Full wwPDB X-ray Structure Validation Report (i)

Jun 16, 2024 – 09:15 PM EDT

PDB ID	:	5N4E
Title	:	Prolyl oligopeptidase B from Galerina marginata bound to 35mer hydrolysis and macrocyclization substrate - H698A mutant
Authors	:	Czekster, C.M.; McMahon, S.A.; Ludewig, H.; Naismith, J.H.
Deposited on	:	2017-02-10
Resolution	:	2.90 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:

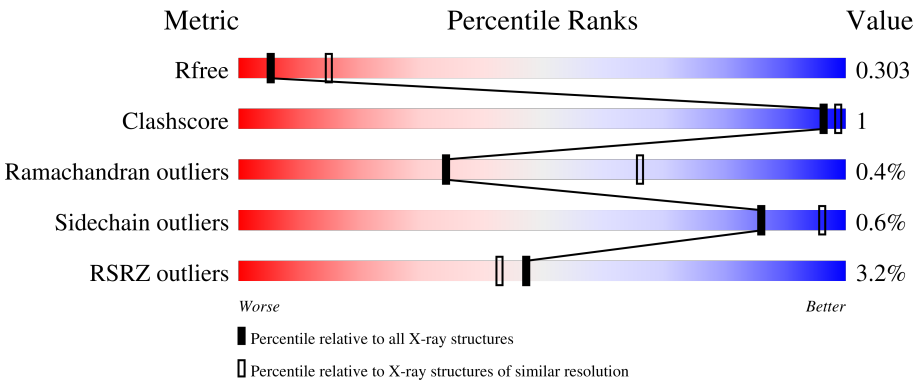
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	765	3%, 90%, • 7%
1	B	765	3%, 89%, • 8%
2	C	35	3%, 83%, • 14%
2	D	35	3%, 83%, 6% 11%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12704 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 Prolyl oligopeptidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	714	5665	3632	945	1077	11	0	1	0
1	B	707	5615	3603	938	1064	10	0	0	0

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	698	ALA	HIS	engineered mutation	UNP H2E7Q8
A	731	MET	-	expression tag	UNP H2E7Q8
A	732	PHE	-	expression tag	UNP H2E7Q8
A	733	ASP	-	expression tag	UNP H2E7Q8
A	734	THR	-	expression tag	UNP H2E7Q8
A	735	ASN	-	expression tag	UNP H2E7Q8
A	736	ALA	-	expression tag	UNP H2E7Q8
A	737	THR	-	expression tag	UNP H2E7Q8
A	738	ARG	-	expression tag	UNP H2E7Q8
A	739	LEU	-	expression tag	UNP H2E7Q8
A	740	PRO	-	expression tag	UNP H2E7Q8
A	741	ILE	-	expression tag	UNP H2E7Q8
A	742	TRP	-	expression tag	UNP H2E7Q8
A	743	GLY	-	expression tag	UNP H2E7Q8
A	744	ILE	-	expression tag	UNP H2E7Q8
A	745	GLY	-	expression tag	UNP H2E7Q8
A	746	CYS	-	expression tag	UNP H2E7Q8
A	747	ASN	-	expression tag	UNP H2E7Q8
A	748	PRO	-	expression tag	UNP H2E7Q8
A	749	TRP	-	expression tag	UNP H2E7Q8
A	750	THR	-	expression tag	UNP H2E7Q8
A	751	ALA	-	expression tag	UNP H2E7Q8
A	752	GLU	-	expression tag	UNP H2E7Q8
A	753	HIS	-	expression tag	UNP H2E7Q8
A	754	VAL	-	expression tag	UNP H2E7Q8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	755	ASP	-	expression tag	UNP H2E7Q8
A	756	GLN	-	expression tag	UNP H2E7Q8
A	757	THR	-	expression tag	UNP H2E7Q8
A	758	LEU	-	expression tag	UNP H2E7Q8
A	759	ALA	-	expression tag	UNP H2E7Q8
A	760	SER	-	expression tag	UNP H2E7Q8
A	761	GLY	-	expression tag	UNP H2E7Q8
A	762	ASN	-	expression tag	UNP H2E7Q8
A	763	ASP	-	expression tag	UNP H2E7Q8
A	764	ILE	-	expression tag	UNP H2E7Q8
A	765	CYS	-	expression tag	UNP H2E7Q8
B	698	ALA	HIS	engineered mutation	UNP H2E7Q8
B	731	MET	-	expression tag	UNP H2E7Q8
B	732	PHE	-	expression tag	UNP H2E7Q8
B	733	ASP	-	expression tag	UNP H2E7Q8
B	734	THR	-	expression tag	UNP H2E7Q8
B	735	ASN	-	expression tag	UNP H2E7Q8
B	736	ALA	-	expression tag	UNP H2E7Q8
B	737	THR	-	expression tag	UNP H2E7Q8
B	738	ARG	-	expression tag	UNP H2E7Q8
B	739	LEU	-	expression tag	UNP H2E7Q8
B	740	PRO	-	expression tag	UNP H2E7Q8
B	741	ILE	-	expression tag	UNP H2E7Q8
B	742	TRP	-	expression tag	UNP H2E7Q8
B	743	GLY	-	expression tag	UNP H2E7Q8
B	744	ILE	-	expression tag	UNP H2E7Q8
B	745	GLY	-	expression tag	UNP H2E7Q8
B	746	CYS	-	expression tag	UNP H2E7Q8
B	747	ASN	-	expression tag	UNP H2E7Q8
B	748	PRO	-	expression tag	UNP H2E7Q8
B	749	TRP	-	expression tag	UNP H2E7Q8
B	750	THR	-	expression tag	UNP H2E7Q8
B	751	ALA	-	expression tag	UNP H2E7Q8
B	752	GLU	-	expression tag	UNP H2E7Q8
B	753	HIS	-	expression tag	UNP H2E7Q8
B	754	VAL	-	expression tag	UNP H2E7Q8
B	755	ASP	-	expression tag	UNP H2E7Q8
B	756	GLN	-	expression tag	UNP H2E7Q8
B	757	THR	-	expression tag	UNP H2E7Q8
B	758	LEU	-	expression tag	UNP H2E7Q8
B	759	ALA	-	expression tag	UNP H2E7Q8
B	760	SER	-	expression tag	UNP H2E7Q8

Continued on next page...

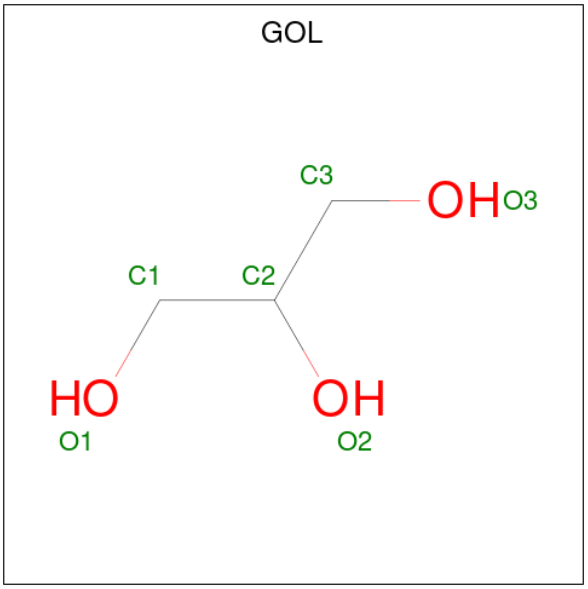
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	761	GLY	-	expression tag	UNP H2E7Q8
B	762	ASN	-	expression tag	UNP H2E7Q8
B	763	ASP	-	expression tag	UNP H2E7Q8
B	764	ILE	-	expression tag	UNP H2E7Q8
B	765	CYS	-	expression tag	UNP H2E7Q8

- Molecule 2 is a protein called Alpha-amanitin proprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	30	236	148	40	47	1	0	0	0
			Total	C	N	O	S			
2	D	31	243	151	42	49	1	0	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	457	457	457	0	0

Continued on next page...

Continued from previous page...

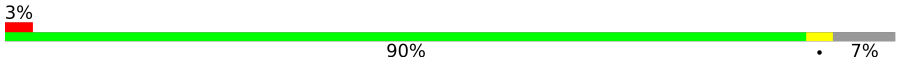
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	B	453	453	453	0	0
			Total	O		
4	C	14	14	14	0	0
			Total	O		
4	D	15	15	15	0	0

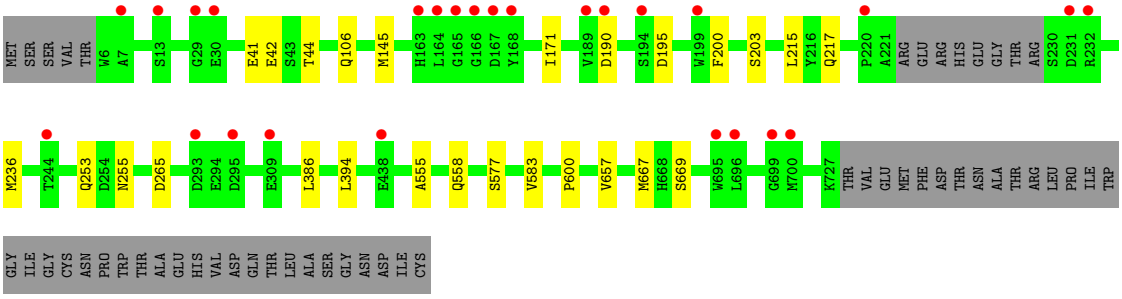
3 Residue-property plots (i)

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: Prolyl oligopeptidase

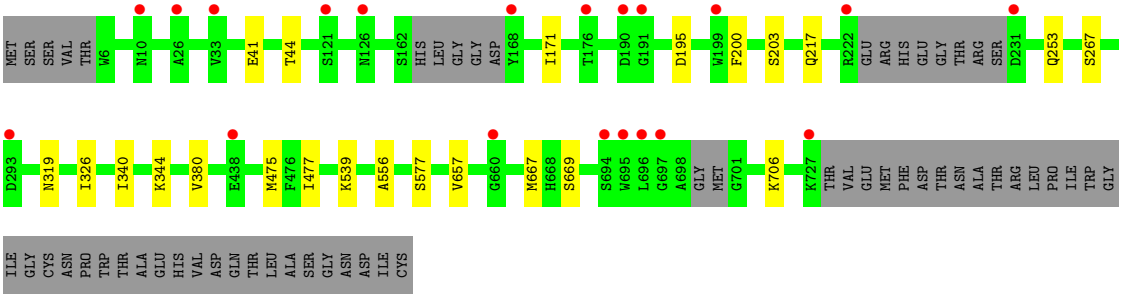
Chain A:





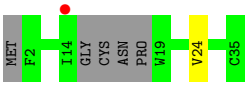
• Molecule 1: Prolyl oligopeptidase

Chain B:



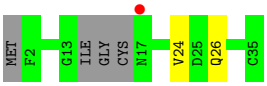
• Molecule 2: Alpha-amanitin proprotein

Chain C:



• Molecule 2: Alpha-amanitin proprotein

Chain D:


MET
F2
G13
ILE
GLY
CYS
N17
V24
D25
Q26
C35

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	99.18Å 115.00Å 141.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	89.23 – 2.90 31.99 – 2.90	Depositor EDS
% Data completeness (in resolution range)	93.9 (89.23-2.90) 94.0 (31.99-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.13 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.249 , 0.295 0.252 , 0.303	Depositor DCC
R_{free} test set	1690 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	34.9	Xtriage
Anisotropy	0.871	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 44.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.54$, $< L^2 > = 0.38$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	12704	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 39.15 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.3149e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: 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	0.36	0/5821	0.57	0/7906
1	B	0.37	0/5768	0.57	0/7832
2	C	0.45	0/241	0.61	0/328
2	D	0.44	0/249	0.58	0/340
All	All	0.37	0/12079	0.57	0/16406

There are no bond length outliers.

There are no bond angle outliers.

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	5665	0	5464	14	0
1	B	5615	0	5425	10	0
2	C	236	0	215	0	0
2	D	243	0	217	0	0
3	B	6	0	8	0	0
4	A	457	0	0	0	0
4	B	453	0	0	0	0
4	C	14	0	0	0	0
4	D	15	0	0	0	0
All	All	12704	0	11329	23	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 1.

All (23) 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:42[B]:GLU:HG3	1:A:44:THR:HG23	1.69	0.72
1:B:475:MET:CE	1:B:556:ALA:HB1	2.24	0.67
1:A:215:LEU:HB3	1:A:236:MET:CE	2.33	0.59
1:B:267:SER:HB2	1:B:319:ASN:HD21	1.68	0.57
1:B:475:MET:HE1	1:B:556:ALA:HB1	1.88	0.55
1:A:106:GLN:HE22	1:A:145:MET:H	1.55	0.53
1:A:386:LEU:CD1	1:A:394:LEU:HD12	2.42	0.50
1:A:253:GLN:HE21	1:A:255:ASN:HD21	1.60	0.49
1:A:555:ALA:HA	1:A:558:GLN:HE21	1.76	0.49
1:B:171:ILE:HD11	1:B:200:PHE:CD1	2.48	0.49
1:A:171:ILE:HD11	1:A:200:PHE:CD1	2.47	0.49
1:B:475:MET:SD	1:B:477:ILE:HD11	2.53	0.48
1:B:41:GLU:HG2	1:B:667:MET:HG3	1.96	0.47
1:A:42[B]:GLU:HB2	1:B:44:THR:HG22	1.98	0.45
1:A:41:GLU:HG2	1:A:667:MET:HG3	1.99	0.45
1:B:203:SER:HB2	1:B:217:GLN:HB3	1.99	0.44
1:A:203:SER:HB2	1:A:217:GLN:HB3	2.00	0.43
1:A:583:VAL:CG2	1:A:600:PRO:HB3	2.48	0.43
1:A:42[B]:GLU:HG3	1:A:44:THR:CG2	2.44	0.43
1:A:583:VAL:HG23	1:A:600:PRO:HB3	2.00	0.43
1:A:657:VAL:HG11	1:A:669:SER:HB2	2.02	0.41
1:B:657:VAL:HG11	1:B:669:SER:HB2	2.02	0.41
1:B:326:ILE:HD12	1:B:340:ILE:HD12	2.04	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	711/765 (93%)	689 (97%)	20 (3%)	2 (0%)	41	71
1	B	699/765 (91%)	680 (97%)	17 (2%)	2 (0%)	41	71
2	C	26/35 (74%)	23 (88%)	2 (8%)	1 (4%)	3	13
2	D	27/35 (77%)	23 (85%)	3 (11%)	1 (4%)	3	13
All	All	1463/1600 (91%)	1415 (97%)	42 (3%)	6 (0%)	34	66

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	24	VAL
2	D	24	VAL
1	A	190	ASP
1	A	577	SER
1	B	577	SER
1	B	380	VAL

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	598/641 (93%)	596 (100%)	2 (0%)	92	98
1	B	593/641 (92%)	588 (99%)	5 (1%)	81	94
2	C	25/29 (86%)	25 (100%)	0	100	100
2	D	26/29 (90%)	25 (96%)	1 (4%)	33	67
All	All	1242/1340 (93%)	1234 (99%)	8 (1%)	86	96

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	195	ASP
1	A	265	ASP
1	B	195	ASP
1	B	253	GLN
1	B	344	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	539	LYS
1	B	706	LYS
2	D	26	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	106	GLN
1	A	187	GLN
1	A	253	GLN
1	A	493	GLN
1	A	558	GLN
1	A	617	GLN
1	B	319	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)

1 ligand is 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	$\#\|Z\| > 2$
3	GOL	B	801	-	5,5,5	0.34	0	5,5,5	0.20	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
3	GOL	B	801	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	714/765 (93%)	0.25	26 (3%)	42	37	19, 30, 53, 98	0
1	B	707/765 (92%)	0.21	20 (2%)	53	49	19, 30, 50, 95	0
2	C	30/35 (85%)	0.47	1 (3%)	46	41	23, 44, 66, 75	0
2	D	31/35 (88%)	0.66	1 (3%)	47	43	27, 47, 74, 80	0
All	All	1482/1600 (92%)	0.24	48 (3%)	47	43	19, 30, 54, 98	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	695	TRP	7.5
1	A	695	TRP	6.8
1	A	700	MET	5.9
1	A	165	GLY	4.9
1	B	190	ASP	4.4
1	A	167	ASP	4.1
1	A	166	GLY	4.0
1	B	696	LEU	3.9
1	B	231	ASP	3.8
1	A	231	ASP	3.8
1	B	438	GLU	3.8
1	B	727	LYS	3.6
1	A	29	GLY	3.5
1	A	163	HIS	3.5
1	B	660	GLY	3.3
1	A	164	LEU	3.2
1	A	295	ASP	3.2
1	B	293	ASP	3.1
1	A	168	TYR	3.1
1	A	189	VAL	3.0
1	A	13	SER	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	220	PRO	2.8
1	B	199	TRP	2.8
1	A	438	GLU	2.6
1	A	699	GLY	2.5
1	A	293	ASP	2.5
1	A	194	SER	2.4
1	A	30	GLU	2.4
1	B	222	ARG	2.3
1	B	33	VAL	2.3
1	B	176	THR	2.3
1	A	199	TRP	2.3
2	C	14	ILE	2.3
1	A	696	LEU	2.3
1	A	244	THR	2.2
1	B	168	TYR	2.2
1	B	26	ALA	2.2
1	B	121	SER	2.2
1	A	309	GLU	2.1
1	A	190	ASP	2.1
1	B	697	GLY	2.1
2	D	17	ASN	2.1
1	B	694	SER	2.1
1	B	10	ASN	2.1
1	B	126	ASN	2.1
1	B	191	GLY	2.1
1	A	7	ALA	2.0
1	A	232	ARG	2.0

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	GOL	B	801	6/6	0.80	0.33	41,43,44,45	0

6.5 Other polymers (i)

There are no such residues in this entry.