



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 12:49 pm GMT

PDB ID : 4U2M
Title : Crystal structure of a complex of the Miz1- and BCL6 POZ domains.
Authors : Stead, M.A.; Wright, S.C.
Deposited on : 2014-07-17
Resolution : 2.23 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
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

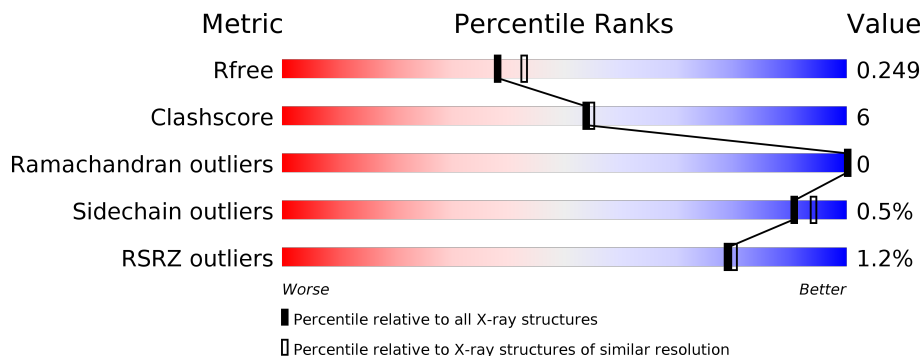
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.23 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	1804 (2.26-2.22)
Clashscore	112137	1957 (2.26-2.22)
Ramachandran outliers	110173	1916 (2.26-2.22)
Sidechain outliers	110143	1917 (2.26-2.22)
RSRZ outliers	101464	1809 (2.26-2.22)

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	256	<div> <div>2%</div> <div>79% 9% 11%</div> </div>
1	B	256	<div> <div>2%</div> <div>88% 6% 6%</div> </div>
1	C	256	<div> <div>2%</div> <div>83% 8% 9%</div> </div>
1	D	256	<div> <div>2%</div> <div>85% 8% 6%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7350 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 Zinc finger and BTB domain-containing protein 17,B-cell lymphoma 6 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	228	Total	C	N	O	S	0	1	0
			1746	1113	298	316	19			
1	B	240	Total	C	N	O	S	0	1	0
			1811	1157	306	331	17			
1	C	233	Total	C	N	O	S	0	2	0
			1772	1130	295	329	18			
1	D	240	Total	C	N	O	S	0	1	0
			1858	1184	314	342	18			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q13105
A	-2	PRO	-	expression tag	UNP Q13105
A	-1	LEU	-	expression tag	UNP Q13105
A	0	GLY	-	expression tag	UNP Q13105
A	1	SER	-	expression tag	UNP Q13105
A	116	ARG	-	linker	UNP Q13105
A	117	SER	-	linker	UNP Q13105
A	118	GLY	-	linker	UNP Q13105
A	119	GLY	-	linker	UNP Q13105
A	120	GLY	-	linker	UNP Q13105
A	121	SER	-	linker	UNP Q13105
A	122	SER	-	linker	UNP Q13105
A	123	GLY	-	linker	UNP Q13105
A	124	GLY	-	linker	UNP Q13105
A	125	SER	-	linker	UNP Q13105
A	126	GLY	-	linker	UNP Q13105
A	127	THR	-	linker	UNP Q13105
B	-3	GLY	-	expression tag	UNP Q13105
B	-2	PRO	-	expression tag	UNP Q13105
B	-1	LEU	-	expression tag	UNP Q13105

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	GLY	-	expression tag	UNP Q13105
B	1	SER	-	expression tag	UNP Q13105
B	116	ARG	-	linker	UNP Q13105
B	117	SER	-	linker	UNP Q13105
B	118	GLY	-	linker	UNP Q13105
B	119	GLY	-	linker	UNP Q13105
B	120	GLY	-	linker	UNP Q13105
B	121	SER	-	linker	UNP Q13105
B	122	SER	-	linker	UNP Q13105
B	123	GLY	-	linker	UNP Q13105
B	124	GLY	-	linker	UNP Q13105
B	125	SER	-	linker	UNP Q13105
B	126	GLY	-	linker	UNP Q13105
B	127	THR	-	linker	UNP Q13105
C	-3	GLY	-	expression tag	UNP Q13105
C	-2	PRO	-	expression tag	UNP Q13105
C	-1	LEU	-	expression tag	UNP Q13105
C	0	GLY	-	expression tag	UNP Q13105
C	1	SER	-	expression tag	UNP Q13105
C	116	ARG	-	linker	UNP Q13105
C	117	SER	-	linker	UNP Q13105
C	118	GLY	-	linker	UNP Q13105
C	119	GLY	-	linker	UNP Q13105
C	120	GLY	-	linker	UNP Q13105
C	121	SER	-	linker	UNP Q13105
C	122	SER	-	linker	UNP Q13105
C	123	GLY	-	linker	UNP Q13105
C	124	GLY	-	linker	UNP Q13105
C	125	SER	-	linker	UNP Q13105
C	126	GLY	-	linker	UNP Q13105
C	127	THR	-	linker	UNP Q13105
D	-3	GLY	-	expression tag	UNP Q13105
D	-2	PRO	-	expression tag	UNP Q13105
D	-1	LEU	-	expression tag	UNP Q13105
D	0	GLY	-	expression tag	UNP Q13105
D	1	SER	-	expression tag	UNP Q13105
D	116	ARG	-	linker	UNP Q13105
D	117	SER	-	linker	UNP Q13105
D	118	GLY	-	linker	UNP Q13105
D	119	GLY	-	linker	UNP Q13105
D	120	GLY	-	linker	UNP Q13105
D	121	SER	-	linker	UNP Q13105

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	122	SER	-	linker	UNP Q13105
D	123	GLY	-	linker	UNP Q13105
D	124	GLY	-	linker	UNP Q13105
D	125	SER	-	linker	UNP Q13105
D	126	GLY	-	linker	UNP Q13105
D	127	THR	-	linker	UNP Q13105

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	40	Total O 40 40	0	0
2	B	56	Total O 56 56	0	0
2	C	16	Total O 16 16	0	0
2	D	51	Total O 51 51	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	91.85Å 97.88Å 119.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.47 – 2.23 58.47 – 2.23	Depositor EDS
% Data completeness (in resolution range)	99.1 (58.47-2.23) 99.1 (58.47-2.23)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.55 (at 2.22Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.211 , 0.240 0.217 , 0.249	Depositor DCC
R_{free} test set	2580 reflections (5.13%)	DCC
Wilson B-factor (Å ²)	37.1	Xtriage
Anisotropy	0.400	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 36.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7350	wwPDB-VP
Average B, all atoms (Å ²)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% 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](#)

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	0/1775	0.80	3/2401 (0.1%)
1	B	0.72	0/1843	0.76	1/2503 (0.0%)
1	C	0.74	0/1803	0.76	0/2447
1	D	0.72	0/1891	0.78	2/2562 (0.1%)
All	All	0.73	0/7312	0.78	6/9913 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	198	ASP	CB-CG-OD1	6.86	124.47	118.30
1	A	167	ARG	CG-CD-NE	-6.29	98.59	111.80
1	D	198	ASP	CB-CG-OD1	5.73	123.45	118.30
1	A	167	ARG	NE-CZ-NH2	-5.66	117.47	120.30
1	A	245	ARG	NE-CZ-NH1	-5.41	117.59	120.30

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	1746	0	1722	29	0
1	B	1811	0	1764	11	0
1	C	1772	0	1726	29	0
1	D	1858	0	1840	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	40	0	0	0	0
2	B	56	0	0	2	0
2	C	16	0	0	0	0
2	D	51	0	0	0	0
All	All	7350	0	7052	87	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 6.

The worst 5 of 87 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:26[B]:CYS:SG	1:C:37:ALA:C	2.46	0.94
1:C:26[B]:CYS:SG	1:C:37:ALA:O	2.28	0.92
1:D:235:LEU:HB3	1:D:237:MET:HE3	1.52	0.92
1:D:235:LEU:CB	1:D:237:MET:HE3	2.14	0.77
1:C:26[B]:CYS:HG	1:C:37:ALA:C	1.87	0.76

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	221/256 (86%)	219 (99%)	2 (1%)	0	100	100
1	B	235/256 (92%)	227 (97%)	8 (3%)	0	100	100
1	C	229/256 (90%)	226 (99%)	3 (1%)	0	100	100
1	D	237/256 (93%)	233 (98%)	4 (2%)	0	100	100
All	All	922/1024 (90%)	905 (98%)	17 (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	187/221 (85%)	186 (100%)	1 (0%)	91	94
1	B	192/221 (87%)	190 (99%)	2 (1%)	80	86
1	C	191/221 (86%)	191 (100%)	0	100	100
1	D	203/221 (92%)	202 (100%)	1 (0%)	91	94
All	All	773/884 (87%)	769 (100%)	4 (0%)	91	94

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	151	ARG
1	B	65	SER
1	B	66	ASN
1	D	65	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	239	HIS

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	228/256 (89%)	-0.05	2 (0%) 84 84	30, 43, 70, 86	2 (0%)
1	B	240/256 (93%)	-0.08	2 (0%) 86 86	31, 46, 77, 96	0
1	C	233/256 (91%)	-0.12	3 (1%) 77 78	28, 43, 67, 102	1 (0%)
1	D	240/256 (93%)	-0.04	4 (1%) 70 71	29, 45, 78, 88	0
All	All	941/1024 (91%)	-0.07	11 (1%) 79 80	28, 44, 75, 102	3 (0%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	60	VAL	3.9
1	D	115	ALA	3.3
1	A	222	GLU	2.8
1	D	1	SER	2.7
1	D	59	VAL	2.7

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.