wwPDB X-ray Structure Validation Summary Report (i)

Oct 8, 2023 – 10:22 AM EDT

PDB ID	:	6EC3
Title	:	Crystal Structure of EvdMO1
Authors	:	McCulloch, K.M.; Iverson, T.M.; Starbird, C.A.; Perry, N.A.; Chen, Q.; Berndt, S.; Yamakawa, I.; Loukachevitch, L.V.
Deposited on	:	2018-08-07
Resolution	:	3.35 Å(reported)

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

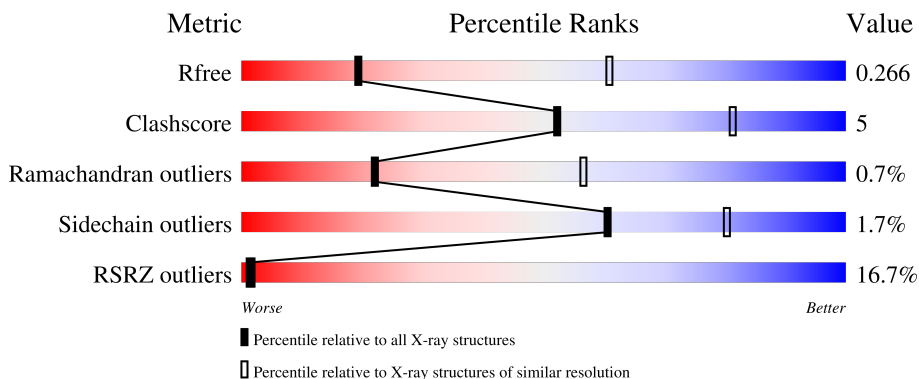|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.35 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	Percentile Ranks	Value
Rfree		0.266
Clashscore		5
Ramachandran outliers		0.7%
Sidechain outliers		1.7%
RSRZ outliers		16.7%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)
RSRZ outliers	127900	1507 (3.42-3.30)

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	471	13% / 75% 13% 12%
1	B	471	15% / 75% 12% 12%
1	C	471	15% / 76% 12% 12%

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:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	BTB	C	501	-	-	-	X

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 9743 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 Methyltransferase domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	414	Total 3227	C 2036	N 581	O 603	S 7	0	0	0
1	B	413	Total 3223	C 2031	N 583	O 602	S 7	0	0	0
1	C	416	Total 3248	C 2047	N 586	O 608	S 7	0	0	0

There are 75 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP A0A1C5AE13
A	-19	GLY	-	expression tag	UNP A0A1C5AE13
A	-18	SER	-	expression tag	UNP A0A1C5AE13
A	-17	SER	-	expression tag	UNP A0A1C5AE13
A	-16	HIS	-	expression tag	UNP A0A1C5AE13
A	-15	HIS	-	expression tag	UNP A0A1C5AE13
A	-14	HIS	-	expression tag	UNP A0A1C5AE13
A	-13	HIS	-	expression tag	UNP A0A1C5AE13
A	-12	HIS	-	expression tag	UNP A0A1C5AE13
A	-11	HIS	-	expression tag	UNP A0A1C5AE13
A	-10	SER	-	expression tag	UNP A0A1C5AE13
A	-9	SER	-	expression tag	UNP A0A1C5AE13
A	-8	GLY	-	expression tag	UNP A0A1C5AE13
A	-7	LEU	-	expression tag	UNP A0A1C5AE13
A	-6	VAL	-	expression tag	UNP A0A1C5AE13
A	-5	PRO	-	expression tag	UNP A0A1C5AE13
A	-4	ARG	-	expression tag	UNP A0A1C5AE13
A	-3	GLY	-	expression tag	UNP A0A1C5AE13
A	-2	SER	-	expression tag	UNP A0A1C5AE13
A	-1	HIS	-	expression tag	UNP A0A1C5AE13
A	0	MET	-	expression tag	UNP A0A1C5AE13
A	100	HIS	ASN	variant	UNP A0A1C5AE13
A	219	THR	LYS	variant	UNP A0A1C5AE13

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	232	ALA	THR	variant	UNP A0A1C5AE13
A	444	SER	ALA	variant	UNP A0A1C5AE13
B	-20	MET	-	expression tag	UNP A0A1C5AE13
B	-19	GLY	-	expression tag	UNP A0A1C5AE13
B	-18	SER	-	expression tag	UNP A0A1C5AE13
B	-17	SER	-	expression tag	UNP A0A1C5AE13
B	-16	HIS	-	expression tag	UNP A0A1C5AE13
B	-15	HIS	-	expression tag	UNP A0A1C5AE13
B	-14	HIS	-	expression tag	UNP A0A1C5AE13
B	-13	HIS	-	expression tag	UNP A0A1C5AE13
B	-12	HIS	-	expression tag	UNP A0A1C5AE13
B	-11	HIS	-	expression tag	UNP A0A1C5AE13
B	-10	SER	-	expression tag	UNP A0A1C5AE13
B	-9	SER	-	expression tag	UNP A0A1C5AE13
B	-8	GLY	-	expression tag	UNP A0A1C5AE13
B	-7	LEU	-	expression tag	UNP A0A1C5AE13
B	-6	VAL	-	expression tag	UNP A0A1C5AE13
B	-5	PRO	-	expression tag	UNP A0A1C5AE13
B	-4	ARG	-	expression tag	UNP A0A1C5AE13
B	-3	GLY	-	expression tag	UNP A0A1C5AE13
B	-2	SER	-	expression tag	UNP A0A1C5AE13
B	-1	HIS	-	expression tag	UNP A0A1C5AE13
B	0	MET	-	expression tag	UNP A0A1C5AE13
B	100	HIS	ASN	variant	UNP A0A1C5AE13
B	219	THR	LYS	variant	UNP A0A1C5AE13
B	232	ALA	THR	variant	UNP A0A1C5AE13
B	444	SER	ALA	variant	UNP A0A1C5AE13
C	-20	MET	-	expression tag	UNP A0A1C5AE13
C	-19	GLY	-	expression tag	UNP A0A1C5AE13
C	-18	SER	-	expression tag	UNP A0A1C5AE13
C	-17	SER	-	expression tag	UNP A0A1C5AE13
C	-16	HIS	-	expression tag	UNP A0A1C5AE13
C	-15	HIS	-	expression tag	UNP A0A1C5AE13
C	-14	HIS	-	expression tag	UNP A0A1C5AE13
C	-13	HIS	-	expression tag	UNP A0A1C5AE13
C	-12	HIS	-	expression tag	UNP A0A1C5AE13
C	-11	HIS	-	expression tag	UNP A0A1C5AE13
C	-10	SER	-	expression tag	UNP A0A1C5AE13
C	-9	SER	-	expression tag	UNP A0A1C5AE13
C	-8	GLY	-	expression tag	UNP A0A1C5AE13
C	-7	LEU	-	expression tag	UNP A0A1C5AE13
C	-6	VAL	-	expression tag	UNP A0A1C5AE13

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-5	PRO	-	expression tag	UNP A0A1C5AE13
C	-4	ARG	-	expression tag	UNP A0A1C5AE13
C	-3	GLY	-	expression tag	UNP A0A1C5AE13
C	-2	SER	-	expression tag	UNP A0A1C5AE13
C	-1	HIS	-	expression tag	UNP A0A1C5AE13
C	0	MET	-	expression tag	UNP A0A1C5AE13
C	100	HIS	ASN	variant	UNP A0A1C5AE13
C	219	THR	LYS	variant	UNP A0A1C5AE13
C	232	ALA	THR	variant	UNP A0A1C5AE13
C	444	SER	ALA	variant	UNP A0A1C5AE13

- Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

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

- Molecule 3 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: BTB) (formula: $C_8H_{19}NO_5$).
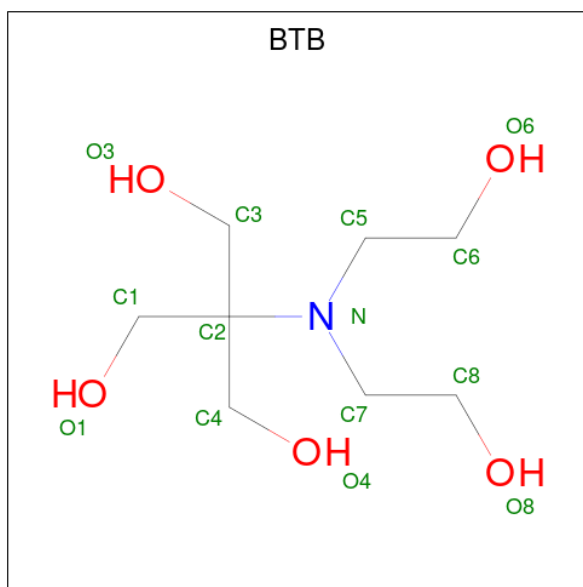
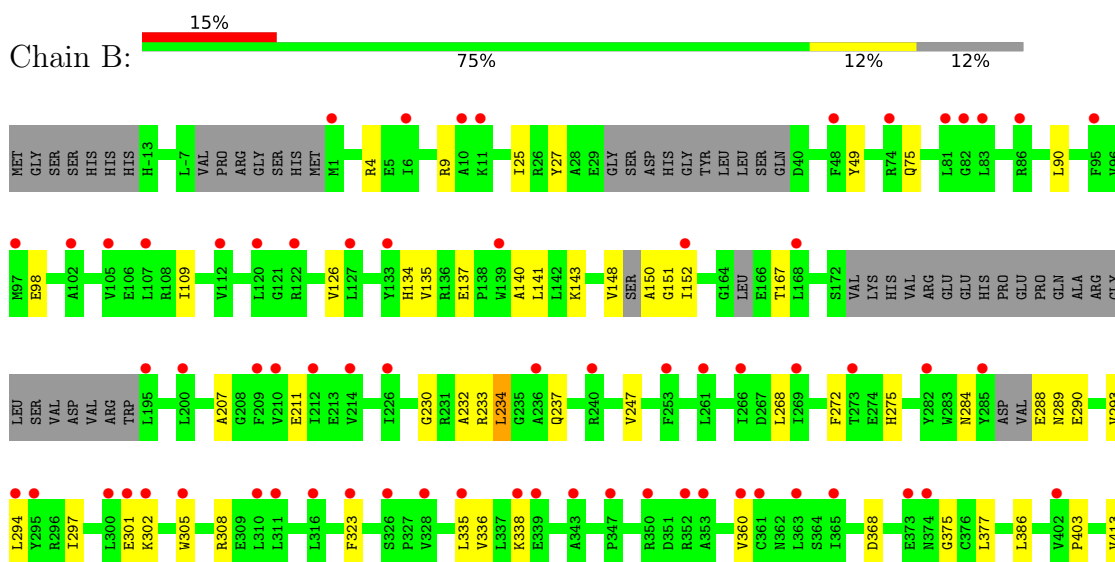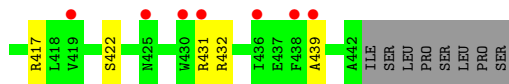| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 14	C 8	N 1	O 5	0	0
3	B	1	Total 14	C 8	N 1	O 5	0	0
3	C	1	Total 14	C 8	N 1	O 5	0	0

3 Residue-property plots ⓘ

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.
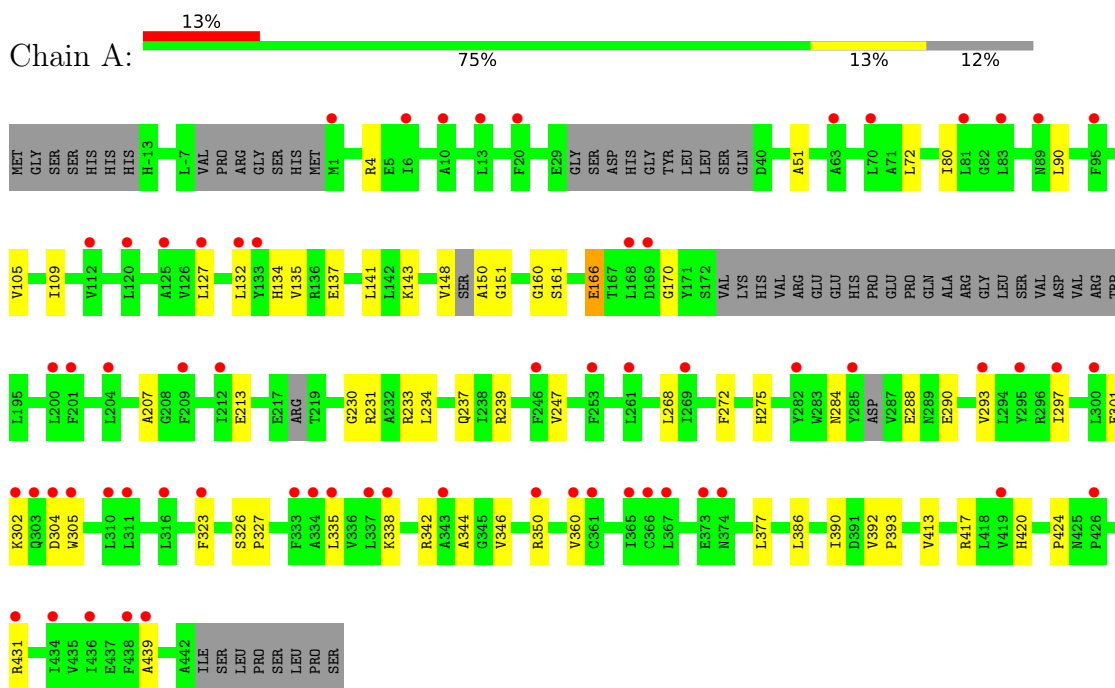
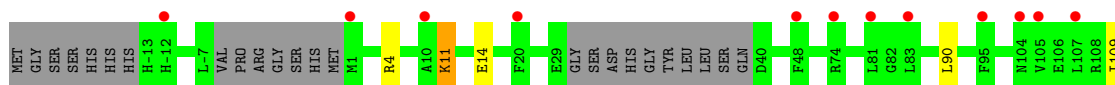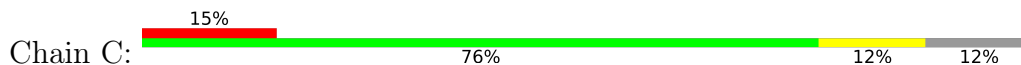• Molecule 1: Methyltransferase domain-containing protein

• Molecule 1: Methyltransferase domain-containing protein

R417 L418 V419 S422 N425 W430 R431 R432 I436 E437 F438 A439 A442 ILE SER LEU PRO SER LEU PRO SER

• Molecule 1: Methyltransferase domain-containing protein

Chain C:

15%

76%   12%   12%

MET GLY SER SER HIS HIS HIS H-13 H-12 L-7 VAL PRO ARG GLY SER HIS MET M1 R4 A10 K11 E14 F20 E29 GLY SER ASP HIS GLY TYR LEU LEU SER GLN D40 F48 R74 L81 G82 L83 L90 F95 N104 V105 E106 L107 R108 I109

V112 E113 T114 L120 GLY R122 F123 L127 L132 Y133 H134 V135 R136 E137 P138 W139 A140 L141 L142 K143 D144 A145 A146 R147 V148 S149 A150 G151 I152 Y153 E166 T167 L168 D169 S172 VAL LYS HIS VAL ARG GLU GLU HIS PRO GLU PRO GLN ALA ARG GLY LEU SER VAL ASP

VAL ARG TRP L195 L200 F201 L204 A207 G208 F209 V210 E211 V214 I215 C224 D225 I226 G230 R233 L234 Q237 R240 F246 V247 F253 L261 F272 H275 Y282 W283 N284 Y285 D286 V287 E288 N289 E290 A291 P292 V293 L294 I297 L300

E301 K302 Q303 D304 W305 R308 E309 L310 L311 L316 F323 L335 V336 L337 K338 R342 A343 R352 A353 N354 V360 C361 N362 L363 S364 I365 C366 L367 P372 E373 N374 G375 C376 L377 L386 V411 I412 V413 V416 R417 N425 D428 Q429 W430 R431 R432 T433

E437 F438 A439 A442 ILE SER LEU PRO SER LEU PRO SER

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, α, β, γ	191.63Å 191.63Å 269.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.64 – 3.35 45.80 – 3.35	Depositor EDS
% Data completeness (in resolution range)	95.9 (45.64-3.35) 94.8 (45.80-3.35)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.88 (at 3.32Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.221 , 0.266 0.221 , 0.266	Depositor DCC
R_{free} test set	1815 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å^2)	121.8	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 104.6	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.377 for -1/2*h+1/2*k-1/2*l,1/2*h-1/2*k-1/2*l,-h-k 0.390 for -1/2*h-1/2*k+1/2*l,-1/2*h-1/2*k-1/2*l,h-k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9743	wwPDB-VP
Average B, all atoms (Å^2)	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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: BTB, NI

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	0.24	0/3298	0.44	0/4480
1	B	0.25	0/3294	0.44	0/4473
1	C	0.24	0/3321	0.43	0/4514
All	All	0.24	0/9913	0.44	0/13467

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	3227	0	3137	36	0
1	B	3223	0	3130	38	0
1	C	3248	0	3158	32	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	14	0	19	5	0
3	B	14	0	19	3	0
3	C	14	0	19	2	0
All	All	9743	0	9482	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 5.

The worst 5 of 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:C:338:LYS:HB2	1:C:431:ARG:HB3	1.66	0.77
1:A:338:LYS:HB2	1:A:431:ARG:HB3	1.68	0.75
1:B:375:GLY:O	1:B:431:ARG:NH2	2.24	0.70
1:A:51:ALA:O	1:A:231:ARG:NH2	2.23	0.69
1:A:377:LEU:HD13	3:A:501:BTB:H82	1.77	0.66

There are no symmetry-related clashes.

5.3　Torsion angles　ⓘ

5.3.1　Protein backbone　ⓘ

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	400/471 (85%)	365 (91%)	32 (8%)	3 (1%)	19	53
1	B	399/471 (85%)	364 (91%)	32 (8%)	3 (1%)	19	53
1	C	406/471 (86%)	367 (90%)	36 (9%)	3 (1%)	22	56
All	All	1205/1413 (85%)	1096 (91%)	100 (8%)	9 (1%)	22	56

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	234	LEU
1	A	160	GLY
1	B	233	ARG
1	C	166	GLU
1	A	161	SER

5.3.2 Protein sidechains ⓘ

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	332/383 (87%)	327 (98%)	5 (2%)	65	82
1	B	331/383 (86%)	327 (99%)	4 (1%)	71	85
1	C	335/383 (88%)	327 (98%)	8 (2%)	49	74
All	All	998/1149 (87%)	981 (98%)	17 (2%)	60	80

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	272	PHE
1	C	431	ARG
1	B	272	PHE
1	B	275	HIS
1	C	4	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 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	$\#|Z| > 2$
3	BTB	A	501	-	13,13,13	0.70	0	7,16,16	0.65	0
3	BTB	B	501	-	13,13,13	0.60	0	7,16,16	0.62	0
3	BTB	C	501	-	13,13,13	0.73	0	7,16,16	0.51	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
3	BTB	A	501	-	-	9/21/21/21	-
3	BTB	B	501	-	-	10/21/21/21	-
3	BTB	C	501	-	-	7/21/21/21	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	501	BTB	O1-C1-C2-N
3	A	501	BTB	C1-C2-C4-O4
3	A	501	BTB	C3-C2-C4-O4
3	A	501	BTB	N-C2-C4-O4
3	A	501	BTB	C6-C5-N-C7

There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	BTB	5	0
3	B	501	BTB	3	0
3	C	501	BTB	2	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(Å²)	Q<0.9
1	A	414/471 (87%)	1.02	63 (15%)	2	2	88, 135, 206, 247	0
1	B	413/471 (87%)	1.05	73 (17%)	1	1	96, 134, 200, 240	0
1	C	416/471 (88%)	1.03	72 (17%)	1	1	86, 132, 199, 262	0
All	All	1243/1413 (87%)	1.03	208 (16%)	1	2	86, 134, 201, 262	0

The worst 5 of 208 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	338	LYS	6.3
1	A	361	CYS	6.0
1	C	431	ARG	5.9
1	C	168	LEU	5.7
1	C	305	TRP	5.3

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	BTB	C	501	14/14	0.76	0.59	151,159,172,173	0
3	BTB	B	501	14/14	0.82	0.38	157,167,179,182	0
3	BTB	A	501	14/14	0.85	0.38	178,189,199,204	0
2	NI	B	500	1/1	0.97	0.32	105,105,105,105	0
2	NI	A	500	1/1	0.97	0.36	110,110,110,110	0
2	NI	C	500	1/1	0.98	0.34	108,108,108,108	0

6.5 Other polymers (i)

There are no such residues in this entry.