Full wwPDB X-ray Structure Validation Report (i)

Aug 5, 2024 – 10:17 am BST

PDB ID	:	8OXS
Title	:	Cholera holotoxin variant (chimera with E. coli heat-labile enterotoxin, 4 C-terminal substitutions)
Authors	:	Kersten, F.; Heim, J.B.; Cordara, G.; Krengel, U.
Deposited on	:	2023-05-02
Resolution	:	1.60 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 (i) 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 (i)) were used in the production of this report:

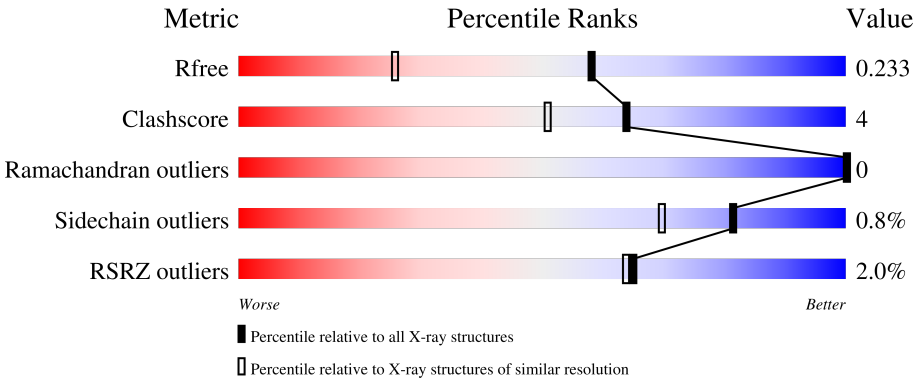
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.60 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.



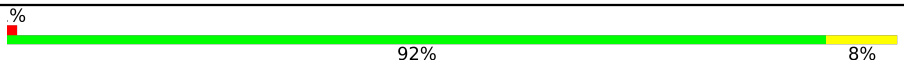
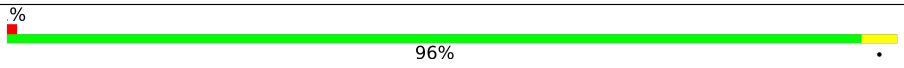
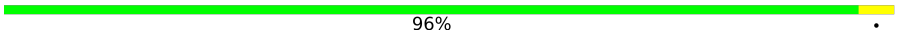
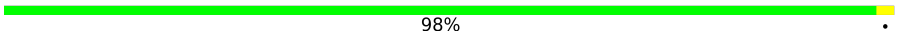
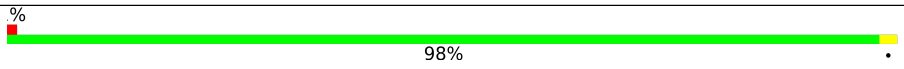
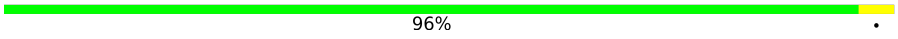
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 <=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	240	5% 80% 15% ·
1	B	240	7% 86% 10% · ·
2	C	103	95% 5%
2	D	103	90% 10%
2	E	103	91% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	103	
2	G	103	
2	H	103	
2	I	103	
2	J	103	
2	K	103	
2	L	103	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 13700 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 Cholera enterotoxin subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	231	Total	C	N	O	S	0	34	0
			2140	1337	389	407	7			
1	B	231	Total	C	N	O	S	0	17	0
			2007	1250	370	380	7			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	229	GLU	ASP	engineered mutation	UNP P01555
A	230	VAL	ILE	engineered mutation	UNP P01555
A	232	ILE	THR	engineered mutation	UNP P01555
A	233	TYR	HIS	engineered mutation	UNP P01555
B	229	GLU	ASP	engineered mutation	UNP P01555
B	230	VAL	ILE	engineered mutation	UNP P01555
B	232	ILE	THR	engineered mutation	UNP P01555
B	233	TYR	HIS	engineered mutation	UNP P01555

- Molecule 2 is a protein called Cholera enterotoxin subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	103	Total	C	N	O	S	0	6	0
			858	537	150	164	7			
2	E	103	Total	C	N	O	S	0	5	0
			855	536	149	164	6			
2	F	103	Total	C	N	O	S	0	7	0
			877	551	155	165	6			
2	G	103	Total	C	N	O	S	0	3	0
			835	523	146	160	6			
2	H	103	Total	C	N	O	S	0	4	0
			846	531	148	160	7			
2	C	103	Total	C	N	O	S	0	4	0
			843	529	147	160	7			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	103	Total	C	N	O	S	0	6	0
			867	543	154	164	6			
2	J	103	Total	C	N	O	S	0	2	0
			828	518	145	159	6			
2	K	103	Total	C	N	O	S	0	4	0
			844	529	147	161	7			
2	L	103	Total	C	N	O	S	0	7	0
			870	547	151	165	7			

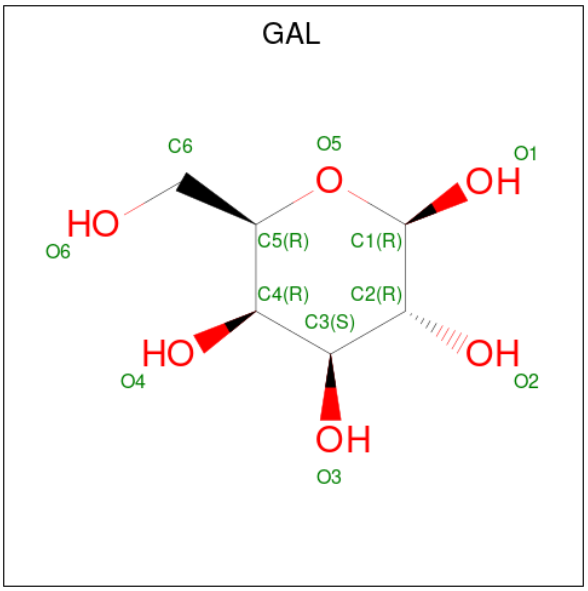
There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	18	HIS	TYR	engineered mutation	UNP P01556
D	47	THR	ILE	engineered mutation	UNP P01556
E	18	HIS	TYR	engineered mutation	UNP P01556
E	47	THR	ILE	engineered mutation	UNP P01556
F	18	HIS	TYR	engineered mutation	UNP P01556
F	47	THR	ILE	engineered mutation	UNP P01556
G	18	HIS	TYR	engineered mutation	UNP P01556
G	47	THR	ILE	engineered mutation	UNP P01556
H	18	HIS	TYR	engineered mutation	UNP P01556
H	47	THR	ILE	engineered mutation	UNP P01556
C	18	HIS	TYR	engineered mutation	UNP P01556
C	47	THR	ILE	engineered mutation	UNP P01556
I	18	HIS	TYR	engineered mutation	UNP P01556
I	47	THR	ILE	engineered mutation	UNP P01556
J	18	HIS	TYR	engineered mutation	UNP P01556
J	47	THR	ILE	engineered mutation	UNP P01556
K	18	HIS	TYR	engineered mutation	UNP P01556
K	47	THR	ILE	engineered mutation	UNP P01556
L	18	HIS	TYR	engineered mutation	UNP P01556
L	47	THR	ILE	engineered mutation	UNP P01556

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

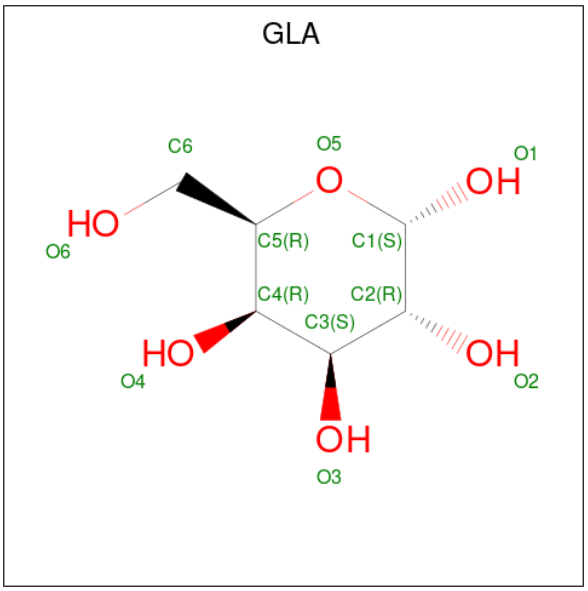
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Na	0	0
			1	1		
3	B	1	Total	Na	0	0
			1	1		

- Molecule 4 is beta-D-galactopyranose (three-letter code: GAL) (formula: $C_6H_{12}O_6$).



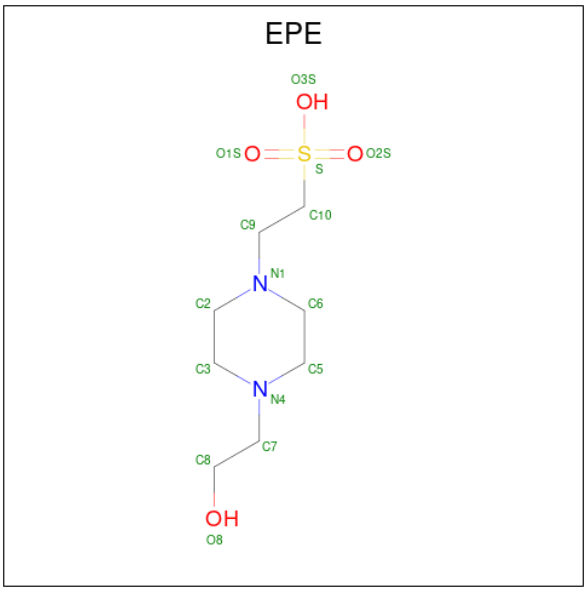
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 12 6 6	0	0
4	D	1	Total C O 12 6 6	0	0
4	E	1	Total C O 12 6 6	0	0
4	F	1	Total C O 12 6 6	0	0
4	G	1	Total C O 12 6 6	0	0
4	H	1	Total C O 12 6 6	0	0
4	C	1	Total C O 12 6 6	0	0
4	I	1	Total C O 12 6 6	0	0
4	K	1	Total C O 12 6 6	0	0
4	L	1	Total C O 12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

- Molecule 5 is alpha-D-galactopyranose (three-letter code: GLA) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	1	Total C O 12 6 6	0	0
5	E	1	Total C O 12 6 6	0	0
5	G	1	Total C O 12 6 6	0	0
5	H	1	Total C O 12 6 6	0	0
5	C	1	Total C O 12 6 6	0	0
5	K	1	Total C O 12 6 6	0	0
5	L	1	Total C O 12 6 6	0	0

- Molecule 6 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	E	1	Total C N O S 15 8 2 4 1	0	0
6	H	1	Total C N O S 15 8 2 4 1	0	0
6	B	1	Total C N O S 15 8 2 4 1	0	0
6	I	1	Total C N O S 15 8 2 4 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	95	Total O 95 95	0	0
7	D	62	Total O 62 62	0	0
7	E	70	Total O 70 70	0	0
7	F	64	Total O 64 64	0	0
7	G	57	Total O 57 57	0	0
7	H	51	Total O 51 51	0	0
7	B	80	Total O 80 80	0	0
7	C	61	Total O 61 61	0	0

Continued on next page...

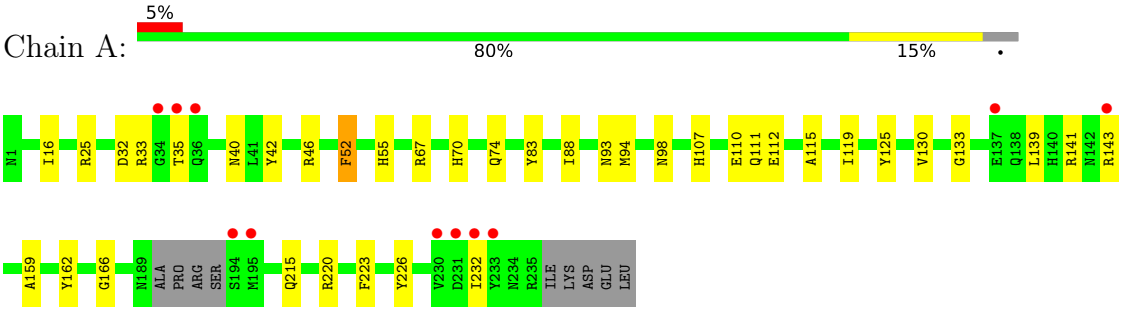
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	I	61	Total O 61 61	0	0
7	J	53	Total O 53 53	0	0
7	K	55	Total O 55 55	0	0
7	L	55	Total O 55 55	0	0

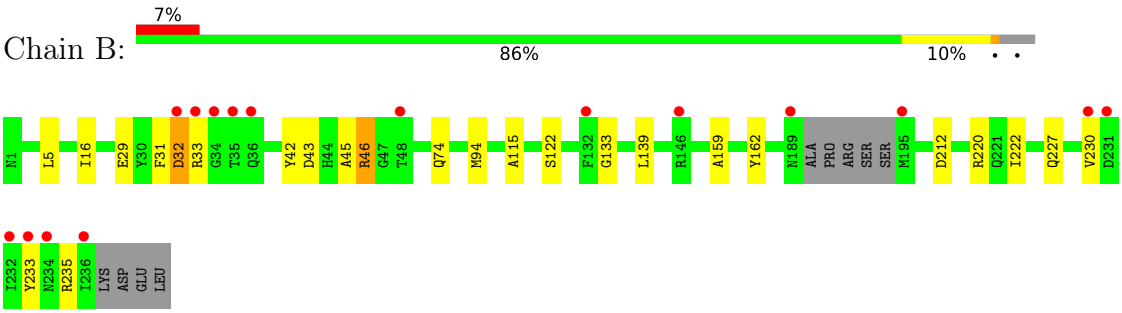
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

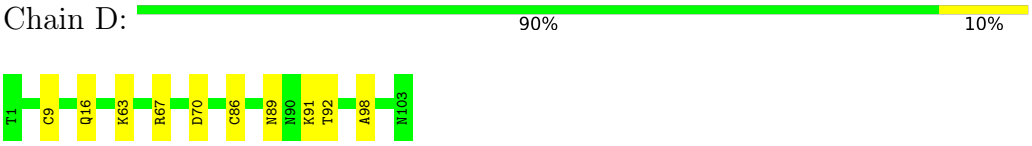
• Molecule 1: Cholera enterotoxin subunit A



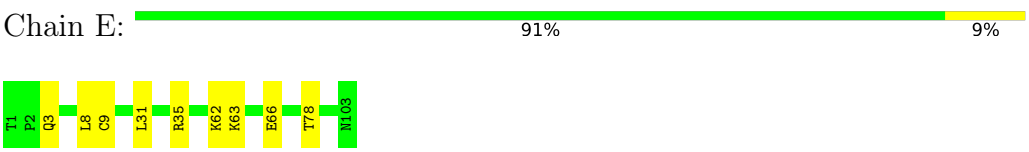
• Molecule 1: Cholera enterotoxin subunit A



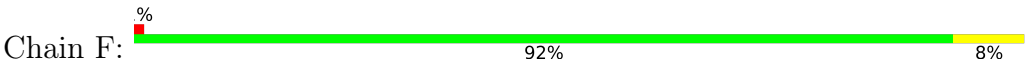
• Molecule 2: Cholera enterotoxin subunit B

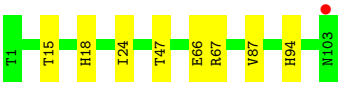


• Molecule 2: Cholera enterotoxin subunit B



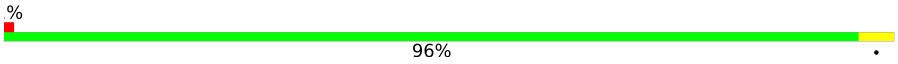
• Molecule 2: Cholera enterotoxin subunit B





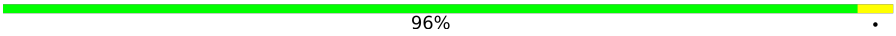
• Molecule 2: Cholera enterotoxin subunit B

Chain G:



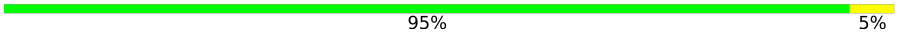
• Molecule 2: Cholera enterotoxin subunit B

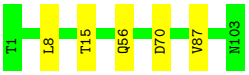
Chain H:



• Molecule 2: Cholera enterotoxin subunit B

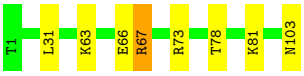
Chain C:





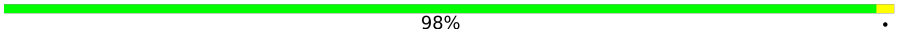
• Molecule 2: Cholera enterotoxin subunit B

Chain I:



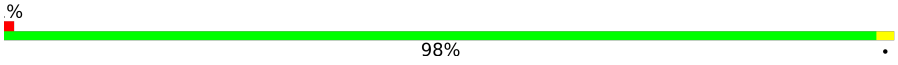
• Molecule 2: Cholera enterotoxin subunit B

Chain J:



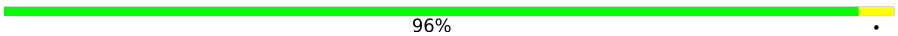
• Molecule 2: Cholera enterotoxin subunit B

Chain K:



• Molecule 2: Cholera enterotoxin subunit B

Chain L:




T1
Q56
R67
I74
W88
N103

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	60.30Å 89.61Å 141.09Å 90.00° 98.19° 90.00°	Depositor
Resolution (Å)	48.90 – 1.60 48.94 – 1.60	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.90-1.60) 99.8 (48.94-1.60)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.8.0405	Depositor
R, R_{free}	0.193 , 0.224 0.203 , 0.233	Depositor DCC
R_{free} test set	2055 reflections (1.05%)	wwPDB-VP
Wilson B-factor (Å^2)	22.8	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 40.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	13700	wwPDB-VP
Average B, all atoms (Å^2)	25.0	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 86.16 % 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 1.0190e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: NA, GLA, GAL, EPE

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	0.41	0/2200	0.70	0/2980
1	B	0.42	0/2062	0.74	0/2791
2	C	0.47	0/857	0.73	0/1156
2	D	0.46	0/873	0.69	0/1179
2	E	0.53	1/869 (0.1%)	0.73	0/1172
2	F	0.47	0/893	0.71	0/1204
2	G	0.47	0/849	0.73	0/1145
2	H	0.42	0/860	0.72	0/1158
2	I	0.54	1/881 (0.1%)	0.73	1/1185 (0.1%)
2	J	0.49	0/842	0.72	0/1137
2	K	0.44	0/858	0.75	0/1158
2	L	0.42	0/884	0.70	0/1191
All	All	0.45	2/12928 (0.0%)	0.72	1/17456 (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
1	A	0	1
1	B	0	1
2	D	0	1
2	G	0	1
2	I	0	1
2	J	0	1
2	K	0	1
All	All	0	7

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	66	GLU	CD-OE2	10.11	1.36	1.25
2	E	66	GLU	CD-OE2	7.90	1.34	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	73	ARG	NE-CZ-NH1	6.12	123.36	120.30

There are no chirality outliers.

All (7) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	67[A]	ARG	Sidechain
1	B	46[A]	ARG	Sidechain
2	D	67	ARG	Sidechain
2	G	73	ARG	Sidechain
2	I	67[A]	ARG	Sidechain
2	J	49	GLN	Peptide
2	K	73	ARG	Sidechain

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	2140	0	1974	44	0
1	B	2007	0	1860	21	0
2	C	843	0	848	4	0
2	D	858	0	858	6	0
2	E	855	0	856	16	0
2	F	877	0	874	8	0
2	G	835	0	837	5	0
2	H	846	0	853	6	0
2	I	867	0	874	5	0
2	J	828	0	826	1	0
2	K	844	0	848	0	0
2	L	870	0	877	7	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	12	0	12	0	0
4	C	12	0	4	1	0
4	D	12	0	4	0	0
4	E	12	0	4	0	0
4	F	12	0	12	0	0
4	G	12	0	2	0	0
4	H	12	0	4	1	0
4	I	12	0	11	0	0
4	K	12	0	4	0	0
4	L	12	0	5	3	0
5	C	12	0	5	0	0
5	D	12	0	3	1	0
5	E	12	0	6	0	0
5	G	12	0	4	0	0
5	H	12	0	3	0	0
5	K	12	0	4	0	0
5	L	12	0	3	1	0
6	B	15	0	17	0	0
6	E	15	0	17	0	0
6	H	15	0	17	0	0
6	I	15	0	18	0	0
7	A	95	0	0	2	0
7	B	80	0	0	3	0
7	C	61	0	0	0	0
7	D	62	0	0	1	0
7	E	70	0	0	2	0
7	F	64	0	0	2	0
7	G	57	0	0	2	0
7	H	51	0	0	2	0
7	I	61	0	0	0	0
7	J	53	0	0	0	0
7	K	55	0	0	0	0
7	L	55	0	0	0	0
All	All	13700	0	12544	108	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 4.

All (108) 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:E:8[B]:LEU:C	2:E:8[B]:LEU:HD13	1.77	1.04
2:E:8[B]:LEU:HD13	2:E:8[B]:LEU:O	1.57	1.01
2:E:8[B]:LEU:C	2:E:8[B]:LEU:CD1	2.33	0.96
1:B:31[A]:PHE:O	1:B:33[A]:ARG:HG2	1.70	0.91
1:A:112[B]:GLU:OE1	7:A:401:HOH:O	1.89	0.90
1:A:226[B]:TYR:CE1	2:H:78:THR:HG21	2.09	0.87
1:B:42:TYR:O	1:B:45[A]:ALA:O	1.99	0.80
1:B:45[A]:ALA:O	1:B:46[A]:ARG:HB2	1.89	0.72
1:A:32[A]:ASP:O	1:A:33[A]:ARG:HD2	1.91	0.70
2:E:62:LYS:HE3	7:E:364:HOH:O	1.92	0.70
1:B:45[A]:ALA:O	1:B:46[A]:ARG:CB	2.41	0.68
2:E:63[B]:LYS:HA	2:E:63[B]:LYS:HE2	1.76	0.68
2:D:16[B]:GLN:OE1	2:D:89:ASN:ND2	2.21	0.68
1:A:94[A]:MET:SD	1:A:115:ALA:HB2	2.35	0.66
7:B:421:HOH:O	2:L:74[B]:ILE:HG12	1.96	0.66
1:B:227[A]:GLN:O	7:B:401:HOH:O	2.12	0.64
2:D:91[A]:LYS:NZ	5:D:201:GLA:O3	2.29	0.64
2:F:24[A]:ILE:HD12	7:F:311:HOH:O	1.97	0.63
1:A:223[B]:PHE:CD1	1:A:226[B]:TYR:CE2	2.87	0.62
2:E:63[B]:LYS:CA	2:E:63[B]:LYS:CE	2.77	0.62
1:A:223[B]:PHE:CE1	1:A:226[B]:TYR:HE2	2.17	0.61
1:A:223[B]:PHE:CD1	1:A:226[B]:TYR:CD2	2.89	0.61
2:F:67[B]:ARG:NH2	2:G:70:ASP:OD1	2.34	0.61
1:A:32[A]:ASP:C	1:A:33[A]:ARG:HD2	2.22	0.60
1:A:226[B]:TYR:CD1	2:H:78:THR:HG21	2.37	0.59
1:A:223[B]:PHE:HD1	1:A:226[B]:TYR:CD2	2.21	0.58
2:E:3[B]:GLN:HG2	2:F:47:THR:HG21	1.86	0.58
1:A:110[A]:GLU:O	1:A:111[A]:GLN:C	2.44	0.56
1:A:74:GLN:HE22	1:A:133:GLY:HA2	1.70	0.56
1:A:143[B]:ARG:NH2	2:G:78:THR:O	2.39	0.55
1:A:42:TYR:CZ	1:A:46[B]:ARG:HD2	2.42	0.55
1:B:94[B]:MET:SD	1:B:115:ALA:HB2	2.47	0.55
1:A:33[A]:ARG:HD3	1:A:215:GLN:OE1	2.07	0.55
1:A:232:ILE:HD11	2:F:66:GLU:HB3	1.88	0.54
2:I:67[B]:ARG:NH1	2:J:70:ASP:OD1	2.40	0.54
2:G:63[A]:LYS:HE2	7:G:308:HOH:O	2.06	0.54
2:C:70:ASP:OD1	2:L:67:ARG:NH1	2.38	0.53
1:A:232:ILE:HB	2:E:63[B]:LYS:HG3	1.90	0.53
1:A:32[A]:ASP:OD2	1:A:35[A]:THR:HG23	2.09	0.53
1:A:223[B]:PHE:CE1	1:A:226[B]:TYR:CE2	2.97	0.52
2:L:88:TRP:CE2	4:L:202:GAL:H5	2.44	0.52
2:H:88:TRP:CE2	4:H:202:GAL:H5	2.44	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:70:ASP:OD1	2:H:67:ARG:NH1	2.34	0.52
1:A:74:GLN:HE21	1:A:74:GLN:HA	1.75	0.51
2:L:56:GLN:HG3	5:L:201:GLA:O4	2.11	0.51
2:E:8[B]:LEU:C	2:E:8[B]:LEU:HD12	2.28	0.51
1:A:223[B]:PHE:HA	1:A:226[B]:TYR:HD2	1.76	0.51
2:C:56:GLN:HG3	4:C:202:GAL:O4	2.12	0.50
2:C:8[A]:LEU:C	2:C:8[A]:LEU:HD23	2.32	0.50
2:E:63[B]:LYS:HE3	2:E:63[B]:LYS:N	2.27	0.49
1:A:42:TYR:CE1	1:A:46[B]:ARG:HD2	2.48	0.49
2:F:24[A]:ILE:CD1	7:F:311:HOH:O	2.56	0.48
2:I:81[B]:LYS:HE3	2:I:103:ASN:HB2	1.95	0.48
1:B:230:VAL:HG12	1:B:235:ARG:HD3	1.95	0.47
1:B:45[A]:ALA:O	1:B:46[A]:ARG:HG3	2.13	0.47
1:A:125:TYR:HE2	1:A:139:LEU:HD21	1.79	0.47
1:A:223[B]:PHE:HD1	1:A:226[B]:TYR:HD2	1.63	0.47
1:B:5:LEU:HD12	1:B:94[A]:MET:SD	2.54	0.47
2:L:88:TRP:CD1	4:L:202:GAL:H3	2.50	0.47
1:A:223[B]:PHE:CD1	1:A:226[B]:TYR:HE2	2.31	0.47
2:D:92[B]:THR:HG22	7:H:332:HOH:O	2.14	0.46
2:E:8[B]:LEU:HD12	2:E:9[B]:CYS:N	2.30	0.46
1:B:29:GLU:O	1:B:32[A]:ASP:HB3	2.16	0.46
1:A:74:GLN:HE22	1:A:133:GLY:CA	2.29	0.46
1:B:220:ARG:HA	2:I:78:THR:HB	1.97	0.46
2:E:63[B]:LYS:CA	2:E:63[B]:LYS:HE3	2.45	0.46
2:F:67[B]:ARG:HH21	2:G:70:ASP:N	2.14	0.46
2:H:67:ARG:NH1	7:H:301:HOH:O	2.33	0.45
1:B:29:GLU:HB2	1:B:32[A]:ASP:HB3	1.98	0.45
1:A:93[B]:ASN:OD1	1:A:94[B]:MET:HG3	2.17	0.45
1:B:43:ASP:OD1	1:B:46[A]:ARG:NH2	2.48	0.45
2:D:63:LYS:NZ	7:D:304:HOH:O	2.50	0.44
1:A:226[B]:TYR:CE1	2:H:78:THR:CG2	2.91	0.44
1:A:107:HIS:HB3	1:A:110[B]:GLU:HG2	1.99	0.44
1:B:29:GLU:HB2	1:B:32[B]:ASP:HB3	2.00	0.44
1:A:223[B]:PHE:CD1	1:A:226[B]:TYR:HD2	2.33	0.43
1:B:5:LEU:HD12	1:B:94[A]:MET:HE1	2.00	0.43
1:A:88:ILE:HG22	1:A:94[B]:MET:HE1	2.00	0.43
1:A:220:ARG:HA	2:E:78:THR:HB	2.01	0.43
1:A:33[A]:ARG:NH1	1:A:215:GLN:OE1	2.52	0.43
1:A:52:PHE:CD1	1:A:55[A]:HIS:CE1	3.07	0.42
7:B:421:HOH:O	2:L:74[B]:ILE:CG1	2.63	0.42
2:L:88:TRP:CH2	4:L:202:GAL:H62	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:86[B]:CYS:HB3	2:D:98:ALA:HB3	2.01	0.42
2:E:63[B]:LYS:HE2	2:E:63[B]:LYS:CA	2.42	0.42
1:B:74[B]:GLN:HE22	1:B:133:GLY:HA3	1.84	0.42
2:F:18:HIS:HE1	2:F:94:HIS:ND1	2.18	0.42
2:E:35:ARG:NH1	7:E:303:HOH:O	2.45	0.42
1:A:70:HIS:O	1:A:74:GLN:HG2	2.20	0.41
1:A:88:ILE:HD13	1:A:119:ILE:HD13	2.02	0.41
1:B:159:ALA:HA	1:B:162:TYR:CD2	2.54	0.41
1:A:223[B]:PHE:HD1	1:A:226[B]:TYR:CE2	2.35	0.41
1:A:223[B]:PHE:HE1	1:A:226[B]:TYR:HE2	1.65	0.41
1:A:83:TYR:CZ	1:A:130:VAL:HG11	2.55	0.41
2:G:63[A]:LYS:CE	7:G:308:HOH:O	2.66	0.41
1:B:33[A]:ARG:NH2	1:B:212:ASP:OD1	2.54	0.41
1:B:45[A]:ALA:O	1:B:46[A]:ARG:CG	2.67	0.41
1:B:233:TYR:HA	2:I:63[B]:LYS:HE3	2.02	0.41
1:A:159:ALA:HA	1:A:162:TYR:CD2	2.56	0.41
1:A:25:ARG:HB2	1:A:55[B]:HIS:CE1	2.56	0.40
1:A:141:ARG:NH1	7:A:404:HOH:O	2.49	0.40
1:B:122:SER:CB	1:B:222:ILE:HD11	2.51	0.40
2:C:15:THR:HA	2:C:87:VAL:O	2.21	0.40
1:A:98:ASN:HD22	1:A:111[A]:GLN:NE2	2.20	0.40
2:E:63[B]:LYS:HA	2:E:63[B]:LYS:CE	2.38	0.40
2:F:15:THR:HA	2:F:87:VAL:O	2.20	0.40
1:A:40:ASN:CG	1:A:166[B]:GLY:HA3	2.41	0.40
2:I:81[B]:LYS:CE	2:I:103:ASN:HB2	2.52	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	261/240 (109%)	252 (97%)	9 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	244/240 (102%)	236 (97%)	8 (3%)	0	100	100
2	C	105/103 (102%)	104 (99%)	1 (1%)	0	100	100
2	D	107/103 (104%)	106 (99%)	1 (1%)	0	100	100
2	E	106/103 (103%)	105 (99%)	1 (1%)	0	100	100
2	F	108/103 (105%)	106 (98%)	2 (2%)	0	100	100
2	G	104/103 (101%)	103 (99%)	1 (1%)	0	100	100
2	H	105/103 (102%)	104 (99%)	1 (1%)	0	100	100
2	I	107/103 (104%)	106 (99%)	1 (1%)	0	100	100
2	J	103/103 (100%)	102 (99%)	1 (1%)	0	100	100
2	K	105/103 (102%)	104 (99%)	1 (1%)	0	100	100
2	L	108/103 (105%)	106 (98%)	2 (2%)	0	100	100
All	All	1563/1510 (104%)	1534 (98%)	29 (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	226/204 (111%)	224 (99%)	2 (1%)	78	65
1	B	211/204 (103%)	207 (98%)	4 (2%)	57	34
2	C	93/89 (104%)	93 (100%)	0	100	100
2	D	95/89 (107%)	93 (98%)	2 (2%)	53	29
2	E	94/89 (106%)	93 (99%)	1 (1%)	73	57
2	F	96/89 (108%)	96 (100%)	0	100	100
2	G	92/89 (103%)	92 (100%)	0	100	100
2	H	93/89 (104%)	91 (98%)	2 (2%)	52	27
2	I	95/89 (107%)	94 (99%)	1 (1%)	73	57
2	J	91/89 (102%)	91 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	K	93/89 (104%)	92 (99%)	1 (1%)	73 57
2	L	96/89 (108%)	96 (100%)	0	100 100
All	All	1375/1298 (106%)	1362 (99%)	13 (1%)	81 65

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	16	ILE
1	A	52	PHE
2	D	9[A]	CYS
2	D	9[B]	CYS
2	E	31	LEU
2	H	9[A]	CYS
2	H	9[B]	CYS
1	B	16	ILE
1	B	32[A]	ASP
1	B	32[B]	ASP
1	B	139	LEU
2	I	31	LEU
2	K	20	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	GLN
1	A	138	GLN
1	A	234	ASN
2	F	18	HIS

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 23 ligands modelled in this entry, 2 are monoatomic - leaving 21 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	$\#\|Z\| > 2$
4	GAL	K	202	5	12,12,12	0.33	0	17,17,17	0.88	1 (5%)
5	GLA	E	202	4	12,12,12	0.52	0	17,17,17	0.84	0
4	GAL	L	202	5	12,12,12	0.45	0	17,17,17	0.85	0
4	GAL	G	201	5	12,12,12	0.20	0	17,17,17	0.62	0
4	GAL	E	203	5	12,12,12	0.60	0	17,17,17	1.12	1 (5%)
5	GLA	K	201	4	12,12,12	0.42	0	17,17,17	0.90	0
6	EPE	B	302	-	15,15,15	0.95	1 (6%)	18,20,20	1.30	3 (16%)
4	GAL	F	201	-	12,12,12	0.57	0	17,17,17	0.77	0
6	EPE	H	201	-	15,15,15	0.95	1 (6%)	18,20,20	1.03	1 (5%)
4	GAL	D	202	5	12,12,12	0.50	0	17,17,17	0.57	0
5	GLA	L	201	4	12,12,12	0.29	0	17,17,17	0.59	0
4	GAL	A	302	-	12,12,12	0.53	0	17,17,17	0.65	0
6	EPE	I	201	-	15,15,15	0.57	0	18,20,20	0.77	1 (5%)
4	GAL	H	202	5	12,12,12	0.33	0	17,17,17	0.80	0
5	GLA	D	201	4	12,12,12	0.47	0	17,17,17	0.64	0
4	GAL	I	202	-	12,12,12	0.60	0	17,17,17	0.70	1 (5%)
6	EPE	E	201	-	15,15,15	1.05	1 (6%)	18,20,20	1.35	2 (11%)
4	GAL	C	202	5	12,12,12	0.51	0	17,17,17	0.65	0
5	GLA	G	202	4	12,12,12	0.39	0	17,17,17	0.71	0
5	GLA	H	203	4	12,12,12	0.43	0	17,17,17	0.96	1 (5%)
5	GLA	C	201	4	12,12,12	0.51	0	17,17,17	0.58	0

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
4	GAL	K	202	5	-	0/2/22/22	0/1/1/1
5	GLA	E	202	4	-	0/2/22/22	0/1/1/1
4	GAL	L	202	5	-	2/2/22/22	0/1/1/1
4	GAL	G	201	5	-	1/2/22/22	0/1/1/1
4	GAL	E	203	5	-	0/2/22/22	0/1/1/1
5	GLA	K	201	4	-	2/2/22/22	0/1/1/1
6	EPE	B	302	-	-	5/9/19/19	0/1/1/1
4	GAL	F	201	-	-	0/2/22/22	0/1/1/1
6	EPE	H	201	-	-	0/9/19/19	0/1/1/1
4	GAL	D	202	5	-	0/2/22/22	0/1/1/1
5	GLA	L	201	4	-	2/2/22/22	0/1/1/1
4	GAL	A	302	-	-	0/2/22/22	0/1/1/1
6	EPE	I	201	-	-	5/9/19/19	0/1/1/1
4	GAL	H	202	5	-	0/2/22/22	0/1/1/1
5	GLA	D	201	4	-	0/2/22/22	0/1/1/1
4	GAL	I	202	-	-	0/2/22/22	0/1/1/1
6	EPE	E	201	-	-	2/9/19/19	0/1/1/1
4	GAL	C	202	5	-	0/2/22/22	0/1/1/1
5	GLA	G	202	4	-	0/2/22/22	0/1/1/1
5	GLA	H	203	4	-	0/2/22/22	0/1/1/1
5	GLA	C	201	4	-	0/2/22/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	E	201	EPE	O1S-S	3.71	1.56	1.45
6	H	201	EPE	O2S-S	3.29	1.54	1.45
6	B	302	EPE	O1S-S	3.21	1.54	1.45

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	E	201	EPE	O1S-S-C10	-4.07	102.01	106.92
6	B	302	EPE	O1S-S-C10	3.72	111.39	106.92
6	H	201	EPE	O3S-S-O1S	3.07	118.78	111.27
4	E	203	GAL	O1-C1-C2	-2.86	100.97	109.03
6	B	302	EPE	O3S-S-O2S	2.77	118.03	111.27
5	H	203	GLA	O2-C2-C3	2.53	116.20	110.35
6	I	201	EPE	O1S-S-C10	2.27	109.65	106.92
4	I	202	GAL	C1-O5-C5	-2.21	109.48	113.66

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	302	EPE	O3S-S-O1S	-2.19	105.93	111.27
4	K	202	GAL	C1-C2-C3	2.11	114.70	110.31
6	E	201	EPE	O3S-S-O2S	2.04	116.26	111.27

There are no chirality outliers.

All (19) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	E	201	EPE	C10-C9-N1-C2
6	B	302	EPE	C9-C10-S-O1S
6	E	201	EPE	N4-C7-C8-O8
5	K	201	GLA	C4-C5-C6-O6
5	L	201	GLA	C4-C5-C6-O6
5	K	201	GLA	O5-C5-C6-O6
6	I	201	EPE	C9-C10-S-O3S
4	L	202	GAL	O5-C5-C6-O6
5	L	201	GLA	O5-C5-C6-O6
6	B	302	EPE	C9-C10-S-O3S
6	I	201	EPE	C10-C9-N1-C2
6	I	201	EPE	C10-C9-N1-C6
6	B	302	EPE	C9-C10-S-O2S
6	I	201	EPE	C9-C10-S-O1S
6	I	201	EPE	C9-C10-S-O2S
4	L	202	GAL	C4-C5-C6-O6
4	G	201	GAL	C4-C5-C6-O6
6	B	302	EPE	C8-C7-N4-C3
6	B	302	EPE	C8-C7-N4-C5

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	L	202	GAL	3	0
5	L	201	GLA	1	0
4	H	202	GAL	1	0
5	D	201	GLA	1	0
4	C	202	GAL	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

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, 95^{th} 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	231/240 (96%)	0.08	11 (4%) 30 28	17, 26, 54, 69	0
1	B	231/240 (96%)	0.12	16 (6%) 16 15	17, 27, 56, 74	0
2	C	103/103 (100%)	-0.49	0 100 100	15, 21, 29, 54	0
2	D	103/103 (100%)	-0.47	0 100 100	15, 21, 30, 50	0
2	E	103/103 (100%)	-0.38	0 100 100	15, 22, 32, 49	0
2	F	103/103 (100%)	-0.35	1 (0%) 82 82	15, 21, 30, 52	0
2	G	103/103 (100%)	-0.11	1 (0%) 82 82	16, 23, 32, 60	0
2	H	103/103 (100%)	-0.36	0 100 100	18, 23, 33, 52	0
2	I	103/103 (100%)	-0.38	0 100 100	16, 22, 32, 55	0
2	J	103/103 (100%)	-0.39	0 100 100	16, 21, 30, 52	0
2	K	103/103 (100%)	-0.30	1 (0%) 82 82	16, 20, 30, 53	0
2	L	103/103 (100%)	-0.49	0 100 100	17, 22, 31, 45	0
All	All	1492/1510 (98%)	-0.23	30 (2%) 65 64	15, 23, 42, 74	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	233	TYR	10.2
1	A	233	TYR	8.6
1	A	36	GLN	6.2
1	B	36	GLN	5.9
1	A	232	ILE	5.6
1	B	132	PHE	5.4
1	A	34[A]	GLY	4.6
1	A	35[A]	THR	4.5
1	A	194	SER	4.5
1	B	189	ASN	4.0
1	B	232	ILE	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	236	ILE	4.0
1	B	35	THR	3.8
1	B	34[A]	GLY	3.6
1	A	195	MET	3.5
1	B	195	MET	3.2
1	B	231	ASP	3.2
1	B	230	VAL	2.6
1	A	230	VAL	2.6
1	A	231	ASP	2.5
2	K	1	THR	2.4
1	B	234	ASN	2.3
1	A	143[A]	ARG	2.3
2	G	9[A]	CYS	2.3
1	B	32[A]	ASP	2.2
1	B	146[A]	ARG	2.2
2	F	103	ASN	2.1
1	B	48	THR	2.1
1	B	33[A]	ARG	2.1
1	A	137	GLU	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 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, 95^{th} 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
6	EPE	E	201	15/15	0.67	0.23	38,48,52,55	15
6	EPE	I	201	15/15	0.83	0.20	34,44,53,54	15
4	GAL	A	302	12/12	0.86	0.14	25,30,33,34	0
6	EPE	B	302	15/15	0.87	0.12	32,35,48,50	15

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GAL	I	202	12/12	0.92	0.10	21,25,29,34	0
4	GAL	F	201	12/12	0.92	0.10	22,24,28,32	0
5	GLA	G	202	12/12	0.93	0.11	18,20,24,27	12
6	EPE	H	201	15/15	0.93	0.09	30,33,37,40	15
5	GLA	E	202	12/12	0.95	0.10	16,18,19,22	12
4	GAL	D	202	12/12	0.95	0.09	13,16,18,18	12
5	GLA	C	201	12/12	0.95	0.09	13,14,16,16	12
4	GAL	G	201	12/12	0.95	0.12	20,23,26,27	12
4	GAL	E	203	12/12	0.95	0.08	16,17,18,20	12
4	GAL	K	202	12/12	0.95	0.08	16,18,20,20	12
5	GLA	D	201	12/12	0.95	0.08	18,20,22,22	12
5	GLA	K	201	12/12	0.96	0.09	19,21,23,23	12
5	GLA	L	201	12/12	0.96	0.09	20,22,23,23	12
4	GAL	L	202	12/12	0.96	0.08	17,18,19,21	12
4	GAL	C	202	12/12	0.97	0.08	13,17,20,21	12
5	GLA	H	203	12/12	0.97	0.08	18,20,22,23	12
4	GAL	H	202	12/12	0.97	0.08	21,23,24,25	12
3	NA	B	301	1/1	0.99	0.07	19,19,19,19	0
3	NA	A	301	1/1	0.99	0.07	15,15,15,15	0

6.5 Other polymers (i)

There are no such residues in this entry.