wwPDB X-ray Structure Validation Summary Report (i)

Jun 19, 2024 – 06:02 AM EDT

PDB ID	:	3WOZ
Title	:	Crystal structure of CLASP2 TOG domain (TOG3)
Authors	:	Hayashi, I.; Maki, T.
Deposited on	:	2014-01-06
Resolution	:	2.20 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

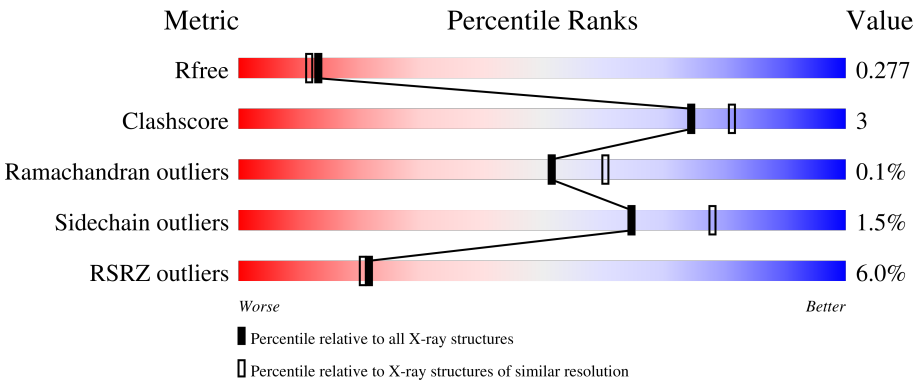
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	232	3%, 88%, 8%
1	B	232	3%, 90%, 9%
1	C	232	6%, 89%, 9%
1	D	232	11%, 86%, 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7722 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 CLIP-associating protein 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	225	1820	1161	315	336	2	6	0	2	0
1	B	229	1842	1174	319	341	2	6	0	0	0
1	C	229	1842	1174	319	341	2	6	0	0	0
1	D	222	1795	1143	309	333	2	8	0	2	0

- Molecule 2 is water.

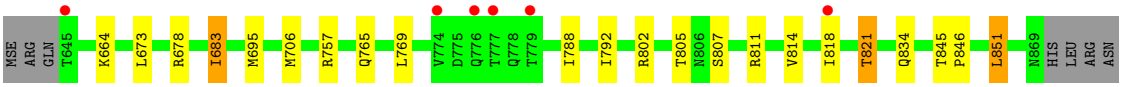
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	115	115	115	0	0
2	B	139	139	139	0	0
2	C	91	91	91	0	0
2	D	78	78	78	0	0

3 Residue-property plots (i)

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: CLIP-associating protein 2

Chain A:

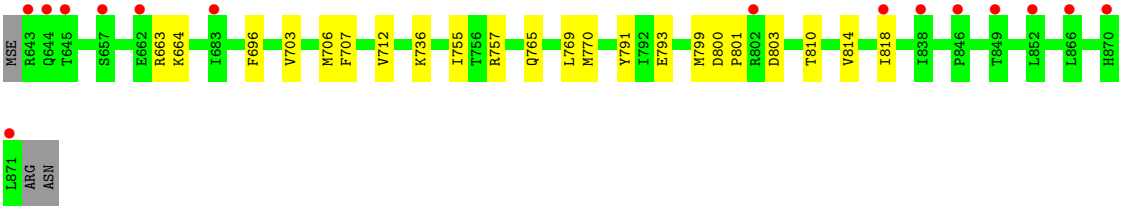


• Molecule 1: CLIP-associating protein 2

Chain B:

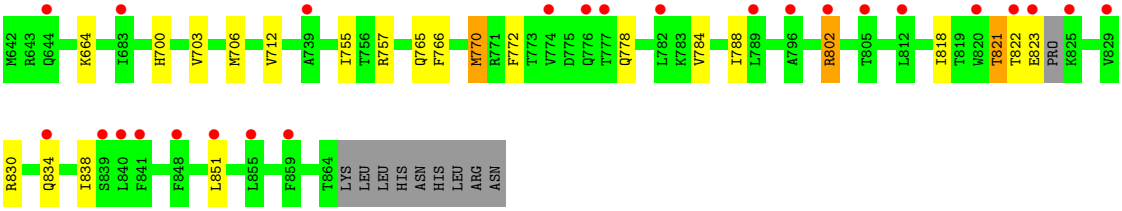
• Molecule 1: CLIP-associating protein 2

Chain C:



• Molecule 1: CLIP-associating protein 2

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	50.90Å 122.32Å 86.51Å 90.00° 94.67° 90.00°	Depositor
Resolution (Å)	50.00 – 2.20 49.88 – 2.15	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.00-2.20) 99.7 (49.88-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.20 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.241 , 0.279 0.238 , 0.277	Depositor DCC
R_{free} test set	2917 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	48.3	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 48.5	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	0.95	EDS
Total number of atoms	7722	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.31% 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	0.35	0/1848	0.50	1/2483 (0.0%)
1	B	0.37	0/1866	0.48	0/2510
1	C	0.35	0/1866	0.49	0/2510
1	D	0.34	0/1821	0.50	1/2444 (0.0%)
All	All	0.36	0/7401	0.49	2/9947 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	851	LEU	CA-CB-CG	5.13	127.10	115.30
1	A	851	LEU	CA-CB-CG	5.05	126.91	115.30

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	1820	0	1886	12	0
1	B	1842	0	1899	12	0
1	C	1842	0	1899	11	0
1	D	1795	0	1855	15	0
2	A	115	0	0	0	0
2	B	139	0	0	1	0
2	C	91	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	78	0	0	1	0
All	All	7722	0	7539	48	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 3.

The worst 5 of 48 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:821:THR:CG2	1:A:834:GLN:HG3	2.09	0.82
1:A:814:VAL:O	1:A:818:ILE:HG12	1.86	0.75
1:B:818:ILE:O	1:B:821:THR:HG22	1.88	0.73
1:A:788:ILE:O	1:A:792:ILE:HG12	1.89	0.71
1:A:769:LEU:HD13	1:A:792:ILE:HD13	1.72	0.70

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	225/232 (97%)	220 (98%)	5 (2%)	0	100 100
1	B	227/232 (98%)	225 (99%)	2 (1%)	0	100 100
1	C	227/232 (98%)	225 (99%)	1 (0%)	1 (0%)	34 37
1	D	219/232 (94%)	216 (99%)	3 (1%)	0	100 100
All	All	898/928 (97%)	886 (99%)	11 (1%)	1 (0%)	51 60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	801	PRO

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	207/205 (101%)	203 (98%)	4 (2%)	57	71
1	B	209/205 (102%)	207 (99%)	2 (1%)	76	86
1	C	209/205 (102%)	206 (99%)	3 (1%)	67	80
1	D	204/205 (100%)	201 (98%)	3 (2%)	65	78
All	All	829/820 (101%)	817 (99%)	12 (1%)	65	80

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	793	GLU
1	C	800	ASP
1	D	821	THR
1	D	770	MSE
1	A	851	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	869	ASN
1	D	700	HIS
1	D	860	GLN
1	D	762	ASN
1	D	765	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 monosaccharides 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	219/232 (94%)	0.34	6 (2%)	54	52	34, 51, 82, 94	0
1	B	223/232 (96%)	0.27	7 (3%)	49	47	35, 48, 82, 91	0
1	C	223/232 (96%)	0.57	15 (6%)	17	16	38, 57, 76, 95	0
1	D	215/232 (92%)	0.74	25 (11%)	4	4	46, 62, 103, 107	0
All	All	880/928 (94%)	0.48	53 (6%)	21	20	34, 54, 96, 107	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	645	THR	7.9
1	D	820	TRP	7.8
1	D	802	ARG	5.9
1	C	871	LEU	5.3
1	D	822	THR	4.9

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 monosaccharides 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.