Full wwPDB X-ray Structure Validation Report (i)

Nov 10, 2024 – 03:03 AM EST

PDB ID	:	4K89
Title	:	Crystal structure of Pseudomonas aeruginosa strain K solvent tolerant elastase
Authors	:	Ali, M.S.M.; Said, Z.S.A.M.; Rahman, R.N.Z.R.A.; Basri, M.; Salleh, A.B.
Deposited on	:	2013-04-18
Resolution	:	1.39 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

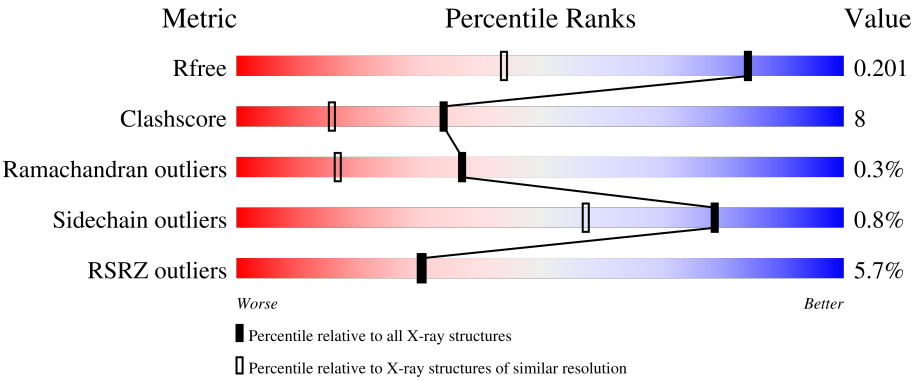
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 1.39 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}	164625	2247 (1.40-1.40)
Clashscore	180529	2446 (1.40-1.40)
Ramachandran outliers	177936	2398 (1.40-1.40)
Sidechain outliers	177891	2397 (1.40-1.40)
RSRZ outliers	164620	2246 (1.40-1.40)

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	301	6% 79% 18%

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
3	PO4	A	406	-	X	-	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 2542 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.

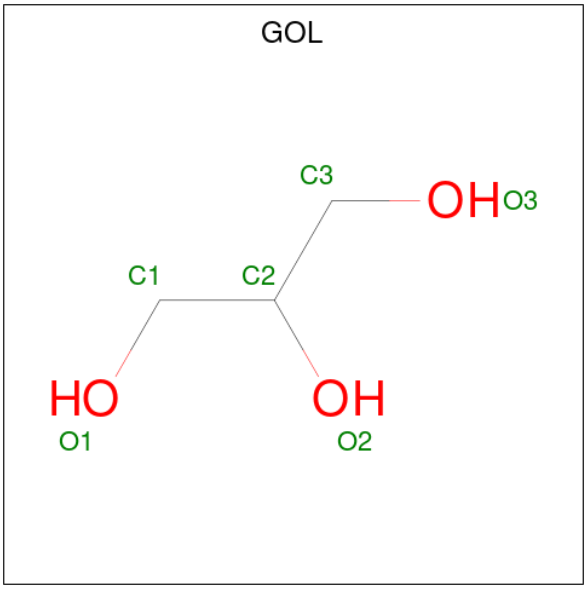
- Molecule 1 is a protein called Organic solvent tolerant elastase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	298	Total	C	N	O	S	0	0	0
			2315	1453	399	450	13			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	116	ASP	GLY	engineered mutation	UNP A7LI11

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



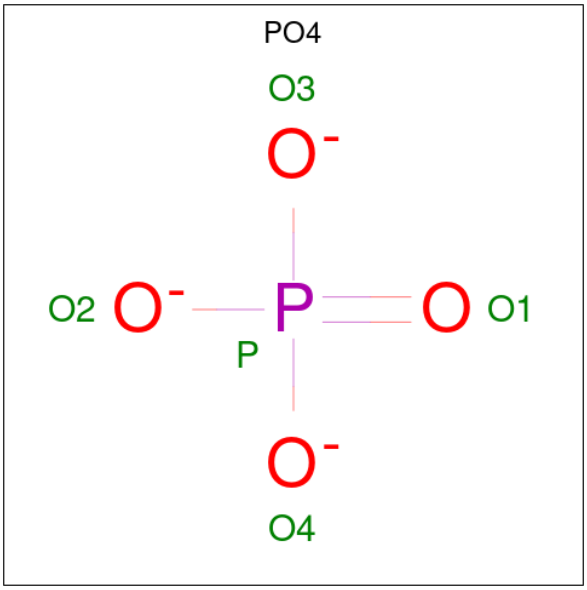
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O		0	0
			6	3	3			
2	A	1	Total	C	O		0	0
			6	3	3			
2	A	1	Total	C	O		0	0
			6	3	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0

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



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

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Ca 1 1	0	0

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Zn 1 1	0	0

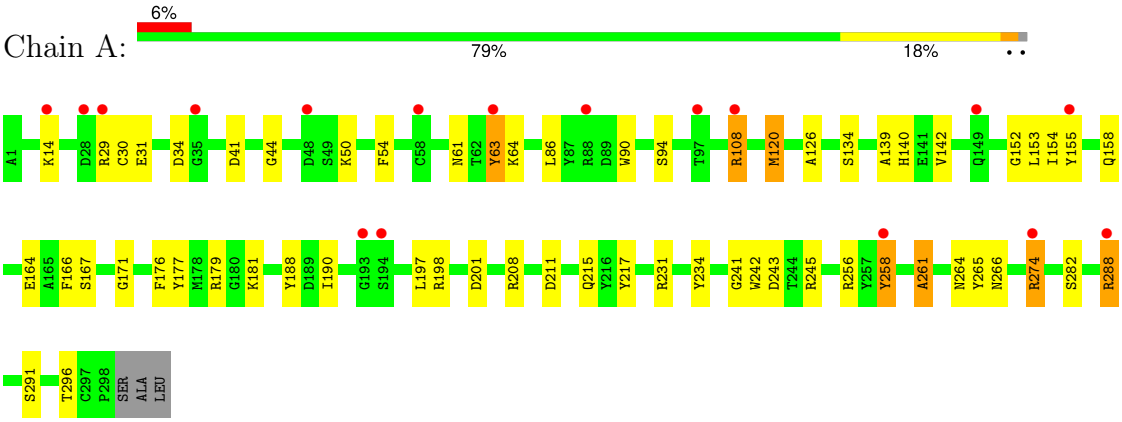
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	186	Total	O	0	0
			186	186		

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: Organic solvent tolerant elastase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	38.99Å 90.17Å 40.60Å 90.00° 113.81° 90.00°	Depositor
Resolution (Å)	45.09 – 1.39 45.09 – 1.39	Depositor EDS
% Data completeness (in resolution range)	99.8 (45.09-1.39) 99.8 (45.09-1.39)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.40 (at 1.39Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.167 , 0.191 0.175 , 0.201	Depositor DCC
R_{free} test set	2610 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	10.8	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 35.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.026 for l,-k,h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2542	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.24% 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

5.1 Standard geometry

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

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	1.54	12/2377 (0.5%)	1.43	27/3221 (0.8%)

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	31	GLU	CD-OE1	-8.39	1.16	1.25
1	A	90	TRP	CB-CG	-6.77	1.38	1.50
1	A	242	TRP	CG-CD1	6.42	1.45	1.36
1	A	177	TYR	CD1-CE1	5.74	1.48	1.39
1	A	291	SER	CB-OG	-5.52	1.35	1.42
1	A	177	TYR	C-O	5.47	1.33	1.23
1	A	94	SER	CB-OG	5.44	1.49	1.42
1	A	63	TYR	CE2-CZ	-5.25	1.31	1.38
1	A	164	GLU	CD-OE2	-5.20	1.20	1.25
1	A	217	TYR	N-CA	5.20	1.56	1.46
1	A	234	TYR	CE1-CZ	-5.13	1.31	1.38
1	A	261	ALA	N-CA	-5.05	1.36	1.46

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	29	ARG	NE-CZ-NH2	9.95	125.28	120.30
1	A	231	ARG	NE-CZ-NH2	-9.90	115.35	120.30
1	A	243	ASP	CB-CG-OD1	9.38	126.74	118.30
1	A	86	LEU	CB-CG-CD1	-8.29	96.92	111.00
1	A	179	ARG	NE-CZ-NH1	7.86	124.23	120.30
1	A	256	ARG	NE-CZ-NH2	-7.85	116.37	120.30
1	A	120	MET	CG-SD-CE	-7.82	87.68	100.20
1	A	274	ARG	NE-CZ-NH1	7.15	123.87	120.30
1	A	245	ARG	NE-CZ-NH1	-6.86	116.87	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	201	ASP	CB-CG-OD1	6.80	124.42	118.30
1	A	41	ASP	CB-CG-OD2	-6.79	112.19	118.30
1	A	243	ASP	CB-CG-OD2	-6.68	112.29	118.30
1	A	217	TYR	CZ-CE2-CD2	-6.26	114.17	119.80
1	A	34	ASP	CB-CG-OD2	6.08	123.77	118.30
1	A	198	ARG	NE-CZ-NH1	-5.90	117.35	120.30
1	A	176	PHE	CB-CG-CD1	-5.84	116.71	120.80
1	A	166	PHE	CB-CG-CD2	-5.81	116.73	120.80
1	A	50	LYS	CD-CE-NZ	5.73	124.88	111.70
1	A	54	PHE	CB-CG-CD2	-5.71	116.81	120.80
1	A	258	TYR	CG-CD2-CE2	5.69	125.85	121.30
1	A	234	TYR	CG-CD2-CE2	-5.61	116.81	121.30
1	A	231	ARG	NE-CZ-NH1	5.53	123.06	120.30
1	A	108	ARG	NE-CZ-NH2	-5.47	117.56	120.30
1	A	30	CYS	CA-CB-SG	-5.46	104.17	114.00
1	A	14	LYS	CD-CE-NZ	5.42	124.18	111.70
1	A	288	ARG	NE-CZ-NH2	5.30	122.95	120.30
1	A	208	ARG	NE-CZ-NH2	-5.02	117.79	120.30

There are no chirality outliers.

There are no planarity outliers.

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	2315	0	2142	34	0
2	A	24	0	32	3	0
3	A	15	0	0	0	0
4	A	1	0	0	0	0
5	A	1	0	0	0	0
6	A	186	0	0	2	0
All	All	2542	0	2174	34	0

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 8.

All (34) 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:44:GLY:HA3	1:A:108:ARG:NH2	1.66	1.09
1:A:44:GLY:HA3	1:A:108:ARG:HH22	1.34	0.91
1:A:241:GLY:H	2:A:403:GOL:H31	1.35	0.88
1:A:44:GLY:CA	1:A:108:ARG:NH2	2.37	0.88
1:A:120:MET:CE	1:A:142:VAL:HG23	2.07	0.83
1:A:44:GLY:CA	1:A:108:ARG:HH22	1.97	0.75
1:A:282:SER:OG	2:A:404:GOL:H2	1.90	0.71
1:A:190:ILE:HD11	1:A:197:LEU:HD11	1.75	0.68
1:A:44:GLY:O	1:A:108:ARG:NH2	2.29	0.65
1:A:120:MET:HE1	1:A:142:VAL:HG23	1.78	0.64
1:A:158:GLN:HE21	1:A:265:TYR:H	1.46	0.64
1:A:44:GLY:C	1:A:108:ARG:NH2	2.56	0.59
1:A:44:GLY:C	1:A:108:ARG:HH22	2.07	0.58
1:A:120:MET:HE3	1:A:142:VAL:HG23	1.85	0.57
1:A:154:ILE:HG23	1:A:261:ALA:HB1	1.89	0.55
1:A:158:GLN:NE2	1:A:264:ASN:HB2	2.22	0.55
1:A:158:GLN:NE2	1:A:265:TYR:H	2.05	0.53
1:A:211:ASP:H	1:A:215:GLN:NE2	2.07	0.52
1:A:266:ASN:ND2	1:A:296:THR:H	2.08	0.51
1:A:158:GLN:HE21	1:A:264:ASN:HB2	1.76	0.51
1:A:211:ASP:H	1:A:215:GLN:HE22	1.59	0.51
1:A:61:ASN:HD21	1:A:63:TYR:HB3	1.76	0.50
1:A:134:SER:HB3	1:A:190:ILE:HG23	1.95	0.48
1:A:139:ALA:HB3	1:A:171:GLY:HA2	1.96	0.48
1:A:44:GLY:CA	1:A:108:ARG:HH21	2.25	0.46
1:A:288:ARG:NE	6:A:520:HOH:O	2.45	0.44
1:A:120:MET:CE	1:A:142:VAL:CG2	2.90	0.43
1:A:258:TYR:CE1	1:A:274:ARG:HD2	2.54	0.42
1:A:140:HIS:CG	1:A:167:SER:HB3	2.55	0.42
1:A:288:ARG:O	1:A:288:ARG:HG2	2.19	0.42
1:A:152:GLY:O	1:A:153:LEU:C	2.58	0.41
1:A:188:TYR:N	6:A:643:HOH:O	2.53	0.41
1:A:63:TYR:HA	1:A:64:LYS:HA	1.88	0.40
1:A:282:SER:HG	2:A:404:GOL:H2	1.85	0.40

There are no symmetry-related clashes.

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	296/301 (98%)	288 (97%)	7 (2%)	1 (0%)	37 17

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	126	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	238/240 (99%)	236 (99%)	2 (1%)	79 57

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	155	TYR
1	A	181	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	43	ASN
1	A	61	ASN
1	A	74	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	149	GLN
1	A	158	GLN
1	A	182	ASN
1	A	215	GLN
1	A	266	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 for Mogul analysis.

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	GOL	A	402	-	5,5,5	0.57	0	5,5,5	1.06	0
2	GOL	A	403	-	5,5,5	0.67	0	5,5,5	1.40	0
2	GOL	A	401	-	5,5,5	0.59	0	5,5,5	1.27	1 (20%)
3	PO4	A	406	-	4,4,4	2.12	1 (25%)	6,6,6	4.37	3 (50%)
3	PO4	A	405	-	4,4,4	0.95	0	6,6,6	0.97	0
3	PO4	A	407	-	4,4,4	2.20	2 (50%)	6,6,6	1.58	1 (16%)
2	GOL	A	404	-	5,5,5	0.32	0	5,5,5	0.77	0

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	GOL	A	402	-	-	2/4/4/4	-
2	GOL	A	403	-	-	4/4/4/4	-
2	GOL	A	401	-	-	0/4/4/4	-
2	GOL	A	404	-	-	0/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	406	PO4	P-O2	-3.83	1.43	1.54
3	A	407	PO4	P-O1	3.23	1.58	1.50
3	A	407	PO4	P-O2	-2.94	1.46	1.54

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	406	PO4	O3-P-O2	-8.07	82.79	107.91
3	A	406	PO4	O4-P-O2	5.12	123.84	107.91
3	A	406	PO4	O4-P-O3	-4.07	95.25	107.91
3	A	407	PO4	O4-P-O2	3.40	118.50	107.91
2	A	401	GOL	C3-C2-C1	-2.11	104.05	111.80

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	403	GOL	O1-C1-C2-C3
2	A	403	GOL	C1-C2-C3-O3
2	A	402	GOL	O1-C1-C2-C3
2	A	403	GOL	O1-C1-C2-O2
2	A	402	GOL	O1-C1-C2-O2
2	A	403	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	403	GOL	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	404	GOL	2	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	298/301 (99%)	0.26	17 (5%) 30 30	5, 12, 23, 28	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	63	TYR	5.5
1	A	155	TYR	4.1
1	A	35	GLY	3.9
1	A	88	ARG	3.8
1	A	28	ASP	3.7
1	A	29	ARG	3.6
1	A	58	CYS	3.6
1	A	193	GLY	3.5
1	A	274	ARG	3.5
1	A	258	TYR	2.8
1	A	194	SER	2.8
1	A	288	ARG	2.7
1	A	14	LYS	2.5
1	A	149	GLN	2.4
1	A	97	THR	2.4
1	A	48	ASP	2.4
1	A	108	ARG	2.2

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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, 95^{th} 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
2	GOL	A	404	6/6	0.77	0.15	35,40,42,43	0
2	GOL	A	403	6/6	0.80	0.15	33,36,37,37	0
2	GOL	A	401	6/6	0.88	0.11	22,26,29,31	0
2	GOL	A	402	6/6	0.89	0.12	16,20,27,31	0
3	PO4	A	405	5/5	0.89	0.10	48,48,50,50	0
3	PO4	A	406	5/5	0.96	0.09	17,19,26,28	0
3	PO4	A	407	5/5	0.97	0.09	16,18,20,22	0
4	CA	A	408	1/1	0.99	0.07	13,13,13,13	0
5	ZN	A	409	1/1	1.00	0.04	9,9,9,9	0

6.5 Other polymers

There are no such residues in this entry.