Antibiotic Structure Determination

Structural and Functional Analysis of Pantocin A: An Antibiotic from *Pantoea agglomerans* Discovered by Heterologous Expression of Cloned Genes**

*Mi Jin, Liang Liu, Sandra A. I. Wright, Steven V. Beer, and Jon Clardy**

The epiphytic bacterium *Erwinia amylovora* causes fire blight, a devastating disease of rosaceous plants such as apple and pear.[1] The closely related but nonpathogenic bacterium *Pantoea agglomerans* (formerly *E. herbicola*), which usually accompanies *E. amylovora* in the wild,[2] produces a family of antibiotics that inhibit *E. amylovora* in a poorly understood strain-dependent fashion.[3,4] The numerous members of this antibiotic family constitute a source of biocontrol agents for fire blight, a pool of new and potentially useful antibiotics against human bacterial pathogens, and a model of nature's ability to produce a "library" of biologically active molecules. The complexity of antibiotic production, especially the multiple antibiotics produced by a given strain and low production in liquid culture, has hampered chemical investigations. We have taken a genomic approach to deconvolute the multiple antibiotics and to address the production problems by a particular strain. Heterologous expression of a genomic DNA library from *P. agglomerans* in *Escherichia coli* provides access to the small-molecule antibiotics. Since both *P. agglomerans* and *E. coli* are members of the widely distributed *Enterobacteriacea*, *E. coli* should be a suitable host for heterologous expression of *P. agglomerans* genes that encode antibiotics. Herein we report the isolation and structure determination of pantocin A, a new peptide-derived antibiotic, and its molecular target.

A genomic library of *P. agglomerans* strain Eh318 was constructed in *E. coli* strain DH5α.[5] Two distinct antibiotic-

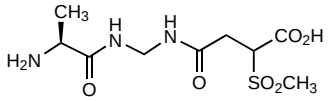
[*] Prof. J. Clardy,[++] M. Jin, L. Liu
Department of Chemistry and Chemical Biology
Cornell University
Ithaca, NY 14853-1301 (USA)
E-mail: jon_clardy@hms.harvard.edu
Prof. S. A. I. Wright,[+] Prof. S. V. Beer
Department of Plant Pathology
Cornell University
Ithaca, NY 14853 (USA)
[+] Current address:
Plant Pathology and Biocontrol Unit
SLU, 750 07 Uppsala (Sweden)
[++] Current address:
Department of Biological Chemistry and Molecular Pharmacology
Harvard Medical School
Boston, MA (USA)
Fax: (+1) 617-432-3702

[**] This work was supported by the National Institutes of Health (National Cancer Institute, R01-CA24487).

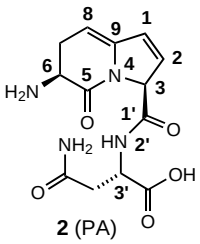
Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

  DOI: 10.1002/ange.200351053

producing cosmid clones, pCPP702 and pCPP719, were identified, and the two antibiotics they produced were trivially named pantocin A and B, respectively. One indication that the two antibiotics were different came from studies which showed that arginine suppresses the antibiotic activity of pantocin B while histidine suppresses the antibiotic activity of pantocin A. In previous work, pantocin B (PB, **1**) was structurally characterized, and its ability to potently inhibit *N*-



1 (PB)



2 (PA)

acetylornithine transaminase, which catalyzes a committed step in arginine biosynthesis, was established.[6–8] The characterization of pantocin A (PA, **2**) was plagued by inconsistent and low levels of production in liquid culture, together with its acid, base, and thermal lability. Significant reproducible production was eventually achieved by culturing *E. coli* XL2Blue harboring a 3.5 kb subclone of pCPP702 in the high copy number plasmid pUC19 (pUC449).

Isolation of **2** was guided by antibiotic activity against the *E. amylovora* strain Ea273 ($IC_{50} = 200$ nM). Pantocin A (**2**) was purified from large-scale fermentation of *E. coli* XL2Blue-pUC449 by four chromatographic separations using anion-exchange, reverse-phase, and hydrophilic interactions. This compound is an optically active small molecule, and its structure was established by spectroscopic methods and synthesis of a hydrogenated derivative.

PA (**2**) has a molecular formula of $C_{13}H_{16}N_4O_5$ as determined from HRFAB-MS studies. The ^{13}C NMR (Table 1) and DEPT spectra indicate four carbonyl, four olefinic, three methine, and two methylene carbon atoms. Three vinylic, a highly deshielded methine, two Cα amino acid, four Cβ amino acid, and three exchangeable protons could be characterized from ^1H and ^1H-^{13}C HMQC NMR spectra recorded in D_2O and [D_6]DMSO. A loss of $m/z = 132$ in the tandem ESI-MS study suggested the presence of an asparagine residue, which was confirmed by 2D NMR studies. Two spin systems (H6/H7a,b/H8 and H1/H2/H3) identified by COSY experiments coupled with HMBC correlations (H6/C5, H8/C9, H1/C9, and H3/C1′) established a continuous nine-carbon chain (C5 to C1′) containing the non-asparagine carbon atoms in **2**. Two additional HMBC correlations (2′NH/C1′ and H3′/C1′) showed that the C_9 and asparagine fragments must be connected by an amide bond.

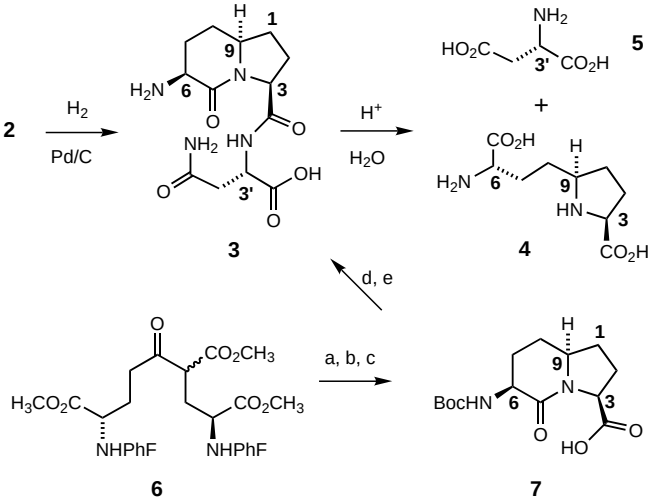
^{15}N NMR experiments on ^{15}N-labeled **2**, which was made using ^{15}N ammonium sulfate in minimal culture media, revealed the nitrogen connectivity. A highly downfield shifted ^{15}N resonance ($\delta = 145.1$ ppm) with correlations to H1, H2, H3, H6, and H8 required N4 to be connected to C3, C5, and C9. A ^{15}N resonance at $\delta = 36.5$ ppm, typical of an amine, and its HMBC correlations with H6 and H7 required an amino group at C6, thus completing the planar structure of **2**.

Table 1: NMR data for **2**.

Atom No.	^{13}C[a]	^1H[b]	^{15}N[c]
1	126.3	6.35 (1H, dd, 6.0, 2.0)	
2	131.5	6.15 (1H, dd, 6.0, 1.7)	
3	66.4	5.21 (1H, brs)	
4-N			145.1 (s)
5	167.9		
6	50.2	3.58 (1H, dd, 13.6, 7.2)	
6-NH_2			36.5 (s)
7	28.6	2.57 (1H, m)	
		2.33 (1H, m)	
8	95.5	5.12 (1H, dd, 6.0, 2.4)	
9	143.3		
1′	166.2		
2′-NH		8.27 (1H, d, 7.5)	126.0 (d, 93.0)
3′	49.8	4.29 (1H, m)	
4′	37.9	2.50 (1H, dd, 15.2, 7.0)	
		2.39 (1H, dd, 15.2, 5.7)	
5′	171.6		
5′-NH_2		6.84, 7.72 (2H, brs)	113.6 (t, 90.0)
6′	172.6		

[a] ^{13}C NMR (100 MHz, 25 °C, [D_6]DMSO). [b] ^1H NMR (500 MHz, 25 °C, [D_6]DMSO) assignments based on ^1H-^{13}C HMQC. [c] ^{15}N NMR (40.5 MHz, 25 °C, D_2O). The ^{15}N chemical shift was calibrated indirectly based on the gyromagnetic ratio of ^{15}N and ^1H according to reference [12].

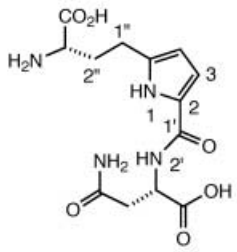
The stereocenters at C3′, C3, and C6 were defined by the hydrogenation of **2**, NOESY experiments, and synthesis. Hydrogenation of **2** gave **3** as a single diastereomer, and NOESY correlations between H3/H9 and H6/H9 require H3 and H6 to be on the same face (Scheme 1). Hydrolysis of **3** gave **4** and L-aspartic acid (**5**), which set the absolute configuration at C3′ as *S*. Both **3** and its 3(*R*),6(*R*),9(*R*) diastereomer were prepared from **6**, which was synthesized from



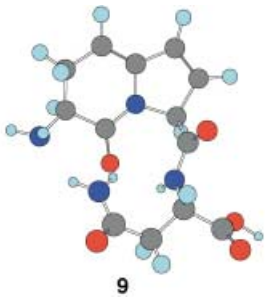
Scheme 1. a) NaOH/dioxane; CH_3I/K_2CO_3/DMSO; b) H_2/Pd-C; c) $(Boc)_2O$/TEA/DMF; NaOH/CH_3OH-H_2O; d) H-Asn-O*t*Bu·HCl, EDC/NMM/DMF; e) TFA/CH_2Cl_2. Boc = *tert*-butoxycarbonyl, TEA = triethylamine, EDC = 3-(3-dimethylaminopropyl)-1-ethylcarbodiimide, NMM = *N*-methylmorpholine, TFA = trifluoroacetic acid.

L- or D-Glu(OMe)-OMe following a literature procedure[9] (Scheme 1). HPLC analysis of synthetic **3** and **3** derived from **2** indicated that the absolute configuration of **3** was 3(*S*),6(*S*). Thus, C3, C6, and C3′ all have the absolute configuration corresponding to an L-amino acid, and the absolute stereostructure of **2** is as shown.

Perhaps the most striking feature of **2**, which to the best of our knowledge has no precedents in reported natural or synthetic compounds, is its very existence given the apparent ease with which it could aromatize to a pyrrole. Removal of a proton from C3, which is α to the carbonyl group at C1′, followed by reprotonation at C8 would generate a pyrrole, and this process likely underlies the lability of **2**. A minor inactive compound was also recovered on isolating **2**, and this was characterized as **8**. Compound **8** arises from the aromatization and hydrolysis of the bicyclic amide bond. No



8

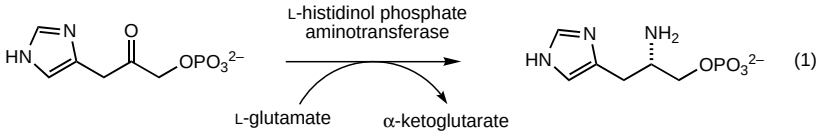


9

material representing aromatization unaccompanied by hydrolysis was isolated—an observation consistent with the known rapid hydrolysis of *N*-acylpyrroles.[10] While **2** is sensitive to acid, base, and elevated temperatures, it survives for days at room temperature in aqueous solution under neutral pH conditions. Two structural features of **2** undoubtedly contribute to its modest persistence: the C3–H bond is shown in energy-minimized structures of **2** not to be aligned with the π bond of the C1′ carbonyl group, and this lack of conjugation creates a kinetic barrier to proton removal (see structure **9**). The out of conjugation conformation is at least partially enforced by hydrogen bonds from asparagine-derived NH groups to the C5 carbonyl oxygen atom. The electron-withdrawing effect of the C5 carbonyl group also diminishes the thermodynamic benefit of aromatization, and this electron-withdrawing effect is enhanced by the same hydrogen bonds.

With the structure of **2** established, we addressed how such a polar molecule gains entry into *E. amylovora* cells and the nature of its molecular target.[11] Addition of the synthetic tripeptide Ala-Gly-Gly to the test medium, in amounts exceeding three times that of **2**, suppresses the antibiotic activity of **2**. Thus, **2** likely gains entry through the tripeptide transporter, and the added tripeptide monopolizes the transporter thereby limiting cellular entry of **2**.[6,11]

Since addition of histidine suppresses the activity of **2**, the cellular target was sought on the histidine biosynthetic pathway.[11] The common strategy of feeding intermediates along a biosynthetic pathway has limited applicability in histidine biosynthesis since many of the intermediates are ionized phosphate esters that have poor cellular access. Using the well-characterized histidine biosynthetic enzymes of *E. coli*, the step catalyzed by L-histidinol phosphate aminotransferase, which converts imidazole acetol phosphate to L-histidinol phosphate [Eq. (1)], was identified as the target of **2**.[11]



L-histidinol phosphate aminotransferase is a pyridoxal-5′-phosphate (PLP) dependent enzyme that couples the reductive amination of imidazole acetol phosphate with the conversion of L-glutamate into α-ketoglutarate.

Pantocin A and B share a close functional relationship that is not apparent from their structures. Both gain access to their cells via the tripeptide transporter and inhibit transaminase-catalyzed amino acid biosynthetic steps that introduce nitrogen functionality.[6,11] Whereas **1** inhibits arginine biosynthesis, **2** inhibits histidine biosynthesis. This functional similarity is unsurprising since they have the common goal of suppressing the growth of bacterial competitors in a nitrogen-poor environment—an environment where import by tripeptide transporters and amino acid biosynthesis are essential.

The structural and functional analysis of **2** (PA) reported here used techniques of small-molecule chemistry and biochemical analysis. The cosmid library/heterologous expression approach provides ready access to the DNA conferring production of **2** on *E. coli* XL2Blue-pUC449. Analysis of this DNA provides insights into the biosynthesis of **2** and identifies strains producing antibiotics related to **2** as discussed in the following communication.

Experimental Section

Cultivation and isolation: *E. coli* XL2Blue-pUC449 was grown in a shaker maintained at 30 °C for 24 h in minimal media supplemented with 100 μg mL^{-1} ampicillin.[5] The culture supernatant was applied to a Dowex 1X8-200 (HCO_3^-) anion-exchange column and eluted with CO_2-saturated H_2O. The active fractions were lyophilized, then applied to a C18 solid-phase extraction column and eluted with H_2O/CH_3CN (100/0 to 90/10). Further fractionation by HPLC, first with a reverse-phase column (Vydac, 218TP510) using H_2O/0.1 M NH_4OAc/CH_3CN (94/5/1), then with a polyhydroxyethyl A column (Poly LC, Inc.) using H_2O/0.1 M NH_4OAc/CH_3CN (20/5/75) gave compounds **2** and **8** in yields of approximately 1 mg L^{-1} and 0.2 mg L^{-1}, respectively.

Cloning of *E. coli*-pUC449: Pantocin A producing cosmid clone pCPP702 was subcloned to generate pCPP1051, from which a XbaI-HindIII fragment was further subcloned into pBluescript to yield plasmid pCPP717.[5] The cloned insert was sequenced (GenBank accession No. U81376). A 3.5 kb DNA fragment was then amplified by a polymerase chain reaction from pCPP717 using primer (5′-CCGCATCTAGAGTAGGTATGAC-3′) and primer (5′-ATACTCTGCAGAGTTGGTGCTCCA-3′), and ligated into pUC19 at the *XbaI* and *PstI* sites. Recombinants were selected

from transformed *E. coli* XL2 Blue cells on Luria–Bertani plates supplemented with ampicillin at 100 μg mL^{-1}, to yield clone *E. coli* XL2 Blue-pUC449.

2: $[\alpha]_D^{25} = -291.1°$ ($c = 1.8$, H_2O); IR (KBr): $\tilde{\nu}_{max} = 3700$–2400 (br), 1676, 1578, 1402 cm^{-1}; UV (H_2O): λ_{max} (ε) = 200 (12089), 272 nm (5447); HRFAB-MS: *m/z*: 331.1018 [*M*+Na]$^+$, calcd *m/z*: 331.1018 for $C_{13}H_{16}N_4O_5Na$; ESI-MS/MS: *m/z*: 309 [*M*+H]$^+$, 177, 159, 151, 149.

8: ^{1}H NMR (500 MHz, D_2O, 25 °C): $\delta = 6.81$ (1 H,d, $J = 3.7$ Hz, H3), 6.12 (1 H, d, $J = 3.7$ Hz, H4), 4.66 (1 H, dd, $J = 9.0$, 4.5 Hz, H3′(Asn αH)), 3.75 (1 H, t, $J = 6.5$ Hz, H3″), 2.85 (1 H, dd, $J = 15.0$, 4.5 Hz, H4′a), 2.80–2.69 (3 H, m, H1″ and H4′b), 2.25–2.10 ppm (2 H, m, H2″); ^{13}C NMR (100 MHz, 5 % CD_3OD in D_2O, 25 °C): $\delta = 178.6$ (C6′, COOH), 177.0 (C5′, $CONH_2$), 175.4 (C4″) 163.4 (C1′), 137.4 (C5), 125.0 (C2), 113.5 (C3), 108.6 (C4), 55.3 (C3″), 52.9 (C3′), 39.1 (C4′), 31.5 (C2″), 23.8 ppm (C1″); ESI-MS/MS/MS: *m/z*: 327 [*M*+H]$^+$, 310, 195, 177, 149, 132, 106; UV (H_2O): λ_{max} (ε) = 200 (6438), 220 (2994, sh), 275 nm (10887); IR (thin film): $\tilde{\nu}_{max} = 3700$–2400 (br), 1668, 1604, 1533, 1404 cm^{-1}.

Preparation of **3** from **2:** A solution of **2** (2.4 mg) in deionized H_2O (300 μL) was treated with Pd/C (10 wt %, 0.66 mg) and stirred under hydrogen at room temperature for 8 h. The reaction mixture was filtered over celite, washed with deionized H_2O (2 mL), and the combined aqueous fractions were dried to give 1.7 mg of **3**. ^{1}H NMR (500 MHz, D_2O, 25 °C): $\delta = 4.46$ (1 H, d, $J = 9.0$ Hz, H3), 4.40 (1 H, dd, $J = 8.0$, 5.0 Hz, H3′), 3.93 (1 H, t, $J = 8.5$ Hz, H6), 3.74 (1 H, m, H9), 2.76 (1 H, dd, $J = 15.0$, 5.0 Hz, H4′a), 2.65 (1 H, dd, $J = 15.0$, 8.0 Hz, H4′b), 2.35 (1 H, m, H7a), 2.25–2.10 (4 H, m, H2a, 1a, 8a, 2b), 1.91 (1 H, m, H7b), 1.67 (1 H, m, H8b), 1.64 ppm (1 H, m, H1b); ^{13}C NMR (100 MHz, D_2O + 5 % CD_3OD, 25 °C): $\delta = 177.8$ (C6′), 176.6 (C5′), 173.6 (C1′), 169.5 (C5), 61.0 (C3), 59.2 (C9), 53.2 (C3′), 49.6 (C6), 38.6 (C4′), 32.3 (C1), 30.2 (C2), 26.2 (C8), 25.6 (C7); $[\alpha]_D^{25} = -57.8°$ ($c = 0.16$, H_2O); UV (H_2O): λ_{max} (ε) = 200 nm (2200); IR (thin film): $\tilde{\nu}_{max} = 3700$–2400 (br), 2954 (sh), 1668, 1529 (sh), 1448, 1392, 1342 cm^{-1}; ESI-MS/MS: *m/z*: 313 [*M*+H]$^+$, 181, 153, 125.

Chiral TLC**:** Chiral TLC analyses were performed on CHIRALPLATE (Macherey-Nagel) using solvent A ($CH_3OH/H_2O/CH_3CN$, 5/5/20) or B (acetone/H_2O/CH_3OH, 10/2/2). **3** (0.15 mg, 0.5 μmol) was heated in 6 M HCl (300 μL) at 85 °C for 19 h to give **4** and **5**. Authentic D- and L- Asn were also treated in the same manner. Compound **5** migrated the same distance as treated L-Asn ($R_f = 0.47$ in A, 0.66 in B), which was easily distinguishable from treated D-Asn ($R_f = 0.40$ in A, 0.58 in B)

3: As shown in Scheme 1, *N*-(PhF) L-or D-dimethyl glutamate was used in a Claisen condensation, reductive amination, and lactam-cyclization sequence to give enantiopure (3*S*,6*S*,9*S*)-**7** or (3*R*,6*R*,9*R*)-**7** following the procedure of Lombart and Lubell.[9] Coupling with *tert*-butyl-L-asparagine ester followed by removal of protecting groups then gave (3*S*,6*S*,9*S*)-**3** or (3*R*,6*R*,9*R*)-**3** in 7 % overall yield.

(3*S*,6*S*,9*S*)-**3**: $[\alpha]_D^{25} = -50.7°$ ($c = 1.01$, H_2O); ^{1}H NMR (500 MHz, D_2O, 25 °C): $\delta = 4.61$ (1 H,dd, $J = 6.0$, 7.0 Hz), 4.51 (1 H, d, $J = 8.8$ Hz), 4.06 (1 H, t, $J = 8.5$ Hz), 3.78 (1 H, m), 2.87–2.77 (2 H, m), 2.42 (1 H, m), 2.31–2.18 (3 H, m), 2.13 (1 H, m), 1.97 (1 H, m), 1.73–1.63 ppm (2 H, m). (The small differences in δ values between (3*S*,6*S*,9*S*)-**3** and **3** derived from **2** are the result of differences in pH values and concentration. Mixing two samples resulted in overlapping peaks). ^{13}C NMR (100 MHz, D_2O + 10 % CD_3OD, 25 °C): $\delta = 175.9$, 175.6, 174.0, 168.0, 60.8, 59.1, 51.1, 49.4, 37.5, 32.3, 30.2, 26.3, 25.0 ppm; IR (thin film): $\tilde{\nu}_{max} = 3700$–2400 (br), 2956 (sh), 1670, 1530 (sh), 1448, 1390, 1340 cm^{-1}; ESI-MS/MS: *m/z*: 313 [*M*+H]$^+$, 181, 153, 125.

(3*R*,6*R*,9*R*)-**3**: $[\alpha]_D^{25} = +45.7°$ ($c = 0.98$, H_2O); ^{1}H NMR (500 MHz, D_2O, 25 °C): $\delta = 4.62$ (1 H, dd, $J = 8.0$, 5.0 Hz), 4.53 (1 H, d, $J = 8.8$ Hz), 4.06 (1 H, t, $J = 8.5$ Hz), 3.78 (1 H, m), 2.84 (1 H, dd, $J = 15.5$, 5.0 Hz), 2.71 (1 H, dd, J = 15.5, 8.0 Hz), 2.42 (1 H, m), 2.30–2.17 (3 H, m), 2.05 (1 H, m), 1.96 (1 H, m), 1.75–1.64 ppm (2 H, m); ^{13}C NMR (100 MHz, D_2O + 10 % CD_3OD, 25 °C): $\delta = 175.9$, 175.3, 173.9, 168.1, 60.8, 59.1, 50.9, 49.5, 37.6, 32.3, 30.6, 26.3, 25.0 ppm; IR (KBr): $\tilde{\nu}_{max} = 3700$–2400 (br), 2955 (sh), 1671, 1528, 1447 cm^{-1}; ESI-MS/MS: *m/z*: 313 [*M*+H]$^+$, 181, 153, 125.

HPLC analysis was performed on an Aqua C18 column (250 × 4.6 mm, 5 μ, 125 Å, Phenomenex) with a mobile phase of H_2O(0.1 % TFA)/CH_3CN(0.1 % TFA) with a gradient of 100/0 (0–5 min); 100/0 to 85/15 (5–20 min) at 1 mL min^{-1}. (3*R*,6*R*,9*R*)-**3** and (3*S*,6*S*,9*S*)-**3** showed base-line separation with retention times of 15.6 and 15.9 min, respectively. **3** derived from **2** co-migrated with (3*S*,6*S*,9*S*)-**3** but not (3*R*,6*R*,9*R*)-**3**.

Received: January 29, 2003 [Z51053]

Keywords: antibiotics · gene expression · natural products · structure elucidation

[1] H. S. Aldwinckle, S. V. Beer, *Hortic. Rev.* **1979**, *1*, 423.
[2] J. M. Erskine, L. E. Lopatecki, *Can. J. Microbiol.* **1975**, *21*, 35.
[3] R. S. Wodzinski, J. P. Paulin, *J. Appl. Bacteriol.* **1994**, *76*, 603.
[4] C. A. Ishimaru, E. J. Klos, R. R. Brubaker, *Phytopathology* **1988**, *78*, 746.
[5] S. A. I. Wright, C. H. Zumoff, L. Schneider, S. V. Beer, *Appl. Environ. Microbiol.* **2001**, *67*, 284.
[6] S. F. Brady, S. A. Wright, J. C. Lee, A. E. Sutton, C. H. Zumoff, R. S. Wodzinski, S. V. Beer, J. Clardy, *J. Am. Chem. Soc.* **1999**, *121*, 11912.
[7] A. E. Sutton, J. Clardy, *Org. Lett.* **2000**, *2*, 319.
[8] A. E. Sutton, J. Clardy, *J. Am. Chem. Soc.* **2001**, *123*, 9935.
[9] H.-G. Lombart, W. D. Lubell, *J. Org. Chem.* **1996**, *61*, 9437.
[10] F. M. Menger, J. A. Donohue, *J. Am. Chem. Soc.* **1973**, *95*, 432.
[11] M. Jin, Ph.D. thesis, Cornell University (Ithaca), **2003**.
[12] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, N. J. Skelton, *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, New York, **1995**.

