



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 12:11 am BST

PDB ID : 6Q8F
Title : Nterminal domain of human SMU1
Authors : Tengo, L.; Le Corre, L.; Fournier, G.; Ashraf, U.; Busca, P.; Rameix-Welti, M.-A.; Gravier-Pelletier, C.; Ruigrok, R.W.H.; Jacob, Y.; Vidalain, P.-O.; Pietrancosta, N.; Naffakh, N.; McCarthy, A.A.; Crepin, T.
Deposited on : 2018-12-14
Resolution : 1.90 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

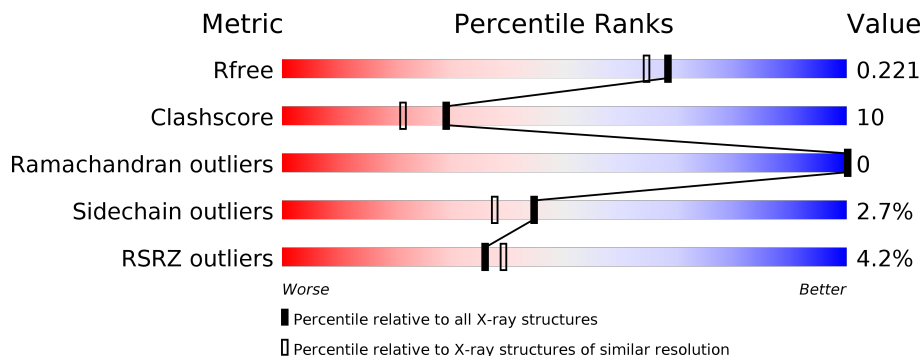
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	514	
1	B	514	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3260 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.

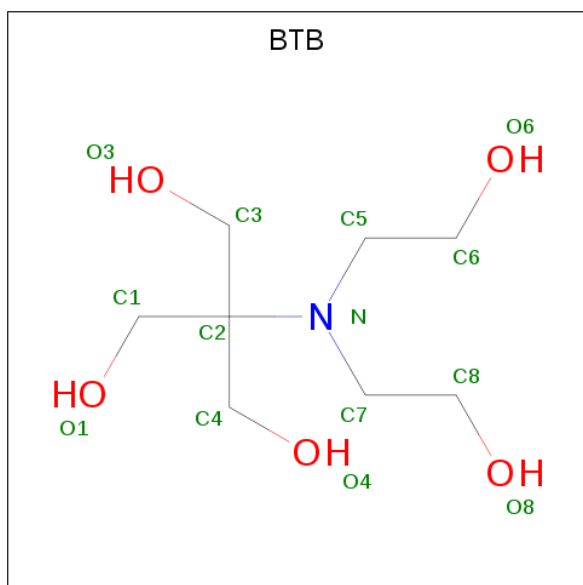
- Molecule 1 is a protein called WD40 repeat-containing protein SMU1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	197	1564	994	270	295	5	0	1	0
1	B	184	1464	928	254	278	4	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q2TAY7
A	1	GLY	-	expression tag	UNP Q2TAY7
B	0	MET	-	initiating methionine	UNP Q2TAY7
B	1	GLY	-	expression tag	UNP Q2TAY7

- Molecule 2 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPAN E-1,3-DIOL (three-letter code: BTB) (formula: C₈H₁₉NO₅).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	114	Total	O	0	0
			114	114		
4	B	103	Total	O	0	0
			103	103		

CYS SER ASN PHE LYS SER LEU GLY THR ALA GLY THR TRP ASP ILE TYR CYS VAL ASN GLY SER VAL ILE LEU LEU PRO LYS ASN PHE HIS THR PHE VAL VAL CYS ASN ARG ARG SER ASN THR VAL VAL HIS VAL PHE VAL VAL CYS ASN ARG ARG THR SER ASN THR THR VAL VAL HIS VAL ILE MET ASN MET ILE GLN GLY GLN ILE ILE VAL ARG HIS SER PHE SER GLN ASN SER LEU LYS ARG GLY TYR

GLY ASP PHE VAL CYS CYS ALA LEU SER PRO ARG GLY GLU TRP ILE TYR CYS VAL GLY GLU ASP PHE VAL LEU TYR CYS PHE SER THR VAL VAL GLY LYS LEU GLU ARG THR LEU THR VAL HIS GLU LYS ASP VAL ILE GLY ILE ALA HIS HIS PRO HIS SER GLN ASN SER LEU ILE ALA THR TYR

SER
GLU
ASP
GLY
LEU
LEU
LYS
LEU
TRP
LYS
PRO

4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	40.18Å 41.23Å 123.60Å 90.00° 91.48° 90.00°	Depositor
Resolution (Å)	39.14 – 1.90 39.11 – 1.90	Depositor EDS
% Data completeness (in resolution range)	97.2 (39.14-1.90) 97.2 (39.11-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.25 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.8.0232	Depositor
R, R_{free}	0.174 , 0.214 0.185 , 0.221	Depositor DCC
R_{free} test set	1610 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å ²)	21.2	Xtriage
Anisotropy	0.060	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 46.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.023 for -k,-h,-l 0.024 for k,h,-l 0.037 for h,-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3260	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.96% 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: CL, BTB

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	0.67	0/1592	0.78	0/2156
1	B	0.66	0/1486	0.79	0/2012
All	All	0.67	0/3078	0.78	0/4168

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
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	179	PRO	Peptide
1	B	65	SER	Peptide

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	1564	0	1609	31	0
1	B	1464	0	1495	35	0
2	A	14	0	19	3	0
3	B	1	0	0	1	0
4	A	114	0	0	2	0
4	B	103	0	0	0	0
All	All	3260	0	3123	59	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 10.

The worst 5 of 59 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:A:601:BTB:O4	2:A:601:BTB:H71	1.68	0.92
1:A:107:LYS:HD2	4:A:790:HOH:O	1.73	0.89
1:B:40:THR:HB	1:B:43:SER:HB2	1.53	0.87
1:B:65:SER:HB2	1:B:66:LEU:HD23	1.55	0.87
1:A:40:THR:HG23	1:B:41:VAL:CG2	2.05	0.86

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	194/514 (38%)	189 (97%)	5 (3%)	0	100	100
1	B	178/514 (35%)	177 (99%)	1 (1%)	0	100	100
All	All	372/1028 (36%)	366 (98%)	6 (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	173/450 (38%)	170 (98%)	3 (2%)	60	57
1	B	162/450 (36%)	156 (96%)	6 (4%)	34	25
All	All	335/900 (37%)	326 (97%)	9 (3%)	44	38

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	40	THR
1	B	185	ASP
1	B	65	SER
1	A	200	LYS
1	B	43	SER

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

Mol	Chain	Res	Type
1	A	39	ASN
1	A	52	ASN
1	A	119	ASN
1	B	99	GLN

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](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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
2	BTB	A	601	-	13,13,13	1.36	2 (15%)	7,16,16	0.56	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
2	BTB	A	601	-	-	6/21/21/21	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	BTB	C2-N	3.64	1.55	1.48
2	A	601	BTB	C5-N	2.65	1.51	1.48

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	601	BTB	C1-C2-C4-O4
2	A	601	BTB	C1-C2-N-C7
2	A	601	BTB	C3-C2-N-C7
2	A	601	BTB	C4-C2-N-C7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	A	601	BTB	C6-C5-N-C7

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	BTB	3	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	35:THR	C	36:VAL	N	2.40

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	197/514 (38%)	-0.13	8 (4%) 37 40	12, 22, 59, 80	0
1	B	184/514 (35%)	-0.09	8 (4%) 35 38	13, 24, 54, 109	0
All	All	381/1028 (37%)	-0.11	16 (4%) 36 39	12, 23, 58, 109	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	40	THR	6.6
1	A	180	PRO	5.5
1	B	39	ASN	4.4
1	B	41	VAL	4.3
1	B	35	THR	3.3

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. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	BTB	A	601	14/14	0.92	0.13	15,20,30,31	0
3	CL	B	601	1/1	0.98	0.06	46,46,46,46	0

6.5 Other polymers [i](#)

There are no such residues in this entry.