



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 09:29 pm BST

PDB ID : 2WVZ
Title : Structure of the Family GH92 Inverting Mannosidase BT3990 from *Bacteroides thetaiotaomicron* VPI-5482
Authors : Suits, M.D.L.; Zhu, Y.; Thompson, A.; Gilbert, H.J.; Davies, G.J.
Deposited on : 2009-10-21
Resolution : 2.40 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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 : 2.11
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) : 2.11

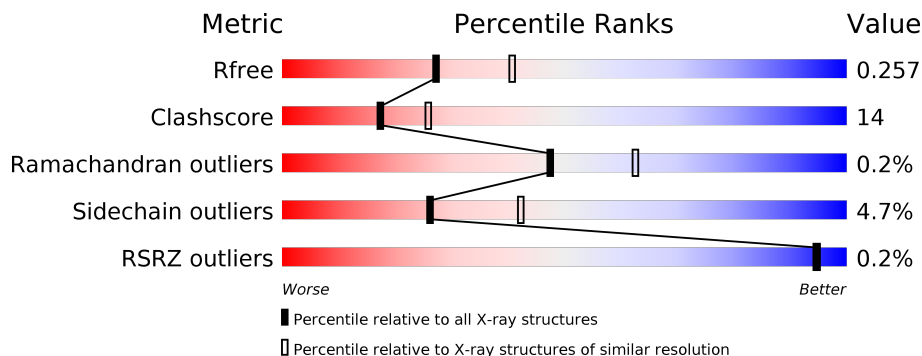
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 on 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 $\leq 5\%$. 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	
1	B	744	
1	C	744	
1	D	744	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 24878 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.

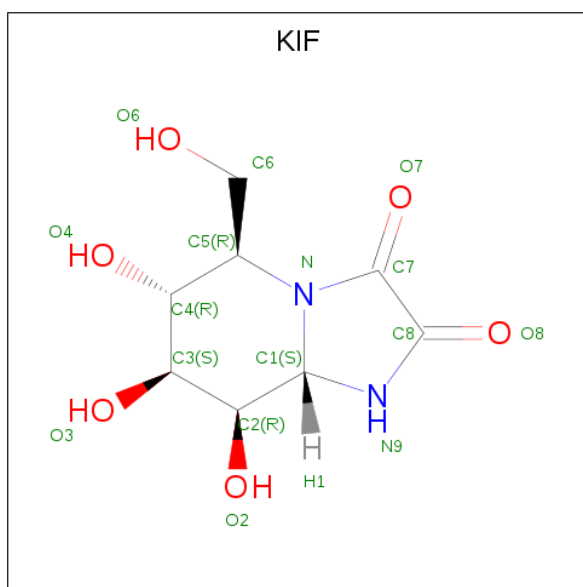
- Molecule 1 is a protein called PUTATIVE ALPHA-1,2-MANNOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	738	5994	3847	986	1127	34	0	2	0
1	B	742	6037	3870	997	1135	35	0	3	0
1	C	737	5987	3839	984	1130	34	0	3	0
1	D	736	5970	3828	983	1125	34	0	1	0

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

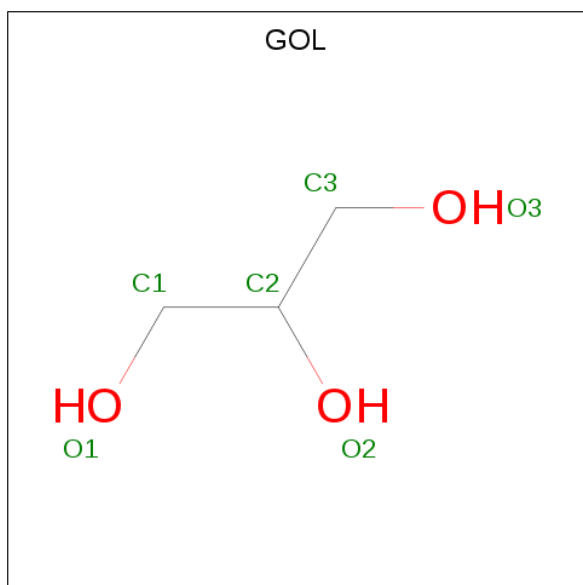
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ca	0	0
			1	1		
2	A	1	Total	Ca	0	0
			1	1		
2	D	1	Total	Ca	0	0
			1	1		
2	C	1	Total	Ca	0	0
			1	1		

- Molecule 3 is KIFUNENSINE (three-letter code: KIF) (formula: C₈H₁₂N₂O₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	Total 16	8	2	6	0	0
3	B	1	Total 16	8	2	6	0	0
3	C	1	Total 16	8	2	6	0	0
3	D	1	Total 16	8	2	6	0	0

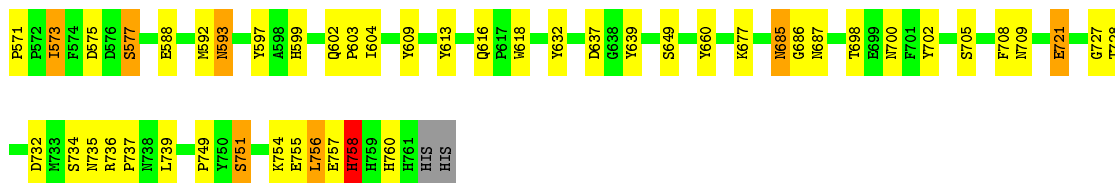
- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



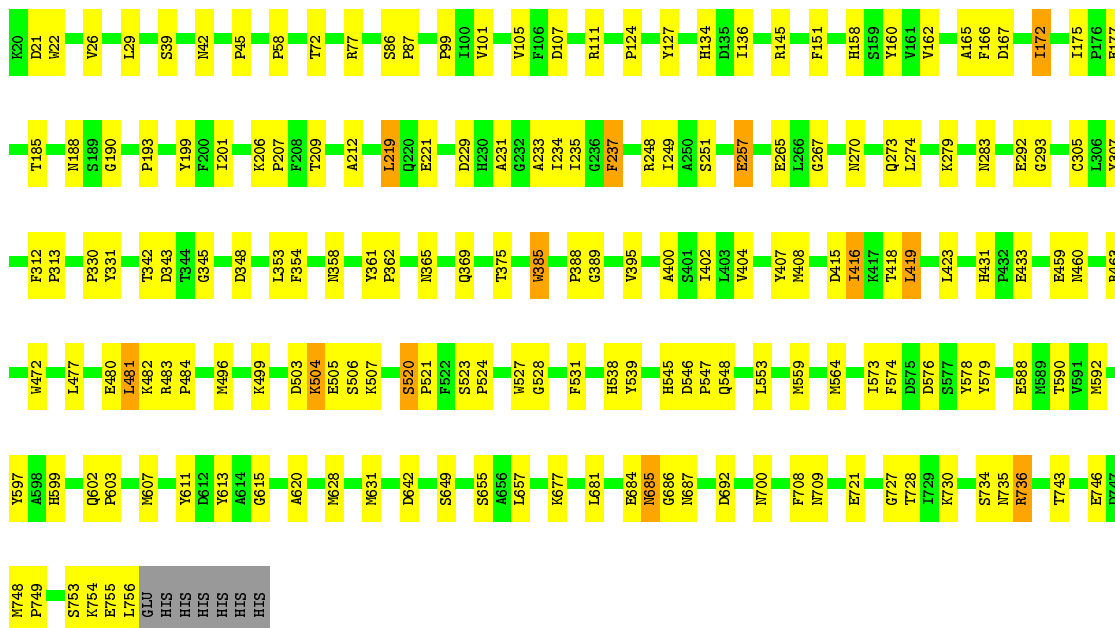
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

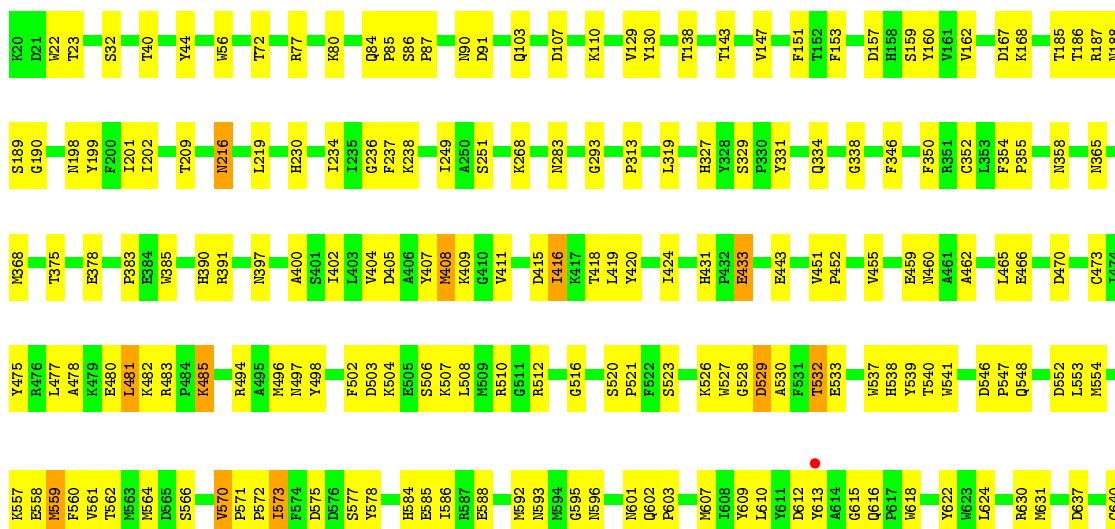
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	244	Total	O	0	0
			244	244		
5	B	179	Total	O	0	0
			179	179		
5	C	203	Total	O	0	0
			203	203		
5	D	166	Total	O	0	0
			166	166		

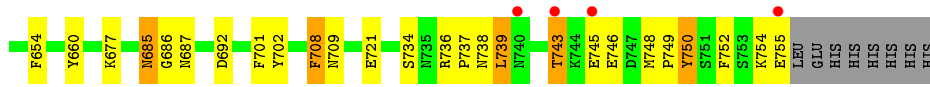


• Molecule 1: PUTATIVE ALPHA-1,2-MANNOSEDASE



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





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	68.90Å 150.19Å 382.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	191.14 – 2.40 59.05 – 2.40	Depositor EDS
% Data completeness (in resolution range)	88.1 (191.14-2.40) 88.1 (59.05-2.40)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.08 (at 2.40Å)	Xtrriage
Refinement program	REFMAC 5.4.0077	Depositor
R, R_{free}	0.201 , 0.257 0.201 , 0.257	Depositor DCC
R_{free} test set	6859 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	33.7	Xtrriage
Anisotropy	0.071	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 35.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	24878	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% of the height of the origin peak. No significant pseudotranslation is detected.*

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

5.1 Standard geometry

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

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	0.59	0/6183	0.66	0/8389
1	B	0.55	0/6231	0.65	0/8452
1	C	0.55	0/6177	0.64	0/8379
1	D	0.53	0/6154	0.64	0/8348
All	All	0.55	0/24745	0.65	0/33568

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	756	LEU	Peptide

5.2 Too-close contacts

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	5994	0	5664	133	0
1	B	6037	0	5699	164	0
1	C	5987	0	5655	152	0
1	D	5970	0	5640	201	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
3	A	16	0	10	0	0
3	B	16	0	10	1	0
3	C	16	0	10	1	0
3	D	16	0	10	1	0
4	A	6	0	8	0	0
4	B	12	0	16	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
5	A	244	0	0	12	0
5	B	179	0	0	19	0
5	C	203	0	0	9	0
5	D	166	0	0	12	0
All	All	24878	0	22738	648	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 14.

The worst 5 of 648 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:C:507:LYS:CD	1:C:559:MET:HE3	1.25	1.56
1:C:507:LYS:HE3	1:C:559:MET:CE	1.32	1.55
1:C:507:LYS:CE	1:C:559:MET:CE	1.88	1.49
1:C:507:LYS:CD	1:C:559:MET:CE	1.94	1.46
1:C:507:LYS:CE	1:C:559:MET:HE2	1.44	1.34

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	738/744 (99%)	707 (96%)	30 (4%)	1 (0%)	51	68
1	B	743/744 (100%)	700 (94%)	39 (5%)	4 (0%)	29	41
1	C	738/744 (99%)	695 (94%)	43 (6%)	0	100	100
1	D	735/744 (99%)	692 (94%)	42 (6%)	1 (0%)	51	68
All	All	2954/2976 (99%)	2794 (95%)	154 (5%)	6 (0%)	47	62

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	757	GLU
1	A	755	GLU
1	B	755	GLU
1	B	758	HIS
1	B	529	ASP

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	637/643 (99%)	616 (97%)	21 (3%)	38	57
1	B	643/643 (100%)	610 (95%)	33 (5%)	24	39
1	C	637/643 (99%)	609 (96%)	28 (4%)	28	45
1	D	635/643 (99%)	597 (94%)	38 (6%)	19	31
All	All	2552/2572 (99%)	2432 (95%)	120 (5%)	26	42

5 of 120 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	72	THR
1	C	419	LEU
1	D	692	ASP
1	C	77	ARG
1	C	237	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 67 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	713	HIS
1	C	283	ASN
1	D	593	ASN
1	B	716	ASN
1	C	216	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 4 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	KIF	A	801	2	16,17,17	2.11	2 (12%)	15,26,26	2.05	3 (20%)
3	KIF	B	801	2	16,17,17	2.02	2 (12%)	15,26,26	2.04	4 (26%)
4	GOL	D	802	-	5,5,5	0.39	0	5,5,5	0.46	0
4	GOL	B	802	-	5,5,5	0.28	0	5,5,5	0.52	0
4	GOL	A	802	-	5,5,5	0.31	0	5,5,5	0.27	0
4	GOL	C	802	-	5,5,5	0.35	0	5,5,5	0.66	0
4	GOL	B	803	-	5,5,5	0.38	0	5,5,5	0.23	0
3	KIF	C	801	2	16,17,17	1.74	2 (12%)	15,26,26	1.91	3 (20%)
3	KIF	D	801	2	16,17,17	1.86	2 (12%)	15,26,26	2.00	3 (20%)

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	KIF	A	801	2	-	0/2/38/38	0/2/2/2
3	KIF	B	801	2	-	2/2/38/38	0/2/2/2
4	GOL	D	802	-	-	0/4/4/4	-
4	GOL	B	802	-	-	2/4/4/4	-
4	GOL	A	802	-	-	2/4/4/4	-
4	GOL	C	802	-	-	2/4/4/4	-
4	GOL	B	803	-	-	4/4/4/4	-
3	KIF	C	801	2	-	1/2/38/38	0/2/2/2
3	KIF	D	801	2	-	2/2/38/38	0/2/2/2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	801	KIF	C7-C8	-6.79	1.38	1.53
3	D	801	KIF	C7-C8	-6.23	1.39	1.53
3	A	801	KIF	C7-C8	-5.85	1.40	1.53
3	C	801	KIF	C7-C8	-5.82	1.40	1.53
3	A	801	KIF	C1-N9	5.28	1.50	1.44

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	801	KIF	C8-C7-N	5.35	110.50	106.20
3	C	801	KIF	C7-C8-N9	5.34	110.46	105.42

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	801	KIF	C7-C8-N9	4.98	110.12	105.42
3	B	801	KIF	C7-C8-N9	4.73	109.89	105.42
3	A	801	KIF	C7-C8-N9	4.66	109.82	105.42

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	801	KIF	C4-C5-C6-O6
4	B	803	GOL	O1-C1-C2-C3
4	B	803	GOL	C1-C2-C3-O3
4	B	803	GOL	O2-C2-C3-O3
4	B	802	GOL	C1-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	801	KIF	1	0
3	C	801	KIF	1	0
3	D	801	KIF	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	738/744 (99%)	-0.56	1 (0%) 95 95	15, 26, 40, 62	0
1	B	742/744 (99%)	-0.43	1 (0%) 95 95	20, 33, 51, 60	0
1	C	737/744 (99%)	-0.57	0 100 100	17, 30, 43, 50	0
1	D	736/744 (98%)	-0.37	5 (0%) 87 86	17, 35, 63, 76	0
All	All	2953/2976 (99%)	-0.48	7 (0%) 95 94	15, 31, 54, 76	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	745	GLU	2.8
1	D	755	GLU	2.8
1	A	756	LEU	2.6
1	D	613	TYR	2.5
1	D	743	THR	2.3

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. 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(\AA^2)	Q<0.9
4	GOL	B	802	6/6	0.77	0.16	55,55,56,56	0
4	GOL	B	803	6/6	0.84	0.22	59,60,61,61	0
4	GOL	A	802	6/6	0.88	0.20	54,57,57,57	0
4	GOL	D	802	6/6	0.91	0.15	44,45,45,45	0
4	GOL	C	802	6/6	0.91	0.16	44,46,46,48	0
3	KIF	B	801	16/16	0.94	0.14	31,36,37,38	0
3	KIF	A	801	16/16	0.94	0.12	21,26,30,30	0
3	KIF	D	801	16/16	0.94	0.13	33,38,40,40	0
3	KIF	C	801	16/16	0.94	0.11	32,35,37,38	0
2	CA	B	800	1/1	0.98	0.08	30,30,30,30	0
2	CA	A	800	1/1	0.98	0.11	32,32,32,32	0
2	CA	D	800	1/1	0.98	0.11	45,45,45,45	0
2	CA	C	800	1/1	0.99	0.08	26,26,26,26	0

6.5 Other polymers [i](#)

There are no such residues in this entry.