



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 06:24 am GMT

PDB ID : 2WW3
Title : STRUCTURE OF THE FAMILY GH92 INVERTING MANNOSIDASE
BT3990 FROM BACTEROIDES THETA IOTAOMICRON VPI-5482 IN
COMPLEX WITH THIOMANNOBIOSIDE
Authors : Suits, M.D.L.; Zhu, Y.; Thompson, A.J.; Gilbert, H.J.; Davies, G.J.
Deposited on : 2009-10-21
Resolution : 2.10 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

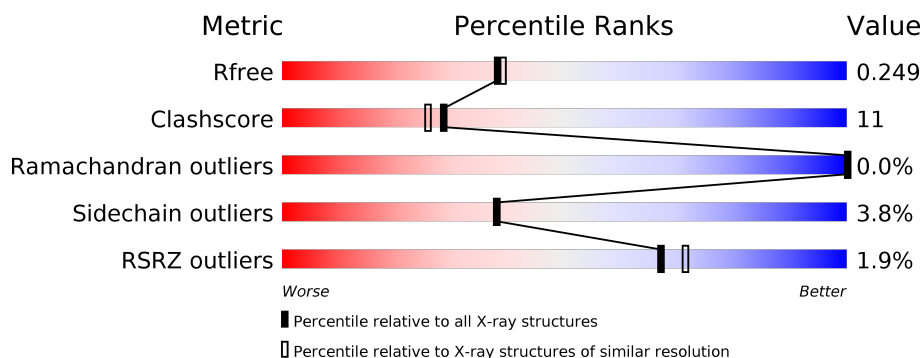
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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}	100719	4243 (2.10-2.10)
Clashscore	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	744	<div> <div>82%</div> <div>15%</div> <div>..</div> </div>
1	B	744	<div> <div>5%</div> <div>72%</div> <div>25%</div> <div>..</div> </div>
1	C	744	<div> <div>84%</div> <div>14%</div> <div>..</div> </div>
1	D	744	<div> <div>2%</div> <div>78%</div> <div>19%</div> <div>..</div> </div>
1	E	744	<div> <div>84%</div> <div>13%</div> <div>..</div> </div>
1	F	744	<div> <div>3%</div> <div>81%</div> <div>17%</div> <div>..</div> </div>

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
4	GOL	A	802	-	-	-	X
4	GOL	C	802	-	-	-	X
4	GOL	E	802	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 37131 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.

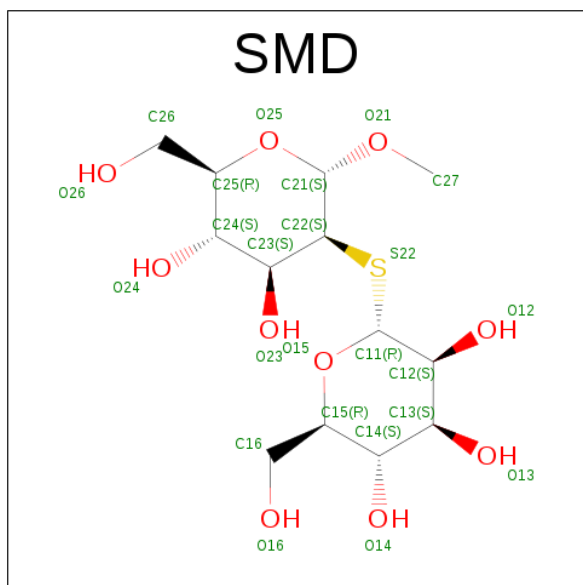
- Molecule 1 is a protein called PUTATIVE ALPHA-1,2-MANNOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	736	Total	C	N	O	S	0	0	0
			5957	3818	981	1125	33			
1	B	736	Total	C	N	O	S	0	0	0
			5937	3806	979	1119	33			
1	C	736	Total	C	N	O	S	0	0	0
			5949	3814	980	1122	33			
1	D	738	Total	C	N	O	S	0	0	0
			5946	3809	977	1127	33			
1	E	736	Total	C	N	O	S	0	1	0
			5945	3811	975	1126	33			
1	F	738	Total	C	N	O	S	0	0	0
			5948	3814	978	1123	33			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Ca	0	0
			1	1		
2	E	1	Total	Ca	0	0
			1	1		
2	B	1	Total	Ca	0	0
			1	1		
2	C	1	Total	Ca	0	0
			1	1		
2	A	1	Total	Ca	0	0
			1	1		
2	F	1	Total	Ca	0	0
			1	1		

- Molecule 3 is METHYL-2-S-(ALPHA-D-MANNOPYRANOSYL)-2-THIO-ALPHA-D-MAN NOPYRANOSIDE (three-letter code: SMD) (formula: C₁₃H₂₄O₁₀S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	S	0	0
			24	13	10	1		
3	B	1	Total	C	O	S	0	0
			24	13	10	1		
3	C	1	Total	C	O	S	0	0
			24	13	10	1		
3	D	1	Total	C	O	S	0	0
			24	13	10	1		
3	E	1	Total	C	O	S	0	0
			24	13	10	1		
3	F	1	Total	C	O	S	0	0
			24	13	10	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

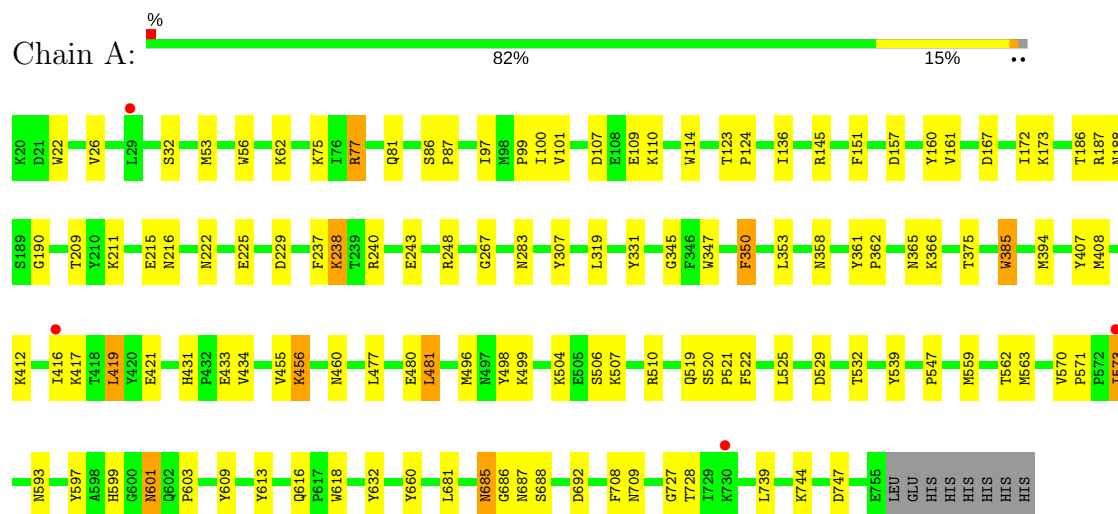
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	230	Total	O	0	0
			230	230		
5	B	162	Total	O	0	0
			162	162		
5	C	322	Total	O	0	0
			322	322		
5	D	155	Total	O	0	0
			155	155		
5	E	217	Total	O	0	0
			217	217		
5	F	195	Total	O	0	0
			195	195		

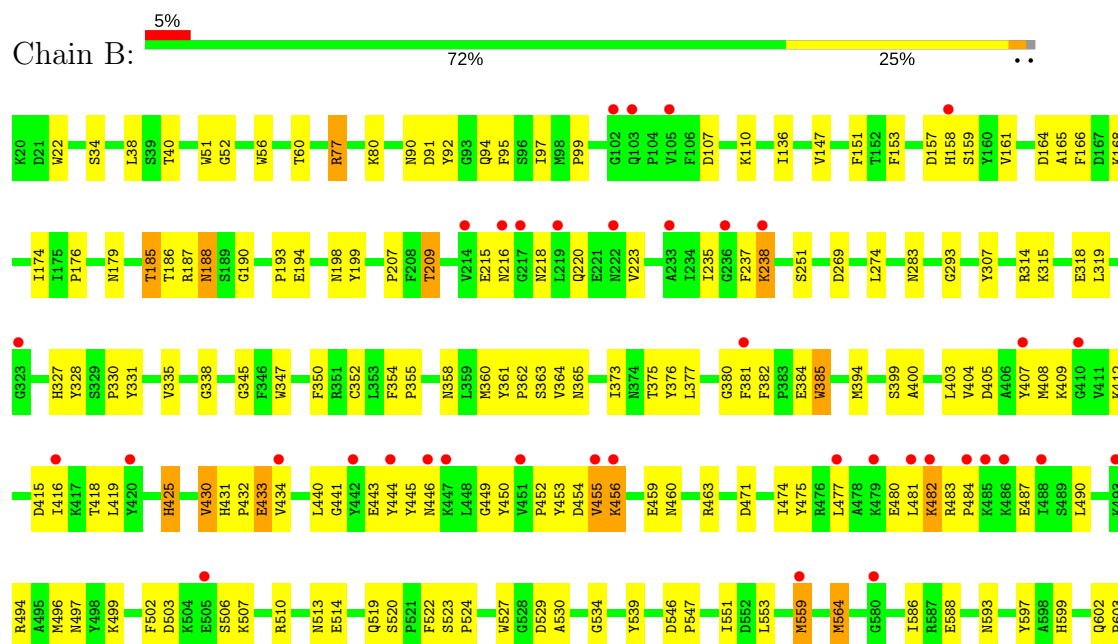
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: PUTATIVE ALPHA-1,2-MANNOSIDASE



• Molecule 1: PUTATIVE ALPHA-1,2-MANNOSIDASE





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	161.26Å 68.59Å 204.01Å 90.00° 94.66° 90.00°	Depositor
Resolution (Å)	203.33 – 2.10 49.90 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.3 (203.33-2.10) 97.3 (49.90-2.10)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.46 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.4.0077	Depositor
R, R_{free}	0.190 , 0.244 0.195 , 0.249	Depositor DCC
R_{free} test set	12627 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	33.6	Xtriage
Anisotropy	0.365	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 41.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	37131	wwPDB-VP
Average B, all atoms (Å ²)	27.0	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 29.05 % 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 1.6577e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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, CA, SMD

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	0.56	0/6138	0.62	0/8330
1	B	0.56	0/6118	0.62	0/8307
1	C	0.57	1/6130 (0.0%)	0.63	0/8320
1	D	0.52	0/6127	0.60	0/8324
1	E	0.55	1/6129 (0.0%)	0.62	1/8322 (0.0%)
1	F	0.48	0/6129	0.59	0/8323
All	All	0.54	2/36771 (0.0%)	0.61	1/49926 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	443	GLU	CB-CG	7.00	1.65	1.52
1	E	443	GLU	CB-CG	5.51	1.62	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	419	LEU	CA-CB-CG	5.40	127.72	115.30

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	5957	0	5610	116	0
1	B	5937	0	5576	218	0
1	C	5949	0	5600	79	0
1	D	5946	0	5557	143	0
1	E	5945	0	5577	83	0
1	F	5948	0	5582	122	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	24	0	23	0	0
3	B	24	0	23	0	0
3	C	24	0	22	0	0
3	D	24	0	22	1	0
3	E	24	0	22	0	0
3	F	24	0	22	0	0
4	A	6	0	8	0	0
4	C	6	0	8	0	0
4	E	6	0	8	0	0
5	A	230	0	0	4	0
5	B	162	0	0	5	0
5	C	322	0	0	3	0
5	D	155	0	0	3	0
5	E	217	0	0	1	0
5	F	195	0	0	4	0
All	All	37131	0	33660	761	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 11.

The worst 5 of 761 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:B:444:TYR:CZ	1:B:455:VAL:HG11	1.50	1.44
1:F:211:LYS:NZ	1:F:222:ASN:OD1	1.63	1.31
1:D:554:MET:O	1:D:559:MET:CE	1.81	1.26
1:D:318:GLU:C	1:D:319:LEU:HD23	1.55	1.26
1:B:446:ASN:HD21	1:B:494:ARG:NH1	1.39	1.21

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	734/744 (99%)	710 (97%)	23 (3%)	1 (0%)	55	57
1	B	734/744 (99%)	693 (94%)	41 (6%)	0	100	100
1	C	734/744 (99%)	711 (97%)	23 (3%)	0	100	100
1	D	736/744 (99%)	700 (95%)	35 (5%)	1 (0%)	55	57
1	E	735/744 (99%)	705 (96%)	30 (4%)	0	100	100
1	F	736/744 (99%)	701 (95%)	35 (5%)	0	100	100
All	All	4409/4464 (99%)	4220 (96%)	187 (4%)	2 (0%)	100	100

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	529	ASP
1	A	529	ASP

5.3.2 Protein sidechains

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	632/643 (98%)	610 (96%)	22 (4%)	41	42
1	B	627/643 (98%)	599 (96%)	28 (4%)	32	30
1	C	630/643 (98%)	609 (97%)	21 (3%)	43	45
1	D	627/643 (98%)	605 (96%)	22 (4%)	41	42
1	E	629/643 (98%)	604 (96%)	25 (4%)	36	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	628/643 (98%)	603 (96%)	25 (4%)	36	36
All	All	3773/3858 (98%)	3630 (96%)	143 (4%)	38	38

5 of 143 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	539	TYR
1	D	319	LEU
1	F	482	LYS
1	C	593	ASN
1	D	77	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 112 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	593	ASN
1	D	264	ASN
1	F	446	ASN
1	C	601	ASN
1	C	713	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 15 ligands modelled in this entry, 6 are monoatomic - leaving 9 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
3	SMD	A	801	2	25,25,25	0.83	1 (4%)	30,36,36	1.01	2 (6%)
4	GOL	A	802	-	5,5,5	0.37	0	5,5,5	0.28	0
3	SMD	B	801	2	25,25,25	0.92	2 (8%)	30,36,36	1.26	3 (10%)
3	SMD	C	801	2	25,25,25	0.87	1 (4%)	30,36,36	1.09	1 (3%)
4	GOL	C	802	-	5,5,5	0.25	0	5,5,5	0.39	0
3	SMD	D	801	2	25,25,25	0.94	3 (12%)	30,36,36	0.82	0
3	SMD	E	801	2	25,25,25	0.86	1 (4%)	30,36,36	0.84	0
4	GOL	E	802	-	5,5,5	0.27	0	5,5,5	0.52	0
3	SMD	F	801	2	25,25,25	0.95	3 (12%)	30,36,36	1.01	1 (3%)

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	SMD	A	801	2	-	0/10/50/50	0/2/2/2
4	GOL	A	802	-	-	0/4/4/4	0/0/0/0
3	SMD	B	801	2	-	0/10/50/50	0/2/2/2
3	SMD	C	801	2	-	0/10/50/50	0/2/2/2
4	GOL	C	802	-	-	0/4/4/4	0/0/0/0
3	SMD	D	801	2	-	0/10/50/50	0/2/2/2
3	SMD	E	801	2	-	0/10/50/50	0/2/2/2
4	GOL	E	802	-	-	0/4/4/4	0/0/0/0
3	SMD	F	801	2	-	0/10/50/50	0/2/2/2

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	801	SMD	C11-S22	-2.60	1.76	1.81
3	E	801	SMD	C22-S22	-2.52	1.77	1.82
3	D	801	SMD	C11-S22	-2.35	1.76	1.81
3	F	801	SMD	C22-S22	-2.26	1.77	1.82
3	D	801	SMD	C22-S22	-2.21	1.78	1.82

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	801	SMD	O12-C12-C11	-2.37	105.55	110.27
3	C	801	SMD	C23-C24-C25	-2.19	106.36	110.22
3	B	801	SMD	O15-C15-C14	2.00	113.35	109.66
3	B	801	SMD	C11-C12-C13	2.16	115.31	110.69
3	A	801	SMD	O15-C11-C12	2.44	113.63	110.28

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	801	SMD	1	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	736/744 (98%)	-0.02	4 (0%) 90 92	17, 23, 33, 40	0
1	B	736/744 (98%)	0.50	40 (5%) 26 33	21, 33, 44, 55	0
1	C	736/744 (98%)	0.01	2 (0%) 93 94	15, 21, 32, 39	0
1	D	738/744 (99%)	0.19	16 (2%) 62 67	18, 28, 38, 45	0
1	E	736/744 (98%)	-0.05	3 (0%) 92 93	17, 23, 33, 40	0
1	F	738/744 (99%)	0.20	19 (2%) 56 62	21, 30, 41, 46	0
All	All	4420/4464 (99%)	0.14	84 (1%) 67 71	15, 26, 40, 55	0

The worst 5 of 84 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	442	TYR	4.8
1	B	444	TYR	4.2
1	B	420	TYR	4.1
1	F	217	GLY	4.0
1	F	172	ILE	3.7

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
4	GOL	C	802	6/6	0.94	0.15	5.56	30,33,34,34	0
4	GOL	A	802	6/6	0.90	0.16	4.11	48,49,49,50	0
4	GOL	E	802	6/6	0.91	0.17	2.65	40,41,42,46	0
3	SMD	D	801	24/24	0.93	0.12	0.19	28,30,32,35	0
3	SMD	B	801	24/24	0.93	0.11	-0.51	34,37,39,40	0
3	SMD	F	801	24/24	0.95	0.10	-1.47	31,34,37,38	0
3	SMD	C	801	24/24	0.96	0.10	-1.53	20,25,28,31	0
2	CA	C	800	1/1	0.99	0.09	-2.46	26,26,26,26	0
2	CA	D	800	1/1	0.98	0.09	-2.92	29,29,29,29	0
2	CA	B	800	1/1	0.99	0.05	-3.05	36,36,36,36	0
3	SMD	A	801	24/24	0.95	0.08	-3.33	22,27,29,31	0
2	CA	F	800	1/1	0.96	0.06	-3.48	31,31,31,31	0
3	SMD	E	801	24/24	0.97	0.07	-4.10	22,27,30,31	0
2	CA	A	800	1/1	0.99	0.06	-6.22	25,25,25,25	0
2	CA	E	800	1/1	0.98	0.06	-8.09	24,24,24,24	0

6.5 Other polymers [i](#)

There are no such residues in this entry.