J Chem Crystallogr (2009) 39:99–103
DOI 10.1007/s10870-008-9431-x

ORIGINAL PAPER

X-ray Structure Investigation of 20-*O*-*β*-D-glucopyranosyl-20(*S*)-protopanaxadiol

Wei Zhou · JiYang Li · MeiQing Feng · Pei Zhou

Received: 11 October 2006 / Accepted: 4 June 2008 / Published online: 27 June 2008

Abstract 20-*O*-*β*-D-glucopyranosyl-20(*S*)-protopanaxadiol **(I)** has been isolated as a metabolite of ginseng saponins by *Paecilomyces bainier*. The crystal structure of **(I)** was also investigated by single crystal X-ray diffraction analysis. It crystallizes in the monoclinic space group $P2_1$, with $a = 15.99\ 2\ (3)$ Å, $b = 11.960\ (19)$ Å, $c = 20.127\ (3)$ Å, $\beta = 101.85\ (4)°$, and $V = 3767.5\ (11)$ Å^3, $Z = 4$, $R_{int} = 0.1129$, $\omega R(F^2) = 0.1749$, $F(000) = 1{,}448$. The title compound is a dammarane-type triterpenoid, and the four rings of this compound connect with each other in *trans* characteristics,with the three-six-membered rings in *chair* conformation and the five-membered ring adopting *envelope* one. A *β*-D-glucopyranosyl group and a 2-methyl-2-pentenyl group connect to C20 of *S* configuration. The space symmetrical motif is adopted and each packing unit-cell contains two symmetrical compound **(I)** molecules and four H_2O molecules. The hydrogen-bonding (both intra- and inter-molecular) interactions play a major role in this structural association.

Keywords Protopanaxadiol · Ginsenoside · Metabolite · Crystal structure

Introduction

Ginseng (the root of *Panax* ginseng C.A.Meyer) is frequently used in Asian countries as a traditional medicine. Ginsenosides, glycosides containing an aglycone (protopanaxadiol and protopanaxatiol), have been regarded as the principal pharmacoactive molecules of ginseng [1–3]. So far, more than 40 ginsenosides have been isolated and their chemical structures have been determined initially on the basis of NMR technique and mass spectrometry [4, 5]. Because these molecules are exceedingly difficult to crystallize in suitable form for structural analysis, the X-ray crystal structures of ginsenosides have never been reported. Compound **(I)** (Fig. 1) is the main metabolite of protopanaxadiol ginsenosides by human intestinal bacteria [6, 7]. This metabolite induces antimetastatic and anticarcinogenic effects and exhibits anti-inflammatory, antiallergic and hepatoprotective activities [8–10]. In this paper, **(I)** was obtained from ginseng saponins transformated by *Paecilomyces bainier* and its single crystal structure was discussed.

W. Zhou · J. Li · M. Feng · P. Zhou (✉)
Department of Biosynthetic Drug, School of Pharmacy, Fudan University, 138 YiXueYuan Road, Shanghai 200032, Peoples' Republic of China
e-mail: pz19444@yahoo.com.cn; pzhou@shmu.edu.cn

Experimental Section

General Methods for **(I)**

The fermentation broth of ginseng saponins (6 g) by *P. bainier* was centrifuged and the precipitation was extracted with EtOH for 48 h. Removal of the EtOH from the extract under reduced pressure gave a syrup, which was further diluted with H_2O and partitioned EtOAc. The combined organic layer was concentrated and subjected to silica gel column chromatography eluting with $CHCl_3$–MeOH–EtOAc–H_2O (2:1:4:1, lower phase) to afford 15 fractions. Fractions 6–8 were repeatedly chromatographed over silica gel and then purified by preparative HPLC (Waters Symmetry C18 column, 7 μm, 7.8 mm × 150 mm, 48% acetonitrile in water) to yield **(I)** (200 mg). The melting point was determined with XT4A micro-melting point apparatus. Optical rotation was detected with Jasco P-1020 automatic polarimeter. MS was recorded with

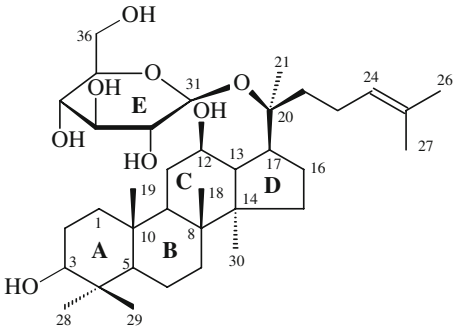

Fig. 1 The chemical structure of **(I)**

Angilent LC-MS instrument using EI method. ^{1}H and ^{13}C NMR were recorded with Mercury plus 400 spectrometer, tetramethylsilane being used as internal standard.

(I), colorless needle, molecular formula: $C_{36}H_{64}O_8 \cdot 2H_2O$, m.p. 162–164 °C; $[\alpha]_D^{22} + 43.1°$ ($c = 0.248$, MeOH); EI-MS *m/z*: 645 $[M + Na]^+$. ^{1}H NMR (400 MHz, pyridine-d_5) data: δ 5.22 (1H, t, $J = 7.1$ Hz, H-24), 5.17 (1H, d, $J = 7.8$ Hz, H-31); 3.93 (1H, ddd-like, H-12-OH), 3.4 (1H, dd, $J = 10.5$, 5.1 Hz, H-3-OH), 1.62 (3H, s, H-21), 1.57 (6H, s, H-26, 27), 1.22 (3H, s, H-28), 1.03 (3H, s, H-29), 0.97 (3H, s, H-30), 0.93 (3H, s, H-18), 0.87 (3H, s, H-19), 0.79 (1H, d, $J = 11.0$ Hz, H-5); ^{13}C NMR (125 MHz, pyridine-d_5) data: δ 130.8 (C, C-25), 125.9 (CH, C-24), 98.4 (CH, C-31), 83.5 (C, C-20), 79.6 (CH, C-33), 78.6 (CH, C-35), 78.3 (CH, C-3), 75.4 (CH, C-32), 72.0 (CH, C-34), 70.5 (CH, C-12), 63.3 (CH_2, C-36), 56.8 (CH, C-5), 52.0 (CH, C-17), 51.9 (C, C-14), 50.8 (CH, C-9), 50.0 (CH, C-13), 40.6 (C, C-8), 40.1 (C, C-4), 39.9 (CH_2, C-1), 37.9 (C, C-10), 36.7 (CH_2, C-22), 35.7 (CH_2, C-7), 31.6 (CH_2, C-15), 31.4 (CH_2, C-11), 29.3 (CH_3, C-28), 28.9 (CH_2, C-2), 27.3 (CH_2, C-16), 26.4 (CH_3, C-26), 23.9 (CH_2, C-23), 23.0 (CH_3, C-21), 19.5 (CH_2, C-6), 18.5 (CH_3, C-27), 18.1 (CH_3, C-30), 17.1 (CH_3, C-18 or C-29), 16.7 (CH_3, C-19).

Crystallography

The single crystal of **(I)** was obtained by the slow evaporation of its solution (MeOH–H_2O, 8:2) at 4 °C after 3 weeks. A suitable colorless crystal of **(I)** with approximate dimensions 0.508 × 0.109 × 0.094 mm was selected. All data were collected on a Bruker Smart Apex-CCD diffractometer with graphite monochromated Mo*K*α radiation ($\lambda =$ 0.71073 Å) at 293 (2) K. A total of 20,757 reflections were collected (13,762 independent collection $R_{int} = 0.1129$). Data reduction, molecular graphics and software used to prepare material for publication were contained in the Bruker SHELXTL program package [11]. The structure were solved by direct methods using SHELXS-97 [12] and refined anisotropically by a full-matrix least-squares based on F^2 using SHELXL-97 [13]. H(3), H(5), H(9), H(12), H(13), H(17), H(31), H(32), H(33), H(34) and H(35) were found from a difference map and refined isotropically. All other H atoms were placed in geometrically calculated positions (C–H = 0.82–0.97 Å) and constrained to ride on their parent atoms. The crystal and structure determination data for **(I)** are summarized in Table 1. Selected bond distance and angles are given in Table 2.

Table 1 Crystal data, data collection and refinement parameters of **(I)**

	(I)
Formula	$C_{36}H_{64}O_8 \cdot 2H_2O$
Formula weight	658.89
Crystal dimensions (mm^3)	0.508 × 0.106 × 0.094
Temperature (K)	293 (2)
Crystal system	Monoclinic
Space group	$P2_1$
Unit cell dimensions	
	$a = 15.992$ (3) Å
	$b = 11.960$ (19) Å
	$c = 20.127$ (3) Å
	$\beta = 101.85$ (4)°
Cell volume ($Å^3$)	3,767.5 (11)
Z	4
D (calcd., g/cm^{-3})	1.162
Radiation	Mo Kα ($\lambda = 0.71073$)
Absorption coefficient (mm^{-1})	0.083
Absorption correction	Empirical
Transmission factors	0.830–1.000
θ range for data collection (°)	1.99–25.99°
Range of h, k, and l	−19/19, −14/14, −21/24
Reflections collected/unique	20,757/13,762 ($R_{int} = 0.1129$)
Reflections with $I > 2\sigma$ (I)	4,748
Data/parameters/restraints	13,762/913/1
Extinction coefficient	0.0032 (5)
Goodness of fit on F^2	0.891
Final R indices [$I > 2\sigma$ (I)]	$R_1 = 0.0898$; $\omega R_2 = 0.1749$
R indices (all data)	$R_1 = 0.2348$; $\omega R_2 = 0.2435$
Max shift/error on final cycle $[(\Delta/\sigma)_{max}]$	0.165
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e $Å^{-3}$)	0.335, −0.274

Results and Discussion

The crystal structure of **(I)** established by X-ray diffraction analysis is shown in Fig. 2. Rings *A* and *B* are almost in chair conformation, with a β-methyl at C10. Ring *C* exhibits slightly distorted chair conformation, which due to the unbalanced existence of 8 β-methyl and 12 β-hydroxyl. The *D* ring is between a twist and a 13

Table 2 Selected geometric parameters (in Å and °)

Bond lengths			
O(1)–C(3)	1.443 (10)	C(4)–C(28)	1.499 (12)
O(2)–C(12)	1.403 (10)	C(4)–C(29)	1.526 (11)
O(3)–C(2O)	1.474 (9)	C(8)–C(18)	1.516 (12)
O(3)–C(31)	1.384 (9)	C(10)–C(19)	1.518 (11)
O(4)–C(31)	1.450 (10)	C(13)–C(17)	1.504 (11)
O(4)–C(35)	1.431 (9)	C(14)–C(30)	1.529 (12)
O(5)–C(32)	1.406 (11)	C(17)–C(20)	1.569 (11)
O(6)–C(33)	1.438 (11)	C(20)–C(21)	1.510 (11)
O(7)–C(34)	1.429 (10)	C(24)–C(25)	1.276 (15)
O(8)–C(36)	1.420 (11)	C(25)–C(26)	1.442 (17)
C(1)–C(2)	1.546 (11)	C(25)–C(27)	1.526 (17)
Bond angles			
O(1)–C(3)–C(2)	108.5 (7)	C(13)–C(14)–C(8)	109.5 (7)
O(2)–C(12)–C(13)	115.6 (7)	C(13)–C(17)–C(16)	105.9 (7)
O(3)–C(31)–C(32)	109.9 (8)	C(13)–C(17)–C(20)	120.4 (7)
O(3)–C(31)–O(4)	108.1 (7)	C(15)–C(14)–C(13)	100.4 (7)
O(3)–C(20)–C(17)	102.4 (6)	C(15)–C(14)–C(30)	107.0 (8)
O(3)–C(20)–C(22)	108.8 (7)	C(16)–C(17)–C(20)	113.6 (7)
O(4)–C(31)–C(32)	105.0 (7)	C(17)–C(13)–C(12)	122.6 (7)
O(4)–C(35)–C(36)	107.5 (7)	C(17)–C(13)–C(14)	104.2 (7)
O(5)–C(32)–C(33)	110.8 (8)	C(18)–C(8)–C(9)	112.4 (7)
O(6)–C(330)–C(34)	109.0 (8)	C(19)–C(10)–C(5)	115.4 (7)
O(7)–C(34)–C(35)	108.9 (8)	C(21)–C(20)–C(17)	109.3 (7)
O(8)–C(36)–C(35)	115.4 (9)	C(24)–C(25)–C(26)	124.1 (13)
C(1)–C(10)–C(9)	105.9 (7)	C(24)–C(25)–C(27)	121.9 (14)
C(6)–C(5)–C(10)	113.3 (7)	C(26)–C(25)–C(27)	113.9 (13)
C(7)–C(8)–C(14)	110.9 (7)	C(28)–C(4)–C(29)	107.4 (7)
C(11)–C(9)–C(10)	113.7 (8)	C(30)–C(14)–C(13)	111.0 (7)
Torsional angles			
C(10)–C(1)–C(2)–C(3)	−55.0 (11)	C(12)–C(13)–C(17)–C(16)	−156.3 (8)
C(1)–C(1)–C(2)–O(1)	−172.0 (8)	C(12)–C(13)–C(17)–C(20)	73.2 (11)
C(2)–C(1)–C(10)–C(19)	−75.1 (9)	C(13)–C(17)–C(20)–O(3)	−71.1 (9)
C(4)–C(5)–C(6)–C(7)	159.9 (7)	C(13)–C(17)–C(20)–C(22)	46.3 (10)
C(4)–C(5)–C(10)–C(9)	−165.6 (7)	C(15)–C(16)–C(17)–C(13)	1.5 (9)
C(6)–C(5)–C(10)–C(1)	171.6 (7)	C(15)–C(16)–C(17)–C(20)	135.8 (8)
C(6)–C(7)–C(8)–C(14)	−166.3 (7)	C(16)–C(17)–C(20)–O(3)	161.8 (7)
C(7)–C(8)–C(14)–C(13)	−177.3 (7)	C(20)–C(22)–C(23)–C(24)	170.7 (8)
C(7)–C(8)–C(9)–C(11)	179.9 (7)	C(20)–O(3)–C(31)–O(4)	−100.8 (8)
C(8)–C(9)–C(10)–C(5)	−52.0 (10)	C(20)–O(3)–C(31)–C(32)	145.1 (7)
C(8)–C(9)–C(11)–C(12)	53.4 (10)	C(23)–C(24)–C(25)–C(26)	6 (3)
C(9)–C(11)–C(12)–O(2)	−174.4 (7)	C(23)–C(24)–C(25)–C(27)	−172.0 (12)
C(10)–C(9)–C(11)–C(12)	−174.4 (7)	C(28)–C(4)–C(5)–C(10)	172.1 (7)
C(11)–C(9)–C(10)–C(5)	178.2 (7)	C(29)–C(4)–C(5)–C(6)	69.1 (10)
O(2)–C(12)–C(13)–C(14)	172.8 (7)	C(31)–C(32)–C(33)–C(34)	−54.7 (10)
O(2)–C(12)–C(13)–C(17)	−61.8 (11)	C(31)–O(4)–C(35)–C(34)	64.1 (9)
C(12)–C(13)–C(14)–C(8)	−58.7 (9)	C(34)–C(35)–C(36)–O(8)	−176.3 (8)
C(12)–C(13)–C(14)–C(15)	178.0 (8)	O(3)–C(31)–C(32)–O(5)	−59.7 (9)

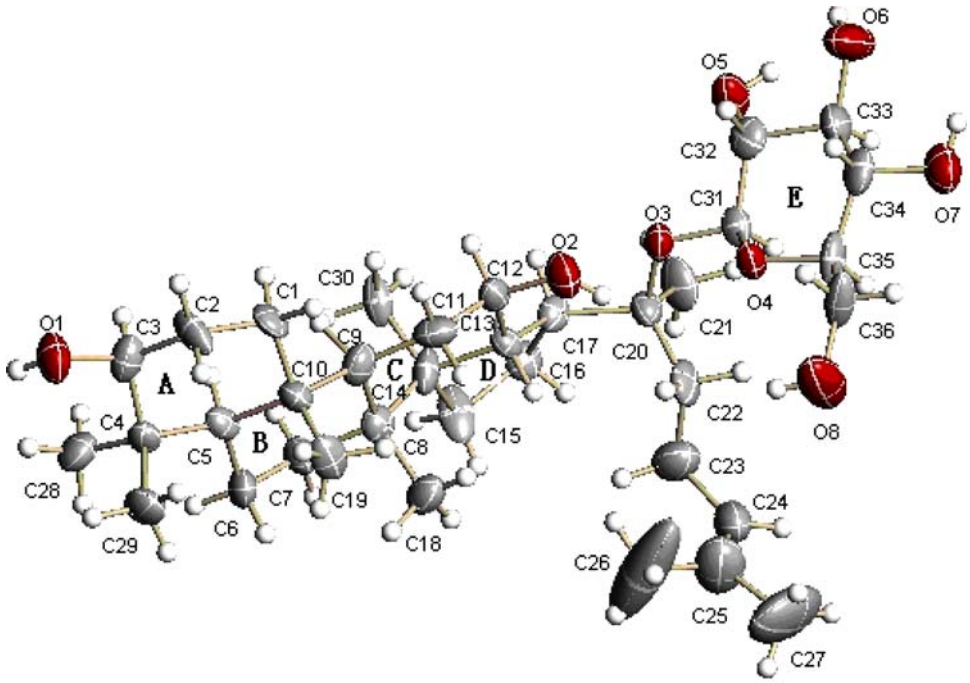

Fig. 2 View of the title molecule **(I)** with the atom-numbering scheme. H atoms are drawn as circles of arbitrary small radius for clarity

β-envelope conformation, with an α-methyl on C14. The *A*/*B*, *B*/*C* and *C*/*D* ring junctions all approach *trans* characteristics. A list of the endocyclic torsion angles about the three ring junctions, which support the above-mentioned ring-junction characteristics, are given in Table 3. The bond distances and angles within the six-membered *A*, *B* and *C* rings and those within the five-membered *D* ring are within the average values found for steroids. The results clearly establish a 20-*O*-*β*-D-glucopyranosyl, as shown by the torsion angles C(20)–O(3)–C(31)–C(32)[145.1(7)], C(20)–O(3)–C(31)–O(4)[−100.8(8)], O(4)–C(31)–C(32)–O(5)[−175.8(7)], O(4)–C(35)–C(36)–O(8)[65.2(10)], O(7)–C(34)–C(35)–O(4)[−171.0(7)], O(6)–C(33)–C(34)–O(7)[−68.9(10)]. The C20 stereocenter is in a synclinal position with respect to both C13 and C16 and has the *S* spatial configuration, with the glucosyl group and a 2-methyl-2-pentenyl group connecting to it. Elucidating form the three-dimensional structure of **(I)**, the glucosyl group and the pentenyl group just locate at the two wings of the rings *A*/*B*/*C*/*D* mean plane, and the glucosyl group is at the side near to the hydroxyl groups at C3 and C12, which may due to the three intramolecular hydrogen bonds [O(2)–H(2)···O(3) 2.722 Å, O(2)–H(2)···O(4) 3.216 Å and O(8)–H(8)···O(1) 2.746 Å [−x + 2, y + 1/2, −z + 2]]. As to the C24=C25 bond of 2-methyl-2-pentenyl group, its length [1.267(15) Å] is much shorter than the typical C_{sp2}=C_{sp2} bond length of 1.341 Å, which is the result of two electron donating methyls on C25.

Table 3 Endocyclic torsion angles (°) about the ring junctions in **(I)**

Junction	Atoms	Angle	Characteristics
A/B	C(4)–C(5)–C(10)–C(1)	−52.8 (10)	*trans*
	C(6)–C(5)–C(10)–C(9)	58.8 (9)	
B/C	C(7)–C(8)–C(9)–C(10)	48.8 (10)	*trans*
	C(14)–C(8)–C(9)–C(11)	−60.8 (9)	
C/D	C(12)–C(13)–C(14)–C(8)	−58.7 (9)	*trans*
	C(17)–C(13)–C(14)–C(8)	166.3 (7)	

The packing diagram of **(I)**, including the hydrogen-bonding network, is shown in Fig. 3. The space symmetrical motif is adopted in its packing and each unit-cell

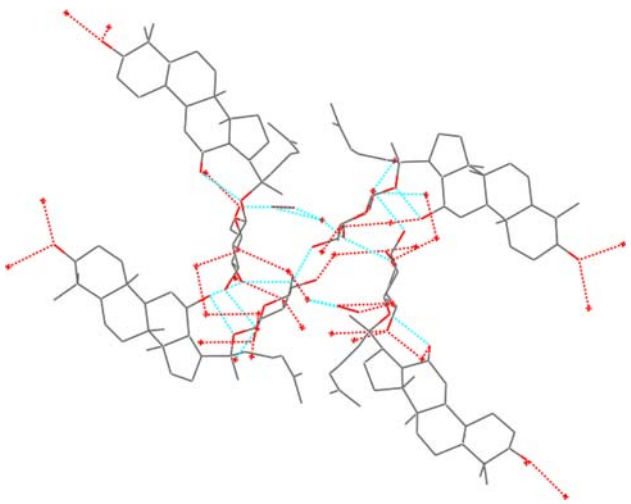

Fig. 3 Partial packing diagram of the molecules (projected along the a-axis). Dashed lines represent the hydrogen bonds

Table 4 Hydrogen-bonding parameters for **(I)**

D–H	d (D–H)	d (H–A)	<DHA	d (D–A)	A
O(5)–H(5)	0.820	2.249	123.13	2.783	O(14)
O(6)–H(6)	0.820	2.501	133.33	3.121	O(18) [−x + 1, y + 1/2, −z + 1]
O(7)–H(7)	0.820				
Alternative approximate positions for H attached to O (7)					
0.4184	1.0440	0.7043	O···O(10) [x, y + 1, z]	2.702	<COH = 112
0.4285	1.0086	0.6394	O···O(16) [x, y + 1, z]	2.773	<COH = 122
0.4126	1.0526	0.6621	O···O(12) [x, y + 1, z]	3.365	<COH = 154
O(13)–H(13)	0.820	2.155	136.40	2.806	O(18) [−x + 1, y + 1/2, −z + 1]
O(14)–H(14)	0.820	2.184	129.71	2.780	O(17)
O(15)–H(15)	0.820	2.598	153.23	3.351	O(17) [−x + 1, y − 1/2, −z + 1]
O(16)–H(16)	0.820	1.983	157.45	2.759	O(15) [−x + 1, y − 1/2, −z + 1]

contains two symmetrical compound **(I)** molecules and four H_2O molecules. Molecules are connected by means of relatively strong intermolecular hydrogen bonds formed within the hydroxyl groups and water of crystallization. Further details of these hydrogen bonds are given in Table 4.

Recently, because of various conspicuous biological actions [8–10], compound **(I)** has received increasing attention. We hope our work would help to clarify the mechanism of its potential bioactivities by the 3D-QSAR analysis.

Acknowledgements We gratefully acknowledge financial support from the National Nature Science Foundation of China (Grant #30772676). We sincerely thank Jie Sun of the Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences (CAS) for X-ray crystallographic assistance.

Supplementary Material

CCDC 621605 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge at www.ccdc.cam.ac.uk/conts/retrieving.html [or from the Cambridge Crystallographic Data Centre (CCDC), 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44-1223-336033; e-mail: deposit@ccdc.cam.ac.uk].

References

1. Lee HU, Bae EA, Han MJ, Kim DH (2005) Biol Pharm Bull 28:1992
2. Park W, Lim W, Cho J, Inoue H, Rhyu MR, Lee Y (2006) Planta Med 72:272
3. Zhang HS, Wang SQ (2006) J Cell Biochem 98:1471
4. Yahara S, Kaji K, Tanaka O (1979) Chem Pharm Bull 27:99
5. Kasai R, Besso H, Tanaka O, Saruwatari YI, Fuwa T (1983) Chem Pharm Bull 31:2120
6. Hasegawa H, Sung JH, Matsumiya S, Uchiyama M (1996) Planta Med 62:453
7. Bae EA, Kim NY, Han MJ, Choo MK, Kim DH (2003) J Microbiol Biotechnol 13:9
8. Lee YJ, Shin JW, Chun KS, Park KK, Chung WY, Bang YJ, Sung JH (2005) Carcinogenesis 26:359
9. Lee HU, Bae EA, Han MJ, Kim DH (2005) Liver Int 25:1069
10. Park EK, Shin YW, Lee HU, Kim SS, Lee YC, Lee BY, Kim DH (2005) Biol Pharm Bull 28:652
11. SMART (Version 5.049). Bruker AXS Inc., Madison, WI, USA
12. Sheldrick GM (1997) SHELXS-97. University of Gottingen, Germany
13. Sheldrick GM (1997) SHELXL 97. University of Göttingen, Germany