Phytochemistry, Vol 23, No 11, pp 2613–2616, 1984
Printed in Great Britain

0031–9422/84 $3 00 + 0 00

SAPOGENINS FROM *GUAIACUM OFFICINALE*

VIQAR UDDIN AHMAD, NASREEN BANO and SHAHEEN BANO

H E J Research Institute of Chemistry, University of Karachi, Karachi-32, Pakistan

(*Revised received* 26 *April* 1984)

Key Word Index—*Guaiacum officinale*, Zygophyllaceae, sapogenins, 3β-hydroxy 30-norolean-12,19-dien-28-oic acid, 3β,20ξ-dihydroxy-30-norolean-12-en-28-oic acid

Abstract—Acid hydrolysis of the saponins from the stem bark of *Guaiacum officinale* yielded the new sapogenin 3β,20ξ-dihydroxy-30-norolean-12-en-28-oic acid and the artifacts 3β-hydroxy-30-norolean-12,19-dien-28-oic acid and its methyl ester Larreagenin, sitosterol and oleanolic acid were also isolated

INTRODUCTION

Guaiacum officinale L is an evergreen tree which is cultivated for ornamental purposes Its wood is used as for the preparation of a diuretic diaphoretic, sudorific, sialagogue and a mouth wash [1]

Earlier investigations on the sapogenins of this plant have led to the isolation of a compound named as guaiagenin [2] which was later identified [3] as oleanolic acid

In this paper we wish to report the isolation of three new triterpenes (**1**, **2** and **3**) besides larreagenin A (**4**) from the acid hydrolysis product of the saponins of *G officinale* The last named sapogenin was first isolated from *Larrea divaricata* [4], a plant belonging to the same family (Zygophyllaceae) Oleanolic acid and sitosterol were also isolated

RESULTS AND DISCUSSION

The crude sapogenins obtained from the alcoholic extract of the bark of the plant were hydrolysed with hydrochloric acid in methanol–water The crude sapogenin mixture thus obtained was chromatographed on a silica gel column to yield the pure sapogenins

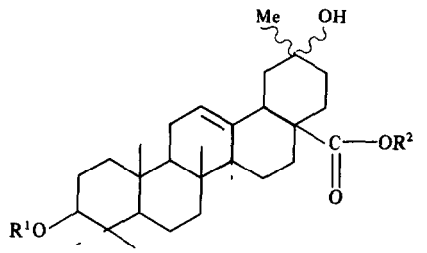
Hexane–benzene (65 35) eluted compound **3**, mp 220–222°, $[\alpha]_D$ +10 72° ($CHCl_3$) which had the formula $C_{30}H_{46}O_3$ Its UV spectrum in methanol showed only end absorption at 209 nm, thus excluding the possibility of conjugated double bonds The IR spectrum (KBr) revealed the presence of hydroxyl (3420 cm^{-1}) and ester (1720 cm^{-1}) groups and tri-substituted double bond(s) (829 cm^{-1}) in the compound The presence of a carbomethoxy group was also indicated by the ^1H NMR spectrum of the compound which showed a three proton singlet at δ3 63 Thus **3** appeared to be the methyl ester of a nortriterpenic acid The ^1H NMR spectrum also showed singlets for six tertiary methyl groups at δ0 77–1 60 The multiplet at 3 25 was assigned to H-3 and H-18 The spectrum showed two olefinic protons signals at 5 12 (*br s*) and 5 40 (*t*, J = 3 75 Hz), the latter was assigned to H-12 which is encountered in the triterpenes of the oleanene and ursene series

On acetylation with acetic anhydride and pyridine a monoacetate (**3a**) (mp 190–193°) was obtained, the IR spectrum of which showed no hydroxyl absorption The ^1H NMR spectrum revealed a double doublet (J_{aa} = 10 0 Hz, J_{ae} = 7 5 Hz) due to H-3 The spectrum also showed a broad singlet at δ3 32 which was assigned to H-18β The downfield shift of this signal as compared to methyl oleanolate (δ2 86) [5] is apparently due to its allylic position to the addition double bond at C-19 In the ^1H NMR spectrum of **3**, this signal was overlapped by the H-3 signal

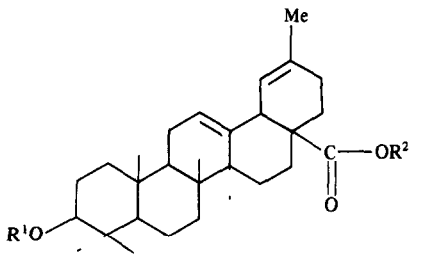
The mass spectra of **3** and **3a** lend support to the proposed structure Besides the molecular peaks at *m/z* 454 and 496 respectively, strong peaks due to retro-Diels–Alder fragmentation [6] at *m/z* 246 and 207 were observed in **3**, the former peak is also present in **3a** Both **3** and **3a** showed a base peak at *m/z* 233 ($C_{15}H_{21}O_2$) to which structure **6** is proposed Fragment **6** is apparently formed through a hydrogen transfer and rearrangement from the ion **5** The formation of ions of type **5** is commonly observed in Δ^{12}-oleanenes and ursene derivatives [6] The high intensity of this ion in **3** and **3a** is apparently due to the presence of a double bond at C-19, increasing the stability of the ion **6** The mass spectrum also showed that the double bond in ring E was not present between C-20 and C-21 because in that case retro-Diels–Alder fragmentation of this ring would give rise to an $[M-68]^+$ peak [6] which was not present in the spectra of **3** and **3a**

All of the spectroscopic data cited above lead to the structure methyl 3β-hydroxy-30-norolean-12,19-dien-28-oate for **3** This structure is also supported by the ^{13}C NMR spectral data (Table 1) of compound **3**, where assignments were made on the basis of known chemical shifts of the β-amyrin series of triterpenes [7] and by gated spin echo (GASPE) measurement after phase correction, when primary and tertiary carbon peaks were upright whereas secondary and quarternary peaks were inverted

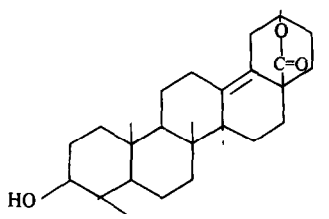
The mother liquors of **3** yielded another compound, mp 137°, $[M]^+$ at *m/z* 414, which was identified as sitosterol by direct comparison Benzene–chloroform (3 2) eluted a triterpene, mp 262° $[\alpha]_D$ −87 5° On the basis of IR, ^1H NMR, ^{13}C NMR and the mass spectra, this compound was identified as larreagenin A (**4**), which was first isolated by Habermehl and Moller from *Larrea divaricata* [4] According to these authors, larreagenin A is not an



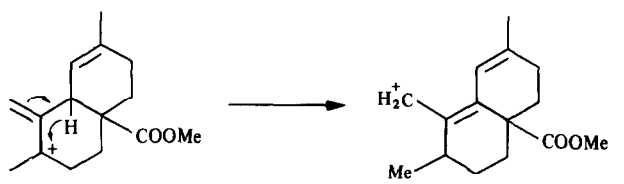
1 $R^1 = R^2 = H$
1a $R^1 = H$ $R^2 = Me$
1b $R^1 = Ac$ $R^2 = Me$



2 $R^1 = H$ $R^2 = H$
3 $R^1 = H$ $R^2 = Me$
3a $R^1 = Ac$ $R^2 = Me$



4



5 **6**

artefact From the fractions eluted with benzene–chloroform (3 7) oleanolic acid, mp 299°, was obtained and it was identified through direct comparison with an authentic sample This compound has already been reported from this plant as a sapogenin

Compound **2** was eluted from the column by chloroform–methanol (19 1) After recrystallization from methanol it was obtained as colourless crystals, mp 215–218°, $[\alpha]_D$ 10 54° (MeOH) The molecular formula was found to be $C_{29}H_{44}O_3$ through the high resolution measurements of the molecular peak On treatment of its methanolic solution with ethereal diazomethane, compound **2** was converted into a compound which was identical in all respects with **3** Therefore compound **2** has

Table 1 ^{13}C NMR chemical shifts of compounds **1a** and **3** (measured in $CDCl_3$ solution)

Carbon No	Chemical shift 1a	Chemical shift 3	Carbon No	Chemical shift 1a	Chemical shift 3
1	38 44*	38 77§	16	23 10	26 19
2	27 15	27 27	17	46 50	45 63‖
3	78 94	78 90	18	43 57	45 53‖
4	38 72*	38 47§	19	46 50	129 24¶
5	55 19	55 43	20	70 93	129 30¶
6	18 29	18 31	21	34 97†	32 40
7	34 03†	33 45	22	32 63	37 05
8	39 26	38 95§	23	28 08	28 14
9	47 55	47 68	24	15 27‡	15 68
10	37 00	37 05	25	15 54‡	15 68
11	23 36	23 36	26	16 82	17 12
12	122 98	122 74	27	25 20	23 10
13	142 64	142 32	28	177 65	177 97
14	41 56	42 72	29	25 80	23 10
15	27 56	27 52	OMe	51 66	51 57

*†‡§‖¶These values are interchangeable

the structure 3β-hydroxy-30-norolean-12,19-dien-28-oic acid It is isomeric to larreic acid [4] from which larreagenin is derived

Compound **1** was also eluted with chloroform–methanol (19 1) After further purification through rechromatography and crystallization, it melted at 234–236°, $[\alpha]_D$ 66 66° (*c* 0 6, MeOH) The mass spectrum showed a molecular ion peak at *m/z* 458, corresponding to the formula $C_{29}H_{46}O_4$ Thus compound **1** differs from **2** only by the elements of a molecule of water The IR spectrum (KBr) of **1** showed bands at 3440 (OH), 2920 (CH_2), 1695 (CO, acid) and 825 cm^{-1} (tri-substituted double bond) The UV spectrum had a maximum at 207 nm (end absorption) On treatment with diazomethane a monomethyl ester (**1a**) was formed The ^1H NMR spectrum of **1a** showed singlets due to six methyl groups at δ0 72–1 27, a multiplet centred at 3 20 (H-3α), a double doublet at 2 80 (*J* = 5 0, 13 75 Hz) attributed to H-18 and a triplet at 5 32 (*J* = 3 75 Hz, H-12) A singlet at 3 63 was due to COOMe The fact that this compound showed a signal for only one carbinylic proton indicated that the other hydroxyl group was tertiary The mass spectrum of **1** showed peaks due to the retro-Diels–Alder fragmentation at *m/z* 207 and 250, the latter losing water easily to yield the base peak at *m/z* 232 This indicated that the second tertiary hydroxyl group was located at C-20

Acetylation of **1a** with acetic anhydride–pyridine yielded the monoacetate (**1b**), $[M]^+$ at *m/z* 514, as the main product, mp 213–215° The IR spectrum ($CHCl_3$) of this compound showed a broad absorption band at 3467 cm^{-1} due to the tertiary hydroxyl group The ^1H NMR spectrum had peaks at δ3 60 (COOMe) and 2 04 (OAc)

On the basis of the above spectra, the structure **1** was suggested for the triterpenoid sapogenin It was supported by the ^{13}C NMR spectrum of **1a** (Table 1) which had signals for a double bond at δ122 98 (C-12) and 142 64 (C-13) which are close to the values in the Δ^{12}-oleanene series of triterpenoids There were signals for two carbon atoms bearing oxygen function at 78 94 (C-3) and 70 93 (C-20) The latter peak was inverted in the GASPE spectrum indicating that this carbon atom was quarternary and the hydroxyl group attached to this carbon atom was tertiary The absolute configuration at C-20 is not known

On refluxing compound **1** in methanolic hydrochloric acid, it was completely converted into a mixture of **2** and **3** It appears therefore that compound **1** represents the genuine sapogenin of *G officinale* whereas **2** and **3** are artefacts formed during acid catalysed hydrolysis of the saponin Larreagenin A (**4**) is not formed in this reaction and therefore cannot be considered as an artefact derived from **1**

EXPERIMENTAL

Melting points were determined on a Gallenkamp apparatus and are uncorr ^1H NMR (100 MHz) and ^{13}C NMR (25 15 MHz) spectra were recorded on a Bruker-FT-WP-100 SY spectrometer in $CDCl_3$ using tetramethylsilane as an internal standard TLC was performed on silica gel plates using the following solvent systems (a) C_6H_6–EtOAc (9 1), (b) $CHCl_3$–MeOH (9 5 0 5), (c) $CHCl_3$–MeOH (9 1) Spots were detected by spraying with ceric sulphate soln in 10% H_2SO_4 followed by heating

Extraction and separation The powdered stem bark of *G officinale* was extracted × 3 with cold MeOH The combined methanolic extracts were evaporated at red pres to afford a gummy residue This residue was partitioned between EtOAc and H_2O The EtOAc layer was separated and the aq layer (containing saponins) was extracted with *n*-BuOH The *n*-BuOH layer was separated and evaporated under red pres to afford crude saponins

Acid hydrolysis of saponins The crude saponins (100 g) were refluxed with methanolic HCl (900 ml MeOH, 100 ml H_2O, 150 ml conc HCl) for 5 hr The MeOH from this reaction mixture was evaporated and H_2O added, whereby a crude ppt of sapogenins was obtained It was filtered and dried to yield 30 g sapogenins This sapogenin mixture was chromatographed on a silica gel column (1 5 kg) which was successively eluted with solvent gradients of increasing polarity in the order of hexane, C_6H_6, $CHCl_3$ and MeOH As a result of this chromatography the following compounds were isolated

Methyl-3β-hydroxy-30-norolean-12,19-dien-28-oate (**3**) Elution with 35% C_6H_6 in hexane afforded a mixture of compound **3** and sitosterol, which on repeated recrystallization from MeOH yielded colourless shining crystals of **3**, mp 220–222°, $[\alpha]_D$ 10 72° (*c* 0 19, $CHCl_3$) UV λ_{max}^{MeOH} 209 nm IR ν_{max}^{KBr} cm^{-1} 3420 (OH), 2940 (C–H), 1720 (ester), 1440, 1385, 1245, 1205, 1112, 1036, 1000, 829 (trisubstituted double bond), 754 ^1H NMR ($CDCl_3$) δ0 77, 0 79, 0 91, 0 96, 1 01, 1 60 (each 3H, *s*, 6 × Me), 3 25 (2H, *m*, H-3 and H-18), 3 63 (3H, *s*, COOMe), 5 12 (1H, *br s*, H-19), 5 4 (1H, *t*, *J* = 3 75 Hz, H-12), ^{13}C NMR (see Table 1) MS *m/z* (rel int) 454 344 $[M]^+$ (Calc for $C_{30}H_{46}O_3$ 454 344674) (36), 439 $[M-Me]^+$ (2), 436 $[M-H_2O]^+$ (2), 395 $[M-COOMe]^+$ (8), 377 $[M-H_2O-COOMe]^+$ (2), 301 (2), 246 16146 [fr *a*]$^+$ (Calc for $C_{16}H_{22}O_2$, 246 161971) (50), 233 15367 [fr **6**]$^+$ (Calc for $C_{15}H_{21}O_2$, 233 154146) (100), 219 (10), 207 [fr *b*]$^+$ (24), 187 14789 [fr *a*−COOMe]$^+$ (Calc for $C_{14}H_{19}$) 187 148668, (62), 173 13278 (Calc for $C_{13}H_{17}$, 173 133019) (68), 160 (17), 145 (16), 130 (18), 117 (20), 105 (30), 91 (22), 79 (28), 69 (29), 55 (36)

Acetylation of compound **3** Compound **3** (20 mg) was treated with Ac_2O (10 ml) in pyridine (1 ml) at room temp overnight Ice was added to the reaction mixture and a white ppt was obtained

The ppt was filtered, washed with H_2O and dried It was crystallized from MeOH yielding colourless needles of monoacetate ester (**3a**), mp 190–193°, $[\alpha]_D$ −10 06° (*c* 0 2, $CHCl_3$) IR ν_{max}^{KBr} cm^{-1} 2940, 1729, 1440, 1365, 1243, 829 ^1H NMR ($CDCl_3$) δ0 80 (3H, *s*, Me), 0 85 (6H, *s*, 2 × Me), 0 96, 1 02, 1 60 (each 3H, *s*, 3 × Me), 2 04 (3H, *s*, OAc), 3 2 (1H, *br s*, H-18), 3 63 (3H, *s*, COOMe), 4 50 (1H, *dd*, J_{ae} = 10 0 Hz, J_{ae} = 7 5 Hz, H-3), 5 13 (1H, *br s*, H-19), 5 40 (1H, *t*, *J* = 3 43 Hz, H-12) MS *m/z* (rel int) 496 $[M]^+$ (32), 454 (1), 436 (33), 421 (4), 393 (2), 377 (8), 367 (4), 299 (2), 257 (14), 246 (56, fr *a*), 233 (100, fr **6**), 218 (18), 201 (26), 190 (58), 187 (58), 173 (60), 160 (14), 149 (18), 130 (18), 119 (20), 105 (20), 95 (28), 89 (38), 69 (57), 57 (95)

Sitosterol This compound was obtained from the mother liquors of **3** and it was purified by repeated recrystallization from Me_2CO as colourless needles, mp 137°, MS *m/z* 414 It was identical (TLC, IR, mmp) with an authentic sample of sitosterol (E Merck, Darmstadt) and purified further through repeated crystallizations

Larreagenin A (**4**) The substance eluted with C_6H_6–$CHCl_3$ (3 2) was purified by repeated recrystallization from MeOH, whereby fine needles were obtained, mp 262°, $[\alpha]_D$ −87 5° (*c* 0 16, $CHCl_3$) This compound was identified as larreagenin A through direct comparison of its spectra with those of larreagenin A [4] ^{13}C NMR ($CDCl_3$) δ15 21, 15 35, 16 07, 17 73, 25 09, 27 90 (methyls), 18 06, 20 51, 24 94, 25 74, 26 71, 27 02, 31 01, 32 72, 34 59, 37 29, 38 60 (methylenes), 50 10 (C-9), 55 08 (C-5), 78 50 (C-3) (methynes), 37 12, 38 68, 40 81, 41 90, 42 79 (C-17), 80 81 (C-20), 123 29 (C-13), 139 51 (C-18), 175 78 (C-28) (quarternary carbons) MS *m/z* (rel int) 440 $[M]^+$ (16), 438 (20), 422 (42), 407 (8), 395 (36), 379 (8), 369 (6), 327 (2), 287 (8), 245 (14), 232 (70), 220 (54), 219 (48), 207 (52), 203 (88), 190 (100), 175 (82), 159 (46), 119 (68), 105 (84), 95 (59), 81 (73), 73 (76), 54 (90)

Oleanolic acid From the C_6H_6–$CHCl_3$ (3 7) eluate, oleanolic acid was obtained and purified by repeated recrystallization from MeOH, mp 299° MS *m/z* 454 $[M]^+$ The IR, ^1H NMR and mass spectra were identical with those of an authentic sample

3β-Hydroxy-30-norolean-12,19-dien-28-oic acid (**2**) Elution with $CHCl_3$–MeOH (19 1) gave a mixture of **2** and **1** which was rechromatographed on silica gel (60 PF_{254} + 366) with a gradient of hexane–EtOAc (EtOAc 10–20%) yielding **2** which was further purified by recrystallization from MeOH as microcrystalline powder, mp 215–218°, $[\alpha]_D$ 10 54° (*c* 0 2, MeOH) MS *m/z* 440 Treatment of its methanolic soln with CH_2N_2 in Et_2O yielded the methyl ester, mp 220–222°, which was identical with compound **3** by TLC, IR, ^1H NMR and mass spectra

3β-20ξ-Dihydroxy-30-norolean-12-en-28-oic acid (**1**) Compound **1** was eluted from the above column with hexane–EtOAc (3 1) and purified by recrystallization from MeOH as a microcrystalline powder, mp 234–236°, $[\alpha]_D$ 66 66° (*c* 0 6, MeOH) UV λ_{max}^{MeOH} nm 207 IR ν_{max}^{KBr} cm^{-1} 3440 (OH), 2920 (C–H), 2870, 1695 (CO, acid), 1468, 1452, 1386, 1308, 1295, 1205, 1120, 1026, 998, 950, 825 (trisubstituted double bond), 742 MS *m/z* (rel int) 458 $[M]^+$ (2), 440 $[M-H_2O]^+$ (12), 425 $[M-H_2O-Me]^+$ (2), 407 (1), 286 (1), 264 (7), 250 [RDA fr *a*]$^+$ (24), 232 $[a\text{-}H_2O]^+$ (100), 219 (6), 207 [RDA fr *b*]$^+$ (32), 187 [fr *a* − H_2O − COOH]$^+$ (84), 173 (15), 160 (10), 140 (16), 132 (16), 119 (18), 107 (18), 95 (18), 80 (19), 69 (23), 55 (22)

Methylation of **1** A soln of **1** (60 mg) in MeOH was treated with an ethereal solution of CH_2N_2 at room temp for 10 min After evaporation of the solvent from reaction mixture, the methyl ester **1a** was crystallized from MeOH as an amorphous powder (55 mg), mp 243–245°, $[\alpha]_D$ 50 0 (*c* 0 2, $CHCl_3$) IR ν_{max}^{KBr} cm^{-1} 3440 (OH), 2940 (C–H), 1705 (COOMe), 1455, 1205, 1112, 1028, ^1H NMR ($CDCl_3$) δ0 72, 0 78, 0 90, 0 98, 1 14, 1 27 (each 3H, *s*, 6 × Me), 2 80 (1H, *dd*, *J* = 5 0, 13 75 Hz, H-18), 3 20 (1H, *m*, H-3), 3 63 (3H, *s*, COOMe), 5 32 (1H, *t*, *J* = 3 75 Hz, H-12) MS *m/z* (rel int) 472 $[M]^+$ (2), 454 $[M-H_2O]^+$ (6), 439 (3), 412 (2), 396 (5), 377 (1), 314 (1), 285 (1), 264 [RDA fr *a*]$^+$ (31), 246 (100), 232 (35), 214 (28), 207 [RDA, fr *b*]$^+$ (30), 187 (80)

Acetylation of **1a** Compound **1a** was treated with Ac_2O and pyridine at room temp overnight Ice was added to the reaction mixture, which yielded a white ppt The ppt was filtered, dried and was purified by prep TLC (solvent $CHCl_3$–MeOH, 9 1) and crystallized with MeOH as colourless microcrystalline powder (**1b**) mp 213–215° IR $\nu_{max}^{CHCl_3}$ cm^{-1} 3647 (OH), 2940, 1705, 1275, 1202, 1112, 1030 ^1H NMR ($CDCl_3$) δ0 74, 0 85, 0 93, 1 14, 1 25, 1 27 (each 3H, *s*, 6 × Me), 2 04 (3H, *s*, OAc), 2 8 (1H, *dd*, *J* = 13 75, 5 Hz, H-18), 3 60 (3H, *s*, COOMe), 4 50 (1H, *t*, *J* = 10 Hz, H-3), 5 32 (1H, *t*, *J* = 3 75 Hz, H-12) MS *m/z* (rel int) 514 $[M]^+$ (2), 496 $[M-H_2O]^+$ (6), 481 (1), 455 (13), 436 (14), 377 (6), 299 (2), 264 [fr *a*]$^+$ (20), 246 (64), 232 (22), 214 (12), 187 (100), 174 (18), 162 (10), 149 (14), 133 (16), 120 (12), 105 (13), 93 (10)

Dehydration of **1** Compound **1** (10 mg) was refluxed with methanolic HCl (10 ml MeOH, 1 ml H_2O, 1 7 ml conc HCl) for 3 hr The MeOH from this reaction mixture was evaporated and the residue was partitioned between EtOAc and H_2O and the organic layer was separated The solvent was evaporated The residue, which consisted of two main products, was separated by prep TLC ($CHCl_3$–MeOH, 9 1) to yield methyl-3β-hydroxy-30-norolean-12,19-dien-28-oate and 3β-hydroxy-30-norolean 12,19-dien-28 oic acid These were identical with the two products (**3** and **2**) already described above

Acknowledgements—We thank the University Grants Commission of Pakistan for financial support

REFERENCES

1 Hoppe, H A (1975) *Drogenkunde, Vol* 1 *Angiosperms*, 8th edn, pp 551–552 Walter de Gruyter Berlin
2 Wedekind, I E and Schicke, W (1931) *Z Physiol Chem* **195**, 132
3 Winterstein, A and Stein, G (1931) *Z Physiol Chem* **199**, 64
4 Habermehl, G and Moeller, H (1974) *Justus Liebigs Ann Chem* **2**, 169
5 Tori, K, Seo, S and Tomita, Y (1975) *Chem Ind* (*London*) 434
6 Budzikiewicz, H, Willson, J M and Djerassi, C (1963) *J Am Chem Soc* **85**, 3688
7 Wehrli, F W and Nishida, T (1979) in *Progress in the Chemistry of Organic Natural Products* (Herz, W, Grisebach, H and Kirby, G W, eds) Vol 36, pp 95–101 Springer, Vienna