



wwPDB X-ray Structure Validation Summary Report

Feb 15, 2017 – 02:50 am GMT

PDB ID : 3OC3
Title : Crystal structure of the Mot1 N-terminal domain in complex with TBP
Authors : Wollmann, P.; Cui, S.; Viswanathan, R.; Berninghausen, O.; Wells, M.N.;
Moldt, M.; Witte, G.; Butryn, A.; Wendler, P.; Beckmann, R.; Auble, D.T.;
Hopfner, K.-P.
Deposited on : 2010-08-09
Resolution : 3.10 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

The following versions of software and data (see [references](#) ) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
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) : recalc28949

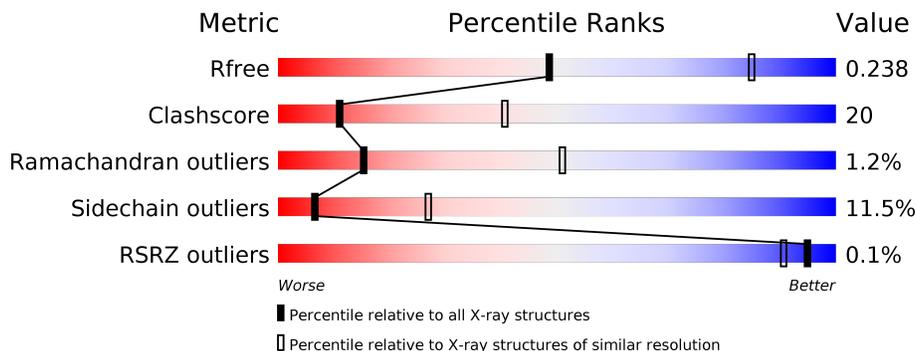
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1001 (3.12-3.08)
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)

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	800	
1	B	800	
2	C	218	
2	D	218	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	MES	D	199	-	-	-	X

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 15037 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 HELICASE MOT1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	752	6086	3926	984	1153	23	104	0	0
1	B	749	6064	3911	981	1149	23	72	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	EXPRESSION TAG	UNP Q8SVZ5
A	-20	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
A	-19	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-18	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-17	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-16	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-15	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-14	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-13	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-12	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-11	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-10	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
A	-9	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
A	-8	GLY	-	EXPRESSION TAG	UNP Q8SVZ5
A	-7	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
A	-6	GLY	-	EXPRESSION TAG	UNP Q8SVZ5
A	-5	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
A	-4	ARG	-	EXPRESSION TAG	UNP Q8SVZ5
A	-3	ASN	-	EXPRESSION TAG	UNP Q8SVZ5
A	-2	MET	-	EXPRESSION TAG	UNP Q8SVZ5
A	-1	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
A	0	SER	-	EXPRESSION TAG	UNP Q8SVZ5
B	-21	MET	-	EXPRESSION TAG	UNP Q8SVZ5
B	-20	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
B	-19	HIS	-	EXPRESSION TAG	UNP Q8SVZ5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-18	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-17	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-16	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-15	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-14	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-13	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-12	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-11	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-10	HIS	-	EXPRESSION TAG	UNP Q8SVZ5
B	-9	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
B	-8	GLY	-	EXPRESSION TAG	UNP Q8SVZ5
B	-7	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
B	-6	GLY	-	EXPRESSION TAG	UNP Q8SVZ5
B	-5	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
B	-4	ARG	-	EXPRESSION TAG	UNP Q8SVZ5
B	-3	ASN	-	EXPRESSION TAG	UNP Q8SVZ5
B	-2	MET	-	EXPRESSION TAG	UNP Q8SVZ5
B	-1	ALA	-	EXPRESSION TAG	UNP Q8SVZ5
B	0	SER	-	EXPRESSION TAG	UNP Q8SVZ5

- Molecule 2 is a protein called TRANSCRIPTION INITIATION FACTOR TFIID (TFIID-1).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	178	1417	919	247	244	7	8	1	0
2	D	178	1417	919	247	244	7	6	1	0

There are 40 discrepancies between the modelled and reference sequences:

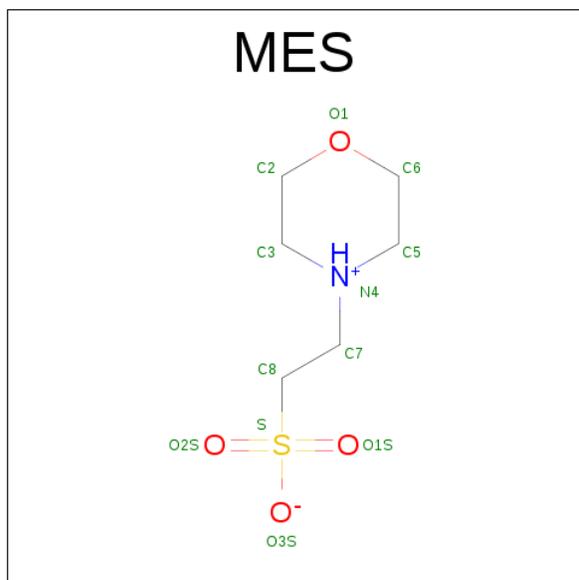
Chain	Residue	Modelled	Actual	Comment	Reference
C	-19	MET	-	EXPRESSION TAG	UNP Q8ST28
C	-18	GLY	-	EXPRESSION TAG	UNP Q8ST28
C	-17	SER	-	EXPRESSION TAG	UNP Q8ST28
C	-16	SER	-	EXPRESSION TAG	UNP Q8ST28
C	-15	HIS	-	EXPRESSION TAG	UNP Q8ST28
C	-14	HIS	-	EXPRESSION TAG	UNP Q8ST28
C	-13	HIS	-	EXPRESSION TAG	UNP Q8ST28
C	-12	HIS	-	EXPRESSION TAG	UNP Q8ST28
C	-11	HIS	-	EXPRESSION TAG	UNP Q8ST28
C	-10	HIS	-	EXPRESSION TAG	UNP Q8ST28
C	-9	SER	-	EXPRESSION TAG	UNP Q8ST28

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	SER	-	EXPRESSION TAG	UNP Q8ST28
C	-7	GLY	-	EXPRESSION TAG	UNP Q8ST28
C	-6	LEU	-	EXPRESSION TAG	UNP Q8ST28
C	-5	VAL	-	EXPRESSION TAG	UNP Q8ST28
C	-4	PRO	-	EXPRESSION TAG	UNP Q8ST28
C	-3	ARG	-	EXPRESSION TAG	UNP Q8ST28
C	-2	GLY	-	EXPRESSION TAG	UNP Q8ST28
C	-1	SER	-	EXPRESSION TAG	UNP Q8ST28
C	0	HIS	-	EXPRESSION TAG	UNP Q8ST28
D	-19	MET	-	EXPRESSION TAG	UNP Q8ST28
D	-18	GLY	-	EXPRESSION TAG	UNP Q8ST28
D	-17	SER	-	EXPRESSION TAG	UNP Q8ST28
D	-16	SER	-	EXPRESSION TAG	UNP Q8ST28
D	-15	HIS	-	EXPRESSION TAG	UNP Q8ST28
D	-14	HIS	-	EXPRESSION TAG	UNP Q8ST28
D	-13	HIS	-	EXPRESSION TAG	UNP Q8ST28
D	-12	HIS	-	EXPRESSION TAG	UNP Q8ST28
D	-11	HIS	-	EXPRESSION TAG	UNP Q8ST28
D	-10	HIS	-	EXPRESSION TAG	UNP Q8ST28
D	-9	SER	-	EXPRESSION TAG	UNP Q8ST28
D	-8	SER	-	EXPRESSION TAG	UNP Q8ST28
D	-7	GLY	-	EXPRESSION TAG	UNP Q8ST28
D	-6	LEU	-	EXPRESSION TAG	UNP Q8ST28
D	-5	VAL	-	EXPRESSION TAG	UNP Q8ST28
D	-4	PRO	-	EXPRESSION TAG	UNP Q8ST28
D	-3	ARG	-	EXPRESSION TAG	UNP Q8ST28
D	-2	GLY	-	EXPRESSION TAG	UNP Q8ST28
D	-1	SER	-	EXPRESSION TAG	UNP Q8ST28
D	0	HIS	-	EXPRESSION TAG	UNP Q8ST28

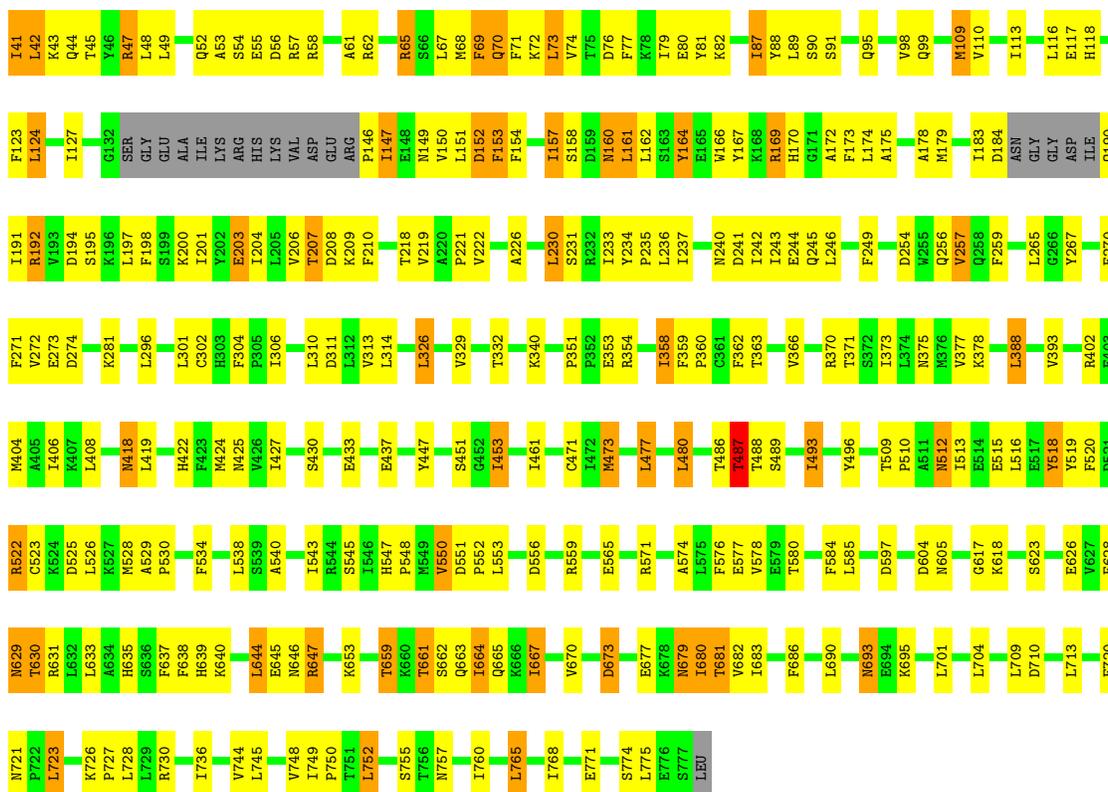
- Molecule 3 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C₆H₁₃NO₄S).



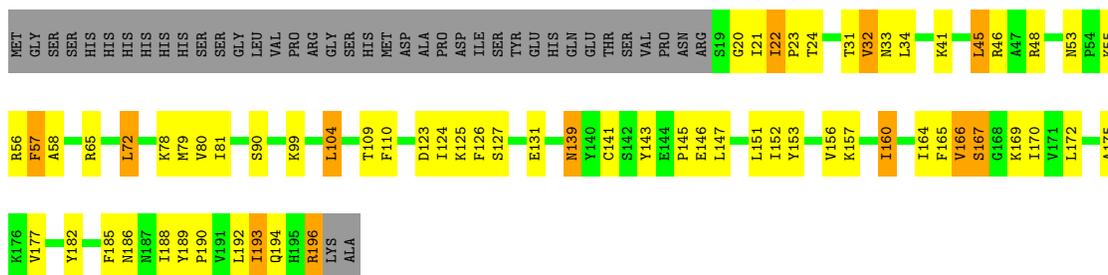
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	C	1	12	6	1	4	1	0	0
3	D	1	12	6	1	4	1	0	0

- Molecule 4 is water.

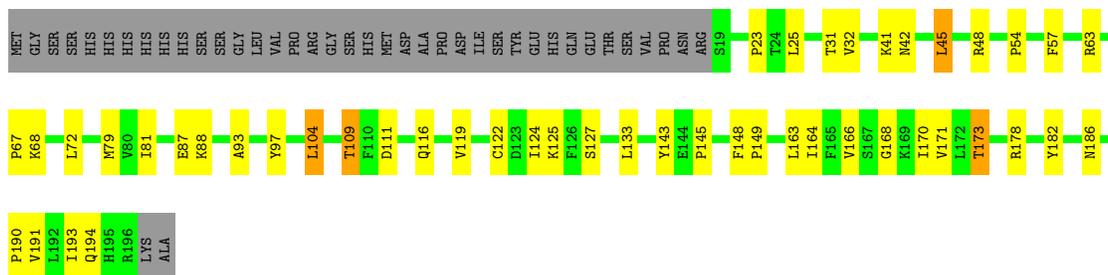
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	12	Total	O	0	0
			12	12		
4	B	11	Total	O	0	0
			11	11		
4	C	4	Total	O	0	0
			4	4		
4	D	2	Total	O	0	0
			2	2		



● Molecule 2: TRANSCRIPTION INITIATION FACTOR TFIID (TFIID-1)



● Molecule 2: TRANSCRIPTION INITIATION FACTOR TFIID (TFIID-1)



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	99.26Å 147.82Å 103.44Å 90.00° 94.52° 90.00°	Depositor
Resolution (Å)	47.28 – 3.10 47.28 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.28-3.10) 99.7 (47.28-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.66 (at 3.12Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, R_{free}	0.187 , 0.242 0.179 , 0.238	Depositor DCC
R_{free} test set	2697 reflections (5.00%)	DCC
Wilson B-factor (Å ²)	76.1	Xtrriage
Anisotropy	0.142	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 69.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.020 for l,-k,h	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	15037	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MES

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	0.27	0/6203	0.46	0/8365
1	B	0.29	0/6181	0.47	0/8335
2	C	0.28	0/1447	0.46	0/1952
2	D	0.30	0/1447	0.46	0/1952
All	All	0.28	0/15278	0.46	0/20604

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	40	TYR	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	6086	0	6157	263	0
1	B	6064	0	6130	274	0
2	C	1417	0	1491	46	0
2	D	1417	0	1491	27	0
3	C	12	0	12	1	0
3	D	12	0	12	2	0
4	A	12	0	0	1	0
4	B	11	0	0	0	0
4	C	4	0	0	0	0
4	D	2	0	0	0	0
All	All	15037	0	15293	598	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 20.

The worst 5 of 598 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:B:79:ILE:HD11	1:B:237:ILE:HD11	1.21	1.14
2:D:42:ASN:HD21	3:D:199:MES:H72	1.17	1.09
1:B:19:LEU:HA	1:B:31:ILE:HD11	1.35	1.09
1:A:33:GLU:HG2	1:A:62:ARG:HH22	1.22	1.02
1:A:149:ASN:HD22	1:A:149:ASN:H	1.09	1.00

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	746/800 (93%)	659 (88%)	76 (10%)	11 (2%)	12 45
1	B	743/800 (93%)	655 (88%)	78 (10%)	10 (1%)	14 48

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	177/218 (81%)	164 (93%)	12 (7%)	1 (1%)	28	67
2	D	177/218 (81%)	167 (94%)	9 (5%)	1 (1%)	28	67
All	All	1843/2036 (90%)	1645 (89%)	175 (10%)	23 (1%)	15	51

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	183	ILE
1	A	629	ASN
1	B	70	GLN
1	B	487	THR
1	A	24	THR

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	690/727 (95%)	616 (89%)	74 (11%)	8	30
1	B	687/727 (94%)	600 (87%)	87 (13%)	5	21
2	C	154/188 (82%)	134 (87%)	20 (13%)	5	21
2	D	154/188 (82%)	142 (92%)	12 (8%)	15	48
All	All	1685/1830 (92%)	1492 (88%)	193 (12%)	6	27

5 of 193 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	69	PHE
1	B	296	LEU
2	C	186	ASN
1	B	87	ILE
1	B	161	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	454	ASN
1	B	679	ASN
2	D	53	ASN
1	B	635	HIS
1	B	684	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

2 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
3	MES	C	199	-	12,12,12	2.15	1 (8%)	14,16,16	2.83	8 (57%)
3	MES	D	199	-	12,12,12	2.14	1 (8%)	14,16,16	2.84	7 (50%)

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
3	MES	C	199	-	-	0/6/14/14	0/1/1/1
3	MES	D	199	-	-	0/6/14/14	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	199	MES	C8-S	-7.14	1.66	1.77
3	D	199	MES	C8-S	-7.10	1.66	1.77

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	199	MES	C6-C5-N4	-4.21	104.21	110.11
3	D	199	MES	C6-C5-N4	-4.06	104.42	110.11
3	D	199	MES	C2-C3-N4	-3.99	104.52	110.11
3	C	199	MES	C2-C3-N4	-3.74	104.87	110.11
3	C	199	MES	O3S-S-C8	2.56	109.20	106.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	199	MES	1	0
3	D	199	MES	2	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	752/800 (94%)	-0.51	1 (0%) 95 90	38, 82, 150, 206	36 (4%)
1	B	749/800 (93%)	-0.58	0 100 100	26, 72, 137, 208	23 (3%)
2	C	178/218 (81%)	-0.57	0 100 100	44, 68, 108, 160	2 (1%)
2	D	178/218 (81%)	-0.59	0 100 100	37, 58, 91, 123	3 (1%)
All	All	1857/2036 (91%)	-0.55	1 (0%) 95 90	26, 73, 138, 208	64 (3%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	76	ASP	2.5

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	MES	D	199	12/12	0.95	0.37	3.89	96,115,130,133	0
3	MES	C	199	12/12	0.95	0.19	0.09	89,103,123,129	0

6.5 Other polymers [i](#)

There are no such residues in this entry.