Full wwPDB X-ray Structure Validation Report (i)

May 15, 2020 – 02:37 pm BST

PDB ID	:	5X9B
Title	:	Crystal structure of the cytosolic domain of human MiD51
Authors	:	Sun, F.; Pang, X.; Ma, J.
Deposited on	:	2017-03-06
Resolution	:	2.70 Å(reported)

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

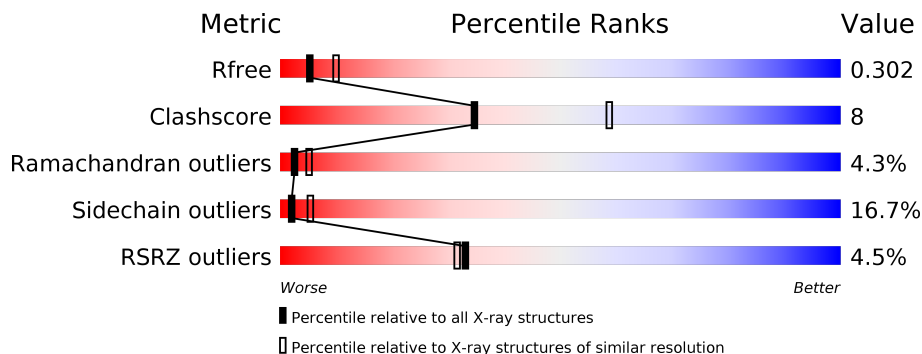|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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		0.302
Clashscore		8
Ramachandran outliers		4.3%
Sidechain outliers		16.7%
RSRZ outliers		4.5%

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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 <=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	336	4% / 71% / 21% / 5% ••

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2604 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 Mitochondrial dynamics protein MID51.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	331	2594	1665	435	482	12	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	128	GLY	-	expression tag	UNP Q9NQG6
A	129	PRO	-	expression tag	UNP Q9NQG6
A	130	LEU	-	expression tag	UNP Q9NQG6
A	131	GLY	-	expression tag	UNP Q9NQG6
A	132	SER	-	expression tag	UNP Q9NQG6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	10	10	10	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: Mitochondrial dynamics protein MID51

Chain A:
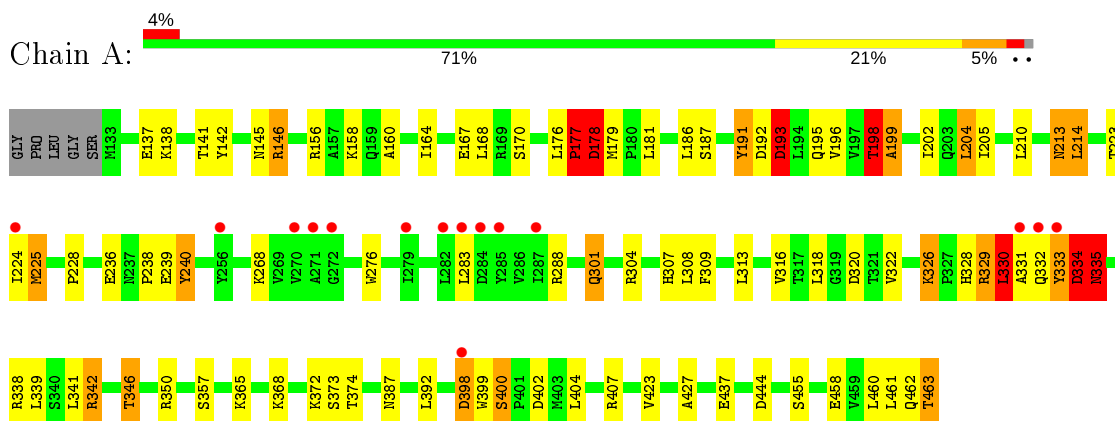
4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	88.84Å 88.84Å 124.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.42 – 2.70 44.42 – 2.69	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.42-2.70) 99.7 (44.42-2.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.50 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.222 , 0.298 0.245 , 0.302	Depositor DCC
R_{free} test set	733 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å²)	71.4	Xtriage
Anisotropy	0.304	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.31 , 56.5	EDS
L-test for twinning[2]	$<|L|> = 0.50$, $<L^2> = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2604	wwPDB-VP
Average B, all atoms (Å²)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.45% 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 (i)

5.1 Standard geometry (i)

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	0.75	0/2653	1.03	6/3618 (0.2%)

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	6

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	330	LEU	CA-CB-CG	7.10	131.62	115.30
1	A	398	ASP	N-CA-C	6.70	129.08	111.00
1	A	326	LYS	C-N-CD	6.23	141.48	128.40
1	A	198	THR	N-CA-C	5.90	126.94	111.00
1	A	146	ARG	NE-CZ-NH2	-5.85	117.37	120.30
1	A	334	ASP	N-CA-C	5.22	125.10	111.00

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	187	SER	Peptide
1	A	196	VAL	Peptide
1	A	198	THR	Peptide
1	A	228	PRO	Peptide
1	A	238	PRO	Peptide
1	A	268	LYS	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	2594	0	2613	43	0
2	A	10	0	0	0	0
All	All	2604	0	2613	43	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 8.

All (43) 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:A:202:ILE:HD12	1:A:308:LEU:HD11	1.59	0.84
1:A:398:ASP:O	1:A:400:SER:N	2.15	0.78
1:A:198:THR:N	1:A:199:ALA:HB2	1.99	0.77
1:A:398:ASP:HB3	1:A:407:ARG:HH21	1.53	0.72
1:A:334:ASP:HA	1:A:335:ASN:HB3	1.74	0.70
1:A:326:LYS:HB3	1:A:338:ARG:HE	1.60	0.67
1:A:331:ALA:HB1	1:A:335:ASN:HA	1.77	0.66
1:A:210:LEU:HD12	1:A:318:LEU:HD21	1.82	0.62
1:A:205:ILE:HD13	1:A:313:LEU:HD21	1.81	0.61
1:A:198:THR:OG1	1:A:346:THR:HG22	2.02	0.60
1:A:328:HIS:HB3	1:A:331:ALA:HB2	1.83	0.60
1:A:333:TYR:CD2	1:A:334:ASP:HB2	2.38	0.58
1:A:334:ASP:HA	1:A:335:ASN:CB	2.34	0.56
1:A:141:THR:HG22	1:A:145:ASN:ND2	2.20	0.56
1:A:176:LEU:HD21	1:A:179:MET:HB3	1.90	0.54
1:A:186:LEU:CD1	1:A:204:LEU:HD12	2.37	0.54
1:A:177:PRO:O	1:A:178:ASP:HB2	2.09	0.52
1:A:301:GLN:HA	1:A:307:HIS:HA	1.92	0.51
1:A:333:TYR:N	1:A:334:ASP:O	2.44	0.51
1:A:387:ASN:HD22	1:A:427:ALA:H	1.57	0.51
1:A:328:HIS:O	1:A:437:GLU:OE2	2.29	0.50
1:A:342:ARG:O	1:A:346:THR:HG23	2.12	0.50
1:A:308:LEU:HD12	1:A:309:PHE:N	2.26	0.49
1:A:176:LEU:HA	1:A:177:PRO:O	2.12	0.49
1:A:186:LEU:HD12	1:A:204:LEU:HD12	1.94	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:308:LEU:HD12	1:A:309:PHE:H	1.78	0.48
1:A:160:ALA:O	1:A:164:ILE:HG12	2.14	0.47
1:A:374:THR:HA	1:A:463:THR:HG21	1.96	0.47
1:A:326:LYS:HB3	1:A:338:ARG:NE	2.29	0.47
1:A:158:LYS:HE3	1:A:192:ASP:HA	1.97	0.46
1:A:330:LEU:O	1:A:333:TYR:O	2.34	0.45
1:A:142:TYR:CD1	1:A:146:ARG:HG3	2.52	0.45
1:A:223:THR:HB	1:A:224:ILE:C	2.36	0.45
1:A:167:GLU:HG3	1:A:283:LEU:HD11	1.98	0.45
1:A:335:ASN:HD22	1:A:335:ASN:C	2.20	0.44
1:A:191:TYR:CD1	1:A:191:TYR:N	2.85	0.43
1:A:213:ASN:HB3	1:A:214:LEU:HD13	2.01	0.42
1:A:330:LEU:HD22	1:A:331:ALA:H	1.83	0.42
1:A:331:ALA:HB1	1:A:335:ASN:CA	2.48	0.42
1:A:328:HIS:O	1:A:329:ARG:CB	2.67	0.41
1:A:192:ASP:O	1:A:193:ASP:HB3	2.20	0.41
1:A:176:LEU:CB	1:A:177:PRO:HA	2.52	0.40
1:A:176:LEU:HB3	1:A:177:PRO:HA	2.04	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	329/336 (98%)	290 (88%)	25 (8%)	14 (4%)	2	5

All (14) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	177	PRO
1	A	199	ALA
1	A	225	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	239	GLU
1	A	329	ARG
1	A	334	ASP
1	A	399	TRP
1	A	193	ASP
1	A	240	TYR
1	A	458	GLU
1	A	320	ASP
1	A	357	SER
1	A	178	ASP
1	A	335	ASN

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	282/288 (98%)	235 (83%)	47 (17%)	2	5

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	137	GLU
1	A	138	LYS
1	A	156	ARG
1	A	168	LEU
1	A	170	SER
1	A	177	PRO
1	A	178	ASP
1	A	181	LEU
1	A	191	TYR
1	A	193	ASP
1	A	195	GLN
1	A	204	LEU
1	A	213	ASN
1	A	214	LEU
1	A	225	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	236	GLU
1	A	240	TYR
1	A	276	TRP
1	A	288	ARG
1	A	301	GLN
1	A	304	ARG
1	A	316	VAL
1	A	322	VAL
1	A	330	LEU
1	A	332	GLN
1	A	333	TYR
1	A	335	ASN
1	A	339	LEU
1	A	341	LEU
1	A	342	ARG
1	A	346	THR
1	A	350	ARG
1	A	365	LYS
1	A	368	LYS
1	A	372	LYS
1	A	373	SER
1	A	392	LEU
1	A	400	SER
1	A	402	ASP
1	A	404	LEU
1	A	423	VAL
1	A	444	ASP
1	A	455	SER
1	A	460	LEU
1	A	461	LEU
1	A	462	GLN
1	A	463	THR

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	145	ASN
1	A	201	HIS
1	A	213	ASN
1	A	387	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)

There are no ligands in this entry.

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	331/336 (98%)	0.23	15 (4%)	33　31	45, 74, 131, 174	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	271	ALA	5.2
1	A	284	ASP	4.7
1	A	272	GLY	4.1
1	A	282	LEU	3.7
1	A	283	LEU	3.5
1	A	279	ILE	3.3
1	A	256	TYR	3.1
1	A	398	ASP	2.9
1	A	333	TYR	2.9
1	A	270	VAL	2.6
1	A	331	ALA	2.6
1	A	285	TYR	2.5
1	A	224	ILE	2.4
1	A	332	GLN	2.2
1	A	287	ILE	2.2

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.