The Biosynthesis of Sulfomycin Elucidated by Isotopic Labeling Studies

Gwendolyn D. Fate,*,† Charles P. Benner,† Stephen H. Grode,‡ and Terry J. Gilbertson†

Contribution from the Animal Health Drug Metabolism and Chemical and Fermentation Operations Structure Laboratory, The Upjohn Company, Kalamazoo Michigan 49001

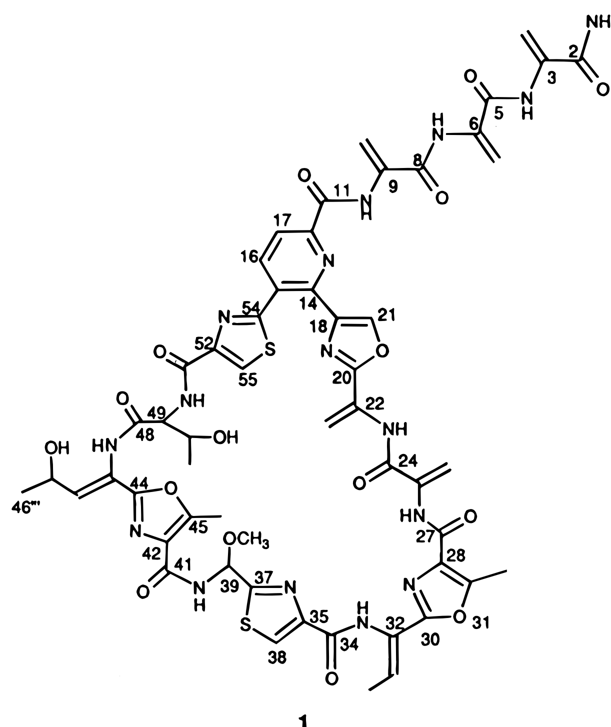
Received June 3, 1996⊗

Abstract: U-102408 has been isolated from *S. arginensis* cultures and fully characterized by NMR and MS. These studies indicate that U-102408 is identical with sulfomycin, and this data confirm the structure of sulfomycin which was assigned by chemical degradation and limited NMR studies. The biosynthetic origins of all U-102408 structural features have been elucidated through isotopic labeling experiments with primarily ^{13}C but also using ^{14}C, deuterium or tritium. Of particular interest, the 4-hydroxy-2-amino-2-pentenoic acid moiety was determined to originate from threonine and a one carbon unit derived from the two position of glycine or the *S*-methyl of methionine. The biosynthetic pathway for U-102408 differs from that for nosiheptide, thiostrepton, and berninamycin in that cysteine and serine do not cross label sites in U-102408, while this has been observed for the other thiopeptides. In U-102408 cysteine and serine are each incorporated into unique sites within the molecule. This indicates that the interconversion of cysteine and serine is not an active pathway for *S. arginensis*. In fact for fermentations in defined medium the addition of cysteine was found to be necessary to achieve higher antibiotic titer. Unlike the other thiopeptides that have been studied, threonine was incorporated into sites labeled by serine though serine did not incorporate into sites labeled by threonine within U-102408. No cross labeling of threonine and serine has been observed in studies of nosiheptide, thiostrepton, or berninamycin. In the current study glycine is investigated as a precursor for U-102408. The interconversion of glycine to serine is a very facile pathway in *S. arginensis*; glycine labels all sites labeled by serine. This data suggests glycine as a possible precursor for nosiheptide, thiostrepton, and berninamycin which has not been investigated for these other thiopeptides. Incorporation of label from 3-^{2}H-serine is facile for thiostrepton and nosiheptide and was used to probe the mechanism of formation of the pyridine ring. However, no incorporation from 3-^{2}H-serine or 3-^{3}H-serine was observable for U-102408. While the lack of incorporation may be due to washout of the label from the precursor, it also may be due to the interconversion of serine and glycine causing loss of the label from the serine pool.

Introduction

Sulfomycin I (**1**) has been isolated from *Streptomyces viridochromogenes* and characterized by Abe.[1] Structural assignments were based on degradation of the macrocycle under acidic or basic conditions and characterization of the resultant fragments. The overall structure of **1** was then deduced from limited NMR studies performed on the intact molecule.[2] U-102408 was subsequently isolated from *S. arginensis*, was fully characterized by NMR and MS, and found to be structurally identical with sulfomycin.

Compound **1** is a member of the thiopeptide family of antibiotics which includes thiostrepton,[3] thiopeptin,[4] nosiheptide,[5,6] micrococcin,[7] siomycin,[8] berninamycin,[9] and sporangio-



1

† Animal Health Drug Metabolism.
‡ Chemical and Fermentation Structure Laboratory.
⊗ Abstract published in *Advance ACS Abstracts,* October 1, 1996.

(1) Abe, H.; Ikeda, M.; Takaishi, T.; Ito, Y.; Okuda, T. *Tetrahedron Lett.* **1977**, 735.
(2) Abe, H.; Kushida, K.; Shiobara,Y.; Kodama, M. *Tetrahedron Lett.* **1988**, 1401−1404.
(3) Anderson, B.; Crowfoot Hodgkin, D. C.; Viswamitra, M. A. *Nature* **1970**, *225*, 233−235.
(4) Hensens, O. D.; Albers-Shönberg G. *J. Antibiot.* **1983**, *36*, 814−831.
(5) Depaire, H.; Thomas, J. P.; Brun, A.; Olesker, A. Lukacs G. *Tetrahedron Lett.* **1977**, *16*, 1397−1400, 1403−1406.
(6) Pascard, C.; Ducroix, A.; Lunel, J.; Prange, T. *J. Am. Chem. Soc.* **1977**, *99*, 6418−6423.
(7) Shoji, J.; Hinoo, H.; Wakisaka, Y.; Koizumi, K. Mayama, M.; Matsuura, S.; Matsumoto, K. *J. Antibiot.* **1976**, *29*, 366−374.

S0002-7863(96)01864-1 CCC: $12.00 © 1996 American Chemical Society

Table 1. ^{1}H-NMR Correlations Observed by COSY

residue	Coupled ^{1}H Resonances			
pyridine	8.71	8.30		
threonine	8.18	4.34	4.15	1.09
				5.23
oxazole C	6.27	4.60	1.21	
			4.94	
thiazole A	8.35	6.46		
oxazole B	6.49	1.76		

mycin.[10] This class of compounds is interesting due to its potent antibiotic activity against Gram-positive bacteria. *In vivo* the compounds inhibit bacterial protein synthesis by binding to the L11 protein 23S RNA complex in the ribosome during transcription.[11]

The thiopeptide antibiotics are highly modified peptides possessing one or two macrocyclic rings with the following structural features: pyridine or dihydropyridine ring, dehydroalanine residues, and thiazole and oxazole rings. Recently the biosynthesis of nosiheptide,[12] thiostrepton,[13] and berninamycin[14] has been investigated by following the incorporation of isotopically labeled amino acid precursors. We now report the results of isotopic labeling studies of U-102408. Incorporation of amino acid precursors has allowed the origin of the structural features of **1** to be determined including a novel 2-amino-4-hydroxy-2-pentenoic acid residue.

Results

Structure Analysis. The electrospray ionized mass spectrum of U-102408 was dominated by an ion at *m/z* 1268.1. This was consistent with the molecular weight of sulfomycin I plus one sodium adduct ($C_{54}H_{52}N_{16}O_{16}S_2$ Na, 1268.2 amu). Characterization of the intact molecule by NMR spectroscopy was accomplished at 500 MHz in DMSO-d_6 utilizing a series of 2D experiments.[15] The ^{1}H-NMR of U-102408 revealed a spectrum that was remarkably free of spectral overlap except in the exomethylene region between δ 5.6 and δ 5.8.[16] Most of the signals were singlets and those that did manifest multiplet structure were readily defined by COSY (*CO*rrelation *S*pectroscop*Y*). For example, a spin system consisting of δ 1.09 (3H, d) −4.15 (1H, m) −5.23 (1H, d)) −4.34 (1H, m) −8.18 (1H, d) was assigned to the threonine residue. Parts of the pyridine, thiazole A, oxazole A, and oxazole B residues were determined similarly. A listing of the proton−proton correlations observed by COSY is presented in Table 1.

The standard method of characterizing a molecule by NMR spectroscopy is to correlate all H−H connectivities by COSY and transfer that information to the carbon skeleton by means of a C−H correlation experiment. Difficulties arise in highly unsaturated molecules because the large number of nonprotonated carbons represent breaks in the sequence. In these cases less sensitive interactions such as NOE and long-range correlations need to be defined in order to bridge the many isolated spin systems.

Table 2. ^{1}H-NMR Correlations Observed by LRCOSY

residue	coupled ^{1}H resonances	
dehydroalanine A	9.07	5.65
		6.12
dehydroalanine B	10.08	5.69
dehydroalanine C	10.42	5.95
oxazole A	9.97	5.79
		5.66
oxazole A	8.64	5.79
		5.66
dehydroalanine D	9.17	5.78
oxazole B	10.00	6.49
oxazole C	9.44	6.27

One such method is to observe long-range H−H correlations by a COSY experiment modified to accentuate small interactions, LRCOSY (Long Range COSY). In this manner it is possible to correlate the N*H*-C═C*H*$_2$ spins and identify the four dehydroalanine residues and the allylic parts of the oxazole residues. In addition, a highly unusual six-bond correlation between an oxazole proton at 8.64 (H21) and exomethylene protons at 5.79/5.66 (H22′) allowed determination of the oxazole A residue. A summary is presented in Table 2.

The correlation of adjacent residues and the determination of the ordering of the allylic systems constituting the dehydroalanine residues (reported above) were resolved by NOESY. The terminal NH_2 (H1, 7.52/7.94) determined dehydroalanine A by interaction with the exomethylene protons at 5.65/6.12 (H3′). The NH of dehydroalanine A (H4, 9.07), in turn, correlated to the dehydroalanine B exomethylene protons at 5.69/5.72 (H6′); and the NH of dehydroalanine B (H7, 10.08) correlated to the dehydroalanine C exomethylene protons at 6.58/5.95 (H9′), completing determination of the "dehydroalanine tail". The dehydroalanine tail could now be tied to the endocyclic ring by means of the NOEs associated with the dehydroalanine C N*H*. The 10.42 −8.64 correlation established the dehydroalanine C(N*H*10)−oxazole A (*H*21) connectivity and the 10.42−8.30 (pyridine *H*17) and 8.71 (pyridine *H*16) −8.18 (threonine N*H*50) established the dehydroalanine C−pyridine−threonine connectivity. Additional NOEs observed within the endocyclic ring helped establish adjacent residues between oxazole A and dehydroalanine D (N*H*(23) 9.97 and exomethylene protons (H25′) 6.40/5.78), oxazole B and thiazole A (N*H*(33) 10.00 and (H38) 8.45), and oxazole C and threonine (N*H*(47) 8.18 and (H49′) 4.15). Also the location of the OCH_3 was fixed on thiazole A by virtue of an NOE between N*H*(40) 8.35 and OC*H*$_3$ 3.24.

At this point, with the exception of the oxazole B and C methyl groups (interchangeably assigned to 2.56 and 2.57), the assignment of all proton resonances had been made (Table 3). What remained were the ^{13}C-NMR assignments and these followed from the HETCOR (for protonated carbons) and from long-range C−H connectivities, for the nonprotonated carbons. A list of long-range C−H correlations derived from the LRHETCOR experiment are presented in Table 4. All ^{13}C-NMR chemical shifts from U-102408 are tabulated in Table 5 and are consistent with the partial assignment of resonances in ref 2.

Biosynthetic Studies. Several functional motifs are repeated throughout structure **1**: (1) Dehydroalanine residues form the side chain and a portion of the macrocyclic ring. (2) Three oxazole rings are present, two bearing a methyl adjacent to the

(8) Tokura, K.; Tori, K.; Yoshimura, Y,; Okabe, K.; Otsuka, H.; Matsushita, K.; Inagaki, K.; Miyazawa, T. *J. Antibiot.* **1980**, *33*, 1563−1567.
(9) Lau, R. C. M.; Rinehart, K. L. *J. Am. Chem. Soc.* **1995**, *117*, 7606−7610.
(10) Thiemann, J. E.; Coronelli, C.; Pagani, H.; Beretta, G.; Tamoni, G.; Arioli, V. *J. Antibiot.* **1968**, *21*, 525−531.
(11) Cundliffe, E.; Thompson, J. *J. Gen. Microbiol.* **1981**, *126*, 185−192.
(12) Mocek, U.; Knaggs, A. R.; Tsuchiya, R.; Nguyen, T.; Beale, J. M.; Floss, H. *J. Am. Chem. Soc.* **1993**, *115*, 7557−7568.
(13) Mocek, U.; Zeng, Z.; O'Hagan, D.; Zhou, P.; Fan, L.-D. G.; Beale, J. M.; Floss, H. *J. Am. Chem. Soc.* **1993**, *115*, 7992−8001.
(14) Lau, R. C. M., Rinehart, K. L. *J. Am. Chem. Soc.* **1995**, *117*, 7606−7610.
(15) The authors of ref 1 report the assignment of ^{1}H- and ^{13}C-NMR resonances of the intact molecule, however the data were not presented.
(16) All NMR peaks are reported in δ units (ppm) and internally referenced to DMSO (^{1}H: DMSO-d_5, δ = 2.49; ^{13}C: DMSO-d_6, δ = 39.5).

Table 3. The ^{1}H-NMR Assignments of U-102408

position	^{1}H assignment	position	^{1}H assignment
1x	7.52 (s)	38	8.45 (s)
1y	7.94 (s)		
3′x	5.65 (s)	39	6.46 (d, J = 9.45)
3′y	6.12 (s)		
4	9.07 (s)	OCH_3	3.24 (s)
6′x	5.69 (s)	40	8.35 (d, J = 9.45)
6′y	5.72 (s)		
7	10.08 (s)	45′	2.56[b] (s)
9′x	6.58 (s)	46′	6.27 (d, J = 8.40)
9′y	5.95 (s)		
10	10.42 (s)	46″	4.46 (ddq, J = 8.40, 4.60, 6.05)
16	8.71 (d, J = 8.25[a])	46″OH	4.94 (d, J = 4.60)
17	8.30 (d, J = 8.25)	46‴	1.21 (d, J = 6.05)
21	8.64 (s)	47	9.44 (s)
22′x	5.79 (s)	49	4.34 (m)
22′y	5.66 (s)		
23	9.97 (s)	49′	4.15 (m)
25′x	6.40 (s)	49′OH	5.23 (d, J = 5.50)
25′y	5.78 (s)		
26	9.17 (s)	49″	1.09 (d, J = 6.05)
31′	2.57[b]	50	8.18 (d, J = 8.00)
32′	6.49 (q, J = 6.85)	55	8.64 (s)
32″	1.76 (d, J = 6.85)		

[a] Coupling constants are recorded in Hz. [b] These assignments are interchangeable.

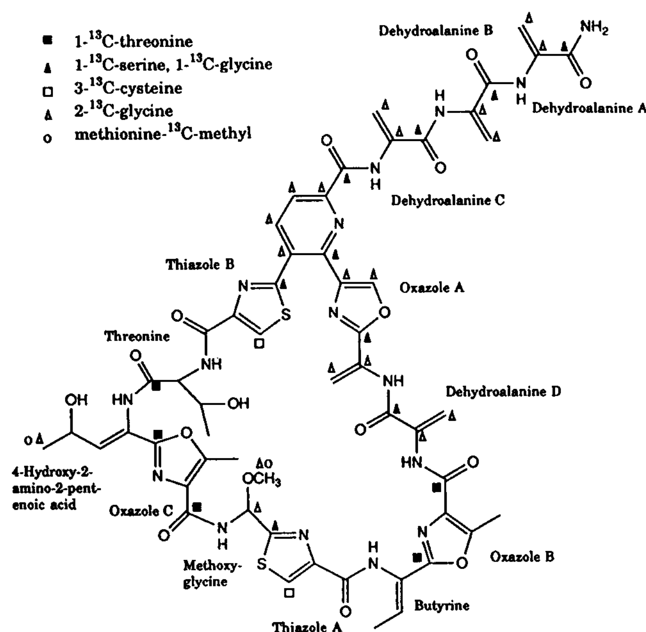


Figure 1. Incorporation of ^{13}C into **1** from isotopically labeled cysteine, glycine, methionine, threonine, and serine.

ring oxygen. (3) There are two thiazoles in the macrocycle. (4) There is an unmodified threonine residue incorporated at position C-48. There were several questions of interest: (1) The origin of the novel 2-amino-4-hydroxy-2-pentenoic acid residue at position C-45 was not clear. (2) Elucidation of the origin of the butyrine residue at position C-30 and the source of the methoxy attached to C-39. (3) The biosynthetic origin and condensation mechanism which give rise to the pyridine ring.

The approach used to study the biosynthesis of **1** was to fortify the fermentation medium with likely amino acid precursors so that these amino acids would be utilized during production of **1**. *Streptomyces arginensis* was cultured as previously described[17] in the presence of 125 mg/L ^{13}C or ^{2}H labeled threonine, serine, glycine, or methionine. The sites of enrichment observed for these precursors is shown in Figure 1 and summarized with percents of incorporation in Table 5.

Table 4. Long-Range ^{1}H- and ^{13}C-NMR Correlations Observed in the LRHETCOR Experiment

^{1}H assignment	^{13}C assignment
dehydroalanine A, H1; 7.52	dehydroalanine A, C3; 134.5
dehydroalanine A, H3′; 5.65/6.12	dehydroalanine A, C2; 165.0
dehydroalanine B, H6′; 5.69	dehydroalanine B, C5; 162.2
dehydroalanine B, H6′; 5.69/5.72	dehydroalanine B, C6; 136.9
dehydroalanine B, H7; 10.08	dehydroalanine C, C8; 162.6
dehydroalanine C, H9′; 6.58	dehydroalanine C, C9; 133.8
dehydroalanine C, H10; 10.42	pyridine, C11; 161.4
pyridine, H16; 8.71	pyridine, C12; 149.0
pyridine, H16; 8.71	pyridine, C14; 146.7
oxazole A, H21; 8.64	oxazole A, C18; 138.7
oxazole A, H22′; 5.66	oxazole A, C22; 129.6
dehydroalanine D, H25′; 6.40	dehydroalanine D, C25; 133.4
oxazole B, H32″; 1.76	oxazole B, C32; 123.5
oxazole B, H33; 10.00	thiazole A, C34; 159.1
thiazole A, H38; 8.45	thiazole A, C35; 148.8
thiazole A, H38; 8.45	thiazole A, C37; 167.5
thiazole A, H40; 8.35	oxazole C, C41; 161.3
oxazole C, H46′; 6.27	oxazole C, C44; 156.3
threonine, H50; 8.18	thiazole B, C51; 160.6
thiazole B, H55; 8.64	thiazole B, C52; 148.8
thiazole B, H55; 8.64	thiazole B, C54; 162.3

U-^{14}C-glucose, 1-^{14}C-acetate, and 1-^{14}C-propionate were also investigated as precursors but gave no detectable label incorporation.

Thiazoles. Thiazoles A and B are derived from the condensation of cysteine with the carbonyl of an adjacent amino acid residue, in each case a modified serine moiety. Incorporation of 3-^{13}C-cysteine was observed into thiazoles A and B at positions C-38 (δ 126.95, 30%) and C-55 (δ 127.5, 24%). Position C-39 (δ 77.1, 8%) adjacent to thiazole A was labeled by 2-^{13}C-glycine so this ring results from the condensation of a glycine with a cysteine residue. In the case of thiostrepton, sites that incorporated 3-^{13}C-cysteine also were enriched by 3-^{13}C-serine.[13] This was not the case for **1** since cross labeling of sites was not observed for the two precursors.

Dehydroalanines. The side chain of the macrocycle consists of three dehydroalanine residues. These residues originate from dehydration of serine as demonstrated by the incorporation of 1-^{13}C-serine into the amide carbon of dehydroalanine A, B, and C at C-2 (δ 165.0, 6%), C-5 (δ 162.2, 8%), and C-8 (δ 162.6, 3%), respectively. Two other dehydroalanine residues are present within the macrocycle and positions C-20 (δ 158.3, 4%) and C-24 (δ 162.9, 6%) are labeled by 1-^{13}C-serine. These correspond to dehydroalanine D and the adjacent dehydroalanine residue in which the carboxyl has condensed with an adjacent serine to form oxazole A.

Addition of 2-^{13}C-glycine to the fermentation medium resulted in the labeling of both methylenic carbons present in all the dehydroalanine residues labeled by serine. In the side chain dehydroalanine carbons C-3 (δ 134.5, 7%), C-3′ (δ 104.2, 3%), C-6 (δ 136.9, 7%), C-6′ (δ 111.8, 2%), C-9 (δ 133.8, 8%), and C-9′ (δ 105.7, 3%) are labeled. In the macrocycle dehydroalanines are labeled at C-25 (δ 133.4, 5%), C-25′ (δ 105.2, 3%), and adjacent to oxazole A at positions C-22 (δ 129.6, 3%) and C-22′ (δ 112.9, 3%). These sites of incorporation suggest that glycine is being used as a precursor for the biosynthesis of serine. This pathway is mediated by serine hydroxymethyltransferase and involves tetrahydrofolate as methylene transfer agent.[18,19] Biosynthesized in this way from 2-^{13}C-glycine, the

(17) Argoudelis, A. D.; Shilladay, F. B.; Laborde, A. L.; Truesdale, S. E. Process for the Preparation of Antibiotics 10381B. *International Patent Application No.* PCT/US87/01450, International Publication No. WO 88/00200, 14 Jan. 1988.

(18) Bender, D. A. *Amino Acid Metabolism*; John Wiley & Sons: New York, 1985, pp 95−118.

Table 5. Incorporation of ^{13}C from Various Isotopically Enriched Amino Acid Precursors Added to the Fermentation Medium into 1

position	^{13}C δ, ppm	% increase in ^{13}C enrichment					
		1-^{13}C-Ser	1-^{13}C-Gly	2-^{13}C-Gly	^{13}C-me-Met	1-^{13}C-Cys	1-^{13}C-Thr
2	165.0	6	10	-			*b*
3	134.5			7			
3′	104.2			3			
5	162.2	8	12				*c*
6	136.9			7			
6′	111.8			2			
8	162.6	3	4				*b*
9	133.8			8			
9′	105.7			3			
11	161.4	8	11				*c*
12	149.0			5			
14	146.7	4	6				*b*
15	130.8			8			
16	140.1			4			
17	121.9			3			
18	138.7			4			
20	158.3	4	4				*b*
21	140.2			4			
22	129.6			3			
22′	112.9			3			
24	162.9	6	9				*b*
25	133.4			5			
25′	105.2			2			
27	159.8						28
28	128.6[a]						
30	156.8						14
31	154.1[a]						
31′	11.5[a]						
32	123.5						
32′	129.7						
32″	13.7						
34	159.1						
35	148.8						
37	167.5	3	5				*b*
38	127.0					30	
39	77.1			8			
OCH_3	55.4			5	7		
41	161.3						12
42	129.1[a]						
44	156.3						15
45	153.5[a]						
45′	11.5[a]						
46	120.6						
46′	136.7						
46″	62.5						
46‴	22.8			4	8		
48	169.7						20
49	59.4						
49′	66.2						
49″	20.2						
51	160.6						
52	148.8						
54	127.5	4	7				
55	162.3					24	*b*

[a] The oxazole B 28, 31, 31′ series is interchangeable with the oxazole C 42, 45, 45′ series. [b] Variable levels of label incorporation were observed at this position. Incorporation was as high as 3% (C-2), 7% (C-5), 12% (C-14), 8% (C-20), 4% (C-24), 10% (C-37), and 15% (C-55). [c] Incorporation was observed into this position, however exact integration of the peak was not possible due to an overlapping resonance.

product is doubly labeled serine which is then available for incorporation into **1**.

Label from 1-^{13}C-threonine was also incorporated into these sites labeled by 1-^{13}C-serine and 1-^{13}C-glycine with the level of incorporation varying from run to run. Incorporation was as high as 3% (C-2), 7% (C-5), 12% (C-14), 8% (C-20), 4% (C-24), 10% (C-37), and 15% (C-55); however, in other runs little or no enrichment in these positions was detectable. Threonine may be converted to glycine by threonine aldolase and subsequently to serine, resulting in the observed cross labeling of sites within **1**. Sites uniquely labeled by threonine (i.e., not cross labeled by glycine or serine) had reproducible levels of incorporation (typically within ±5%), results are reported in Table 5.

No incorporation of 3-^{3}H-serine was observable presumably due to loss of the label from the precursor prior to or during incorporation. The loss of label may occur through the same pathway of serine and glycine interconversion, though this is not a mechanistic imperative, or may simply be due to washout of the label from the precursor.

Pyridine. The labeling pattern of the pyridine ring shows that it is formed by condensation of two molecules of serine in a tail-to-tail orientation with the carboxyl of a third serine residue. This is suggested by incorporation from 1-^{13}C-serine at positions C-11 (δ 161.4, 8%) of the side chain, C-14 (δ 146.7, 4%) of the pyridine, and C-54 of thiazole B (δ 162.3, 4%). A complimentary labeling pattern is obtained when 2-^{13}C-glycine is used as a precursor. Glycine is incorporated after conversion to doubly labeled serine to give ^{13}C enrichment of the pyridine positions arising from two serine residues in a tail-to-tail orientation. Thus, positions C-12 (δ 149.0, 5%), C-17 (δ 121.9, 3%), C-15 (δ 130.8, 8%), and C-16 (δ 140.1, 4%) are labeled. The third doubly labeled serine contributes a carboxyl to the formation of the pyridine ring and the side chain to oxazole A to give enrichment from 2-^{13}C-glycine at positions C-18 (δ 138.7, 4%) and C-21 (δ 140.2, 4%). The tetrahydropyridine ring of thiostrepton and the 3-hydroxypyridine ring of nosiheptide are also derived from the condensation of three serine residues. A mechanism has been proposed for the ring forming condensation reaction in this class of antibiotics.[13] The labeling pattern observed for U-102408 is also consistent with this mechanism.

Oxazoles. Label from 1-^{13}C-threonine was incorporated into C-48, the amide carbon δ 169.71 (20%) of an unmodified threonine residue in the macrocycle. Oxazole B and C are also labeled by threonine. Oxazole B results from the condensation of two adjacent threonine residues to give incorporation at positions C-27 (δ 159.8, 28%) and C-30 (δ 156.8, 14%). Labeling of C-30 demonstrates that the adjacent 2-butene side chain arises from the dehydration of the threonine precursor. Labeling at C-41 (δ 161.3, 12%) and C-44 (δ 156.3, 15%) of Oxazole C suggests it is also formed from the dimerization of two threonines, one modified to an unusual 2-amino 4-hydroxy-2-pentenoic acid moiety.

2-Amino-4-hydroxy-2-pentenoic Acid. The 2-amino-4-hydroxy-2-pentenoic acid residue originates from threonine and an additional terminal carbon appended from the one carbon pool. Threonine is also a precursor for thiostreptone residue in thiostrepton and is a precursor to isoleucine biosynthesis.[13] A 15% enrichment of this residue from threonine was observed, this was similar to that for other sites of threonine incorporation which averaged 18 percent enrichment. The additional terminal carbon of the side chain (δ 22.8) originates from the one carbon pool and was incorporated from 2-^{13}C-glycine (4%) or methionine-*methyl*-^{13}C (8%).

Methoxy Glycine. 1-^{13}C-Serine or 1-^{13}C-Glycine were incorporated into the methoxy glycine residue at C-37 (δ 167.5, 3% and 5% enrichment, respectively). 2-^{13}C-glycine was incorporated into the endocyclic methoxy glycine residue at

(19) Stauffer, G. V., Regulation of Serine, Glycine, and One-Carbon Biosynthesis. In *Amino Acids: Biosynthesis and Genetic Regulation;* Herrmann, K. M., Somerville, R. L., Eds.; Addison-Wesley Publishing Co.: Reading, MA, 1983; pp 103−113.

position C-39 (δ 77.1, 8%) and in the methoxyl group (δ 55.4, 5%). Position C-39 was also labeled by methionine-*methyl*-^{13}C (7%). This indicates that C-2 of glycine enters the one carbon pool and is available for incorporation as a single carbon unit.

Experimental Section

Mass Spectrometry. Electrospray-MS data were collected and processed on a Finnigan TSQ-70 triple stage quadrupole mass spectrometer. The instrument was tuned by extrapolation of an FC-43 generated electron impact tune table to electrospray by flow injection of conceral (1153.3 average MW). The sample was dissolved in methanol:water 1:1 and flow infused into the source at a rate of 3 μL per min. The interface was operated at a needle voltage of 3500 V and a bias voltage of 50 V. The probe, collar, and ion source temperatures were set to 210 °C, 45 °C and 200 °C, respectively. The mass spectrometer was programmed to pass ions through Q1 and Q2 while scanning Q3 from 250 to 2000 amu in 2 s. The gain was set to 10^{-8}, the merge width to 200%, the fragment width to 90%, the minimum peak width to 150%, the minimum area to 500, the zero to −8, the electron multiplier to 1350V, and the dynodes to −15 kV. The resolution was opened, and the calibration adjusted, such that the observed ion corresponded to the average molecular weight of the standard.

Nuclear Magnetic Resonance. NMR spectral data were recorded on a Bruker AMX500 operating at 500.13 MHz for the observation of ^{1}H and 125.77 MHz for the observation of ^{13}C. Samples were dissolved and internally referenced to DMSO-d_6 (^{1}H DMSO-d_5, $\delta = 2.49$; ^{13}C DMSO-d_6, $\delta = 39.5$). One-dimensional NMR data were recorded as a 32k complex point data table with a 10 600 Hz sweep width for proton and 20 800 Hz sweep width for carbon. The number of transients acquired was dependent upon the concentration of the sample. ^{1}H experiments were processed with a 0.3 Hz exponential multiplication and ^{13}C with 2 Hz.

A number of two-dimensional NMR experiments were performed to establish scalar and dipolar correlations (COSY, NOESY, HETCOR, LRCOSY, and LRHETCOR). The proton experiments (COSY, NOESY, and LRCOSY) were usually acquired with 1024 data points in t_2 and 256 points in t_1. The sweep width selected encompassed approximately 0.5 ppm beyond the most high field and low field resonance. The t_1 dimension was zero-filled to 512 points and the 2-D data table processed with $\pi/32$ (t_2) and $\pi/16$ (t_1) shifted sinebell squared apodization factors. The carbon experiments (HETCOR and LRHETCOR) were acquired with 2048 data points in t_2 and 256 points in t_1. The t_1 dimension was zero-filled to 512 points, and the 2-D table was processed with $\pi/5$ (t_2) and $\pi/3$ (t_1) shifted sinebell squared apodization factors for HETCOR and $\pi/2$ in both the t_1 and t_2 dimensions for the LRHETCOR experiment. The number of transients acquired in the 2-D experiments varied with sample concentration.

Percents of incorporation were determined by comparing the integrated peak intensities of resonances in the labeled spectrum with resonances in a natural abundance spectrum. The methyl resonances of oxazoles B and C fortuitously overlapped and were used as an internal reference according to the formula below where P_l is labeled peak intensity, P_n is the intensity of the peak in the natural abundance spectrum, R_l is the intensity of the oxazole B/C methyl in the labeled spectrum, and R_n is the intensity of the oxazole B/C methyl in the natural abundance spectrum.

$$\text{Enrichment} = \frac{P_l}{P_n} \times \frac{R_n}{R_l}$$

Fermentation. U-102408 is produced when *S. arginensis* Deitz sp. NRRL-15941 is grown in an aqueous nutrient medium under submerged aerobic conditions. Typically the microorganism is grown in a nutrient medium containing a carbon source and an assimilable nitrogen compound or proteinaceous material. Initially a primary seed culture is prepared. Conditions for fermentation of the seed medium culture have been described previously.[17] Subsequently, a defined medium was used for labeling experiments which per liter of medium consisted of the following: sonic yeast (0.075 g), magnesium sulfate (1.0 g), monopotassium phosphate (0.05 g), yeast extract (3.0 g), and mineral solution (40 mL). Mineral solution consisted of ferrous sulfate (0.075 g), cobalt chloride (0.06 g), manganese sulfate (0.06 g), and zinc sulfate (0.065 g) diluted to 200 mL with distilled water and filter sterilized (0.2 μ). The appropriate amount of water is added to the media and the pH is adjusted to 6.5 with sodium hydroxide. The medium was aliquoted into shaker flasks, capped, and sterilized by autoclave. After cooling sterile 50% fructose w/v is added at 150 mL/L medium along with any filter sterilized amino acid solutions to be added to the labeling experiment. Amino acids were added at a rate of 0.3 g/L (added as 30g/L sterile solutions in deionized water). Unlabeled cysteine was added to the medium at 0.3 g/L in all growths unless labeled cysteine was to be added. Typically, fermentations were incubated for 7 days at 30 °C on a New Brunswick shaker with a 3/4″ throw. Solutions were sterilized by filtration through a 0.2 μ filter. The medium is then inoculated with the primary seed to give a 2.5% v/v inoculum.

Stable isotope labeled amino acids were obtained from Cambridge Isotope Laboratories: L-threonine (1-^{13}C, 99%), L-serine (1-^{13}C, 99%), L-serine (2-^{13}C, 99%), L-serine (3,3-D_2, 98%), L-glycine (1-^{13}C, 99%), L-glycine (2-^{13}C, 99%), L-methionine (methyl-^{13}C, 98%). The following radiolabeled amino acids were obtained from Amersham: L-[3-^{3}H]serine (20−40 Ci/mmol), D-[U-^{14}C]glucose (230−370 mCi/mmol), 1-^{14}C-acetate (50−62 mCi/mmol). 1-^{14}C-propionate (40−60 mCi/mmol) was obtained from New England Nuclear. Radiolabeled compounds were added at a rate of 1 μCi/mL fermentation medium. Sonic yeast was obtained from N.P.C. Inc. and yeast extract from Difco, other media components were reagent grade materials purchased from Aldrich.

Isolation of 1 (U-102408). Typically the product was isolated by washing the beer with 10 volumes of methanol or ethyl acetate followed by concentration and crystallization of the crude product. Crystals were precipitated from methanol by addition of one volume of water and out of ethyl acetate with hexane. For large scale purification needed for the initial structure elucidation studies the product was then purified by reverse phase chromatography on a Zorbax C8 column (21 × 250 mm). The column, typically stored in 80% aqueous acetonitrile (ACN), was eluted with 5 mL of dimethyl formamamide and mobile phase 1:1 and then equilibrated with mobile phase, water:tetrahydrofuran:ACN (60:27:3), using a Varian Vista 5000 liquid chromatograph. The pump was set to 8 mL/min and the detector to 248 nm.

Crude U-102408 was dissolved in DMF (50 mg/mL) and centrifuged at 1000G for 5 min, and 200 μL supernatant was injected per chromatographic run. U-102408 was collected, and fractions pooled and stored at −20 °C until all of the material was chromatographed. The pooled material was then concentrated to an aqueous solution, and the purified material was extracted into methylene chloride. The organic extract was evaporated, the residue dissolved in about 10 mL *tert*-butyl alcohol and lyophillized. For the relatively small scale isolation of isotopically labeled U-102408 the concentrated organic beer extract was subsequently applied to a varian C18 solid phase extraction column and **1** was eluted with methanol.

Discussion

There are a wide variety of structures in the thiopeptide family of antibiotics which is composed of at least ten members. This diversity arises from different arrangement and repetition of several common structural motifs including a pyridine or dihydropyridine ring, dehydroalanine residues, and thiazole or oxazole rings incorporated into a macrocycle. The biosynthetic origin of each of these motifs for **1**, may be compared with data available for other compounds.

The incorporation of serine into the dehydroalanine residues has been observed in the related thiopeptide antibiotics thiostrepton, nosiheptide, and berninamycin. The dehydroalanine residues of thiostrepton and are enriched by 1- or 3-^{13}C-serine and 3-deuteroserine.[14] Similar results were obtained for nosiheptide and berninamycin.[12,14] The incorporation of two molecules of glycine into the dehydroalanines as observed for **1** was not investigated in either of these other systems. The pathway for the interconversion of glycine and serine appears to be more prevalent for **1** than thiostrepton, berninamycin, or nosiheptide. Cross labeling of serine and glycine sites is

observed for **1** and label in the three position of serine is not retained in **1**, while incorporation of serine labeled in the side chain was observed for the other antibiotics.

The thiazole rings in the macrocycle were found to originate from the condensation of the side chain of cysteine with the carboxyl group of the adjacent residue. For **1** the carboxyl originated from serine or glycine. These two precursors were apparently interchangeable due to their facile interconversion by the organism. Thiazole rings present in thiostrepton and nosiheptide also originated from the dimerization of the side chain of cysteine with the carboxyl of serine. The ability of the organism to use glycine as a precursor for the serine derived carboxyl positions was not determined. However, it was observed for berninamycin, thiostrepton, and nosiheptide that ^{13}C label from serine was also incorporated into the positions labeled by cysteine. This suggests that in *S. azureus* and *S. laurentii* which produce thiostrepton and *S. bernensis* and *S. actuosus* which produce berninamycin and nosiheptide, respectively, the interconversion of serine and cysteine is an active metabolic pathway, while for *S. arginensis* the pathway is less prominent under the fermentation conditions employed. For *S. arginensis* cysteine was added as a medium component in order to improve the antibiotic titer of U-102408. Improved titer upon cysteine addition may be due to the lower interconversion of serine and cysteine in *S. arginensis*.

In contrast, the interconversion of threonine and glycine to serine appears to be more facile in *Streptomyces arginensis* than in the other *Streptomycetes*: U-102408 is the only thiopeptide for which cross labeling of serine sites by threonine has been observed. Incorporation of threonine into the sites labeled by serine and glycine occurred with varying levels of enrichment from run to run, while sites uniquely labeled by threonine were enriched reproducibly. This may indicate run to run differences in the metabolism of threonine by threonine aldolase.

The pyridine ring of **1** originates from the condensation of three molecules of serine. Two molecules in a tail-to-tail manner condense with the carboxyl carbon of the third serine. Similar labeling patterns have been observed for the pyridine ring of berninamycin, the tetrahydropyridine ring of thiostrepton, and the hydroxypyridine ring of nosiheptide. This common labeling pattern suggests that the cyclization proceeds through a common pathway. A mechanism for the condensation has been proposed.[13] However, for thiostrepton and nosiheptide the carboxyl of the third residue may be from either a serine or a cysteine precursor. This is consistent with the conversion of serine to cysteine indicated by the pattern of incorporation into the thiazole rings. However, the analogous incorporation of the carboxyl carbon of cysteine into the pyridine ring of berninamycin was not observed.

Similar to the dehydroalanine residues which originate from serine, the butyrine moiety in U-102408 arises from the dehydration, in this case form the precursor threonine. Among the amino acids explored, threonine was the only precursor from which label was incorporated. Thiostrepton, berninamycin, and nosiheptide each contain one butyrine group which also originated from threonine.

In *S. arginensis* the methylene group of glycine rapidly enters the one carbon pool and is a source of the methylating equivalents for methionine. This is evidenced by the facile incorporation of ^{13}C from the two position of glycine into the methoxyl of the methoxy glycine residue. The percent label incorporation into this position from glycine is about 70% of that from methionine-*methyl*-^{13}C.

The 2-amino-4-hydroxy-2-pentenoic acid present in the macrocycle of **1** originates from threonine and an additional terminal carbon appended from the one carbon pool. Precursors feeding into the one carbon pool which gave incorporation were 2-^{13}C-glycine and methionine-*methyl*-^{13}C. In comparison, the unusual amino acid thiostreptone in the macrocycle of thiostrepton is derived from either isoleucine or threonine. With threonine as a precursor an additional one carbon unit is required to give rise to the thiostreptone sidechain. The origin of the terminal carbon in thiostreptone was not investigated but if similar to **1** could be envisioned to originate from glycine or methionine-*methyl*-^{13}C through the one carbon pool.

Conclusions

The biosynthetic origin of all the structural components of **1** has been elucidated. By comparison of these results with the incorporation of labeled precursors into other macrocyclic antibiotics several metabolic similarities, and differences were noted. Specifically differences are evident in the metabolic interconversion of threonine, glycine, and serine, and serine and cysteine among the various *Streptomycetes*. Threonine was found to be a biosynthetic precursor to the unusual 2-amino-4-hydroxy-2-pentenoic acid moiety present in the macrocycle of **1**. The terminal carbon of this residue originated from the one carbon pool through glycine or methionine.

Supporting Information Available: Figures 2−8 contain ^{1}H, COSY, long-range COSY, NOESY, long-range HETCOR, ^{13}C and HETCOR NMR spectra of U-102408; Figures 9 and 10 contain electrospray and MS/MS daughter ion spectra for U-102408; Figure 11−13 contain the ^{13}C NMR spectra of U-102408 labeled with 3-^{13}C-cysteine or 1-^{13}C-serine (Figure 11), 1-^{13}C-glycine or 2-^{13}C-glycine (Figure 12), and 1-^{13}C-threonine (Figure 13) (12 pages). See any current masthead page for ordering and Internet instructions.

JA961864G