Pergamon

Tetrahedron Letters 43 (2002) 2351–2354

TETRAHEDRON LETTERS

Integrastatins: structure and HIV-1 integrase inhibitory activities of two novel racemic tetracyclic aromatic heterocycles produced by two fungal species

Sheo B. Singh,[a,*] Deborah L. Zink,[a] Donette S. Quamina,[a] Fernando Pelaez,[b] Ana Teran,[b] Peter Felock[c] and Daria J. Hazuda[c]

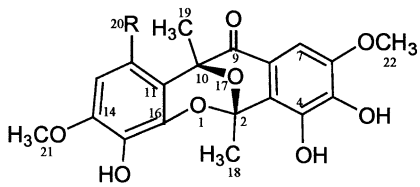
[a]*Merck Research Laboratories, RY80Y-355, PO Box 2000, Rahway, NJ 07065, USA*
[b]*CIBE, Merck Sharp & Dohme de Espana, S. A. Josefa Valcárcel 38, 28027 Madrid, Spain*
[c]*Merck Research Laboratories, West Point, PA 19486, USA*

Received 10 January 2002; revised 3 February 2002; accepted 6 February 2002

Abstract—Integrastatins are two novel aromatic natural products derived from fungal fermentations who possess a novel [6/6/6/6]-ring system and are racemic despite having two asymmetric centers. These compounds inhibited the strand transfer reaction of HIV-1 integrase with IC_{50} values of 1.1–2.5 μM. © 2002 Published by Elsevier Science Ltd.

Integration is an essential event in the life cycle of retroviruses including HIV and is a three-step process that includes assembly of a proviral DNA on integrase, endonucleolytic processing of the proviral DNA, and strand transfer of the proviral DNA into the host cell DNA.[1] This is a unique process by which the virus proliferates and the entire event is catalyzed by a single viral enzyme, HIV-1 integrase. This process as well as the enzyme is absent in the host, therefore, its intervention presents a safe target for development of a novel anti-HIV therapy that can be used in combination with existing (protease and reverse transcriptase inhibitor) therapies. The recent discovery of diketo acid (DKA) based inhibitors of integrase that showed anti-viral inhibitory potency comparable to clinically effective HIV protease inhibitors has validated the approach and the likelihood of developing integrase inhibitors as useful therapeutic agents.[2] Even though DKAs are valid small molecule chemical leads, which could potentially be refined into clinical agents, the unpredictable nature of the drug development process makes it important to identify alternative structures and provide structurally diverse scaffolds.

Natural product extracts have historically been used for the discovery of leads for a variety of biological targets. Screening of such extracts against recombinant HIV-1 integrase led us to the discovery of several novel natural product inhibitors including equisetin,[3] integric acid[4] and complestatin.[5] Continued screening of fungal extracts has more recently led to the discovery of two racemic compounds named herein integrastatin A (**1a**) and B (**1b**). These tetracyclic aromatic [6/6/6/6/]-heterocycles of polyketide origin inhibit the strand transfer reaction of recombinant HIV-1 integrase with an IC_{50} value of 1.1 and 2.5 μM, respectively. The bioassay guided isolation, structure elucidation, and the biological activities of integrastatins A (**1a**) and B (**1b**) are herein described.



1a: R = CH_2OH (Integrastatin A)
1b: R = CHO (Integrastatin B)

A sterile unidentified fungus (ATCC74478) isolated from herbivore dung collected in New Mexico was grown on a brown rice-based liquid medium and was extracted with 1.2 volumes of methyl ethyl ketone which was concentrated and dissolved in a 1:6 ratio of methanol–water that was washed with hexane and the activity extracted with ethyl acetate. Size exclusion (Sephadex LH 20) chromatography of the latter extract followed by reversed-phase HPLC (Zorbax RX C-8)

* Corresponding author. Fax: (732) 594-6880; e-mail: sheo_singh@merck.com

0040-4039/02/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0040-4039(02)00265-4

afforded integrastatin A (**1a**, 40 mg/L) and B (**1b**, 70 mg/L) as brown powders. Integrastatin B (**1b**) was also isolated from an endophytic *Ascochyta* sp. (ATCC74477) isolated from leaves of *Urtica urens* collected in Ontígola, near Madrid, Spain. Both of these compounds were optically inactive and likewise did not exhibit absorption bands in CD spectra.[6]

Integrastatin A (1a): high-resolution EIMS analysis of integrastatin A (**1a**) provided a molecular formula of $C_{20}H_{20}O_9$ (found m/z 404.1110, calcd m/z 404.1107) that was corroborated by the ^{13}C NMR spectrum (Table 1) and indicated that it has 11 degrees of unsaturation. The ^{13}C NMR spectrum of **1a** revealed the presence of two methyls, two aromatic methoxy groups, an oxy-methylene, two aromatic methines, two quaternary carbons each connected with one or two oxygen atoms, 10 non-protonated aromatic carbons including six attached to oxygen atoms, and a conjugated keto group. The UV spectrum showed absorption bands at λ_{max} 212 (log ε=4.4), 262 (3.83) and 310 (sh) nm. The IR spectrum of **1a** showed absorption bands characteristic of hydroxyl (3394 cm^{-1}), conjugated ketone (1776 cm^{-1}) and aromatic (1610 cm^{-1}) groups. The ^1H NMR spectrum of **1a** was rather simple and showed two shielded singlets for aromatic protons (δ 6.70 and 7.03), singlets for two methoxy groups (δ 3.75 and 3.85), two downfield shifted singlets for methyl groups (δ 2.09 and 1.76) and an AB doublet (J=12.8 Hz) of an oxy-methylene group (δ 4.40 and 4.58). The ^1H NMR shifts were assigned to respective ^{13}C NMR shifts by an HMQC experiment. The elucidation of the structure of integrastatin A was severely hampered by the presence of a disproportionally high numbers of oxygen atoms, 13 of 20 quaternary carbons, and only a limited number of proton bearing carbons. However, HMBC experiments were very helpful in defining the substitution pattern of the two aromatic rings and deducing the structure of **1a**. The aromatic singlet H-13 (δ 6.70) produced strong three-bond HMBC correlations (Fig. 1) to C-11, C-15, C-20 and a weak two-bond correlation to C-14, which showed an HMBC correlation with one of the methoxy groups (δ 3.75). The methylene protons H-20 displayed HMBC correlations to C-12 and C-13. The substitution pattern in the left aromatic ring was deduced on the basis of these correlations, and, strongly supported by the calculated ^{13}C chemical shifts of the respective carbons. Similar two and three-bond HMBC correlations from the other aromatic proton H-7 (δ 7.03) to C-3, C-5, C-6 and from methoxy group (δ 3.83) to C-6 established the substitution pattern of the other aromatic ring, and was supported by NOEDS experiments of congener **1b** (vide infra). In addition, H-7 showed an HMBC correlation to the C-9 keto group which was also correlated to the methyl group H_3-19 (δ 1.76). This methyl group also exhibited HMBC correlations to C-10 and C-11, and enabled connection of C-8 and C-11 of the two aromatic rings through a two-carbon bridge. The remaining methyl group H_3-18 showed HMBC correlations to doubly oxygenated carbon C-2 and aromatic carbon C-3 thus placing a phenethyl type substitution at C-3 of the other aromatic ring. The pyranone and the 1,3-dioxene rings were assembled to meet the requirement of the remaining two degrees of unsaturation and to fulfil the ^{13}C chemical shift requirement of C-2. Based on these data a novel [6/6/6/6]-tetracyclic heterocycle structure **1a** was established for integrastatin A.

Table 1. ^1H and ^{13}C NMR spectral assignments of integrastatin A (**1a** in CD_3CN) and B (**1b** in 1:1 CD_3CN+$CDCl_3$)

Position	δC (**1a**)	δH (**1a**)	δC (**1b**)	δH (**1b**)
2	97.8		97.8	
3	121.3		120.7	
4	140.1[a]		140.1[a]	
5	140.4		142.2	
6	148.6		148.7	
7	101.4	7.03, s	101.8	7.06, s
8	121.1		120.4	
9	193.8		193.7	
10	77.4		77.1	
11	114.1		121.3	
12	130.5		126.0	
13	107.3	6.70, s	105.7	7.14, s
14	147.7		147.8	
15	134.4		140.9	
16	142.0[a]		140.8[a]	
18	26.4	2.09, s	26.5	2.15, s
19	22.2	1.76, s	25.8	1.87, s
20	60.9	4.58, d, 12.8 4.40, d, 12.8	190.6	10.21, s
21	56.4	3.75, s	56.6	3.75, s
22	56.6	3.83, s	56.7	3.85, s

[a] Assignments in the columns could be interchanged.

Integrastatin B (1b): The high-resolution EIMS of **1b** gave a molecular formula of $C_{20}H_{18}O_9$ (found 402.0960, calcd 402.0950) and indicated that it was a dehydro derivative of **1a**. Comparison of the ^1H NMR spectrum of **1b** with that of **1a** (Table 1) showed the absence of the oxy-methylene protons and the presence of an aldehyde group (δ 10.21) which was corroborated by the corresponding signals in the ^{13}C NMR spectrum. The structural assignment was supported by analogous HMBC correlations and by NOEDS experiments (Fig. 2). This was further supported by the bathochromic shift and hyperchromic effect of the second absorbance band that appeared at λ_{max} 315 (log ε=4.09) nm in the

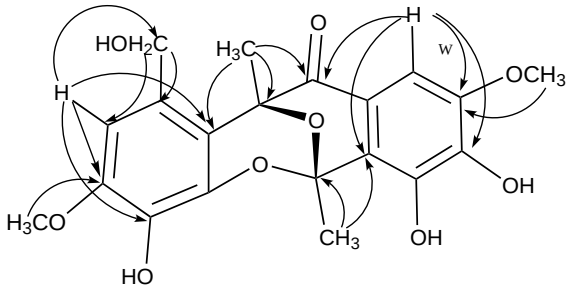


Figure 1. HMBC ($^nJ_{CH}$=7 Hz) correlations of **1a** in CD_3CN.

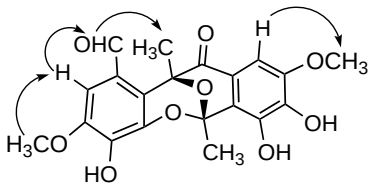
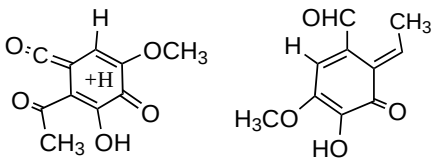


Figure 2. NOEDS of **1b** in $CDCl_3$–CD_3CN.

UV spectrum of **1b** due to additional conjugation caused by the aldehyde group. The EI mass spectrum of both **1a** and **1b** produced two major fragment ions A and B observed at *m*/*z* 209 (97%) and *m*/*z* 194 (100%) (Fig. 3). The EIMS spectrum of **1a** produced expected ion A (85%) and surprisingly produced unexpected ions B (50%) and another ion at *m*/*z* 195 (50%) instead of an expected ion at *m*/*z* 196 (equivalent of B ion). Presumably, the ions at *m*/*z* 195 and 194 in **1a** appear as a result of either further rearrangements or simply by a loss of one and two hydrogens, respectively from unobserved ion *m*/*z* 196. These fragment ions originate from the concomitant



A *m/z* 209.0450 $C_{10}H_9O_5$

B *m/z* 194.0576 $C_{10}H_{10}O_4$

Figure 3. EIMS fragment ions of **1a** and **1b**.

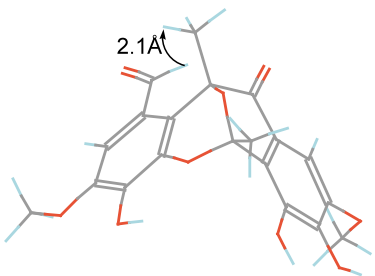


Figure 4. ChemDraw 3D model of **1b**.

cleavage of both heterocyclic rings. The structure of fragment ions were verified by HREIMS.

Relative stereochemistry: the relative stereochemistry and conformation of integrastatins **1a** and **1b** was deduced as *R*,*R* or *S*,*S* by ChemDraw 3D modelling (Fig. 4) and was supported by the analysis of different possibilities by the Dreiding model. These models indicated that the bridgehead oxygen-17 connected C-2 and C-10 through di-equatorial bonds to attain the lowest energy (12 kcal/mol) conformation and thus leading to the diaxial fusion of the pyranone ring producing a inverted-V-shaped (*R*,*R*) or a V-shaped (*S*,*S*) conformation. This would allow only the equatorial orientations of C-18 and C-19 methyl groups leading to a wing like structures. These structural conformations supported the inter-atomic distance (2.1 Å) and observed NOEs between the aldehyde proton and the H_3-19 of **1b**.

Biogenesis. Integrastatins are likely produced by traditional polyketide pathway originating from the condensation of nine acetate units as shown in Fig. 5. The aromatization, decarboxylation, methylation and oxidations of the putative intermediate would presumably result in the formation of integrastatins. The introduction of the C-19 methyl at the C-10 keto group could potentially precede the hetero-cyclization, which would result in the elimination of the C-10 *tert*-hydroxy group.

Biological activities. Integrastatins A (**1a**) and B (**1b**) were evaluated in the coupled and the strand transfer HIV-1 integrase assays using recombinant enzyme.[2a] Integrastatin A exhibited IC_{50} values of 0.6 and 1.1 μM in coupled and strand transfer assays, respectively. Integrastatin B (**1b**) was two-fold less active and exhibited corresponding IC_{50} values of 1.04 and 2.5 μM. These compounds were about 5- to 10-fold selective for HIV-1 integrase when compared with DNAase (IC_{50} = 12 μM).

In summary, we have described the discovery and structure elucidation of integrastatins A and B possessing a novel [6/6/6/6] tetracyclic heterocyclic skeleton that are potent inhibitors of HIV-1 integrase. These two microbial natural products each contain two chiral centers, but occur in the racemic forms in nature, which is a rare occurrence.

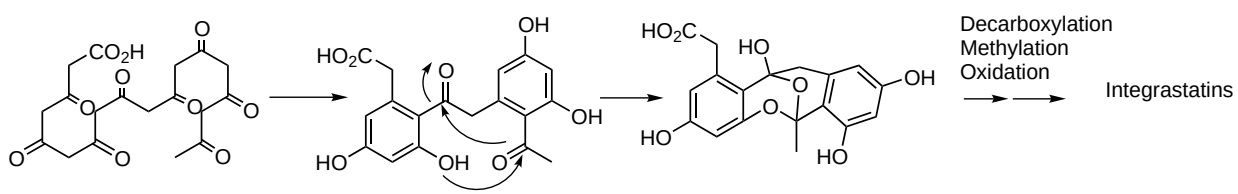


Figure 5. Proposed biogenesis of integrastatin B (**1b**).

References

1. For recent reviews on HIV integrase, see: (a) Craigie, R. *J. Biol. Chem*. **2001**, *276*, 23213; (b) Esposito, D.; Craigie, R. *Adv. Virus Res*. **1999**, *52*, 319.
2. (a) Hazuda, D. J.; Felock, P.; Witmer, M.; Wolfe, A.; Stillmock, K.; Grobler, J. A.; Espeseth, A.; Gabryelski, L.; Schleif, W.; Blau, C.; Miller, M. D. *Science* **2000**, *287*, 646; (b) Wai, J. S.; Egbertson, M. S.; Payne, L. S.; Fisher, T. E.; Embrey, M. W.; Tran, L. O.; Melamed, J. Y.; Langford, H. M.; Guare, J. P.; Zhuang, L.; Grey, V. E.; Vacca, J. P.; Holloway, M. K.; Naylor-Olsen, A. M.; Hazuda, D. J.; Felock, P. J.; Wolfe, A. L.; Stillmock, K. A.; Schleif, W.; Gabryelski, L. J.; Young, S. D. *J. Med. Chem*. **2000**, *43*, 4923; (c) Neamati, N. *Exp. Opin. Invest. Drugs* **2001**, *10*, 281.
3. Singh, S. B.; Zink, D. L.; Goetz, M. A.; Dombrowski, A. W.; Polishook, J. D.; Hazuda, D. L. *Tetrahedron Lett*. **1998**, *39*, 2243.
4. (a) Singh, S. B.; Zink, D.; Polishook, J.; Valentino, D.; Shafiee, A.; Silverman, K.; Felock, P.; Teran, A.; Vilella, D.; Hazuda, D. J.; Lingham, R. B. *Tetrahedron Lett*. **1999**, *40*, 8775; (b) Singh, S. B.; Felock, P.; Hazuda, D. J. *Bioorg. Med. Chem. Lett*. **2000**, *10*, 235.
5. Singh, S. B.; Jayasuriya, H.; Salituro, G. M.; Zink, D. L.; Shafiee, A.; Heimbuch, B.; Silverman, K. C.; Lingham, R. B.; Genilloud, O.; Teran, A.; Vilella, D.; Felock, P.; Hazuda, D. *J. Nat. Prod*. **2001**, *64*, 874.
6. Compound **1a**: $[\alpha]_D^{23}=0°$ (*c*, 0.09, CH_3OH); **1b**: $[\alpha]_D^{23}=0°$ (*c*, 0.56, CH_3OH). The CD spectra of **1a** and **1b** were measured at 0.14–4.9 μM concentrations.