



Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 02:30 pm BST

PDB ID : 4RU1
Title : Crystal structure of carbohydrate transporter ACEI_1806 from *Acidothermus cellulolyticus* 11B, TARGET EFI-510965, in complex with myo-inositol
Authors : Patskovsky, Y.; Toro, R.; Bhosle, R.; Al Obaidi, N.; Morisco, L.L.; Wasserman, S.R.; Chamala, S.; Attonito, J.D.; Scott Glenn, A.; Chowdhury, S.; Lafleur, J.; Hillerich, B.; Siede, R.D.; Love, J.; Whalen, K.L.; Gerlt, J.A.; Almo, S.C.; Enzyme Function Initiative (EFI)
Deposited on : 2014-11-17
Resolution : 1.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

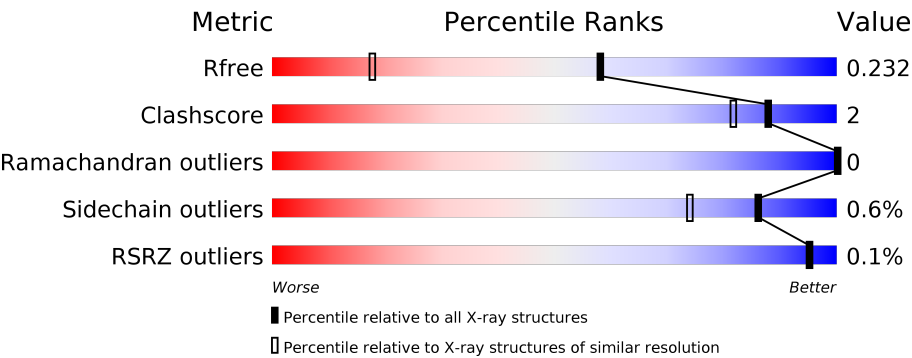
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.50 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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	304	<div><div>89%</div><div><div></div><div></div><div></div><div></div></div><div>6% 5%</div></div>
1	B	304	<div><div>90%</div><div><div></div><div></div><div></div><div></div></div><div>5% 5%</div></div>
1	C	304	<div><div>87%</div><div><div></div><div></div><div></div><div></div></div><div>7% 6%</div></div>
1	D	304	<div><div>90%</div><div><div></div><div></div><div></div><div></div></div><div>• 6%</div></div>
1	E	304	<div><div>90%</div><div><div></div><div></div><div></div><div></div></div><div>• 6%</div></div>
1	F	304	<div><div>87%</div><div><div></div><div></div><div></div><div></div></div><div>7% 6%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	304	<div><div></div><div>88%</div><div>6%6%</div></div>
1	H	304	<div><div></div><div>90%</div><div>•6%</div></div>
1	I	304	<div><div></div><div>88%</div><div>7%6%</div></div>
1	J	304	<div><div></div><div>88%</div><div>6%6%</div></div>
1	K	304	<div><div></div><div>88%</div><div>5%6%</div></div>
1	L	304	<div><div></div><div>87%</div><div>7%6%</div></div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 29498 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 Monosaccharide ABC transporter substrate-binding protein, CUT2 family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	0	4	0
			2184	1384	371	425	4			
1	B	288	Total	C	N	O	S	0	1	0
			2164	1369	369	422	4			
1	C	286	Total	C	N	O	S	0	3	0
			2161	1368	368	421	4			
1	D	287	Total	C	N	O	S	0	1	0
			2156	1364	367	421	4			
1	E	286	Total	C	N	O	S	0	2	0
			2154	1366	366	418	4			
1	F	286	Total	C	N	O	S	0	7	0
			2190	1384	373	427	6			
1	G	286	Total	C	N	O	S	0	4	0
			2164	1369	368	421	6			
1	H	286	Total	C	N	O	S	0	1	0
			2150	1362	366	418	4			
1	I	286	Total	C	N	O	S	0	3	0
			2156	1365	366	421	4			
1	J	286	Total	C	N	O	S	0	1	0
			2149	1362	366	417	4			
1	K	286	Total	C	N	O	S	0	2	0
			2144	1362	364	414	4			
1	L	286	Total	C	N	O	S	0	6	0
			2170	1376	366	422	6			

There are 24 discrepancies between the modelled and reference sequences:

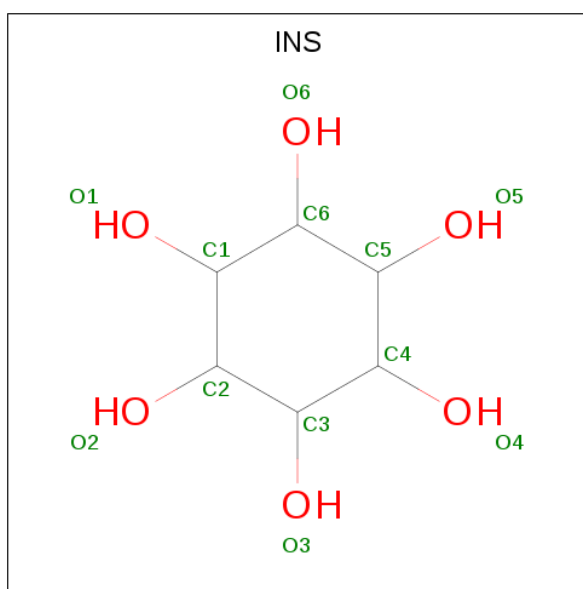
Chain	Residue	Modelled	Actual	Comment	Reference
A	30	SER	-	EXPRESSION TAG	UNP A0LVW8
A	31	MET	-	EXPRESSION TAG	UNP A0LVW8
B	30	SER	-	EXPRESSION TAG	UNP A0LVW8
B	31	MET	-	EXPRESSION TAG	UNP A0LVW8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	30	SER	-	EXPRESSION TAG	UNP A0LVW8
C	31	MET	-	EXPRESSION TAG	UNP A0LVW8
D	30	SER	-	EXPRESSION TAG	UNP A0LVW8
D	31	MET	-	EXPRESSION TAG	UNP A0LVW8
E	30	SER	-	EXPRESSION TAG	UNP A0LVW8
E	31	MET	-	EXPRESSION TAG	UNP A0LVW8
F	30	SER	-	EXPRESSION TAG	UNP A0LVW8
F	31	MET	-	EXPRESSION TAG	UNP A0LVW8
G	30	SER	-	EXPRESSION TAG	UNP A0LVW8
G	31	MET	-	EXPRESSION TAG	UNP A0LVW8
H	30	SER	-	EXPRESSION TAG	UNP A0LVW8
H	31	MET	-	EXPRESSION TAG	UNP A0LVW8
I	30	SER	-	EXPRESSION TAG	UNP A0LVW8
I	31	MET	-	EXPRESSION TAG	UNP A0LVW8
J	30	SER	-	EXPRESSION TAG	UNP A0LVW8
J	31	MET	-	EXPRESSION TAG	UNP A0LVW8
K	30	SER	-	EXPRESSION TAG	UNP A0LVW8
K	31	MET	-	EXPRESSION TAG	UNP A0LVW8
L	30	SER	-	EXPRESSION TAG	UNP A0LVW8
L	31	MET	-	EXPRESSION TAG	UNP A0LVW8

- Molecule 2 is 1,2,3,4,5,6-HEXAHYDROXY-CYCLOHEXANE (three-letter code: INS) (formula: C₆H₁₂O₆).



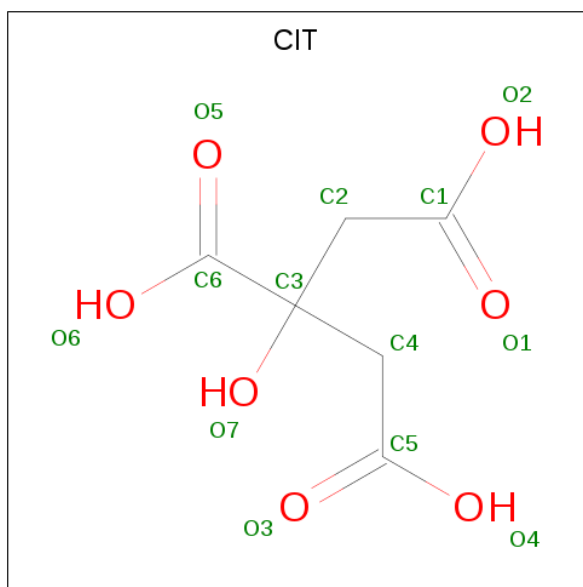
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	C	O	0	0
			12	6	6		
2	C	1	Total	C	O	0	0
			12	6	6		
2	D	1	Total	C	O	0	0
			12	6	6		
2	E	1	Total	C	O	0	0
			12	6	6		
2	F	1	Total	C	O	0	0
			12	6	6		
2	G	1	Total	C	O	0	0
			12	6	6		
2	H	1	Total	C	O	0	0
			12	6	6		
2	I	1	Total	C	O	0	0
			12	6	6		
2	J	1	Total	C	O	0	0
			12	6	6		
2	K	1	Total	C	O	0	0
			12	6	6		
2	L	1	Total	C	O	0	0
			12	6	6		

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			13	6	7		
3	E	1	Total	C	O	0	0
			13	6	7		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	287	Total	O	0	0
			287	287		
4	B	278	Total	O	0	0
			278	278		
4	C	303	Total	O	0	0
			303	303		
4	D	315	Total	O	0	0
			315	315		
4	E	293	Total	O	0	0
			293	293		
4	F	305	Total	O	0	0
			305	305		
4	G	280	Total	O	0	0
			280	280		
4	H	237	Total	O	0	0
			237	237		
4	I	251	Total	O	0	0
			251	251		
4	J	314	Total	O	0	0
			314	314		
4	K	298	Total	O	0	0
			298	298		
4	L	225	Total	O	0	0
			225	225		

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: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain A: 




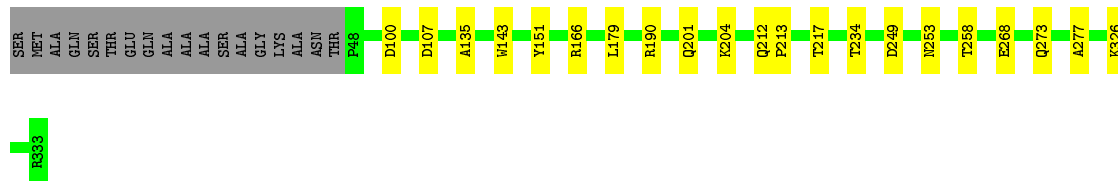
- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain B: 




- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain C: 



- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain D: 




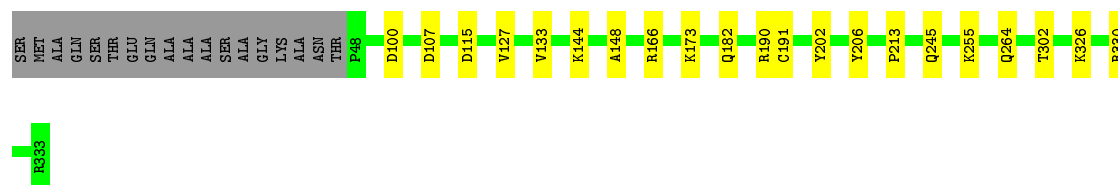
- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain E: 




- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain F: 




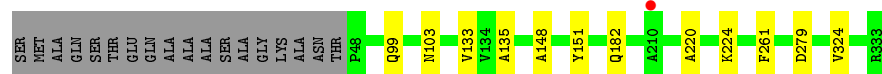
- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain G: 




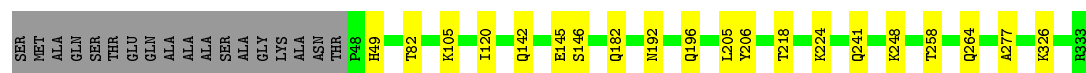
- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain H: 




- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain I: 




- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain J: 




- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain K: 



- Molecule 1: Monosaccharide ABC transporter substrate-binding protein, CUT2 family

Chain L: 

SER	MET	ALA	GLN	SER	THR	GLU	GLN	ALA	ALA	ALA	SER	ALA	GLY	LYS	ALA	ASN	THR	P48	H49	I55	T82	D80	I98	Q125	A135	Y151	R166	C191	F198	Y202	R216	Q241	S252	Q264	E268	Q273	I274	Q275	Y283	R333
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	52.54Å 149.98Å 221.63Å 90.00° 91.85° 90.00°	Depositor
Resolution (Å)	50.00 – 1.50 25.26 – 1.50	Depositor EDS
% Data completeness (in resolution range)	89.5 (50.00-1.50) 89.6 (25.26-1.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.52 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.185 , 0.226 0.195 , 0.232	Depositor DCC
R_{free} test set	14974 reflections (3.08%)	wwPDB-VP
Wilson B-factor (Å ²)	13.2	Xtriage
Anisotropy	0.281	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 29.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.065 for h,-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	29498	wwPDB-VP
Average B, all atoms (Å ²)	20.0	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 55.29 % 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 3.2324e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical 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: CIT, INS

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	0/2231	0.96	4/3039 (0.1%)
1	B	0.92	0/2208	0.95	3/3008 (0.1%)
1	C	0.98	1/2208 (0.0%)	0.99	5/3006 (0.2%)
1	D	0.92	0/2197	0.97	6/2993 (0.2%)
1	E	0.96	0/2198	0.97	3/2994 (0.1%)
1	F	0.99	0/2234	1.00	5/3042 (0.2%)
1	G	0.97	0/2211	0.97	3/3010 (0.1%)
1	H	0.94	0/2191	0.93	1/2984 (0.0%)
1	I	0.91	0/2203	0.94	0/3000
1	J	0.97	0/2193	0.99	5/2985 (0.2%)
1	K	0.92	0/2191	0.94	1/2986 (0.0%)
1	L	0.97	2/2226 (0.1%)	1.00	6/3031 (0.2%)
All	All	0.95	3/26491 (0.0%)	0.97	42/36078 (0.1%)

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	H	0	1
1	K	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	283	TYR	CE2-CZ	-5.35	1.31	1.38
1	L	252	SER	CB-OG	-5.33	1.35	1.42
1	C	143	TRP	CE3-CZ3	-5.01	1.29	1.38

All (42) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	211	ASP	CB-CG-OD1	7.73	125.25	118.30
1	D	115	ASP	CB-CG-OD2	-7.46	111.59	118.30
1	F	190	ARG	NE-CZ-NH2	-7.44	116.58	120.30
1	L	90	ASP	CB-CG-OD1	7.13	124.72	118.30
1	L	333	ARG	NE-CZ-NH1	-7.04	116.78	120.30
1	J	216	ARG	NE-CZ-NH2	-6.84	116.88	120.30
1	A	279	ASP	CB-CG-OD1	6.78	124.40	118.30
1	F	166	ARG	NE-CZ-NH2	-6.64	116.98	120.30
1	L	198	PHE	CB-CG-CD1	-6.60	116.18	120.80
1	F	100	ASP	CB-CG-OD2	-6.57	112.39	118.30
1	J	100	ASP	CB-CG-OD2	-6.57	112.39	118.30
1	G	90	ASP	CB-CG-OD1	6.36	124.02	118.30
1	K	249	ASP	CB-CG-OD1	6.16	123.85	118.30
1	C	107	ASP	CB-CG-OD1	-6.15	112.77	118.30
1	E	115	ASP	CB-CG-OD2	-5.94	112.95	118.30
1	C	100	ASP	CB-CG-OD2	-5.89	113.00	118.30
1	D	279	ASP	CB-CG-OD1	5.88	123.60	118.30
1	C	190	ARG	NE-CZ-NH1	5.88	123.24	120.30
1	D	249	ASP	CB-CG-OD1	5.81	123.53	118.30
1	B	100	ASP	CB-CG-OD2	-5.79	113.09	118.30
1	A	279	ASP	CB-CG-OD2	-5.65	113.22	118.30
1	D	166	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	E	249	ASP	CB-CG-OD1	5.53	123.27	118.30
1	D	249	ASP	CB-CG-OD2	-5.51	113.34	118.30
1	G	100	ASP	CB-CG-OD2	-5.48	113.37	118.30
1	J	211	ASP	CB-CG-OD2	-5.48	113.36	118.30
1	L	166	ARG	NE-CZ-NH2	-5.44	117.58	120.30
1	B	279	ASP	CB-CG-OD1	5.43	123.19	118.30
1	D	211	ASP	CB-CG-OD1	5.40	123.16	118.30
1	L	333	ARG	NE-CZ-NH2	5.34	122.97	120.30
1	F	107	ASP	CB-CG-OD1	-5.33	113.50	118.30
1	A	216	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	A	90	ASP	CB-CG-OD1	5.24	123.02	118.30
1	E	288	GLU	OE1-CD-OE2	5.21	129.56	123.30
1	L	216	ARG	NE-CZ-NH1	5.21	122.90	120.30
1	H	279	ASP	CB-CG-OD1	5.17	122.95	118.30
1	G	115	ASP	CB-CG-OD2	-5.16	113.66	118.30
1	F	115	ASP	CB-CG-OD2	-5.16	113.66	118.30
1	J	64	PHE	CB-CG-CD1	5.14	124.40	120.80
1	C	166	ARG	NE-CZ-NH2	-5.11	117.74	120.30
1	C	234	THR	CA-CB-CG2	-5.10	105.25	112.40
1	B	211	ASP	CB-CG-OD1	5.10	122.89	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	261	PHE	Peptide
1	K	261	PHE	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	2184	0	2177	11	0
1	B	2164	0	2150	7	0
1	C	2161	0	2150	8	0
1	D	2156	0	2140	5	0
1	E	2154	0	2145	5	0
1	F	2190	0	2169	19	0
1	G	2164	0	2149	9	0
1	H	2150	0	2138	5	0
1	I	2156	0	2146	13	0
1	J	2149	0	2143	12	0
1	K	2144	0	2138	11	0
1	L	2170	0	2166	12	0
2	A	12	0	12	0	0
2	B	12	0	12	0	0
2	C	12	0	12	0	0
2	D	12	0	12	0	0
2	E	12	0	12	0	0
2	F	12	0	12	0	0
2	G	12	0	12	0	0
2	H	12	0	12	0	0
2	I	12	0	12	0	0
2	J	12	0	12	0	0
2	K	12	0	12	0	0
2	L	12	0	12	0	0
3	C	13	0	5	0	0
3	E	13	0	5	0	0
4	A	287	0	0	5	0
4	B	278	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	303	0	0	0	0
4	D	315	0	0	3	0
4	E	293	0	0	1	0
4	F	305	0	0	9	0
4	G	280	0	0	0	0
4	H	237	0	0	0	0
4	I	251	0	0	2	0
4	J	314	0	0	1	0
4	K	298	0	0	1	0
4	L	225	0	0	0	0
All	All	29498	0	25965	105	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 2.

All (105) 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:L:49:HIS:CE1	1:L:82[A]:THR:HG23	1.96	1.00
1:A:206:TYR:H	1:I:182:GLN:HE22	1.23	0.85
1:F:191[A]:CYS:SG	1:F:202:TYR:OH	2.24	0.85
1:L:191[B]:CYS:HG	1:L:202:TYR:HH	0.96	0.84
1:F:206:TYR:H	1:J:182:GLN:HE22	1.24	0.84
1:L:49:HIS:HE1	1:L:82[A]:THR:HG23	1.38	0.82
1:F:191[A]:CYS:HG	1:F:202:TYR:HH	1.11	0.82
1:A:133[A]:VAL:HG13	1:A:148:ALA:HA	1.60	0.81
1:L:191[B]:CYS:SG	1:L:202:TYR:OH	2.28	0.79
1:B:206:TYR:H	1:G:182:GLN:HE22	1.28	0.78
1:I:49:HIS:CE1	1:I:82[A]:THR:HG23	2.18	0.78
1:A:145:GLU:HG3	4:A:663:HOH:O	1.85	0.76
1:F:182:GLN:HE22	1:J:206:TYR:H	1.32	0.75
1:A:182:GLN:HE22	1:I:206:TYR:H	1.36	0.74
1:B:182:GLN:HE22	1:G:206:TYR:H	1.35	0.74
1:A:281:GLN:HE22	1:D:302:THR:H	1.37	0.72
1:I:49:HIS:HE1	1:I:82[A]:THR:HG23	1.55	0.71
1:J:302:THR:H	1:K:281:GLN:HE22	1.39	0.70
1:E:133[A]:VAL:HG13	1:E:148:ALA:HA	1.74	0.69
1:F:302[A]:THR:HG22	4:F:609:HOH:O	1.95	0.67
1:D:96:VAL:HG22	4:D:641:HOH:O	1.94	0.67
1:J:179:LEU:CD1	1:J:204[B]:LYS:HG2	2.26	0.66
1:L:49:HIS:CE1	1:L:82[A]:THR:CG2	2.78	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:224:LYS:HD3	4:I:693:HOH:O	1.94	0.66
1:I:120:ILE:HD12	1:I:146:SER:HB3	1.79	0.65
4:A:787:HOH:O	1:K:201:GLN:HG2	1.97	0.64
1:C:213:PRO:HB3	1:F:245[A]:GLN:HE21	1.62	0.63
1:L:49:HIS:HE1	1:L:82[A]:THR:CG2	2.11	0.63
1:K:201:GLN:NE2	4:K:668:HOH:O	2.22	0.63
1:E:145:GLU:HG3	4:E:592:HOH:O	1.99	0.61
1:I:120:ILE:CD1	1:I:146:SER:HB3	2.31	0.60
1:B:120:ILE:HD12	1:B:146:SER:HB3	1.84	0.59
1:F:330:ARG:HD3	4:F:744:HOH:O	2.02	0.59
1:A:125:GLN:NE2	4:A:739:HOH:O	2.34	0.59
1:F:264[A]:GLN:NE2	4:F:793:HOH:O	2.38	0.57
1:C:179:LEU:HD11	1:C:204:LYS:HG2	1.86	0.57
1:A:133[A]:VAL:CG1	1:A:148:ALA:HA	2.33	0.56
1:J:179:LEU:CD1	1:J:204[B]:LYS:CG	2.84	0.56
1:G:191[A]:CYS:SG	1:G:204:LYS:HE2	2.46	0.55
1:C:212:GLN:OE1	1:F:264[A]:GLN:NE2	2.41	0.54
1:J:179:LEU:HD11	1:J:204[B]:LYS:CG	2.37	0.54
1:H:133[A]:VAL:HG13	1:H:148:ALA:HA	1.91	0.53
1:G:120:ILE:CD1	1:G:146:SER:HB3	2.40	0.52
1:L:241:GLN:HG3	1:L:264:GLN:HE22	1.75	0.52
1:F:173:LYS:HD3	4:F:792:HOH:O	2.10	0.51
1:G:261:PHE:HZ	1:G:316:VAL:HG21	1.75	0.51
1:D:145:GLU:HG3	4:D:783:HOH:O	2.10	0.51
1:H:99:GLN:NE2	1:H:103:ASN:OD1	2.44	0.50
1:F:133[A]:VAL:HG13	1:F:148:ALA:HA	1.94	0.49
1:K:247:VAL:HG21	1:K:256:ILE:HD11	1.95	0.49
1:C:253:ASN:HD21	1:L:275:GLN:NE2	2.11	0.49
1:I:248:LYS:NZ	4:I:666:HOH:O	2.47	0.48
1:L:268:GLU:HG2	1:L:273:GLN:HE21	1.78	0.48
1:F:302[A]:THR:CG2	4:F:609:HOH:O	2.59	0.48
1:K:144:LYS:HE2	1:K:306:GLY:O	2.13	0.48
1:I:192:ASN:O	1:I:196:GLN:HG3	2.14	0.47
1:A:258:THR:O	1:A:277:ALA:HA	2.14	0.47
1:F:326:LYS:HG2	4:F:744:HOH:O	2.14	0.47
1:I:205:LEU:HD21	1:I:218:THR:HG22	1.96	0.47
1:E:135:ALA:O	1:E:151:TYR:HA	2.15	0.46
1:G:191[B]:CYS:HB3	1:G:202:TYR:OH	2.15	0.46
1:I:142:GLN:O	1:I:145:GLU:HG2	2.16	0.46
1:F:127:VAL:HG22	1:F:133[A]:VAL:HG12	1.98	0.46
1:J:59:GLN:NE2	1:J:60:PRO:HD2	2.30	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:241:GLN:HG3	1:I:264:GLN:HE22	1.81	0.46
1:B:258:THR:O	1:B:277:ALA:HA	2.15	0.45
1:K:237:ALA:N	1:K:238:PRO:CD	2.80	0.45
1:C:135:ALA:O	1:C:151:TYR:HA	2.17	0.45
1:G:142:GLN:O	1:G:145:GLU:HG2	2.15	0.45
1:K:144:LYS:CE	1:K:306:GLY:O	2.64	0.45
1:E:215:VAL:HB	1:E:242:LEU:HD13	1.98	0.45
1:J:179:LEU:HD13	1:J:204[B]:LYS:HG2	1.99	0.45
1:C:249:ASP:HA	1:F:213:PRO:HG3	1.99	0.44
1:C:258:THR:O	1:C:277:ALA:HA	2.17	0.44
1:J:135:ALA:O	1:J:151:TYR:HA	2.18	0.44
1:J:270:GLU:OE2	4:J:623:HOH:O	2.21	0.44
1:J:247:VAL:HG21	1:J:256:ILE:HD11	2.00	0.44
1:E:120:ILE:CD1	1:E:146:SER:HB3	2.48	0.43
1:L:191[A]:CYS:HB3	1:L:202:TYR:HH	1.82	0.43
1:B:120:ILE:CD1	1:B:146:SER:HB3	2.47	0.43
1:F:191[B]:CYS:HB3	1:F:202:TYR:HH	1.83	0.43
1:D:192:ASN:ND2	4:D:710:HOH:O	2.52	0.43
1:F:302[A]:THR:HG21	4:F:555:HOH:O	2.17	0.43
1:K:261:PHE:CZ	1:K:324:VAL:HG11	2.54	0.43
1:L:55:ILE:HD13	1:L:98:ILE:HA	1.99	0.43
1:H:135:ALA:O	1:H:151:TYR:HA	2.19	0.42
1:F:144:LYS:NZ	4:F:722:HOH:O	2.50	0.42
1:J:241:GLN:OE1	1:J:268:GLU:OE2	2.38	0.42
1:L:135:ALA:O	1:L:151:TYR:HA	2.20	0.42
1:B:268:GLU:HB3	1:B:274:LEU:HD12	2.01	0.41
1:D:102:VAL:HA	1:D:131:ILE:HD12	2.01	0.41
1:G:258:THR:O	1:G:277:ALA:HA	2.20	0.41
1:A:50:LEU:HB2	1:A:81:VAL:HG22	2.02	0.41
1:K:226:ASP:HB2	1:K:229:ILE:HD12	2.01	0.41
1:B:241:GLN:NE2	1:B:268:GLU:OE2	2.27	0.41
1:G:188:GLU:HA	1:G:191[A]:CYS:SG	2.60	0.41
1:C:268:GLU:HG2	1:C:273:GLN:OE1	2.21	0.41
1:A:99:GLN:HG2	4:A:704:HOH:O	2.21	0.41
1:K:175[A]:VAL:HG12	1:K:198:PHE:CD2	2.56	0.41
1:K:258:THR:O	1:K:277:ALA:HA	2.21	0.41
1:A:310:LYS:HG2	4:A:613:HOH:O	2.21	0.41
1:H:220:ALA:O	1:H:224:LYS:HG3	2.21	0.40
1:I:258:THR:O	1:I:277:ALA:HA	2.21	0.40
1:H:324:VAL:HG22	1:H:324:VAL:O	2.22	0.40
1:F:245[A]:GLN:CD	4:F:670:HOH:O	2.60	0.40

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	290/304 (95%)	286 (99%)	4 (1%)	0	100	100
1	B	287/304 (94%)	283 (99%)	4 (1%)	0	100	100
1	C	287/304 (94%)	281 (98%)	6 (2%)	0	100	100
1	D	286/304 (94%)	281 (98%)	5 (2%)	0	100	100
1	E	286/304 (94%)	282 (99%)	4 (1%)	0	100	100
1	F	291/304 (96%)	287 (99%)	4 (1%)	0	100	100
1	G	288/304 (95%)	285 (99%)	3 (1%)	0	100	100
1	H	285/304 (94%)	281 (99%)	4 (1%)	0	100	100
1	I	287/304 (94%)	278 (97%)	9 (3%)	0	100	100
1	J	285/304 (94%)	280 (98%)	5 (2%)	0	100	100
1	K	286/304 (94%)	282 (99%)	4 (1%)	0	100	100
1	L	290/304 (95%)	286 (99%)	4 (1%)	0	100	100
All	All	3448/3648 (94%)	3392 (98%)	56 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	226/231 (98%)	223 (99%)	3 (1%)	69	44
1	B	223/231 (96%)	221 (99%)	2 (1%)	78	61
1	C	223/231 (96%)	220 (99%)	3 (1%)	69	44
1	D	222/231 (96%)	221 (100%)	1 (0%)	88	78
1	E	222/231 (96%)	221 (100%)	1 (0%)	88	78
1	F	227/231 (98%)	226 (100%)	1 (0%)	91	82
1	G	224/231 (97%)	223 (100%)	1 (0%)	91	82
1	H	221/231 (96%)	220 (100%)	1 (0%)	88	78
1	I	223/231 (96%)	221 (99%)	2 (1%)	78	61
1	J	221/231 (96%)	221 (100%)	0	100	100
1	K	220/231 (95%)	220 (100%)	0	100	100
1	L	226/231 (98%)	225 (100%)	1 (0%)	91	82
All	All	2678/2772 (97%)	2662 (99%)	16 (1%)	86	74

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	145	GLU
1	A	154	GLN
1	A	326	LYS
1	B	154	GLN
1	B	326	LYS
1	C	201	GLN
1	C	217	THR
1	C	326	LYS
1	D	47	THR
1	E	248	LYS
1	F	255	LYS
1	G	222	LYS
1	H	182	GLN
1	I	105	LYS
1	I	326	LYS
1	L	125	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (44) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	182	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	241	GLN
1	A	245	GLN
1	A	281	GLN
1	B	182	GLN
1	B	253	ASN
1	C	192	ASN
1	C	212	GLN
1	C	241	GLN
1	C	264	GLN
1	D	192	ASN
1	D	212	GLN
1	D	241	GLN
1	D	245	GLN
1	D	253	ASN
1	E	212	GLN
1	E	241	GLN
1	F	182	GLN
1	F	192	ASN
1	F	241	GLN
1	F	262	ASN
1	G	99	GLN
1	G	182	GLN
1	H	99	GLN
1	H	103	ASN
1	H	212	GLN
1	H	245	GLN
1	H	264	GLN
1	I	182	GLN
1	I	196	GLN
1	I	245	GLN
1	I	264	GLN
1	J	59	GLN
1	J	125	GLN
1	J	182	GLN
1	J	192	ASN
1	J	241	GLN
1	K	281	GLN
1	L	49	HIS
1	L	125	GLN
1	L	192	ASN
1	L	264	GLN
1	L	273	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	275	GLN

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 ⓘ

14 ligands are modelled in this entry.

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	$\# Z > 2$
2	INS	B	401	-	12,12,12	0.86	0	18,18,18	1.42	2 (11%)
2	INS	D	401	-	12,12,12	0.76	0	18,18,18	1.95	3 (16%)
2	INS	F	401	-	12,12,12	0.57	0	18,18,18	1.82	3 (16%)
2	INS	H	401	-	12,12,12	0.83	0	18,18,18	1.61	4 (22%)
2	INS	K	401	-	12,12,12	0.71	0	18,18,18	1.69	5 (27%)
2	INS	A	401	-	12,12,12	0.72	0	18,18,18	1.37	3 (16%)
2	INS	G	401	-	12,12,12	0.85	0	18,18,18	1.50	3 (16%)
2	INS	C	402	-	12,12,12	0.95	0	18,18,18	1.21	2 (11%)
3	CIT	C	401	-	3,12,12	1.00	0	3,17,17	3.39	2 (66%)
2	INS	J	401	-	12,12,12	1.16	2 (16%)	18,18,18	1.69	3 (16%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	CIT	E	401	-	3,12,12	0.74	0	3,17,17	1.74	2 (66%)
2	INS	L	401	-	12,12,12	1.08	1 (8%)	18,18,18	1.72	6 (33%)
2	INS	E	402	-	12,12,12	0.86	0	18,18,18	1.46	4 (22%)
2	INS	I	401	-	12,12,12	0.85	0	18,18,18	1.82	4 (22%)

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
2	INS	B	401	-	-	-	0/1/1/1
2	INS	D	401	-	-	-	0/1/1/1
2	INS	F	401	-	-	-	0/1/1/1
2	INS	H	401	-	-	-	0/1/1/1
2	INS	K	401	-	-	-	0/1/1/1
2	INS	A	401	-	-	-	0/1/1/1
2	INS	G	401	-	-	-	0/1/1/1
2	INS	C	402	-	-	-	0/1/1/1
3	CIT	C	401	-	-	2/6/16/16	-
2	INS	J	401	-	-	-	0/1/1/1
3	CIT	E	401	-	-	0/6/16/16	-
2	INS	L	401	-	-	-	0/1/1/1
2	INS	E	402	-	-	-	0/1/1/1
2	INS	I	401	-	-	-	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	401	INS	C6-C5	-2.39	1.46	1.52
2	J	401	INS	O4-C4	-2.11	1.38	1.43
2	L	401	INS	C6-C5	-2.08	1.47	1.52

All (46) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	401	INS	C6-C1-C2	5.33	120.13	110.82
2	D	401	INS	C6-C1-C2	5.08	119.69	110.82
2	F	401	INS	C6-C1-C2	4.64	118.92	110.82
2	D	401	INS	O2-C2-C1	-4.52	99.91	110.35
3	C	401	CIT	C3-C4-C5	-4.42	107.91	114.98

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	401	INS	C6-C1-C2	4.28	118.29	110.82
2	F	401	INS	O2-C2-C1	-3.71	101.76	110.35
3	C	401	CIT	C4-C3-C2	3.63	119.03	109.33
2	L	401	INS	O2-C2-C1	-3.57	102.10	110.35
2	A	401	INS	C6-C1-C2	3.54	117.00	110.82
2	B	401	INS	C6-C1-C2	3.47	116.88	110.82
2	H	401	INS	C6-C1-C2	3.43	116.81	110.82
2	L	401	INS	O4-C4-C3	-3.39	102.50	110.35
2	I	401	INS	O2-C2-C1	-3.35	102.61	110.35
2	G	401	INS	C6-C1-C2	3.28	116.55	110.82
2	K	401	INS	O1-C1-C6	3.13	117.58	110.35
2	K	401	INS	C6-C1-C2	3.04	116.13	110.82
2	H	401	INS	O2-C2-C1	-3.02	103.36	110.35
2	G	401	INS	O2-C2-C1	-3.01	103.38	110.35
2	F	401	INS	C4-C3-C2	2.91	115.91	110.82
2	J	401	INS	O2-C2-C1	-2.90	103.64	110.35
2	H	401	INS	O1-C1-C2	2.85	116.95	110.35
2	C	402	INS	C6-C1-C2	2.77	115.67	110.82
2	K	401	INS	C5-C6-C1	2.62	115.40	110.82
2	E	402	INS	O3-C3-C2	-2.61	104.31	110.35
2	C	402	INS	O4-C4-C3	-2.59	104.35	110.35
2	I	401	INS	O4-C4-C5	-2.59	104.36	110.35
2	E	402	INS	C6-C1-C2	2.55	115.28	110.82
2	B	401	INS	O2-C2-C1	-2.51	104.55	110.35
2	I	401	INS	C4-C3-C2	2.48	115.16	110.82
2	K	401	INS	O2-C2-C1	-2.45	104.69	110.35
2	K	401	INS	C4-C3-C2	2.41	115.04	110.82
2	E	402	INS	C3-C2-C1	-2.35	106.72	110.82
2	G	401	INS	C4-C3-C2	2.34	114.90	110.82
2	H	401	INS	O5-C5-C6	-2.26	105.12	110.35
2	L	401	INS	O3-C3-C2	-2.23	105.20	110.35
2	E	402	INS	O1-C1-C6	2.19	115.41	110.35
2	D	401	INS	O4-C4-C3	-2.19	105.29	110.35
2	L	401	INS	O1-C1-C2	2.09	115.19	110.35
2	J	401	INS	C3-C2-C1	-2.09	107.17	110.82
3	E	401	CIT	C3-C2-C1	-2.07	111.67	114.98
2	L	401	INS	C6-C1-C2	2.04	114.39	110.82
3	E	401	CIT	C3-C4-C5	-2.04	111.72	114.98
2	A	401	INS	O3-C3-C2	-2.04	105.64	110.35
2	L	401	INS	O1-C1-C6	2.03	115.04	110.35
2	A	401	INS	O2-C2-C1	-2.01	105.70	110.35

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	401	CIT	C2-C3-C4-C5
3	C	401	CIT	C1-C2-C3-O7

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	288/304 (94%)	-0.56	0 100 100	13, 17, 30, 72	0
1	B	288/304 (94%)	-0.52	1 (0%) 94 95	12, 19, 34, 86	0
1	C	286/304 (94%)	-0.58	0 100 100	12, 17, 29, 44	0
1	D	287/304 (94%)	-0.57	0 100 100	13, 18, 34, 62	0
1	E	286/304 (94%)	-0.61	0 100 100	13, 18, 29, 49	0
1	F	286/304 (94%)	-0.55	0 100 100	12, 17, 34, 57	0
1	G	286/304 (94%)	-0.53	0 100 100	12, 18, 31, 52	0
1	H	286/304 (94%)	-0.56	1 (0%) 94 95	13, 18, 33, 60	0
1	I	286/304 (94%)	-0.47	0 100 100	13, 20, 37, 62	0
1	J	286/304 (94%)	-0.63	0 100 100	12, 17, 27, 62	0
1	K	286/304 (94%)	-0.53	0 100 100	13, 19, 31, 59	0
1	L	286/304 (94%)	-0.61	0 100 100	11, 17, 30, 52	0
All	All	3437/3648 (94%)	-0.56	2 (0%) 95 95	11, 18, 33, 86	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	46	ASN	2.1
1	H	210	ALA	2.0

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

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
3	CIT	C	401	13/13	0.88	0.12	28,37,65,69	0
3	CIT	E	401	13/13	0.94	0.08	19,23,31,33	0
2	INS	I	401	12/12	0.95	0.07	14,16,21,23	0
2	INS	D	401	12/12	0.96	0.07	12,15,19,20	0
2	INS	J	401	12/12	0.96	0.07	11,14,20,22	0
2	INS	F	401	12/12	0.96	0.06	11,14,19,20	0
2	INS	A	401	12/12	0.96	0.07	10,14,20,20	0
2	INS	K	401	12/12	0.97	0.05	13,15,19,20	0
2	INS	B	401	12/12	0.97	0.06	13,14,21,24	0
2	INS	G	401	12/12	0.97	0.06	11,13,19,20	0
2	INS	L	401	12/12	0.97	0.07	13,15,18,23	0
2	INS	E	402	12/12	0.97	0.06	13,14,19,21	0
2	INS	C	402	12/12	0.97	0.06	11,13,17,24	0
2	INS	H	401	12/12	0.98	0.05	14,16,20,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.