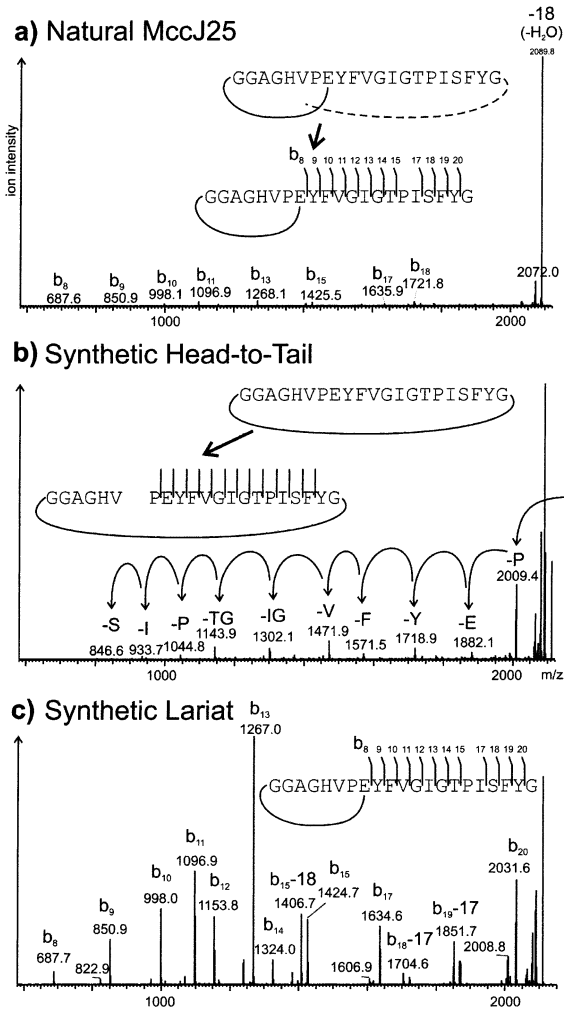

Figure 2. Comparison of MALDI-ion trap MS^2 spectra obtained from (a) natural MccJ25 and the synthetic microcin-isomers (b) **3** and (c) **4**. In each case, the parent ion at *m*/*z* 2107.0 was selected and fragmented using identical instrument settings. The arrows in spectrum (b) indicate sequential residue losses.

NMR Spectroscopy. MccJ25 (4 mg, 1.9 μmole) and synthetic lariat (4 mg, 1.9 μmole) were dissolved in 500 μL *d*6-DMSO because the peptides were sparingly soluble in aqueous buffers, and spectra were acquired on a Bruker DMX500 spectrometer. Spectra were processed using NMRPipe[17] and analyzed using NMRView.[18] The synthetic linear peptide (2.2 mg, 1.0 μmole)

(14) Krutchinsky, A. N.; Kalkum, M.; Chait, B. T. *Anal. Chem.* **2001**, *73*, 5066−5077.
(15) Krutchinsky, A. N.; Zhang, W.; Chait, B. T. *J. Am. Soc. Mass Spect.* **2000**, *11*, 493−504.
(16) Kalkum, M.; Lyon, G. J.; Chait, B. T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2795−2800.
(17) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(18) Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603−614.

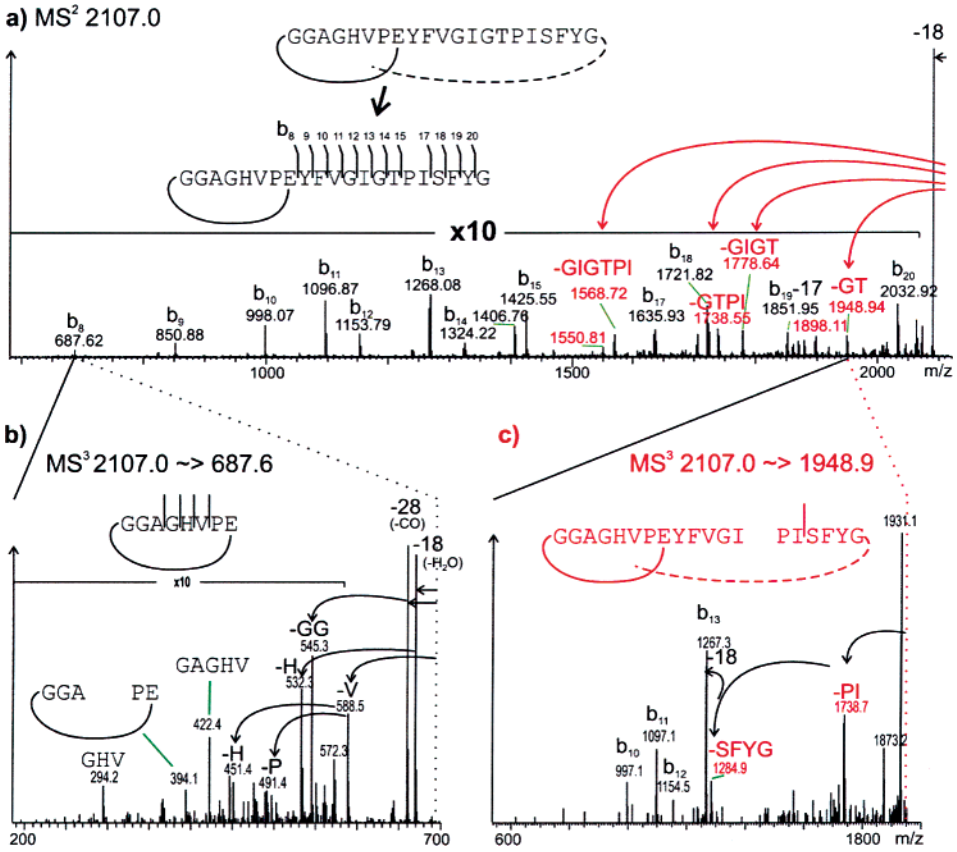

Figure 3. (a) MS^2 spectrum of MccJ25 from Figure 2a with peak intensities amplified by a factor of 10. Signals that do not belong to the b-ion series are labeled in red. (b) MS^3 spectrum of the b8-fragment, characterized by nonsequential residue losses that provide evidence for the existence of the lariat-ring. (c) MS^3 spectrum of a nonb-ion fragment that originates from the loss of G14T15 from the MccJ25-parent ion, indicating the existence of the noncovalent connection between the N-to-C-terminal parts of the molecule.

was dissolved in 500 μL *d*3-CD_3OD, and spectra were acquired on a Bruker DPX400 spectrometer, and processed using XwinNMR. All spectra were acquired at 298K.

Sequence-specific backbone and side chain assignments for MccJ25 were derived from TOCSY and NOESY (250 ms, 175 ms, 100 ms mixing time) data. A total of 274 NOESY peaks were assigned. Cross-peak intensities were measured at a NOESY mixing time of 175 ms, and NOE-derived constraints were generated and sorted into strong (1.8−2.8 Å), medium (1.8−3.4 Å), or weak (1.8−5.0 Å) bins. Protons were not assigned stereospecifically; spectrally differentiable protons were collapsed to pseudoatoms, yielding 192 constraints (see Supplemental Data), and distance corrections were handled internally within the software suite CNS.[19]

Structure Determination. A total of 192 distance constraints [of which 79 were intraresidue, 59 sequential, 9 medium range ($|i - j| < 4$ amino acids) and 45 long range ($|i - j| \geq 4$ amino acids)] were utilized in distance geometry simulated annealing calculations[20] using the standard CNS protocol, with a modified protein parameter file to allow for the isopeptide bond between Gly1 and Glu8. A total of 75 embedded substructures were generated from the lariat peptide in an extended starting conformation. All embedded starting structures, irrespective of energies, were used to generate 75 final structures by NOE-restrained molecular dynamics simulated annealing, followed by an additional minimization step with reduced dependence on NOE constraints. The 20 lowest energy structures were selected for analysis. No NMR constraint violations > 0.5 Å were observed in these 20 structures. A Ramachandran plot revealed that 46.2% of the dihedral angles fell in the most favored region, 46.1% fell in additionally allowed regions, and the remaining 7.7% fell in the generously allowed region.

Results and Discussion

The Structure of MccJ25 Requires Reassessment. The biological activity of totally synthetic versions of MccJ25 was assessed to evaluate methods for the preparation of large amounts of highly pure, biologically active MccJ25 required for structural studies of the RNAP/MccJ25 complex (Figure 1a). An intramolecular version of native chemical ligation was used to synthesize peptide **2** (Figure 1a), where Ser18 of the coded MccJ25 sequence (**1**, Figure 1a) was mutated to provide the required N-terminal Cys residue.[21] In addition, cyclic peptide **3** (Figure 1a), with the coded MccJ25 sequence obtained by a native chemical ligation scheme followed by desulfurization,[22] was a generous gift of P. E. Dawson.

Two tests of biological activity were used, in vivo (killing of sensitive bacterial cells) and in vitro (inhibition of transcription activity by purified *E. coli* RNAP).[5] Each of the synthetic peptides **1**, **2**, and **3**, were totally inactive in vivo (not shown) and in vitro (Figure 1b). For instance, 10 μM MccJ25 (purified

(19) Adams, P. D.; Pannu, N. S.; Read, R. J.; Brunger, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5018−5023.
(20) Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1988**, *229*, 317−324.
(21) Camarero, J. A.; Muir, T. W. *Chem. Commun.* **1997**, *15*, 1369−1370.
(22) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526−533.

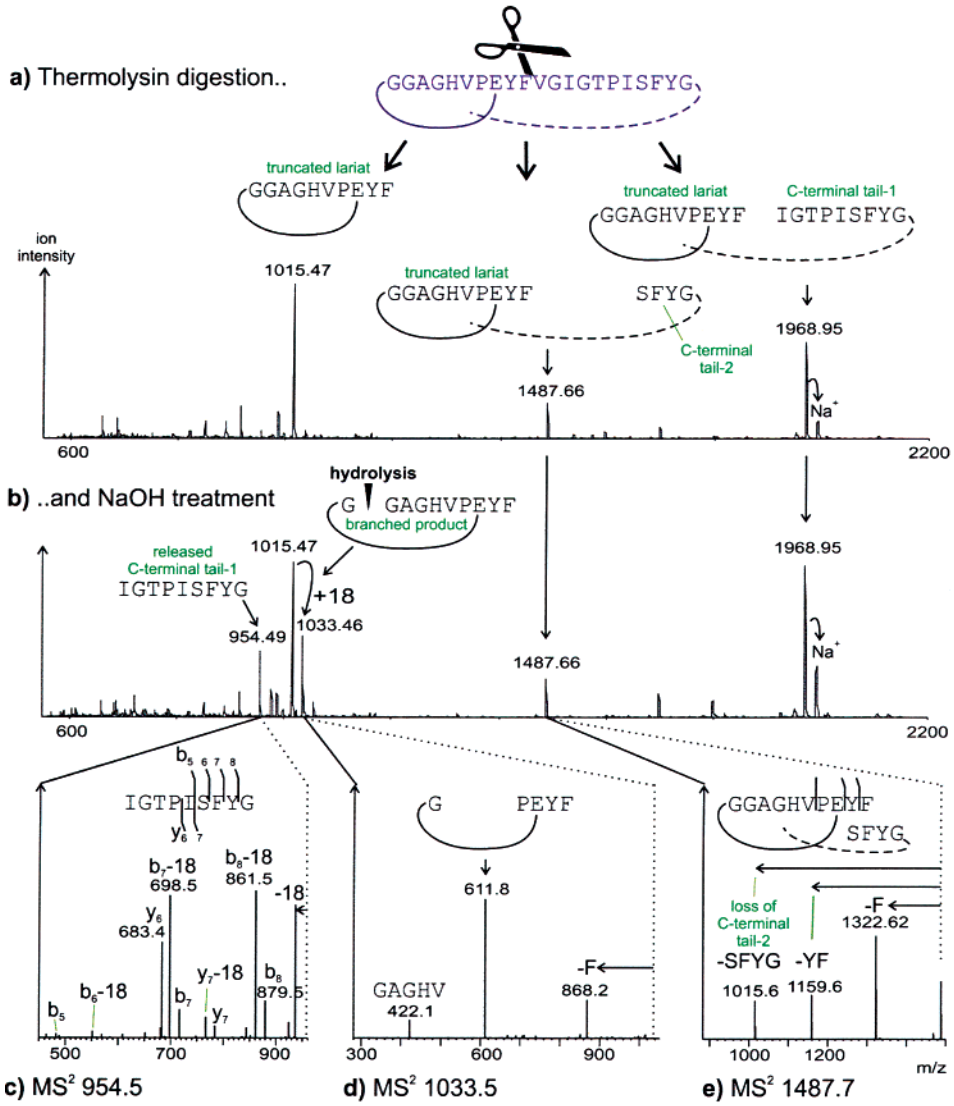

Figure 4. MALDI-QqTOF single-stage mass spectra of thermolysin digested MccJ25 − (a) before and (b) after NaOH treatment. The ion trap MS^2 spectra below provide evidence for the release of the C-terminal tail (shown in (c)) due to hydrolysis of the circular portion (as revealed by spectrum (d)). The MS^2 analysis of the smallest proteolytic product that bears a portion of the trapped C-terminus is shown in (e).

from MccJ25 producing *E. coli*) [7] reduced transcription to 43% of the level in the absence of MccJ25 (Figure 1b, lanes 1 and 2), whereas 50 μM MccJ25 nearly completely inhibited transcription (19%, lane 3). In contrast, concentrations of **2** and **3** up to 400 μM (the highest concentration tested) showed no inhibition.

The behavior of MccJ25 and **3** was also compared by analytical reversed phase HPLC. The retention time of MccJ25 was 13.8 min on a 45 min, 50−65% methanol gradient using a C18 column, whereas the retention time of **3** was 21.0 min.

Finally, MccJ25 and **3** were compared by mass spectrometric analysis (Figure 2). The *m*/*z* of MccJ25 and **3** were both measured to be 2107.0, which is 18 Da lower than that of the linear form (peptide **1**). The MALDI-ion trap mass spectrometric fragmentation spectrum (MS^2 spectrum) of MccJ25 was characterized by an intense peak arising through the loss of water from the parent ion and a series of low intensity b-ions (b_8 to b_{20}; Figure 2a). This spectra was compared to the MS^2 spectrum of **3** (Figure 2b), which displayed a low intensity series of fragment ions characteristic of a cyclic peptide,[16] with fragment masses distinct from those of MccJ25.

In summary, peptide **3** did not behave like MccJ25 in three significant tests. First, **3** was inactive in in vivo and in vitro assays of biological activity at concentrations up to 400 μM, whereas MccJ25 showed significant transcription inhibition at 10 μM (Figure 1b). Second, **3** and MccJ25 had significantly different retention times by reversed phase HPLC. And third, the MS^2 spectra of **3** and MccJ25 were very different (Figures 2a, b). Together, these observations conclusively establish that the structure of MccJ25 is not **3**.

Mass Spectrometric Analysis of MccJ25. Further fragmentation of the smallest observable ion of the MccJ25 MS^2 series, the b_8-ion (Figure 2a, magnified in Figure 3a), led to an MS^3 spectrum characteristic of the unbranched, 8-residue cyclic moiety shown in Figure 3b.[16] The data show that Gly1 is condensed with either Pro7 or Glu8. By far the simplest way of achieving this connectivity involves an isopeptide bond between the α-amine of Gly1 and the side chain of Glu8. Moreover, this conclusion is supported in all subsequent MS and NMR experiments (see later). We therefore synthesized a peptide that incorporated this putative Gly1 to Glu8 amide linkage ("lariat", peptide **4** of Figure 1a). However, the lariat

peptide **4** showed no in vivo (not shown) or in vitro activity (Figure 1b). MS^2 fragmentation of **4** was compared to that of MccJ25 taken under identical instrumental settings (Figure 2a versus 2c). The same b-ion series was observed for both peptides, but, surprisingly, the intensity of the lariat-derived series was approximately 2 orders of magnitude higher (when normalized to the intense water loss peak) than that obtained from MccJ25. Comparison of the spectra in Figure 2, all of which were derived from peptides with identical molecular masses, demonstrated that the synthetic peptides, **3** and **4**, were structurally distinct from MccJ25.

More detailed comparison between the MS^2 spectra of natural MccJ25 and **4** revealed several peaks unique to MccJ25 (colored red in Figure 3a). For example, the unique peak at *m*/*z* 1948.9 appeared to arise from the loss of $G^{14}T^{15}$ from the parent ion at *m*/*z* 2107.0. This loss can only be explained if residues 16−21 were connected in some way to the N-terminal part of the peptide. We indicate this unknown N- to C-terminal "connection" by the dashed line in Figures 2, 3, and 4. Further fragmentation of the MS^2 fragment at *m*/*z* 1948.9 (MS^3 experiment) confirmed this connectivity through observation of the sequential losses of $P^{16}I^{17}$ and $S^{18}FYG^{21}$ (Figure 3c). These data demonstrate that the C-terminal part of MccJ25 (beyond residue 17) is attached to the N-terminal part (before residue 12) firmly enough to survive both the MALDI process and the subsequent collision induced dissociation. The observation of other unique MS^2 fragments at *m*/*z* 1778.6 (loss of $G^{12}IGT^{15}$), *m*/*z* 1738.6 (loss of $G^{14}TPI^{17}$), and *m*/*z* 1568.7 (loss of G^{12}-$IGTPI^{17}$) further supports the existence of the putative connection between the C- and N-terminal parts. This connection is mysterious, as it results in no apparent change in mass.

Further insight into structural differences between MccJ25 and **4** was obtained by mass spectrometric analysis of proteolytic fragments derived from thermolysin digestion. Treatment of MccJ25 with low quantities of thermolysin resulted in cleavage of the peptide bond between Phe10 and Val11, yielding a single chemical species with *m*/*z* 2125.0, as reported previously (in the context of the erroneous structure **3**).[7,23] This result is inconsistent with the primary structure of **4**, because cleavage between Phe10/Val11 would yield two peptide fragments. Experiments with larger amounts of thermolysin produced three major degradation products at *m*/*z* 1968.9, 1487.7, and 1015.5 (Figure 4a). The ion at *m*/*z* 1015.5 corresponds to the truncated lariat moiety containing the first eight amino acids and the short Y^9F^{10}-tail, $\overline{G^1GAGHVPEYF}^{10}$. The ion at *m*/*z* 1487.7 corresponds to the protonated truncated lariat moiety, $\overline{G^1GAGHVPEYF}$,[10] plus the peptide $S^{18}FYG$.[21] The latter 4-residue peptide contains a free amino- and carboxyl-terminus, as verified by subjecting the peptide at *m*/*z* 1487.7 to mass spectrometric fragmentation (Figure 4e). Likewise, the ion at *m*/*z* 1968.9 arose from protonated $\overline{G^1GAGHVPEYF}^{10}$ + I^{13}-$GTPISFYG$.[21] Thus, we find that the unique connection between the C- and N-terminal portions of MccJ25 is at least partially stable under the conditions of enzymatic digestion. The nature of this connection was further investigated by treatment of the thermolysin digested peptide with denaturing agents. Treatment

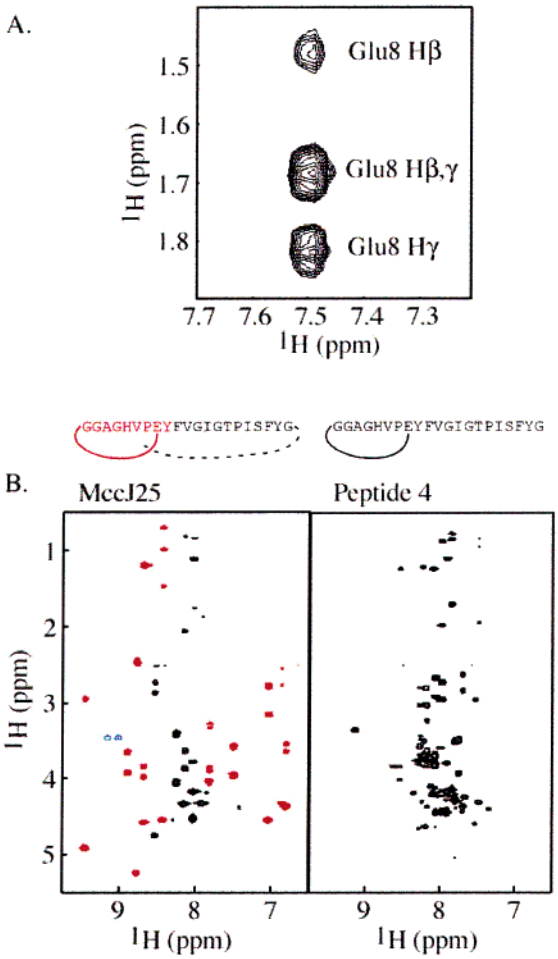

Figure 5. (a) NOE cross-peaks between Gly1 HN and Glu8 side chain indicating the isopeptide bond forming the lariat ring. (b) Amide regions of the TOCSY spectra of MccJ25 (left) and synthetic lariat peptide **4** (right). For MccJ25, peaks corresponding to residues 1−9 and 18−21 are in red; residues 10−17 are in black, and the residual solvent ridge is in blue.

with 1% SDS, 6 M guanidine-HCl, or 0.1% TFA had no effect on the connection, even after 1 h incubations at 95 °C, implying the connection was stable to these denaturing treatments.

Additional information on the chemical nature of the mysterious N- to C-terminal connection was obtained by subjecting the thermolysin generated MccJ25-digest to partial hydrolysis with NaOH (Figure 4b). The peptide bond between Gly1 and Gly2, within the lariat ring, was labile to base treatment, leaving a branched product with a mass 18 Da higher than that of the truncated lariat (Figure 4b, peak doublet at *m*/*z* 1015.47 and 1033.46). The identity of this branched product was verified by MS^2 fragmentation of the *m*/*z* 1033.46 ion (Figure 4d). In addition a peak at *m*/*z* 954.49 appeared, which we identified as the C-terminal tail, $I^{13}GTPISFYG^{21}$ (Figure 4c). These observations are consistent with hydrolysis of the ring, allowing release of the C-terminal tail. This hypothesis was further supported by the lack of peak doublets for the products at *m*/*z* 1968.95 and 1487.66 (Figure 4b), indicating that these structures retained the intact ring and associated C-terminal tail. Finally, under MS^2 conditions, the labile bond between Val6 and Pro7 dissociated (verified by MS^3 experiments), leading to release of $S^{18}FYG^{21}$ (Figure 4e).

NMR Analysis of MccJ25. The structure of MccJ25 was further analyzed using ^1H NMR experiments. No sequential

(23) Blond, A.; Cheminant, M.; Destoumieux-Garzon, D.; Segalas-Milazzo, I.; Peduzzi, J.; Goulard, C.; Rebuffat, S. *Eur. J. Biochem.* **2002**, *269*, 6212−6222.

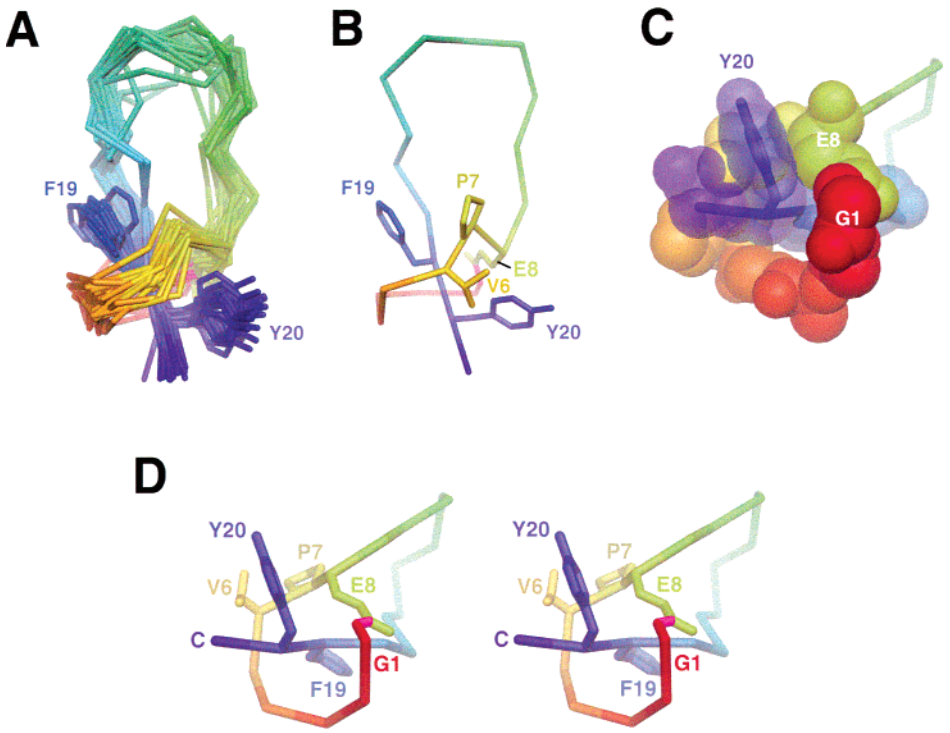

Figure 6. NMR-based structural model of MccJ25. In each view, the peptide chain is colored in a ramp (red-orange-yellow-green-blue) from the N-terminal Gly1 (red) to the C-terminal Gly21 (blue), with the isopeptide bond linking the α-amino group of Gly1 with the γ-carboxyl of Glu8 colored magenta. (a) Ensemble of 20 models resulting from distance geometry simulated annealing calculations. The side chains of E8, F19, and Y20 are shown and labeled. (b) Representative structure, shown in the same view as part a. The side-chains of V6, P7, E8, F19, and Y20 are shown. (c) The lariat ring (G1-E8) and the C-terminal tail (S18-G21) are shown as CPK atoms, with the C-terminal tail rendered transparent so the backbone is visible. Selected side-chains are labeled. (d) Stereoview with selected side-chains shown and labeled.

NOE cross-peaks were observed between Gly21 and Gly1, either in our experiments or in previously reported experiments,[7] an observation inconsistent with structure **3**. However, strong NOEs between Gly1 NH and the side chain protons of Glu8 were observed in NOESY spectra of MccJ25 (Figure 5a), consistent with the formation of the lactam linkage, as shown above by MS analysis (Figure 3b).

The ^{1}H NMR spectra of MccJ25 displayed significant chemical shift dispersion, typical of a highly structured peptide (Figure 5b). Although the chemical shifts of the NH protons ranged from 6.8 to 9.4 ppm, the NH chemical shifts of residues 10−17 were all clustered between 7.9 and 8.5; the large chemical shift dispersion was due primarily to residues 1−9 and 18−21. This disparity is especially interesting in light of the MS^2 analysis of the thermolysin cleavage experiments, which indicated an association between the lariat residues and the C-terminus of the peptide (Figure 4). Although peptide **4** also contains the lactam linkage between Gly1 and Glu8, the NH chemical shifts ranged only from 7.3 to 8.6 ppm, typical of a random coil structure (Figure 5b). Similarly, the spectra of linear peptide **1** showed NH chemical shifts ranging from 7.7 to 8.5 ppm (data not shown).

The wide dispersion of chemical shifts allowed for the unambiguous assignment of 274 peaks in the NOESY spectrum of MccJ25. Several interesting long-range interactions were observed between the C-terminal tail of the peptide and the N-terminal lariat ring. In particular, Phe19 and Tyr20 define two independent clusters of long-range NOEs. For example, NOE connectivities were observed between the ring protons of Tyr20 and the side chains of Val6 and Glu8 but not Pro7, and between the Tyr20 NH and backbone protons of "lariat" residues 5−8. On the other hand, the ring protons of Phe19 interacted solely with the side chains of "tail" residues 16−18 and "lariat" residue Pro7 (but not Val6 or Glu8), and the β protons of Phe19 interacted with backbone protons of 'lariat' residues 3−7. NOE cross-peaks between the side-chains of Phe19 and Tyr20 were conspicuously absent.

A model structure consistent with the NMR distance information was calculated. Specifically, 192 NOE-derived distance constraints were utilized in hybrid distance geometry simulated annealing, using the covalent structure of **4** as the starting point. An ensemble of 20 structures was generated that converged to a common fold with RMSD for backbone atoms of 1.73 ± 0.54 Å, and for heavy atoms of 2.38 ± 0.55 Å (Figure 6a). Residues 18−20 and 1−8 represented an area of well-defined structure, with backbone (including Glu8 side chain) RMSD of 0.91 ± 0.24 Å and heavy atom RMSD of 1.36 ± 0.27 Å.

The calculated structure closest to the mean of the ensemble was selected as a representative structure (Figure 6, parts b, c, and d). The clusters of interactions involving Phe19 and Tyr20 define an unusual structure, in which the tail residues 18−21 are threaded through the lariat ring formed by the backbone of residues 1−7 plus the side chain of Glu8. The bulky side chains of Phe19 and Tyr20 straddle the lariat, protruding on either face of the loop.

Two distinct topologies are possible for the tightened noose structure, depending on which direction the C-terminal tail passes through the lariat ring. In the low-energy structures, only one topology was observed (Figure 6). This topology is consistent with the NOE clusters described above, where interactions between tail residues Phe19 and Tyr20 with lariat

residues Val6 and Pro7, respectively, define opposite faces of the lariat ring (Figure 6b).

Conclusions

Here, we establish that the previously proposed structure for MccJ25 [7,11] is incorrect. Synthetic peptides based on this structure (**2**, with a single Ser-Cys mutation at position 18, and **3**, identical to the proposed structure) differ from MccJ25 in three important ways. First, they are biologically inactive (Figure 1b). Second, they behave differently by reversed phase HPLC. Third, they behave differently upon mass spectrometric fragmentation analysis (Figure 2).

We show that MccJ25 contains a lactam linkage between the α-amino group of Gly1 and the γ-carboxyl of Glu8, forming an 8-residue ring (Gly1-Glu8, the lariat ring). Our data suggests that the C-terminal "tail" of MccJ25 (Tyr9-Gly21) is trapped within the lariat ring. Release of the C-terminus is prevented by steric hindrance and requires the cleavage of at least one peptide bond, which must occur either within the ring or, based on the NMR data, between residues Phe19 and Tyr20. The tail-free lariat observed at *m*/*z* 1015.47 (Figure 4, parts a and b) may have arisen from thermolysin digestion of the bond between Phe19 and Tyr20, but additional experiments are required to establish this hypothesis.

Our MccJ25 model (Figure 6) is consistent with all of our biochemical and mass spectrometric fragmentation results, and is strongly supported by the NMR analysis. In retrospect, the previously published experimental support for **3** as the structure of MccJ25 was not conclusive.[7] We re-evaulate the previous interpretations in light of our new results in the following four points:

(1) The N- and C-termini of MccJ25 were blocked to Edman and carboxypeptidase Y sequencing, respectively.[6,7] In fact, the blocked N-terminus is due to the lactam linkage with the γ-carboxyl of Glu8. The inability to sequence from the C-terminus by carboxypeptidase Y is not evidence that the C-terminus is engaged in a chemical linkage, but can also be explained by steric considerations since the large protease must be able to bind the C-terminus of its substrate. In our model of the MccJ25 structure, the two C-terminal residues of the peptide are not freely accessible, being engaged in noncovalent interactions with the lariat ring (Figure 6).

(2) Partial acid hydrolysis of MccJ25 yielded peptide fragments, two of which were analyzed by N-terminal sequencing (fragment H-9^{II}, corresponding to G^1GAGHVPEYF^{10}VG, and H-10^{II}, corresponding to Y^9FVGIG) and these results were cited in support of **3**.[7] Because these peptide fragments did not contain a peptide bond linking Gly21 with Gly1 (Figure 1a), they neither support nor contradict **3**. However, the observed fragments are consistent with our model with the assumption that the harsh acid hydrolysis conditions cleaved the lariat -ring lactam linkage.

(3) Thermolysin digestion, which cleaved between Phe10 and Val11, yielded a single peptide,[7] as expected from **3**. However, a more thorough analysis of the results of thermolysin digestion (Figure 4) supports our model.

(4) The most compelling evidence that the structure of MccJ25 corresponded to **3** was the previously published N-terminal sequence analysis, by Edman degradation, of thermolysin cleaved MccJ25.[7] The proteolytic cleavage of the peptide bond between Phe10 and Val11 was interpreted to generate a linear peptide with a free N-terminus, allowing N-terminal sequencing to begin at Val11. Blond et al. (1999) [7] report the following sequencing results: V^{11}GIGTPISFY^{20}G^{21}G^1-GAG.[4] Sequencing through the C-terminal Gly21 to Gly1 and on to Gly4 supports **3** and contradicts our model. For this reason, we analyzed our own sample of thermolysin-cleaved MccJ25 by Edman degradation. We obtained the sequence from Val11 through Gly21, followed by several cycles indicating Gly, which is expected behavior for the C-terminal residue of a peptide (J. S. Smith, personal communication). Importantly, in cycles after Gly21, none of the amino acids expected in the N-terminal region of MccJ25 (other than Gly) were observed. For example, Ala or His (expected at positions 3 and 5 of the coded MccJ25 sequence, Figure 1a) were not observed. Thus, we believe our Edman degradation results are consistent with our model, because the identification of the C-terminal residue by this technique is imprecise.

Although the previously published NMR structure for MccJ25[11,23] does not have the correct covalent linkages, the previous structure and our structure (Figure 6) are quite similar when considering the well-defined portion of our model. The α-carbon backbone of residues 4−8 and 19−20 of our model can be superimposed onto the corresponding atoms of the previous structure with an RMSD of 1.13 Å.

In principle, the final test of the MccJ25 structure should come from its total chemical synthesis and comparison of its biological activity with MccJ25. Due to the nature of the structure, with its sterically trapped tail, total chemical synthesis has so far proven unsuccessful (data not shown). In this case, the pathway to the final structure must be critical, because the MccJ25 precursor must fold into a pre-defined conformation prior to the lariat ring cyclization to allow steric trapping of the C-terminal tail (Figure 6). Pre-folding of the MccJ25 structure may be a function of the McJA pro-protein, or may be assisted by one of the processing enzymes.

Synthesis of active MccJ25 in vivo requires the action of two processing enzymes, McJB and McJC,[8,24] which act on the McJA pro-protein to produce the mature MccJ25 peptide. Little is currently known about the structure or function of these enzymes, or the MccJ25 processing pathway. Interestingly, McJC shows sequence and predicted structural homology[25] with a group of related ATP/Mg^{2+}-dependent enzymes including class B asparagine synthetases (that convert aspartic acid to asparagine;[26,27]) and β-lactam synthetase (catalyzes the formation of the β-lactam ring in clavulanic acid),[28,29] suggesting McJC is an ATP/Mg^{2+}-dependent enyzme that catalyzes the formation of the lactam ring found in MccJ25. McJB shows no apparent homology with any known sequences, but one possibility is that it prepares the fold of the MccJ25 pro-protein (McJA) to allow McJC to catalyze formation of the lariat ring around the C-terminal tail appropriately, and/or it may cleave the final mature peptide from the McJA pro-protein.

(24) Solbiati, J. O.; Ciaccio, M.; Farias, R. N.; Salomon, R. A. *J. Bacteriol.* **1996**, *178*, 3661−3663.
(25) Rost, B.; Sander, C. *Proteins* **1994**, *19*, 55−72.
(26) Richards, N. G.; Schuster, S. M. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1998**, *72*, 145−198.
(27) Scofield, M. A.; Lewis, W. S.; Schuster, S. S. *J. Biol. Chem.* **1990**, *265*, 12895−12902.
(28) Miller, M. T.; Bachmann, B. O.; Townsend, C. A.; Rosenzweig, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14752−14757.
(29) Miller, M. T.; Bachmann, B. O.; Townsend, C. A.; Rosenzweig, A. C. *Nat. Struct. Biol.* **2001**, *8*, 684−689.

The lassoed-tail structure has been observed previously in small peptide inhibitors.[30] In some cases, this architecture appears to be supported by a network of covalent linkages, including disulfide bonds,[30,31] whereas in one case the tail appears to be sterically trapped within the ring, as is the case with MccJ25.[32]

With the structure of MccJ25 in hand, future studies can focus on its total chemical synthesis, as well as the structure/activity relationship in its interactions with RNAP. Modeling studies indicate MccJ25 (Figure 6b) is an appropriate size and shape to bind in and block the RNAP secondary channel, as suggested from biochemical and genetic experiments.[5] Further insights will also come from investigations into the processing pathway, as well as structural studies of the MccJ25 complex with RNAP.

(30) Katahira, R.; Yamasaki, M.; Matsuda, Y.; Yoshida, M. *Bioorg. Med. Chem.* **1996**, *4*, 121−129.
(31) Frechet, D.; Guitton, J. D.; Herman, F.; Faucher, D.; Helynck, G.; Monegier du Sorbier, B.; Ridoux, J. P.; James-Surcouf, E.; Vuilhorgne, M. *Biochemistry* **1994**, *33*, 42−50.
(32) Katahira, R.; Shibata, K.; Yamasaki, M.; Matsuda, Y.; Yoshida, M. *Bioorg. Med. Chem.* **1995**, *3*, 1273−1280.

Acknowledgment. We are indebted to D. Craik and J. Rosengren for major contributions to our thinking on this problem. We thank R. Salomon for providing plasmids for MccJ25 production, and P. E. Dawson for the gift of cyclized peptide. N-terminal sequencing by Edman degradation was done by J. S. Smith at the University of Texas Medical Branch Biomolecular Resource Facility Protein Chemistry Lab, Galveston, TX. Figure 6 was generated using DINO (http://www.dino3d.org). J. O. was supported by The Rockefeller University's Merck Postdoctoral Fellowship Program. This work was supported by NIH Grant Nos. RR00862 and CA89810 (B.T.C.), GM38660 (R.L.), GM59908 and GM55843 (T.M.), GM64530 (K.S.), and GM61898 (S.A.D.).

Supporting Information Available: Tables containing NMR assignment data and the restraints list used to calculate the MccJ25 structural model (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA036756Q