



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 24, 2024 – 11:03 PM EDT

PDB ID : 6IWQ
Title : Crystal structure of GalNAc-T7 with Mn²⁺
Authors : Yu, C.; Yin, Y.X.
Deposited on : 2018-12-06
Resolution : 2.95 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

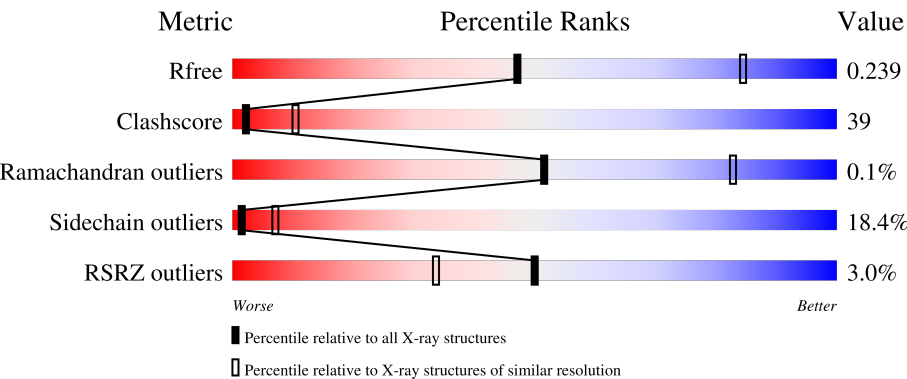
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.95 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	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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 $\leq 5\%$. 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	597	
1	B	597	
1	C	597	
1	D	597	
1	E	597	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	597	<div><div></div><div>4%</div><div>42%</div><div>39%</div><div>10%</div><div>9%</div></div>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 26834 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 N-acetylgalactosaminyltransferase 7.

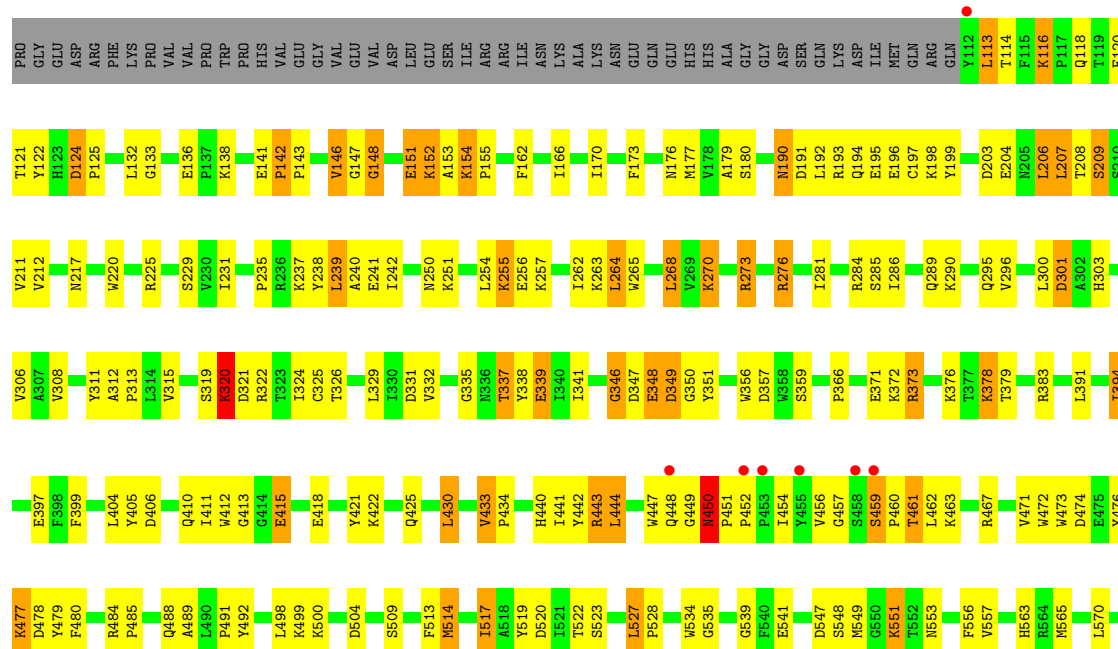
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	546	Total	C	N	O	S	0	0	0
			4440	2855	757	805	23			
1	B	546	Total	C	N	O	S	0	0	0
			4440	2855	757	805	23			
1	C	546	Total	C	N	O	S	0	0	0
			4440	2855	757	805	23			
1	D	546	Total	C	N	O	S	0	0	0
			4440	2855	757	805	23			
1	E	546	Total	C	N	O	S	0	0	0
			4440	2855	757	805	23			
1	F	546	Total	C	N	O	S	0	0	0
			4440	2855	757	805	23			

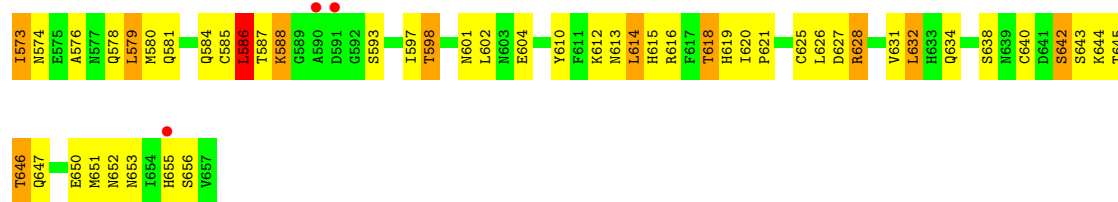
- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

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

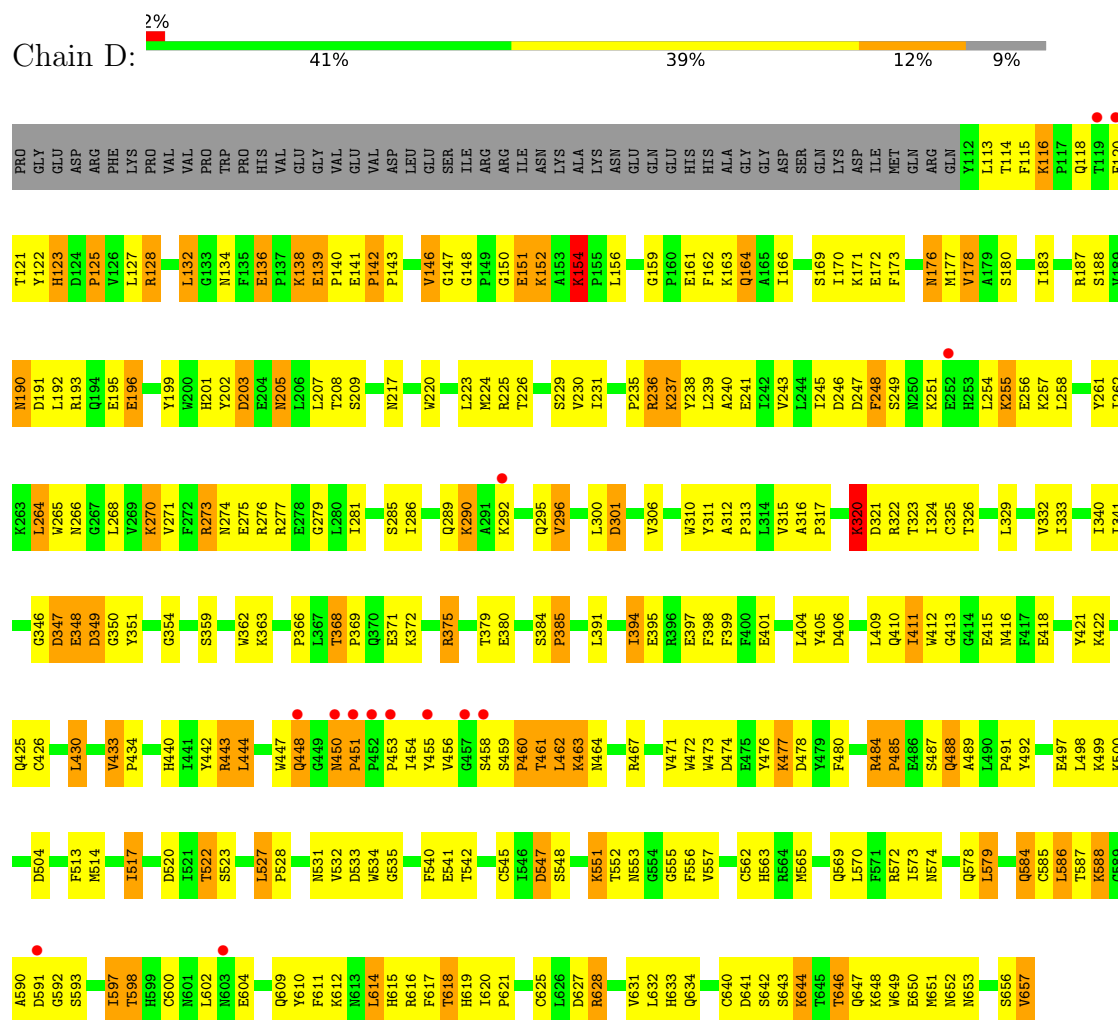
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	42	Total 42	O 42	0	0
3	B	31	Total 31	O 31	0	0
3	C	42	Total 42	O 42	0	0
3	D	25	Total 25	O 25	0	0
3	E	30	Total 30	O 30	0	0
3	F	18	Total 18	O 18	0	0

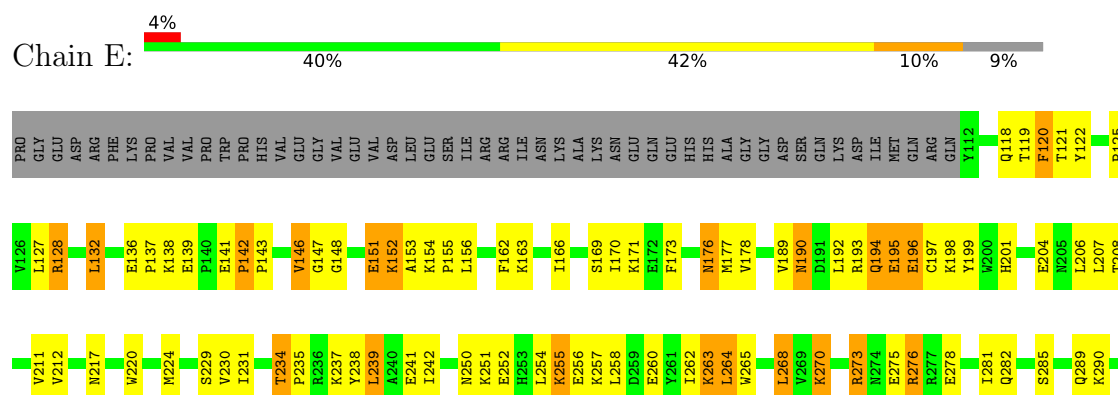




• Molecule 1: N-acetylgalactosaminyltransferase 7



• Molecule 1: N-acetylgalactosaminyltransferase 7





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	137.69Å 158.23Å 251.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.25 – 2.95 49.27 – 2.95	Depositor EDS
% Data completeness (in resolution range)	88.9 (49.25-2.95) 89.0 (49.27-2.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.66 (at 2.96Å)	Xtriage
Refinement program	PHENIX (dev_2400: ???)	Depositor
R, R_{free}	0.225 , 0.239 0.225 , 0.239	Depositor DCC
R_{free} test set	5179 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	55.2	Xtriage
Anisotropy	0.251	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 34.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	26834	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.92	9/4569 (0.2%)	0.86	19/6192 (0.3%)
1	B	0.80	9/4569 (0.2%)	0.85	17/6192 (0.3%)
1	C	0.82	5/4569 (0.1%)	0.84	15/6192 (0.2%)
1	D	0.76	5/4569 (0.1%)	0.82	15/6192 (0.2%)
1	E	0.74	7/4569 (0.2%)	0.80	17/6192 (0.3%)
1	F	0.70	6/4569 (0.1%)	0.81	11/6192 (0.2%)
All	All	0.79	41/27414 (0.1%)	0.83	94/37152 (0.3%)

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	2
1	C	0	2
All	All	0	4

The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	137	PRO	N-CD	9.43	1.61	1.47
1	E	137	PRO	N-CD	8.29	1.59	1.47
1	E	212	VAL	CB-CG2	-7.39	1.37	1.52
1	A	212	VAL	CB-CG2	-7.38	1.37	1.52
1	F	212	VAL	CB-CG2	-7.37	1.37	1.52

The worst 5 of 94 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	133	GLY	N-CA-C	-8.75	91.23	113.10
1	C	586	LEU	CA-CB-CG	7.68	132.98	115.30
1	E	586	LEU	CA-CB-CG	7.68	132.96	115.30
1	F	586	LEU	CA-CB-CG	7.67	132.94	115.30
1	C	459	SER	C-N-CD	6.37	141.79	128.40

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	553	ASN	Peptide
1	B	645	THR	Peptide
1	C	346	GLY	Peptide
1	C	348	GLU	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	4440	0	4329	269	0
1	B	4440	0	4329	441	0
1	C	4440	0	4327	268	0
1	D	4440	0	4331	374	0
1	E	4440	0	4329	387	0
1	F	4440	0	4331	356	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	42	0	0	9	0
3	B	31	0	0	10	0
3	C	42	0	0	7	0
3	D	25	0	0	4	0
3	E	30	0	0	6	0
3	F	18	0	0	8	0
All	All	26834	0	25976	2060	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 39.

The worst 5 of 2060 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:E:545:CYS:HB3	1:E:562:CYS:SG	1.57	1.43
1:B:614:LEU:CD1	1:B:616:ARG:HD3	1.55	1.36
1:F:128:ARG:HG3	1:F:201:HIS:CD2	1.62	1.32
1:A:332:VAL:HG13	1:A:442:TYR:CD2	1.65	1.30
1:E:513:PHE:CE1	1:E:517:ILE:HD11	1.67	1.30

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	544/597 (91%)	515 (95%)	29 (5%)	0	100	100
1	B	544/597 (91%)	518 (95%)	25 (5%)	1 (0%)	47	79
1	C	544/597 (91%)	517 (95%)	27 (5%)	0	100	100
1	D	544/597 (91%)	521 (96%)	22 (4%)	1 (0%)	47	79
1	E	544/597 (91%)	520 (96%)	24 (4%)	0	100	100
1	F	544/597 (91%)	518 (95%)	25 (5%)	1 (0%)	47	79
All	All	3264/3582 (91%)	3109 (95%)	152 (5%)	3 (0%)	51	83

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	459	SER
1	F	451	PRO
1	D	491	PRO

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	480/525 (91%)	381 (79%)	99 (21%)	1	4
1	B	480/525 (91%)	395 (82%)	85 (18%)	2	8
1	C	480/525 (91%)	406 (85%)	74 (15%)	2	11
1	D	480/525 (91%)	389 (81%)	91 (19%)	1	6
1	E	480/525 (91%)	387 (81%)	93 (19%)	1	6
1	F	480/525 (91%)	391 (82%)	89 (18%)	1	7
All	All	2880/3150 (91%)	2349 (82%)	531 (18%)	1	7

5 of 531 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	225	ARG
1	F	277	ARG
1	F	221	SER
1	F	614	LEU
1	C	207	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 100 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	524	HIS
1	E	425	GLN
1	F	653	ASN
1	D	568	ASN
1	D	653	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](#)

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

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	546/597 (91%)	-0.23	7 (1%) 77 61	19, 33, 61, 95	0
1	B	546/597 (91%)	0.19	24 (4%) 34 21	21, 42, 89, 153	0
1	C	546/597 (91%)	-0.03	10 (1%) 68 51	26, 38, 70, 109	0
1	D	546/597 (91%)	0.08	14 (2%) 56 39	27, 52, 81, 121	0
1	E	546/597 (91%)	0.14	22 (4%) 38 25	27, 48, 86, 142	0
1	F	546/597 (91%)	0.32	22 (4%) 38 25	37, 65, 91, 129	0
All	All	3276/3582 (91%)	0.08	99 (3%) 50 34	19, 46, 85, 153	0

The worst 5 of 99 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	657	VAL	11.1
1	B	456	VAL	8.4
1	B	656	SER	8.2
1	B	455	TYR	6.1
1	A	458	SER	5.7

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 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
2	MN	A	701	1/1	0.55	0.25	87,87,87,87	0
2	MN	B	701	1/1	0.88	0.19	80,80,80,80	0
2	MN	E	701	1/1	0.88	0.18	90,90,90,90	0
2	MN	D	701	1/1	0.92	0.28	74,74,74,74	0
2	MN	F	701	1/1	0.92	0.19	82,82,82,82	0
2	MN	C	701	1/1	0.95	0.29	73,73,73,73	0

6.5 Other polymers [i](#)

There are no such residues in this entry.