



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 23, 2017 – 10:30 PM EST

PDB ID : 5T4Y
Title : Crystal structure of BT1762-1763
Authors : van den Berg, B.
Deposited on : 2016-08-30
Resolution : 3.10 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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
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)	:	rb-20028442

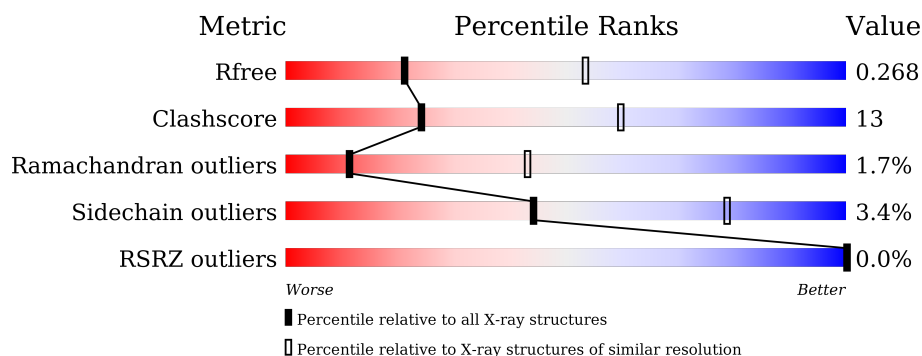
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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}	91344	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	576	<div> <div>76%</div> <div>19%</div> <div>• •</div> </div>
1	B	576	<div> <div>76%</div> <div>20%</div> <div>• •</div> </div>
2	C	1041	<div> <div>48%</div> <div>27%</div> <div>•</div> <div>22%</div> </div>
2	D	1041	<div> <div>48%</div> <div>27%</div> <div>•</div> <div>22%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 21727 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 SusD homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	552	Total	C	N	O	S	0	2	0
			4450	2818	737	874	21			
1	B	553	Total	C	N	O	S	0	2	0
			4460	2824	740	875	21			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	553	HIS	-	expression tag	UNP Q8A6W4
A	554	HIS	-	expression tag	UNP Q8A6W4
A	555	HIS	-	expression tag	UNP Q8A6W4
A	556	HIS	-	expression tag	UNP Q8A6W4
A	557	HIS	-	expression tag	UNP Q8A6W4
A	558	HIS	-	expression tag	UNP Q8A6W4
B	553	HIS	-	expression tag	UNP Q8A6W4
B	554	HIS	-	expression tag	UNP Q8A6W4
B	555	HIS	-	expression tag	UNP Q8A6W4
B	556	HIS	-	expression tag	UNP Q8A6W4
B	557	HIS	-	expression tag	UNP Q8A6W4
B	558	HIS	-	expression tag	UNP Q8A6W4

- Molecule 2 is a protein called SusC homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	807	Total	C	N	O	S	0	0	0
			6406	4057	1089	1241	19			
2	C	807	Total	C	N	O	S	0	0	0
			6406	4057	1089	1241	19			

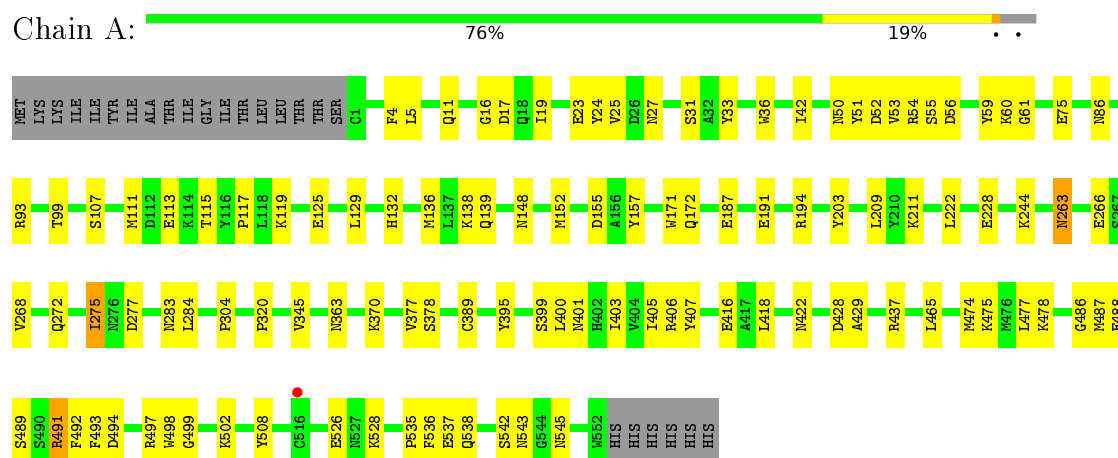
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total 2	Mg 2	0	0
3	A	2	Total 2	Mg 2	0	0
3	C	1	Total 1	Mg 1	0	0

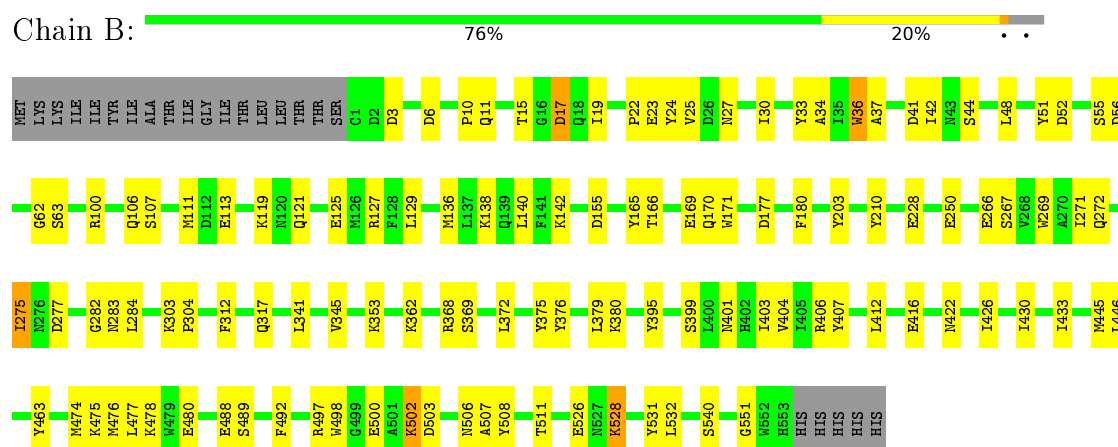
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

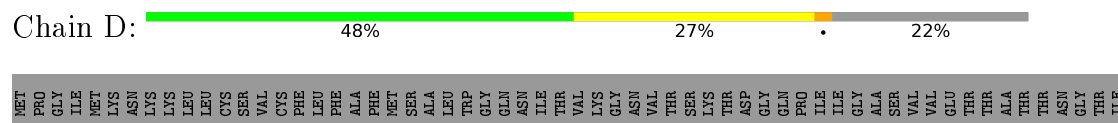
• Molecule 1: SusD homolog

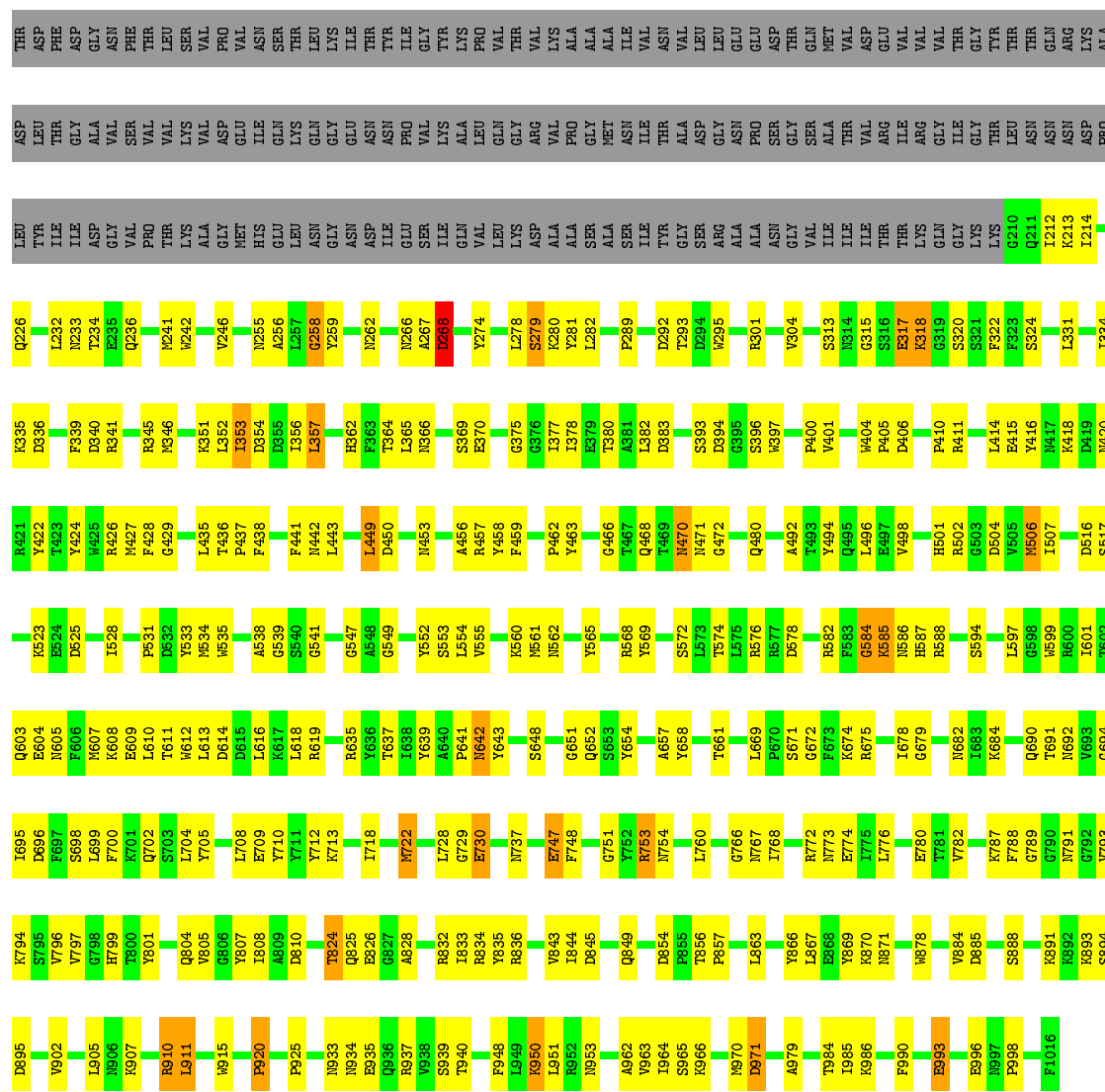


• Molecule 1: SusD homolog

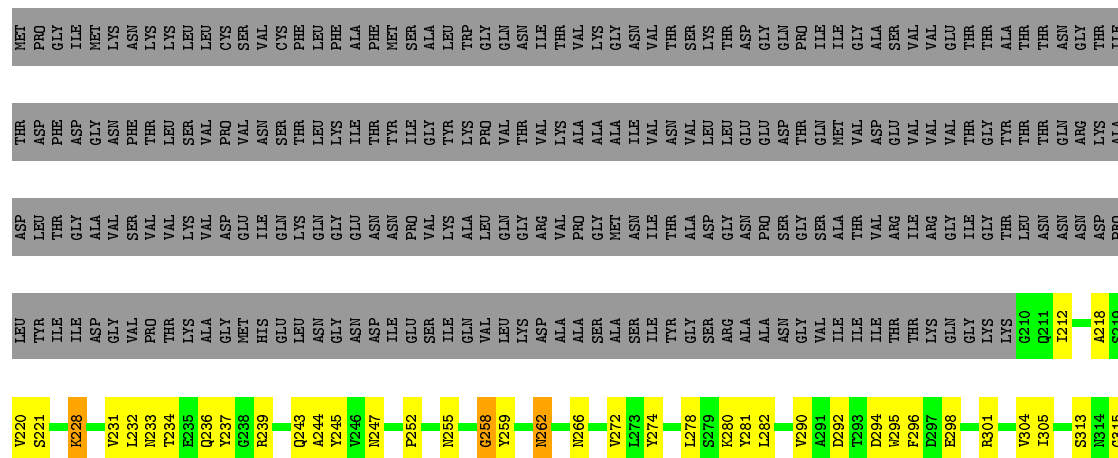


• Molecule 2: SusC homolog





- Molecule 2: SusC homolog



T1009	L911	D817	S703	E604	R502	G403	K318
L912	E818	L704	L705	M605	G503	D406	E319
N913	H822	N706	G706	K608	H505	R407	S320
A914	A823	S707	L708	D614	H506	R408	S321
P925	T824	E709	E709	D615	L507	M409	F322
A926	Q825	G827	G827	L616	D516	P410	L325
L927	E828	A828	Y712	L617	S520	R411	L328
T928	A828		L718	L618		E415	Y328
N933	R832		L719	Q624	D525	Y422	L331
N934	R833		T720	Q628	L528	M425	L334
E935	R834		G731		L529	R426	K335
Q936	Y835		G732	M632	T530	M427	D336
R937	R836		S733	L633	Y533	F428	T337
V938	D837		R734	T637	G429	G429	D338
V939	L838		G735	L638	H534	D430	
T940			L736	Y639	H535	A431	R345
V943	D845		G739	A640	G541	Y432	R346
K950	E846		K742	P641	Y546	V433	D349
L951	R847		E747	M642	G547	L435	Y350
	Q849		F749		L546	T436	K351
Q955	N850		G751	D647	G549	P437	L352
L956	L852		Y752	S648		F438	L353
S965	P857		L764	G651	Y562	K439	D354
K966	L865		L768	T656	S564	G440	D355
	F872		L776	A657	Y565	F441	L356
	D873		L777	Y658	R568	M442	
	L874		E777	I660	Y569	L449	G360
R974	T875			L669		D450	L361
	V884			P670	T574	Y451	H362
	D885				L575	K454	F363
	L886			F673	R576	Q455	T364
	L887			K674	R577	A456	R367
	S888			R675	D578	P462	E370
	D889			M676		Y463	V371
	V890			Q677	R582		
K997	K891			I678	K585	G466	P374
K988	R892			G679	H586	T467	G375
N999	K893				H587	Q468	G376
F990	S894			M682	R588	T469	L377
T991	D895			I683	Y589	M470	L378
G992				K684	A590	N471	E379
E993				M685		G472	
	N901				P593		D383
	V902				S594	M483	
	G903			T691			V390
	F904			L695	L597	M487	S396
	L905			D696	G598	E497	L397
	N906			T697		V498	
	K907			K701	T601	G499	P400
	G908			R813	T602	K500	V401
V1007	T909			R814	Q702	H501	G402
I1008	R910						

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	110.83Å 152.09Å 253.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	130.35 – 3.10 130.35 – 2.99	Depositor EDS
% Data completeness (in resolution range)	98.9 (130.35-3.10) 98.6 (130.35-2.99)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.96 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, R_{free}	0.198 , 0.270 0.194 , 0.268	Depositor DCC
R_{free} test set	3942 reflections (5.09%)	DCC
Wilson B-factor (Å ²)	68.3	Xtriage
Anisotropy	0.382	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 56.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	21727	wwPDB-VP
Average B, all atoms (Å ²)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% 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: MG

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	0.44	0/4562	0.62	2/6190 (0.0%)
1	B	0.49	1/4573 (0.0%)	0.66	1/6205 (0.0%)
2	C	0.59	8/6571 (0.1%)	0.76	4/8909 (0.0%)
2	D	0.52	0/6571	0.76	3/8909 (0.0%)
All	All	0.52	9/22277 (0.0%)	0.71	10/30213 (0.0%)

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
2	D	0	4

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	709	GLU	CG-CD	-6.93	1.41	1.51
2	C	565	TYR	CE1-CZ	-6.63	1.29	1.38
2	C	565	TYR	CE2-CZ	-6.15	1.30	1.38
2	C	565	TYR	CG-CD1	-6.06	1.31	1.39
1	B	36	TRP	CB-CG	-5.99	1.39	1.50

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	911	LEU	CA-CB-CG	6.11	129.35	115.30
2	D	911	LEU	CB-CG-CD2	5.82	120.89	111.00
1	A	275	ILE	CG1-CB-CG2	-5.81	98.62	111.40
2	D	449	LEU	CA-CB-CG	5.42	127.77	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	565	TYR	CA-CB-CG	5.39	123.65	113.40

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	317	GLU	Peptide
2	D	353	ILE	Peptide
2	D	470	ASN	Peptide
2	D	584	GLY	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	4450	0	4219	80	0
1	B	4460	0	4227	81	0
2	C	6406	0	6058	208	0
2	D	6406	0	6058	222	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	1	0	0	0	0
All	All	21727	0	20562	549	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 13.

The worst 5 of 549 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)
2:D:427:MET:CE	2:C:427:MET:SD	2.35	1.12
2:D:427:MET:HE2	2:C:427:MET:SD	1.95	1.04
2:D:427:MET:HE3	2:C:427:MET:SD	2.00	1.00
1:B:275:ILE:HG22	1:B:284:LEU:HD11	1.47	0.95
2:D:365:LEU:HD13	2:D:427:MET:SD	2.12	0.90

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	552/576 (96%)	500 (91%)	48 (9%)	4 (1%)	26	65
1	B	553/576 (96%)	514 (93%)	36 (6%)	3 (0%)	34	72
2	C	805/1041 (77%)	716 (89%)	68 (8%)	21 (3%)	7	32
2	D	805/1041 (77%)	714 (89%)	72 (9%)	19 (2%)	7	33
All	All	2715/3234 (84%)	2444 (90%)	224 (8%)	47 (2%)	11	43

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	608	LYS
2	D	826	GLU
2	D	993	GLU
2	C	732	GLY
2	C	971	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	475/495 (96%)	469 (99%)	6 (1%)	76	91
1	B	476/495 (96%)	469 (98%)	7 (2%)	72	90
2	C	673/869 (77%)	644 (96%)	29 (4%)	35	72

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	D	673/869 (77%)	637 (95%)	36 (5%)	28 64
All	All	2297/2728 (84%)	2219 (97%)	78 (3%)	44 79

5 of 78 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	801	TYR
1	B	17	ASP
2	C	969	ARG
2	D	834	ARG
2	D	911	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	110	GLN
2	C	913	ASN
2	C	791	ASN
2	D	849	GLN
2	C	442	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 5 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	552/576 (95%)	-0.38	1 (0%) 95 91	47, 83, 111, 146	0
1	B	553/576 (96%)	-0.33	0 100 100	37, 65, 95, 124	0
2	C	807/1041 (77%)	-0.43	0 100 100	37, 64, 104, 150	0
2	D	807/1041 (77%)	-0.39	0 100 100	35, 71, 111, 146	0
All	All	2719/3234 (84%)	-0.39	1 (0%) 100 100	35, 70, 108, 150	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	516	CYS	2.1

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(\AA^2)	Q<0.9
3	MG	C	1101	1/1	0.96	0.22	1.24	47,47,47,47	0
3	MG	A	601	1/1	0.97	0.16	-0.39	28,28,28,28	0
3	MG	B	602	1/1	0.94	0.17	-0.93	113,113,113,113	0
3	MG	A	602	1/1	0.83	0.09	-1.63	84,84,84,84	0
3	MG	B	601	1/1	0.96	0.12	-1.83	4,4,4,4	0

6.5 Other polymers [i](#)

There are no such residues in this entry.