



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 8, 2017 – 11:24 AM EDT

PDB ID : 5NNP
Title : Structure of Naa15/Naa10 bound to HypK-THB
Authors : Weyer, F.A.; Gumiero, A.; Kopp, J.; Sinning, I.
Deposited on : 2017-04-10
Resolution : 2.60 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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029077
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20029077

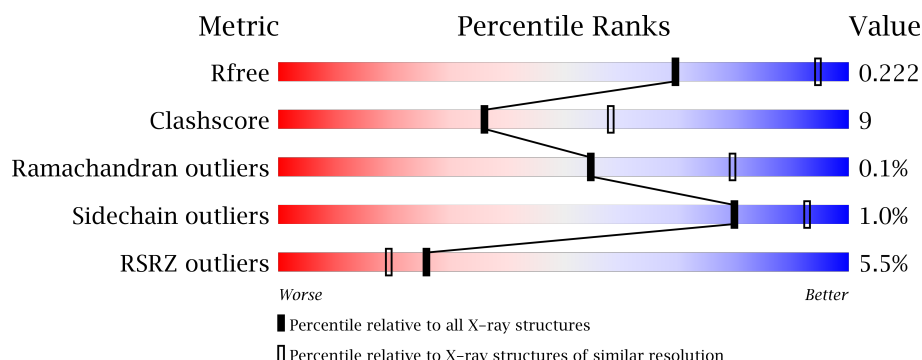
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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}	100719	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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 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	745	<div> <div>5%</div> <div> <div></div> <div>74%</div> <div>16%</div> <div>11%</div> </div> </div>
1	E	745	<div> <div>6%</div> <div> <div></div> <div>73%</div> <div>15%</div> <div>12%</div> </div> </div>
2	B	195	<div> <div>4%</div> <div> <div></div> <div>78%</div> <div>14%</div> <div>8%</div> </div> </div>
2	F	195	<div> <div>2%</div> <div> <div></div> <div>75%</div> <div>16%</div> <div>9%</div> </div> </div>
3	C	62	<div> <div>2%</div> <div> <div></div> <div>60%</div> <div>13%</div> <div>27%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	62	
4	I	4	
4	L	4	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	GOL	A	801	-	-	X	X
5	GOL	A	804	-	-	-	X
5	GOL	A	805	-	-	-	X
5	GOL	A	807	-	-	-	X
5	GOL	A	808	-	-	-	X
5	GOL	A	809	-	-	-	X
5	GOL	A	810	-	-	-	X
5	GOL	A	812	-	-	-	X
5	GOL	B	201	-	-	-	X
5	GOL	C	201	-	-	-	X
5	GOL	C	202	-	-	-	X
5	GOL	E	801	-	-	-	X
5	GOL	E	802	-	-	-	X
5	GOL	E	803	-	-	-	X
5	GOL	E	805	-	-	-	X
5	GOL	E	806	-	-	-	X
5	GOL	F	201	-	-	-	X
5	GOL	F	202	-	-	-	X
5	GOL	F	203	-	-	-	X
5	GOL	F	204	-	-	-	X
5	GOL	F	205	-	-	-	X
5	GOL	F	206	-	-	-	X
6	PO4	E	807	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 15101 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called N-terminal acetyltransferase-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	665	Total	C	N	O	S	0	0	0
			5390	3428	943	997	22			
1	E	658	Total	C	N	O	S	0	1	0
			5346	3400	934	990	22			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP G0S4M4
A	1	GLY	-	expression tag	UNP G0S4M4
A	707	LYS	ASN	conflict	UNP G0S4M4
A	724	ASP	GLY	conflict	UNP G0S4M4
E	0	MET	-	initiating methionine	UNP G0S4M4
E	1	GLY	-	expression tag	UNP G0S4M4
E	707	LYS	ASN	conflict	UNP G0S4M4
E	724	ASP	GLY	conflict	UNP G0S4M4

- Molecule 2 is a protein called Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	179	Total	C	N	O	S	0	0	0
			1458	930	251	270	7			
2	F	178	Total	C	N	O	S	0	0	0
			1450	926	250	267	7			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	FME	-	expression tag	UNP G0SEE8
B	190	HIS	-	expression tag	UNP G0SEE8
B	191	HIS	-	expression tag	UNP G0SEE8
B	192	HIS	-	expression tag	UNP G0SEE8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	193	HIS	-	expression tag	UNP G0SEE8
B	194	HIS	-	expression tag	UNP G0SEE8
B	195	HIS	-	expression tag	UNP G0SEE8
F	1	FME	-	expression tag	UNP G0SEE8
F	190	HIS	-	expression tag	UNP G0SEE8
F	191	HIS	-	expression tag	UNP G0SEE8
F	192	HIS	-	expression tag	UNP G0SEE8
F	193	HIS	-	expression tag	UNP G0SEE8
F	194	HIS	-	expression tag	UNP G0SEE8
F	195	HIS	-	expression tag	UNP G0SEE8

- Molecule 3 is a protein called Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	45	Total	C	N	O	S	0	0	0
			337	214	55	66	2			
3	G	45	Total	C	N	O	S	0	0	0
			337	214	55	66	2			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	79	GLY	-	expression tag	UNP G0SCY6
C	80	ALA	-	expression tag	UNP G0SCY6
C	81	MET	-	expression tag	UNP G0SCY6
C	82	GLY	-	expression tag	UNP G0SCY6
C	127	GLY	-	expression tag	UNP G0SCY6
C	128	SER	-	expression tag	UNP G0SCY6
C	129	GLY	-	expression tag	UNP G0SCY6
C	130	SER	-	expression tag	UNP G0SCY6
C	131	GLY	-	expression tag	UNP G0SCY6
C	132	SER	-	expression tag	UNP G0SCY6
C	133	TRP	-	expression tag	UNP G0SCY6
C	134	SER	-	expression tag	UNP G0SCY6
C	135	HIS	-	expression tag	UNP G0SCY6
C	136	PRO	-	expression tag	UNP G0SCY6
C	137	GLN	-	expression tag	UNP G0SCY6
C	138	PHE	-	expression tag	UNP G0SCY6
C	139	GLU	-	expression tag	UNP G0SCY6
C	140	LYS	-	expression tag	UNP G0SCY6
G	79	GLY	-	expression tag	UNP G0SCY6
G	80	ALA	-	expression tag	UNP G0SCY6

Continued on next page...

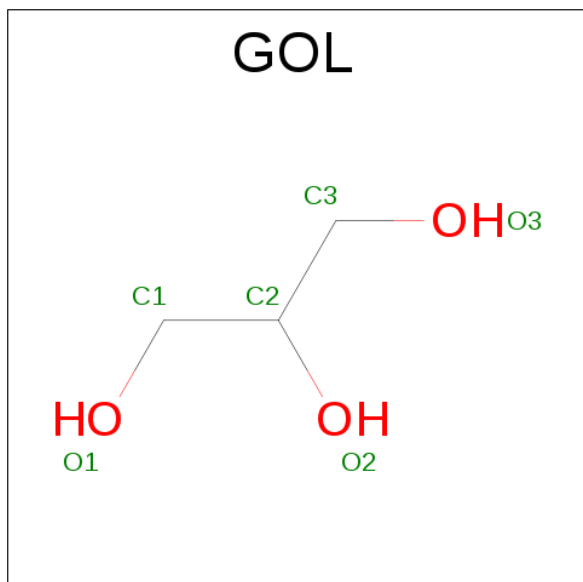
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	81	MET	-	expression tag	UNP G0SCY6
G	82	GLY	-	expression tag	UNP G0SCY6
G	127	GLY	-	expression tag	UNP G0SCY6
G	128	SER	-	expression tag	UNP G0SCY6
G	129	GLY	-	expression tag	UNP G0SCY6
G	130	SER	-	expression tag	UNP G0SCY6
G	131	GLY	-	expression tag	UNP G0SCY6
G	132	SER	-	expression tag	UNP G0SCY6
G	133	TRP	-	expression tag	UNP G0SCY6
G	134	SER	-	expression tag	UNP G0SCY6
G	135	HIS	-	expression tag	UNP G0SCY6
G	136	PRO	-	expression tag	UNP G0SCY6
G	137	GLN	-	expression tag	UNP G0SCY6
G	138	PHE	-	expression tag	UNP G0SCY6
G	139	GLU	-	expression tag	UNP G0SCY6
G	140	LYS	-	expression tag	UNP G0SCY6

- Molecule 4 is a protein called Ser-Glu-Ser-Ser.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	I	4	Total	C	N	O	0	0	0
			28	14	4	10			
4	L	4	Total	C	N	O	0	0	0
			28	14	4	10			

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



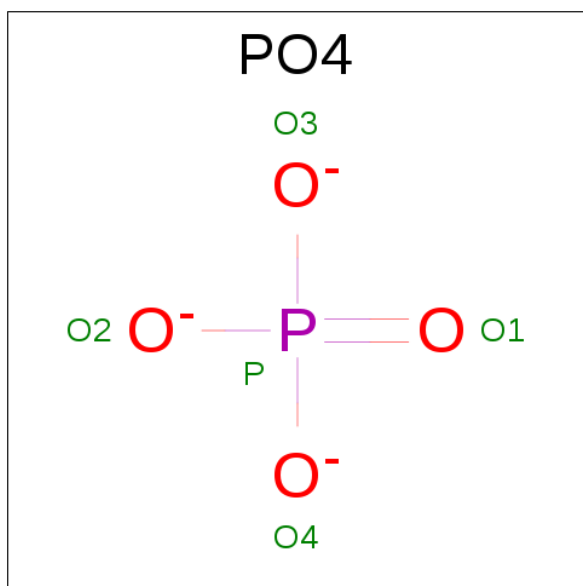
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	F	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

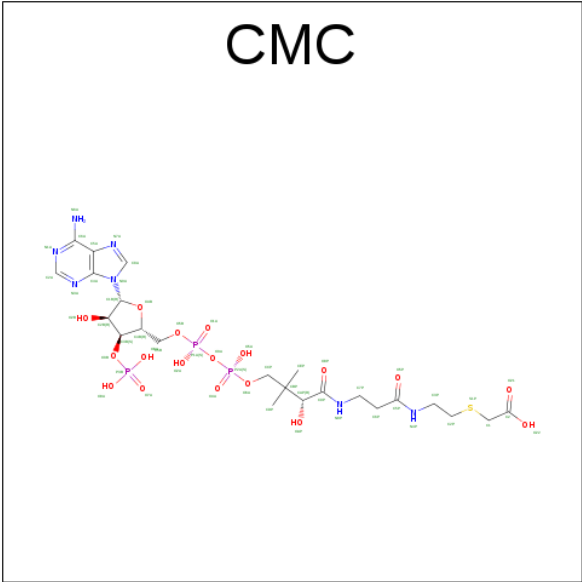
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	F	1	Total	C	O	0	0
			6	3	3		
5	F	1	Total	C	O	0	0
			6	3	3		
5	F	1	Total	C	O	0	0
			6	3	3		
5	F	1	Total	C	O	0	0
			6	3	3		
5	F	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	P	0	0
			5	4	1		
6	E	1	Total	O	P	0	0
			5	4	1		
6	E	1	Total	O	P	0	0
			5	4	1		
6	F	1	Total	O	P	0	0
			5	4	1		

- Molecule 7 is CARBOXYMETHYL COENZYME *A (three-letter code: CMC) (formula: C₂₃H₃₈N₇O₁₈P₃S).

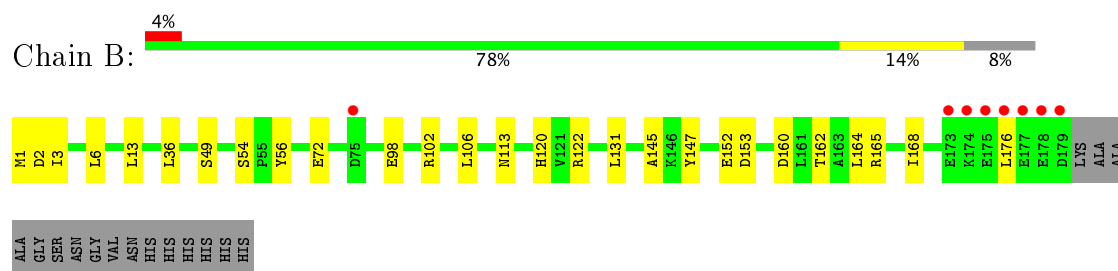


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
7	I	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		
7	L	1	Total	C	N	O	P	S	0	0
			51	23	7	17	3	1		

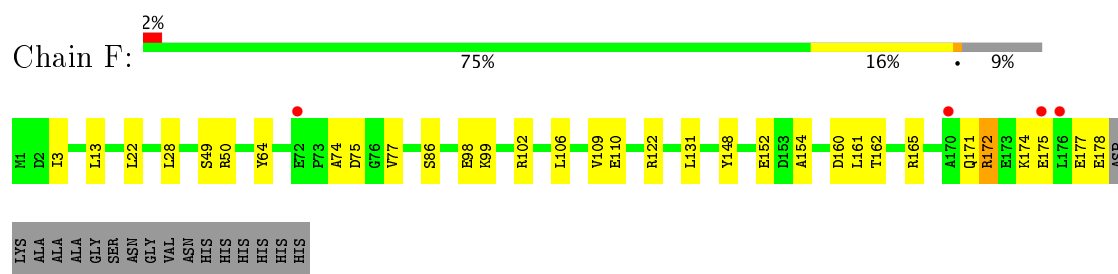
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	155	Total	O	0	0
			155	155		
8	B	52	Total	O	0	0
			52	52		
8	C	4	Total	O	0	0
			4	4		
8	E	159	Total	O	0	0
			159	159		
8	F	54	Total	O	0	0
			54	54		
8	G	6	Total	O	0	0
			6	6		
8	I	8	Total	O	0	0
			8	8		
8	L	5	Total	O	0	0
			5	5		

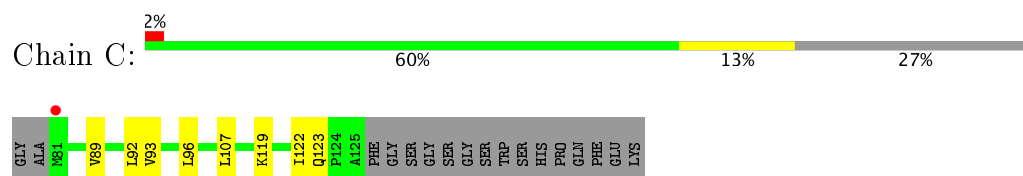
- Molecule 2: Putative uncharacterized protein



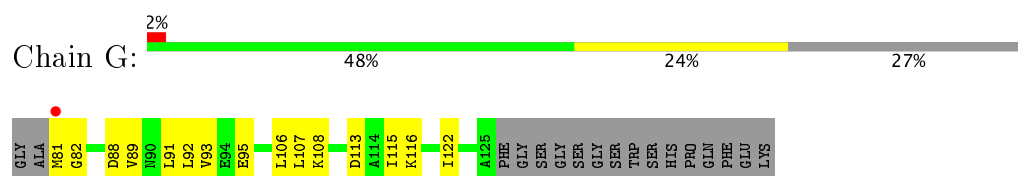
- Molecule 2: Putative uncharacterized protein



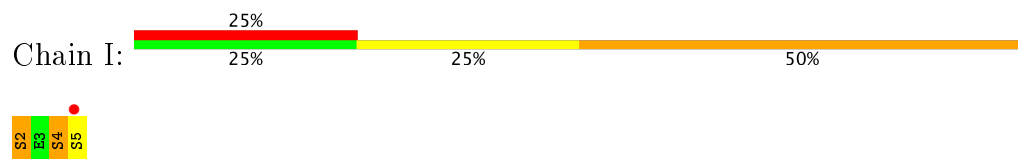
- Molecule 3: Putative uncharacterized protein



- Molecule 3: Putative uncharacterized protein



- Molecule 4: Ser-Glu-Ser-Ser



- Molecule 4: Ser-Glu-Ser-Ser





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	93.32Å 106.04Å 130.94Å 90.00° 94.98° 90.00°	Depositor
Resolution (Å)	49.32 – 2.60 49.32 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.32-2.60) 99.7 (49.32-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.75 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.161 , 0.220 0.163 , 0.222	Depositor DCC
R_{free} test set	3925 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	44.1	Xtriage
Anisotropy	0.274	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 54.5	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	0.95	EDS
Total number of atoms	15101	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: GOL, PO4, FME, CMC

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	0.38	0/5501	0.49	0/7413
1	E	0.38	0/5455	0.49	0/7348
2	B	0.40	0/1482	0.53	0/2008
2	F	0.40	0/1474	0.54	0/1997
3	C	0.35	0/340	0.47	0/457
3	G	0.31	0/340	0.47	0/457
4	I	2.03	1/27 (3.7%)	1.44	0/33
4	L	2.07	1/27 (3.7%)	1.44	0/33
All	All	0.40	2/14646 (0.0%)	0.50	0/19746

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	L	4	SER	C-N	5.88	1.47	1.34
4	I	4	SER	C-N	5.63	1.47	1.34

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	5390	0	5391	104	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	5346	0	5339	107	0
2	B	1458	0	1449	23	0
2	F	1450	0	1445	23	0
3	C	337	0	352	5	0
3	G	337	0	352	13	0
4	I	28	0	20	2	0
4	L	28	0	20	3	0
5	A	72	0	96	14	0
5	B	6	0	8	2	0
5	C	12	0	16	0	0
5	E	36	0	48	5	0
5	F	36	0	48	4	0
6	A	5	0	0	0	0
6	E	10	0	0	3	0
6	F	5	0	0	1	0
7	I	51	0	33	4	0
7	L	51	0	33	1	0
8	A	155	0	0	9	0
8	B	52	0	0	1	0
8	C	4	0	0	0	0
8	E	159	0	0	4	0
8	F	54	0	0	5	0
8	G	6	0	0	0	0
8	I	8	0	0	2	0
8	L	5	0	0	0	0
All	All	15101	0	14650	267	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 9.

The worst 5 of 267 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:19:HIS:CB	1:E:24:GLN:HG3	1.73	1.16
1:E:19:HIS:O	1:E:24:GLN:HG2	1.52	1.10
1:E:654:LEU:HD12	1:E:681:VAL:HG13	1.33	1.08
1:E:19:HIS:HB3	1:E:24:GLN:HG3	1.01	1.00
1:E:19:HIS:HB3	1:E:24:GLN:CG	1.91	0.99

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	659/745 (88%)	648 (98%)	11 (2%)	0	100	100
1	E	653/745 (88%)	638 (98%)	13 (2%)	2 (0%)	44	70
2	B	177/195 (91%)	176 (99%)	1 (1%)	0	100	100
2	F	176/195 (90%)	175 (99%)	1 (1%)	0	100	100
3	C	43/62 (69%)	43 (100%)	0	0	100	100
3	G	43/62 (69%)	43 (100%)	0	0	100	100
4	I	2/4 (50%)	2 (100%)	0	0	100	100
4	L	2/4 (50%)	2 (100%)	0	0	100	100
All	All	1755/2012 (87%)	1727 (98%)	26 (2%)	2 (0%)	55	79

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	23	LYS
1	E	17	ILE

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	560/614 (91%)	555 (99%)	5 (1%)	82	93
1	E	554/614 (90%)	550 (99%)	4 (1%)	87	96
2	B	155/166 (93%)	154 (99%)	1 (1%)	89	96

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	F	154/166 (93%)	152 (99%)	2 (1%)	73	90
3	C	35/47 (74%)	35 (100%)	0	100	100
3	G	35/47 (74%)	35 (100%)	0	100	100
4	I	4/4 (100%)	2 (50%)	2 (50%)	0	0
4	L	4/4 (100%)	3 (75%)	1 (25%)	1	1
All	All	1501/1662 (90%)	1486 (99%)	15 (1%)	80	93

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	76	SER
1	E	319	LYS
4	I	2	SER
2	B	54	SER
2	F	172	ARG

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	19	HIS
1	E	132	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	FME	B	1	2	9,9,10	0.95	0	7,9,11	1.79	1 (14%)
2	FME	F	1	2	9,9,10	1.09	1 (11%)	7,9,11	1.25	1 (14%)

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	FME	B	1	2	-	0/6/9/11	0/0/0/0
2	FME	F	1	2	-	0/6/9/11	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1	FME	CA-C	2.34	1.53	1.50

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1	FME	CB-CA-C	-4.09	104.91	111.65
2	F	1	FME	CB-CA-C	-2.58	107.40	111.65

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1	FME	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

33 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	$\# Z > 2$
5	GOL	A	801	-	5,5,5	0.34	0	5,5,5	0.42	0
5	GOL	A	802	-	5,5,5	0.38	0	5,5,5	0.21	0
5	GOL	A	803	-	5,5,5	0.33	0	5,5,5	0.46	0
5	GOL	A	804	-	5,5,5	0.31	0	5,5,5	0.34	0
5	GOL	A	805	-	5,5,5	0.28	0	5,5,5	0.34	0
5	GOL	A	806	-	5,5,5	0.33	0	5,5,5	0.29	0
5	GOL	A	807	-	5,5,5	0.33	0	5,5,5	0.20	0
5	GOL	A	808	-	5,5,5	0.36	0	5,5,5	0.35	0
5	GOL	A	809	-	5,5,5	0.32	0	5,5,5	0.25	0
5	GOL	A	810	-	5,5,5	0.45	0	5,5,5	0.25	0
5	GOL	A	811	-	5,5,5	0.31	0	5,5,5	0.26	0
5	GOL	A	812	-	5,5,5	0.33	0	5,5,5	0.62	0
6	PO4	A	813	-	4,4,4	0.69	0	6,6,6	0.50	0
5	GOL	B	201	-	5,5,5	0.42	0	5,5,5	0.12	0
5	GOL	C	201	-	5,5,5	0.35	0	5,5,5	0.27	0
5	GOL	C	202	-	5,5,5	0.31	0	5,5,5	0.16	0
5	GOL	E	801	-	5,5,5	0.32	0	5,5,5	0.62	0
5	GOL	E	802	-	5,5,5	0.35	0	5,5,5	0.46	0
5	GOL	E	803	-	5,5,5	0.49	0	5,5,5	0.20	0
5	GOL	E	804	-	5,5,5	0.32	0	5,5,5	0.33	0
5	GOL	E	805	-	5,5,5	0.40	0	5,5,5	0.19	0
5	GOL	E	806	-	5,5,5	0.41	0	5,5,5	0.44	0
6	PO4	E	807	-	4,4,4	0.76	0	6,6,6	0.50	0
6	PO4	E	808	-	4,4,4	0.72	0	6,6,6	0.52	0
5	GOL	F	201	-	5,5,5	0.32	0	5,5,5	0.28	0
5	GOL	F	202	-	5,5,5	0.35	0	5,5,5	0.31	0
5	GOL	F	203	-	5,5,5	0.42	0	5,5,5	0.26	0
5	GOL	F	204	-	5,5,5	0.39	0	5,5,5	0.24	0
5	GOL	F	205	-	5,5,5	0.36	0	5,5,5	0.39	0
5	GOL	F	206	-	5,5,5	0.36	0	5,5,5	0.26	0
6	PO4	F	207	-	4,4,4	0.78	0	6,6,6	0.40	0
7	CMC	I	101	4	46,53,54	1.78	9 (19%)	51,78,80	1.94	9 (17%)
7	CMC	L	101	4	46,53,54	1.72	9 (19%)	51,78,80	1.71	5 (9%)

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
5	GOL	A	801	-	-	0/4/4/4	0/0/0/0
5	GOL	A	802	-	-	0/4/4/4	0/0/0/0
5	GOL	A	803	-	-	0/4/4/4	0/0/0/0
5	GOL	A	804	-	-	0/4/4/4	0/0/0/0
5	GOL	A	805	-	-	0/4/4/4	0/0/0/0
5	GOL	A	806	-	-	0/4/4/4	0/0/0/0
5	GOL	A	807	-	-	0/4/4/4	0/0/0/0
5	GOL	A	808	-	-	0/4/4/4	0/0/0/0
5	GOL	A	809	-	-	0/4/4/4	0/0/0/0
5	GOL	A	810	-	-	0/4/4/4	0/0/0/0
5	GOL	A	811	-	-	0/4/4/4	0/0/0/0
5	GOL	A	812	-	-	0/4/4/4	0/0/0/0
6	PO4	A	813	-	-	0/0/0/0	0/0/0/0
5	GOL	B	201	-	-	0/4/4/4	0/0/0/0
5	GOL	C	201	-	-	0/4/4/4	0/0/0/0
5	GOL	C	202	-	-	0/4/4/4	0/0/0/0
5	GOL	E	801	-	-	0/4/4/4	0/0/0/0
5	GOL	E	802	-	-	0/4/4/4	0/0/0/0
5	GOL	E	803	-	-	0/4/4/4	0/0/0/0
5	GOL	E	804	-	-	0/4/4/4	0/0/0/0
5	GOL	E	805	-	-	0/4/4/4	0/0/0/0
5	GOL	E	806	-	-	0/4/4/4	0/0/0/0
6	PO4	E	807	-	-	0/0/0/0	0/0/0/0
6	PO4	E	808	-	-	0/0/0/0	0/0/0/0
5	GOL	F	201	-	-	0/4/4/4	0/0/0/0
5	GOL	F	202	-	-	0/4/4/4	0/0/0/0
5	GOL	F	203	-	-	0/4/4/4	0/0/0/0
5	GOL	F	204	-	-	0/4/4/4	0/0/0/0
5	GOL	F	205	-	-	0/4/4/4	0/0/0/0
5	GOL	F	206	-	-	0/4/4/4	0/0/0/0
6	PO4	F	207	-	-	0/0/0/0	0/0/0/0
7	CMC	I	101	4	-	0/46/67/68	0/3/3/3
7	CMC	L	101	4	-	0/46/67/68	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	I	101	CMC	C2B-C1B	-3.63	1.47	1.53
7	L	101	CMC	C2B-C3B	-3.33	1.45	1.53
7	I	101	CMC	C2B-C3B	-3.21	1.45	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	L	101	CMC	C2B-C1B	-2.76	1.49	1.53
7	L	101	CMC	OAP-CAP	-2.11	1.38	1.42

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	I	101	CMC	N3A-C2A-N1A	-9.03	120.99	128.86
7	L	101	CMC	N3A-C2A-N1A	-8.45	121.50	128.86
7	I	101	CMC	C4A-C5A-N7A	-3.25	106.27	109.41
7	I	101	CMC	C7P-C6P-C5P	-3.12	107.20	112.22
7	L	101	CMC	C1B-N9A-C4A	-3.06	121.35	126.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

17 monomers are involved in 34 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	801	GOL	4	0
5	A	805	GOL	2	0
5	A	806	GOL	1	0
5	A	808	GOL	3	0
5	A	809	GOL	1	0
5	A	810	GOL	2	0
5	A	811	GOL	1	0
5	B	201	GOL	2	0
5	E	803	GOL	3	0
5	E	806	GOL	2	0
6	E	807	PO4	3	0
5	F	201	GOL	1	0
5	F	204	GOL	1	0
5	F	206	GOL	2	0
6	F	207	PO4	1	0
7	I	101	CMC	4	0
7	L	101	CMC	1	0

5.7 Other polymers

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 ⓘ

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	665/745 (89%)	-0.03	39 (5%) 23 17	28, 49, 104, 151	0
1	E	658/745 (88%)	0.04	43 (6%) 20 14	28, 51, 122, 184	0
2	B	178/195 (91%)	0.07	8 (4%) 34 26	27, 37, 123, 153	0
2	F	177/195 (90%)	-0.07	4 (2%) 61 54	25, 38, 121, 179	0
3	C	45/62 (72%)	-0.19	1 (2%) 62 56	47, 57, 89, 107	0
3	G	45/62 (72%)	-0.10	1 (2%) 62 56	50, 65, 83, 113	0
4	I	4/4 (100%)	0.70	1 (25%) 1 0	41, 46, 70, 99	0
4	L	4/4 (100%)	0.21	0 100 100	35, 42, 56, 90	0
All	All	1776/2012 (88%)	-0.00	97 (5%) 26 19	25, 49, 113, 184	0

The worst 5 of 97 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	25	TYR	7.5
1	E	573	PRO	6.6
1	E	32	ALA	5.9
1	E	36	LEU	5.5
2	B	176	LEU	5.4

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

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
2	FME	B	1	10/11	0.98	0.16	-	26,49,60,69	0
2	FME	F	1	10/11	0.97	0.20	-	39,59,78,79	0

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
5	GOL	A	805	6/6	0.75	0.41	26.57	79,84,92,99	0
5	GOL	A	801	6/6	0.78	0.54	18.02	109,111,112,113	0
5	GOL	E	802	6/6	0.92	0.36	17.34	93,94,95,98	0
5	GOL	F	201	6/6	0.60	0.51	16.72	104,107,109,110	0
5	GOL	C	202	6/6	0.78	0.41	12.51	82,90,98,108	0
5	GOL	B	201	6/6	0.83	0.35	11.04	96,100,101,101	0
5	GOL	E	803	6/6	0.85	0.39	10.19	92,96,99,100	0
5	GOL	A	807	6/6	0.74	0.42	9.12	89,90,96,101	0
5	GOL	E	805	6/6	0.70	0.30	8.72	67,73,83,84	0
5	GOL	A	804	6/6	0.84	0.35	8.38	51,72,79,80	0
5	GOL	E	806	6/6	0.73	0.30	8.10	62,70,72,76	0
5	GOL	F	206	6/6	0.86	0.24	7.28	72,73,78,84	0
5	GOL	A	808	6/6	0.88	0.32	5.47	85,95,95,102	0
5	GOL	E	801	6/6	0.79	0.29	5.37	60,64,70,70	0
5	GOL	A	809	6/6	0.85	0.27	5.04	80,92,95,104	0
5	GOL	F	204	6/6	0.80	0.29	4.51	72,81,83,84	0
5	GOL	F	205	6/6	0.78	0.32	4.36	58,69,79,81	0
5	GOL	A	812	6/6	0.88	0.24	4.17	49,67,71,74	0
5	GOL	F	203	6/6	0.92	0.26	4.02	63,67,69,71	0
5	GOL	A	810	6/6	0.93	0.26	4.01	61,67,70,73	0
5	GOL	F	202	6/6	0.80	0.23	3.61	57,77,80,81	0
5	GOL	C	201	6/6	0.89	0.26	3.27	63,70,72,73	0
6	PO4	E	808	5/5	0.88	0.28	1.73	109,113,118,120	0
7	CMC	I	101	51/52	0.97	0.19	0.69	18,44,106,112	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
7	CMC	L	101	51/52	0.97	0.17	-0.03	24,43,84,88	0
6	PO4	E	807	5/5	0.94	0.14	-0.66	94,101,103,105	0
5	GOL	A	806	6/6	0.89	0.21	-	79,80,84,88	0
6	PO4	A	813	5/5	0.97	0.11	-	90,91,92,96	0
5	GOL	E	804	6/6	0.93	0.09	-	88,90,91,92	0
5	GOL	A	802	6/6	0.84	0.22	-	70,80,86,86	0
5	GOL	A	811	6/6	0.83	0.17	-	85,89,90,90	0
6	PO4	F	207	5/5	0.76	0.27	-	165,167,168,168	0
5	GOL	A	803	6/6	0.86	0.22	-	62,69,78,88	0

6.5 Other polymers [i](#)

There are no such residues in this entry.