



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 2, 2019 – 07:54 PM EST

PDB ID : 6ED1
Title : Bacteroides dorei Beta-glucuronidase
Authors : Biernat, K.A.; Redinbo, M.R.
Deposited on : 2018-08-08
Resolution : 2.90 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
Xtriage (Phenix)	:	1.13
EDS	:	rb-20031633
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031633

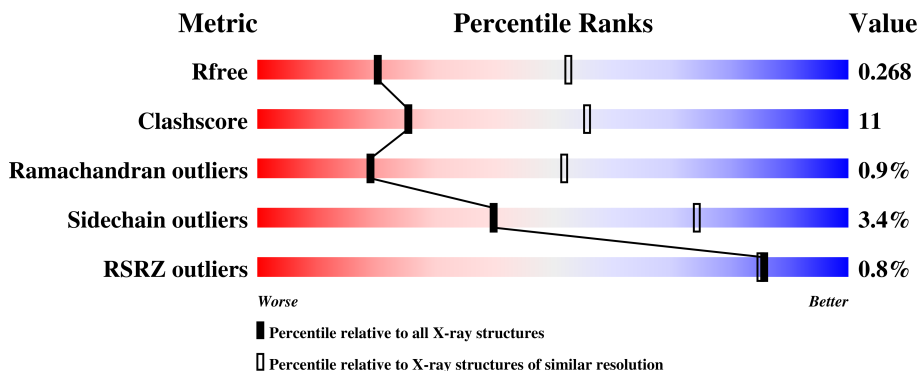
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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}	111664	1716 (2.90-2.90)
Clashscore	122126	1924 (2.90-2.90)
Ramachandran outliers	120053	1884 (2.90-2.90)
Sidechain outliers	120020	1886 (2.90-2.90)
RSRZ outliers	108989	1669 (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 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	879	<div> <div>2%</div> <div>69% 27% ..</div> </div>
1	B	879	<div> <div>71% 26% ..</div> </div>
1	C	879	<div> <div>72% 24% ..</div> </div>
1	D	879	<div> <div>2%</div> <div>65% 31% ..</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 27956 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 Glycosyl hydrolase family 2, sugar binding domain protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	860	Total	C	N	O	S	0	0	0
			6901	4411	1179	1281	30			
1	B	860	Total	C	N	O	S	0	0	0
			6914	4416	1181	1287	30			
1	C	860	Total	C	N	O	S	0	0	0
			6909	4413	1178	1289	29			
1	D	860	Total	C	N	O	S	0	0	0
			6857	4382	1167	1279	29			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP C3R9X4
A	2	HIS	-	expression tag	UNP C3R9X4
A	3	HIS	-	expression tag	UNP C3R9X4
A	4	HIS	-	expression tag	UNP C3R9X4
A	5	HIS	-	expression tag	UNP C3R9X4
A	6	HIS	-	expression tag	UNP C3R9X4
A	7	HIS	-	expression tag	UNP C3R9X4
A	8	SER	-	expression tag	UNP C3R9X4
A	9	SER	-	expression tag	UNP C3R9X4
A	10	GLY	-	expression tag	UNP C3R9X4
A	11	VAL	-	expression tag	UNP C3R9X4
A	12	ASP	-	expression tag	UNP C3R9X4
A	13	LEU	-	expression tag	UNP C3R9X4
A	14	GLY	-	expression tag	UNP C3R9X4
A	15	THR	-	expression tag	UNP C3R9X4
A	16	GLU	-	expression tag	UNP C3R9X4
A	17	ASN	-	expression tag	UNP C3R9X4
A	18	LEU	-	expression tag	UNP C3R9X4
A	19	TYR	-	expression tag	UNP C3R9X4
A	20	PHE	-	expression tag	UNP C3R9X4
A	21	GLN	-	expression tag	UNP C3R9X4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	SER	-	expression tag	UNP C3R9X4
A	23	ASN	-	expression tag	UNP C3R9X4
B	1	MET	-	initiating methionine	UNP C3R9X4
B	2	HIS	-	expression tag	UNP C3R9X4
B	3	HIS	-	expression tag	UNP C3R9X4
B	4	HIS	-	expression tag	UNP C3R9X4
B	5	HIS	-	expression tag	UNP C3R9X4
B	6	HIS	-	expression tag	UNP C3R9X4
B	7	HIS	-	expression tag	UNP C3R9X4
B	8	SER	-	expression tag	UNP C3R9X4
B	9	SER	-	expression tag	UNP C3R9X4
B	10	GLY	-	expression tag	UNP C3R9X4
B	11	VAL	-	expression tag	UNP C3R9X4
B	12	ASP	-	expression tag	UNP C3R9X4
B	13	LEU	-	expression tag	UNP C3R9X4
B	14	GLY	-	expression tag	UNP C3R9X4
B	15	THR	-	expression tag	UNP C3R9X4
B	16	GLU	-	expression tag	UNP C3R9X4
B	17	ASN	-	expression tag	UNP C3R9X4
B	18	LEU	-	expression tag	UNP C3R9X4
B	19	TYR	-	expression tag	UNP C3R9X4
B	20	PHE	-	expression tag	UNP C3R9X4
B	21	GLN	-	expression tag	UNP C3R9X4
B	22	SER	-	expression tag	UNP C3R9X4
B	23	ASN	-	expression tag	UNP C3R9X4
C	1	MET	-	initiating methionine	UNP C3R9X4
C	2	HIS	-	expression tag	UNP C3R9X4
C	3	HIS	-	expression tag	UNP C3R9X4
C	4	HIS	-	expression tag	UNP C3R9X4
C	5	HIS	-	expression tag	UNP C3R9X4
C	6	HIS	-	expression tag	UNP C3R9X4
C	7	HIS	-	expression tag	UNP C3R9X4
C	8	SER	-	expression tag	UNP C3R9X4
C	9	SER	-	expression tag	UNP C3R9X4
C	10	GLY	-	expression tag	UNP C3R9X4
C	11	VAL	-	expression tag	UNP C3R9X4
C	12	ASP	-	expression tag	UNP C3R9X4
C	13	LEU	-	expression tag	UNP C3R9X4
C	14	GLY	-	expression tag	UNP C3R9X4
C	15	THR	-	expression tag	UNP C3R9X4
C	16	GLU	-	expression tag	UNP C3R9X4
C	17	ASN	-	expression tag	UNP C3R9X4

Continued on next page...

Continued from previous page...

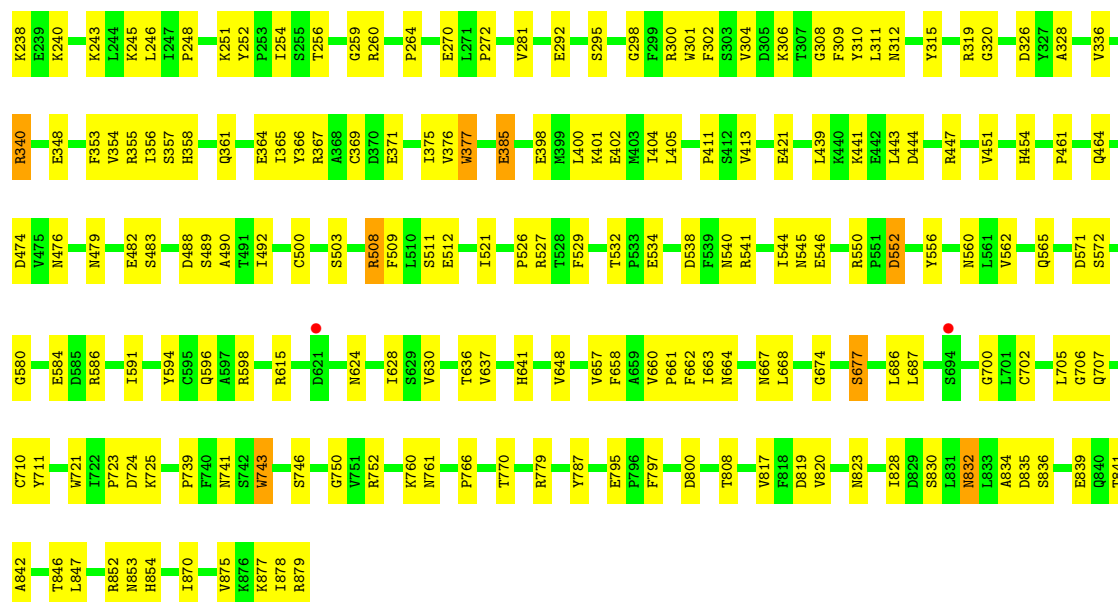
Chain	Residue	Modelled	Actual	Comment	Reference
C	18	LEU	-	expression tag	UNP C3R9X4
C	19	TYR	-	expression tag	UNP C3R9X4
C	20	PHE	-	expression tag	UNP C3R9X4
C	21	GLN	-	expression tag	UNP C3R9X4
C	22	SER	-	expression tag	UNP C3R9X4
C	23	ASN	-	expression tag	UNP C3R9X4
D	1	MET	-	initiating methionine	UNP C3R9X4
D	2	HIS	-	expression tag	UNP C3R9X4
D	3	HIS	-	expression tag	UNP C3R9X4
D	4	HIS	-	expression tag	UNP C3R9X4
D	5	HIS	-	expression tag	UNP C3R9X4
D	6	HIS	-	expression tag	UNP C3R9X4
D	7	HIS	-	expression tag	UNP C3R9X4
D	8	SER	-	expression tag	UNP C3R9X4
D	9	SER	-	expression tag	UNP C3R9X4
D	10	GLY	-	expression tag	UNP C3R9X4
D	11	VAL	-	expression tag	UNP C3R9X4
D	12	ASP	-	expression tag	UNP C3R9X4
D	13	LEU	-	expression tag	UNP C3R9X4
D	14	GLY	-	expression tag	UNP C3R9X4
D	15	THR	-	expression tag	UNP C3R9X4
D	16	GLU	-	expression tag	UNP C3R9X4
D	17	ASN	-	expression tag	UNP C3R9X4
D	18	LEU	-	expression tag	UNP C3R9X4
D	19	TYR	-	expression tag	UNP C3R9X4
D	20	PHE	-	expression tag	UNP C3R9X4
D	21	GLN	-	expression tag	UNP C3R9X4
D	22	SER	-	expression tag	UNP C3R9X4
D	23	ASN	-	expression tag	UNP C3R9X4

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

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

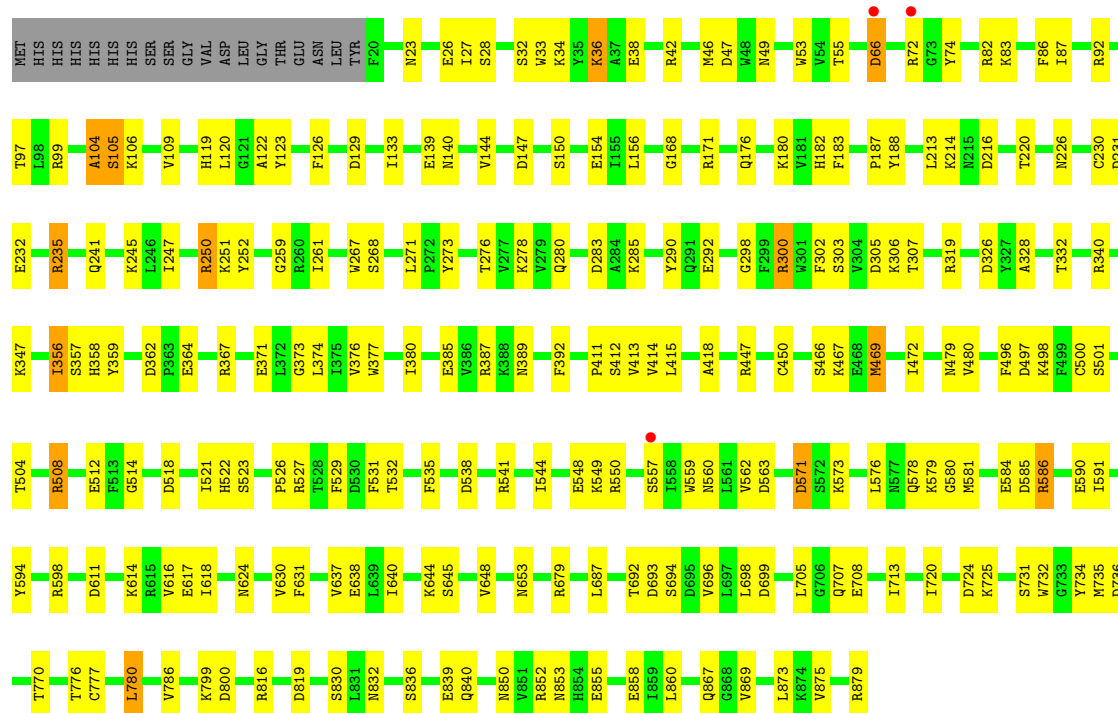
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	98	Total 98	O 98	0	0
3	B	123	Total 123	O 123	0	0
3	C	80	Total 80	O 80	0	0
3	D	70	Total 70	O 70	0	0



- Molecule 1: Glycosyl hydrolase family 2, sugar binding domain protein

Chain C: 72% 24% ..



- Molecule 1: Glycosyl hydrolase family 2, sugar binding domain protein

Chain D: 2% 65% 31% ..



G838	K685	S572	I492	Q406	I294	V210	V103
T841	L686	G580	K498	K407	R300	A211	A104
L847	L687	M581	S501	Y408	W301	F212	S105
V848	L691	L582	Y502	S412	F302	L213	K106
V849	D693	R586	S503	L415	V304	K214	Y110
R852	S694	R587	T504	A418	D305	D216	L115
D695	V696	K588	A505	M419	K306	F217	L116
L860	L705	Y592	K506	N420	T307	T219	K117
L861	G706	Y593	P507	E421	G308	T220	T118
S862	Q707	Y594	S511	D424	F309	H119	H119
I870	Y711	R598	E512	Y425	L311	L120	G121
V875	I720	D601	D518	H426	N312	G122	A122
K876	P603	I602	R520	Q428	F228	V227	I130
K877	P603	I521	I521	A429	L229	L229	T131
I878	K725	D611	H522	L432	I236	V237	V136
R879	R752	S623	S623	E435	R323	G137	G137
	K760	V616	Y524	Q436	H324	Q241	E138
	N761	E617	T525	L436	D326	L242	E139
	T762	I618	P526	K441	T332	K243	N140
	F768	C619	T528	D444	R340	L141	L141
	T769	D620	F529	D444	R341	L244	L142
	T770	D621	F531	Y446	N352	I247	K145
	R779	S622	T532	R447	F353	P248	V146
	L780	N624	P533	L448	V854	P253	D147
	D781	R625	E534	S449	R355	I254	N148
	D784	G626	F535	C450	I356	S255	I155
	F791	S629	Q536	H454	I365	T256	L156
	F797	V630	L537	A455	R260	S257	P157
	R802	N633	N540	F456	D370	V258	F166
	K803	V637	R541	T457	E371	G259	F166
	T808	H641	I544	Y462	I375	I261	I169
	V817	R649	E546	T463	V376	E262	Y170
	F818	V657	K549	Q464	W377	P272	R171
	D819	P661	R550	S465	S378	Y273	R172
	V820	F662	P551	S466	E379	L274	K180
	S821	E666	Y553	K467	I380	V277	Y188
	V822	S673	I554	E468	K388	V279	V191
	D829	G674	F555	N469	V281	Q280	P192
	N832	A675	G556	F470	H394	V282	V193
	D835	L676	Y557	S471	N395	C282	R194
	S836	S677	I558	V475	E398	D283	I195
	Y837	K681	W559	N476	A284	K285	E199
			N560	G477	N286	K285	E199
			D567	V478	L400	N287	A205
			G570	M479	K401	G287	S206
			D571	E482	E402	E288	P207
				S483	M403	E292	Q208
					L405	V293	L209

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	242.66Å 101.36Å 168.60Å 90.00° 94.97° 90.00°	Depositor
Resolution (Å)	29.54 – 2.90 29.54 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.5 (29.54-2.90) 99.5 (29.54-2.90)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.14 (at 2.90Å)	Xtriage
Refinement program	PHENIX (1.14_3260)	Depositor
R, R_{free}	0.178 , 0.268 0.178 , 0.268	Depositor DCC
R_{free} test set	1988 reflections (2.21%)	wwPDB-VP
Wilson B-factor (Å ²)	33.2	Xtriage
Anisotropy	0.665	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 54.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	27956	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.46% 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:
NA

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	0.48	0/7071	0.62	0/9580
1	B	0.50	0/7084	0.64	0/9596
1	C	0.46	0/7079	0.61	0/9592
1	D	0.46	0/7027	0.62	0/9532
All	All	0.47	0/28261	0.62	0/38300

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	D	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	223	VAL	Peptide
1	D	303	SER	Peptide

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	6901	0	6764	152	2
1	B	6914	0	6771	142	0
1	C	6909	0	6760	156	0
1	D	6857	0	6660	183	2
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	98	0	0	6	0
3	B	123	0	0	6	0
3	C	80	0	0	7	0
3	D	70	0	0	8	0
All	All	27956	0	26955	621	2

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 621 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:668:LEU:HD12	1:A:681:LYS:HG2	1.50	0.93
1:B:159:VAL:HG11	1:B:385:GLU:HG2	1.50	0.93
1:D:617:GLU:OE1	1:D:626:ARG:NH1	2.07	0.86
1:D:220:THR:HG22	1:D:247:ILE:HA	1.58	0.86
1:B:852:ARG:NH1	1:C:292:GLU:OE1	2.10	0.85

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:292:GLU:OE1	1:D:852:ARG:NH1[1_565]	2.06	0.14
1:A:855:GLU:OE2	1:D:22:SER:OG[1_565]	2.11	0.09

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	858/879 (98%)	783 (91%)	69 (8%)	6 (1%)	24	58
1	B	858/879 (98%)	786 (92%)	63 (7%)	9 (1%)	17	49
1	C	858/879 (98%)	780 (91%)	69 (8%)	9 (1%)	17	49
1	D	858/879 (98%)	771 (90%)	79 (9%)	8 (1%)	19	52
All	All	3432/3516 (98%)	3120 (91%)	280 (8%)	32 (1%)	19	52

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	830	SER
1	B	830	SER
1	C	104	ALA
1	C	105	SER
1	B	32	SER

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	750/775 (97%)	717 (96%)	33 (4%)	31	65
1	B	752/775 (97%)	727 (97%)	25 (3%)	41	75
1	C	752/775 (97%)	731 (97%)	21 (3%)	47	80
1	D	739/775 (95%)	717 (97%)	22 (3%)	44	78
All	All	2993/3100 (96%)	2892 (97%)	101 (3%)	40	74

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	529	PHE
1	B	819	ASP
1	D	471	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	552	ASP
1	B	677	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	575	ASN
1	D	215	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 4 ligands modelled in this entry, 4 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 ⓘ

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	860/879 (97%)	-0.39	5 (0%) 89 89	11, 27, 44, 65	0
1	B	860/879 (97%)	-0.44	2 (0%) 94 95	11, 23, 41, 57	0
1	C	860/879 (97%)	-0.35	3 (0%) 93 94	14, 27, 48, 61	0
1	D	860/879 (97%)	-0.16	19 (2%) 62 59	15, 33, 59, 77	0
All	All	3440/3516 (97%)	-0.34	29 (0%) 86 85	11, 27, 51, 77	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	465	SER	3.0
1	D	72	ARG	3.0
1	D	218	THR	3.0
1	C	66	ASP	2.8
1	D	263	ASN	2.8

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
2	NA	A	901	1/1	0.88	0.12	27,27,27,27	0
2	NA	C	901	1/1	0.89	0.15	21,21,21,21	0
2	NA	D	901	1/1	0.91	0.13	17,17,17,17	0
2	NA	B	901	1/1	0.99	0.17	15,15,15,15	0

6.5 Other polymers [i](#)

There are no such residues in this entry.