wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 11:17 PM EDT

PDB ID	:	3VVG
Title	:	The Crystal Structure of Cellulase-Inhibitor Complex.
Authors	:	Ishikawa, K.; Maeno, Y.; Kataoka, M.
Deposited on	:	2012-07-24
Resolution	:	1.90 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

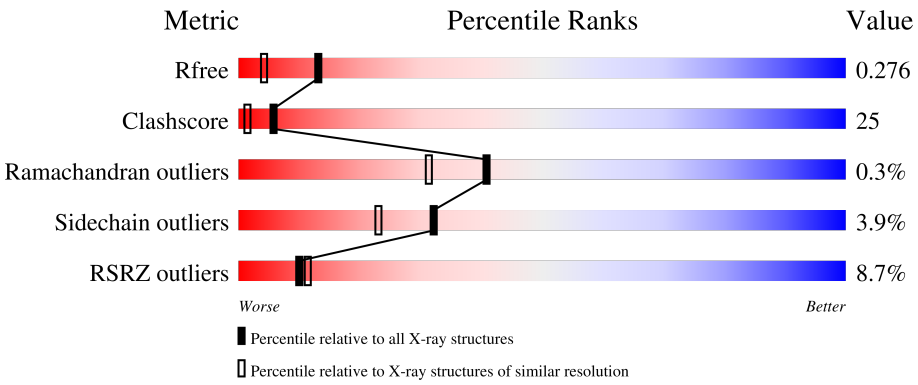
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	377	2%, 80%, 18%, •
1	B	377	19%, 53%, 43%, 5%
1	C	377	4%, 72%, 27%, •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ZGB	A	501	-	-	X	-
2	ZGB	C	501	-	-	X	-
3	PO4	A	502	-	-	X	-
3	PO4	B	502	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9943 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

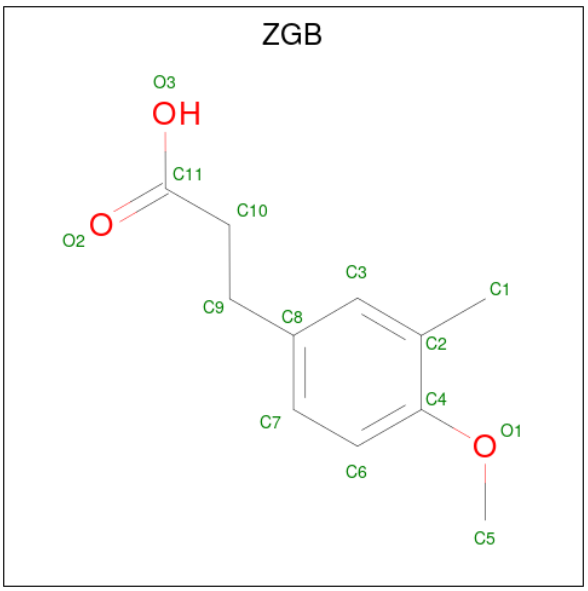
- Molecule 1 is a protein called 458aa long hypothetical endo-1,4-beta-glucanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	377	3069	2002	500	557	10	0	0	0
1	B	377	3069	2002	500	557	10	0	0	0
1	C	377	3069	2002	500	557	10	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

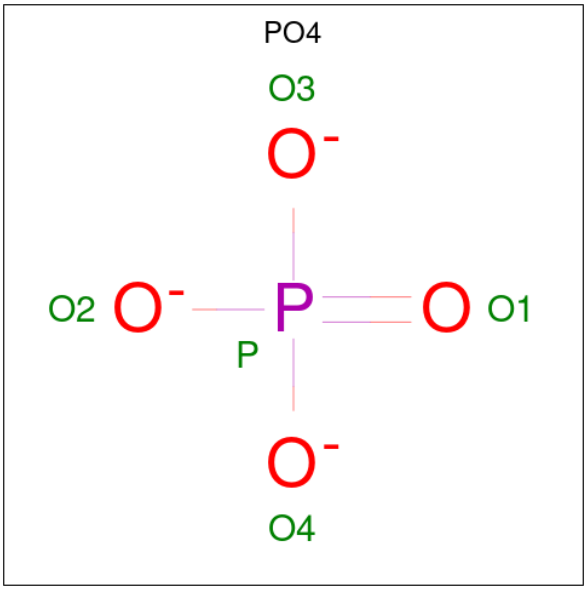
Chain	Residue	Modelled	Actual	Comment	Reference
A	289	LYS	ARG	CONFLICT	UNP O58925
B	289	LYS	ARG	CONFLICT	UNP O58925
C	289	LYS	ARG	CONFLICT	UNP O58925

- Molecule 2 is 3-(4-methoxy-3-methylphenyl)propanoic acid (three-letter code: ZGB) (formula: $C_{11}H_{14}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	12	10	2	0	0
			Total	C	O		
2	A	1	12	10	2	0	0
			Total	C	O		
2	C	1	12	10	2	0	0

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
3	A	1	5	4	1	0	0
			Total	O	P		
3	B	1	5	4	1	0	0
			Total	O	P		
3	B	1	5	4	1	0	0
			Total	O	P		
3	B	1	5	4	1	0	0

- Molecule 4 is water.

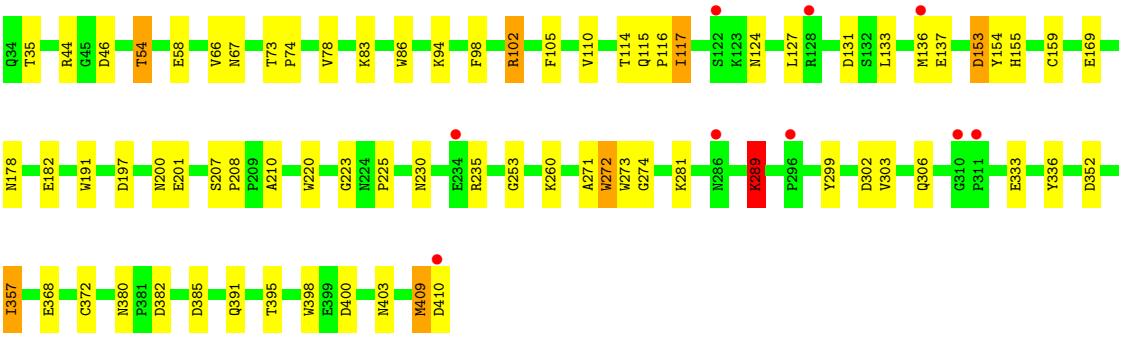
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	304	304	304	0	0
			Total	O		
4	B	146	146	146	0	0
			Total	O		
4	C	230	230	230	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

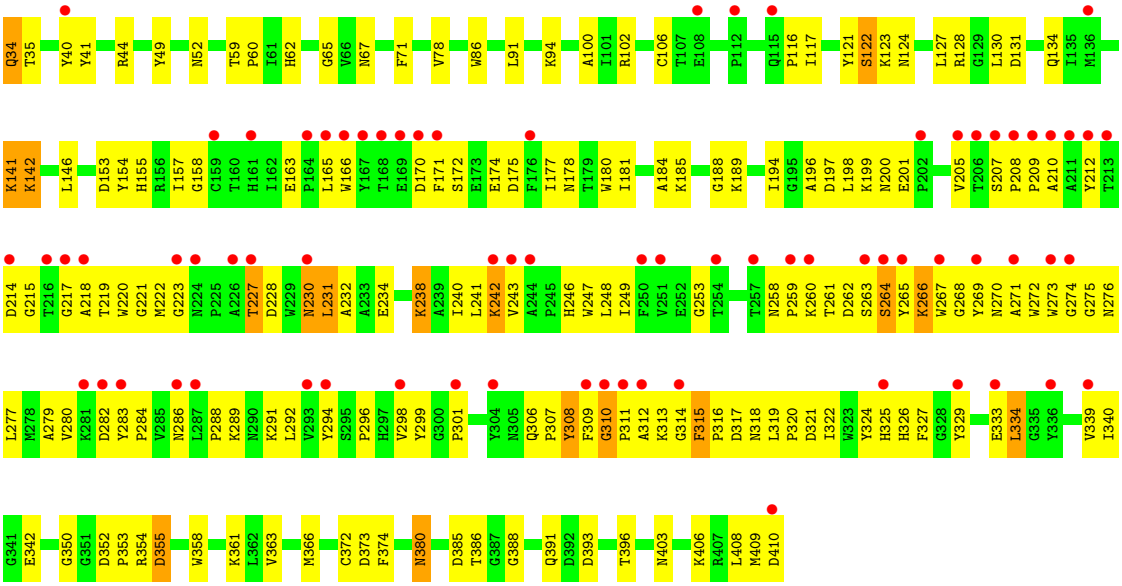
• Molecule 1: 458aa long hypothetical endo-1,4-beta-glucanase

Chain A:



• Molecule 1: 458aa long hypothetical endo-1,4-beta-glucanase

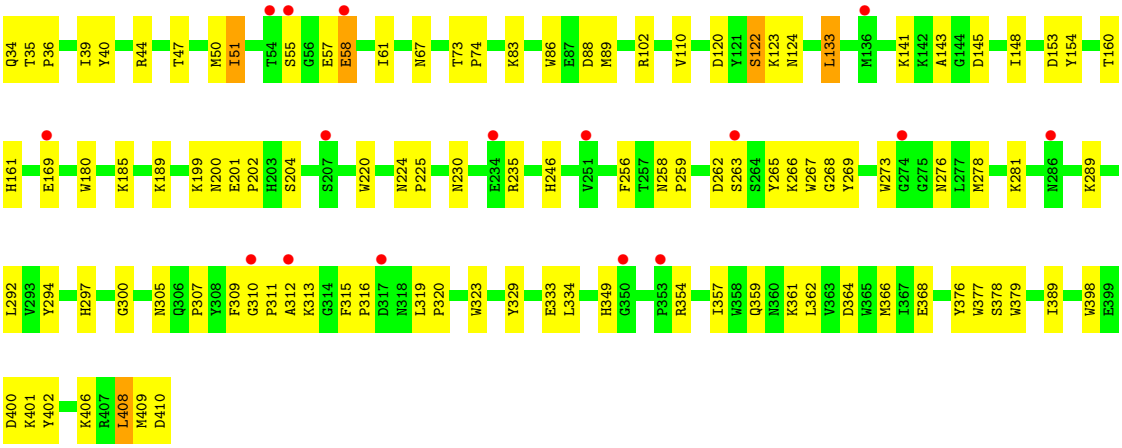
Chain B:



• Molecule 1: 458aa long hypothetical endo-1,4-beta-glucanase

Chain C:

4%
72%
27%


Q34 T35 P36 I39 Y40 R44 T47 M50 I51 T54 S55 G56 E57 E58 I61 N67 T73 P74 K83 W86 E87 D88 M89 R102 V110 D120 Y121 S122 K123 N124 L133 M136 K141 K142 A143 G144 D145 I148 D153 Y154 T160
H161 E169 W180 K185 K189 K199 N200 E201 P202 H203 S204 S207 W220 N224 P225 N230 E234 R235 H246 V251 F256 T257 N258 P259 D262 S263 S264 Y265 K266 W267 G268 Y269 W273 G274 G275 N276 L277 M278 K281 N286 K289
L292 V293 Y294 H297 G300 N305 Q306 P307 Y308 F309 G310 P311 A312 K313 G314 F315 P316 D317 N318 L319 P320 W323 Y329 E333 L334 H349 G350 P353 R354 I357 W358 Q359 N360 K361 L362 V363 D364 W365 M366 I367 E368 Y376 W377 S378 W379 I389 W398 E399
D400 K401 Y402 K406 R407 L408 M409 D410

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	163.57Å 58.43Å 138.01Å 90.00° 109.01° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 28.45 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.2 (50.00-1.90) 99.2 (28.45-1.90)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.19 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.220 , 0.286 0.216 , 0.276	Depositor DCC
R_{free} test set	4855 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	23.2	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 56.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9943	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.54% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZGB, PO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
1	A	1.27	4/3180 (0.1%)	1.05	7/4341 (0.2%)
1	B	0.98	1/3180 (0.0%)	0.91	4/4341 (0.1%)
1	C	1.03	0/3180	0.91	2/4341 (0.0%)
All	All	1.10	5/9540 (0.1%)	0.96	13/13023 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	272	TRP	CE3-CZ3	7.31	1.50	1.38
1	A	272	TRP	CE2-CZ2	5.41	1.49	1.39
1	A	271	ALA	CA-CB	5.41	1.63	1.52
1	B	358	TRP	CB-CG	5.11	1.59	1.50
1	A	289	LYS	CD-CE	5.06	1.63	1.51

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	235	ARG	NE-CZ-NH1	8.19	124.40	120.30
1	C	153	ASP	CB-CG-OD1	7.93	125.44	118.30
1	A	400	ASP	CB-CG-OD1	7.88	125.39	118.30
1	A	352	ASP	CB-CG-OD1	6.83	124.44	118.30
1	B	197	ASP	CB-CG-OD1	6.34	124.01	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	310	GLY	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3069	0	2910	85	2
1	B	3069	0	2907	260	0
1	C	3069	0	2908	111	0
2	A	24	0	20	14	2
2	C	12	0	10	8	0
3	A	5	0	0	3	0
3	B	15	0	0	3	0
4	A	304	0	0	25	0
4	B	146	0	0	36	0
4	C	230	0	0	17	0
All	All	9943	0	8755	458	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 25.

The worst 5 of 458 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:191:TRP:CD1	3:A:502:PO4:O3	1.73	1.40
1:B:234:GLU:HB3	4:B:713:HOH:O	1.16	1.34
1:B:217:GLY:O	1:B:227:THR:CG2	1.85	1.25
1:B:34:GLN:HG3	1:B:40:TYR:CD1	1.72	1.25
1:A:159:CYS:HB3	4:A:887:HOH:O	1.07	1.25

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:403:ASN:ND2	2:A:503:ZGB:C1[4_546]	1.19	1.01
1:A:403:ASN:ND2	2:A:503:ZGB:C2[4_546]	1.96	0.24

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	375/377 (100%)	364 (97%)	11 (3%)	0	100	100
1	B	375/377 (100%)	341 (91%)	32 (8%)	2 (0%)	29	18
1	C	375/377 (100%)	358 (96%)	16 (4%)	1 (0%)	41	31
All	All	1125/1131 (100%)	1063 (94%)	59 (5%)	3 (0%)	41	31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	264	SER
1	B	308	TYR
1	C	312	ALA

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	324/324 (100%)	313 (97%)	11 (3%)	37	28
1	B	324/324 (100%)	305 (94%)	19 (6%)	19	10
1	C	324/324 (100%)	316 (98%)	8 (2%)	47	41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	972/972 (100%)	934 (96%)	38 (4%)	32 23

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	334	LEU
1	C	169	GLU
1	B	380	ASN
1	C	58	GLU
1	C	408	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	391	GLN
1	C	161	HIS
1	C	359	GLN
1	C	200	ASN
1	A	403	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

7 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	ZGB	C	501	-	12,12,14	1.60	2 (16%)	15,15,18	2.11	5 (33%)
3	PO4	B	502	-	4,4,4	0.86	0	6,6,6	0.67	0
3	PO4	B	503	-	4,4,4	0.86	0	6,6,6	0.58	0
3	PO4	A	502	-	4,4,4	0.98	0	6,6,6	0.51	0
3	PO4	B	501	-	4,4,4	1.16	0	6,6,6	1.04	1 (16%)
2	ZGB	A	503	-	11,12,14	1.46	2 (18%)	15,15,18	1.88	4 (26%)
2	ZGB	A	501	-	11,12,14	1.51	2 (18%)	15,15,18	2.02	6 (40%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ZGB	C	501	-	-	2/4/4/7	0/1/1/1
2	ZGB	A	503	-	-	4/4/4/7	0/1/1/1
2	ZGB	A	501	-	-	0/4/4/7	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	ZGB	C9-C10	-3.66	1.34	1.52
2	C	501	ZGB	C10-C9	-3.64	1.35	1.52
2	A	503	ZGB	C9-C10	-3.51	1.35	1.52
2	C	501	ZGB	O3-C11	-3.29	1.25	1.42
2	A	503	ZGB	C3-C2	2.63	1.43	1.39

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	501	ZGB	C10-C9-C8	4.68	131.12	113.67
2	C	501	ZGB	C9-C10-C11	4.01	134.83	113.45
2	A	503	ZGB	C1-C2-C4	-3.65	116.94	120.69
2	A	501	ZGB	C7-C6-C4	-3.32	117.18	120.50
2	A	501	ZGB	C1-C2-C4	3.20	123.98	120.69

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	503	ZGB	C11-C10-C9-C8
2	A	503	ZGB	C9-C10-C11-O2
2	C	501	ZGB	C11-C10-C9-C8
2	A	503	ZGB	C3-C8-C9-C10
2	A	503	ZGB	C7-C8-C9-C10

There are no ring outliers.

5 monomers are involved in 30 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	501	ZGB	8	0
3	B	502	PO4	3	0
3	A	502	PO4	3	0
2	A	503	ZGB	3	2
2	A	501	ZGB	11	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	377/377 (100%)	-0.13	9 (2%)	59	62	5, 16, 31, 40	8 (2%)
1	B	377/377 (100%)	1.14	73 (19%)	1	1	18, 35, 56, 64	10 (2%)
1	C	377/377 (100%)	0.22	16 (4%)	36	39	11, 26, 43, 54	9 (2%)
All	All	1131/1131 (100%)	0.41	98 (8%)	10	11	5, 26, 48, 64	27 (2%)

The worst 5 of 98 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	226	ALA	7.6
1	B	311	PRO	6.8
1	B	283	TYR	5.8
1	B	169	GLU	5.6
1	B	213	THR	5.5

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	PO4	A	502	5/5	0.70	0.23	69,72,72,72	0
2	ZGB	C	501	12/14	0.72	0.23	34,39,41,45	0
2	ZGB	A	501	12/14	0.77	0.15	18,35,40,44	0
2	ZGB	A	503	12/14	0.81	0.14	35,39,41,44	0
3	PO4	B	503	5/5	0.88	0.23	49,49,51,53	0
3	PO4	B	502	5/5	0.89	0.17	43,46,51,52	0
3	PO4	B	501	5/5	0.95	0.11	30,33,37,37	0

6.5 Other polymers (i)

There are no such residues in this entry.