wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 02:47 AM EDT

PDB ID	:	3K8P
Title	:	Structural basis for vesicle tethering by the Dsl1 complex
Authors	:	Ren, Y.; Jeffrey, P.D.; Hughson, F.M.
Deposited on	:	2009-10-14
Resolution	:	2.60 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:

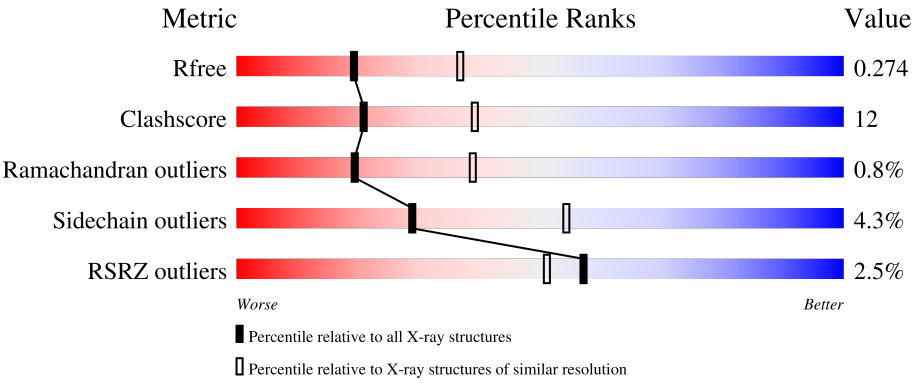
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	C	357	2% 62% 19% • 17%
2	D	709	2% 59% 22% • 16%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7376 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 Dsl1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	C	295	2464	1594	404	458	2	6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	330	GLY	-	EXPRESSION TAG	UNP Q6CUS2
C	331	SER	-	EXPRESSION TAG	UNP Q6CUS2

- Molecule 2 is a protein called Protein transport protein SEC39.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
2	D	594	4909	3182	775	931	7	14	0	0	0

- Molecule 3 is water.

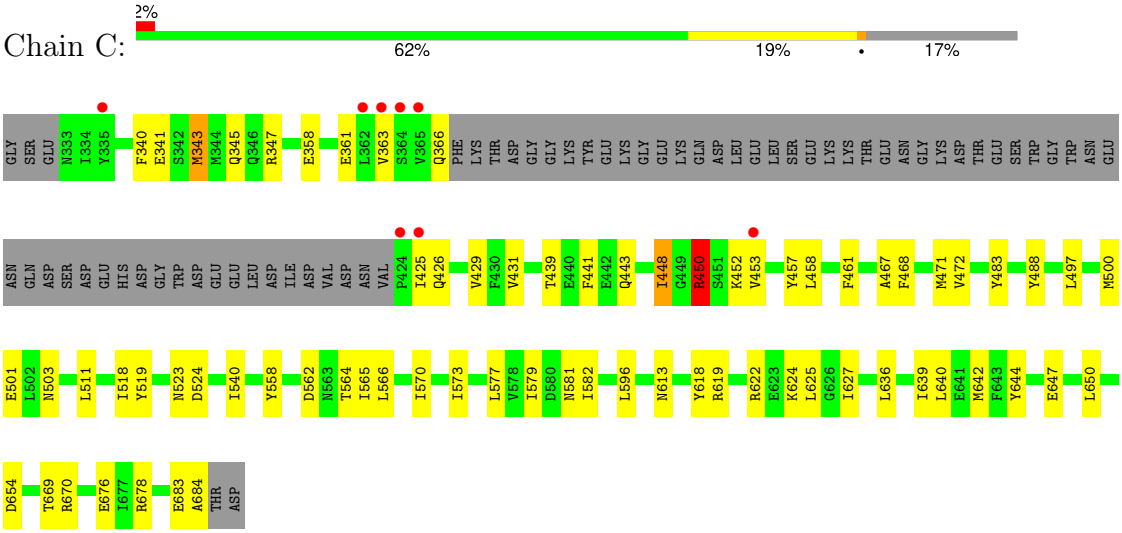
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	D	3	3	3	0	0

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.

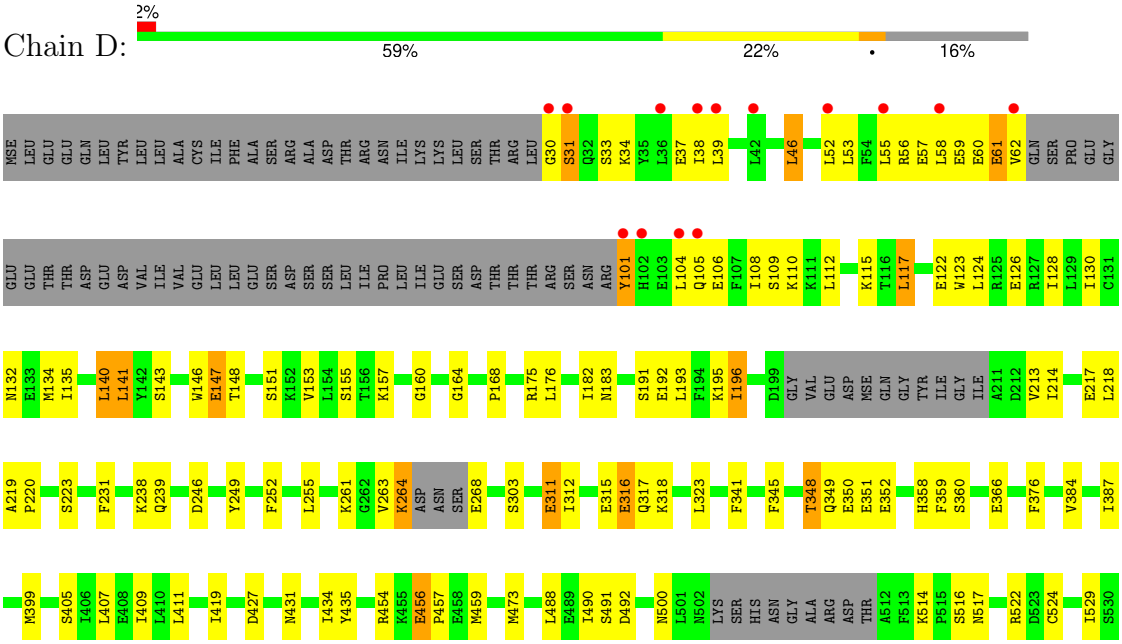
• Molecule 1: Dsl1

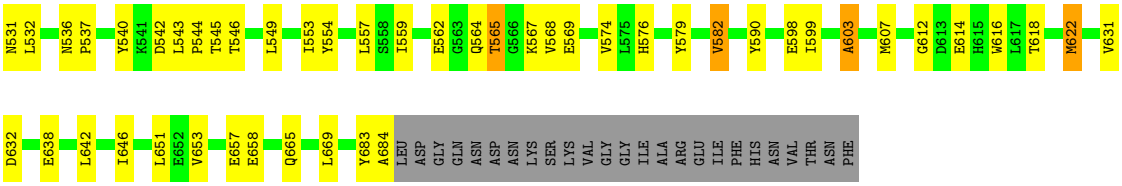
Chain C:



• Molecule 2: Protein transport protein SEC39

Chain D:




N531 L532 N536 P537 Y540 K541 D542 L543 P544 T545 T546 L549 I553 Y554 L557 S558 I559 E562 G563 Q564 T565 G566 K567 V568 E569 V574 L575 H576 Y579 V582 Y590 E598 I599 A603 M607 G612 D613 E614 H615 W616 L617 T618 M622 V631
D632 E638 L642 I646 L651 E652 V653 E657 E658 Q665 L669 Y683 A684 LEU ASP GLY GLN ASN ASP ASN LYS SER LYS VAL GLY GLY ILE ALA ARG GLU ILE PHE HIS ASN VAL THR ASN PHE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.02Å 90.82Å 213.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	31.20 – 2.60 46.08 – 2.60	Depositor EDS
% Data completeness (in resolution range)	94.5 (31.20-2.60) 97.8 (46.08-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	5.08 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.205 , 0.271 0.213 , 0.274	Depositor DCC
R_{free} test set	2212 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	49.1	Xtriage
Anisotropy	0.397	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 48.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7376	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.27% 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

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	C	0.43	0/2503	0.59	1/3376 (0.0%)
2	D	0.41	0/4998	0.55	0/6724
All	All	0.42	0/7501	0.57	1/10100 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	450	ARG	N-CA-C	-5.56	95.99	111.00

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	C	2464	0	2516	46	0
2	D	4909	0	4886	136	0
3	D	3	0	0	0	0
All	All	7376	0	7402	180	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 12.

The worst 5 of 180 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:D:435:TYR:HB2	2:D:473:MSE:HE2	1.42	0.97
2:D:316:GLU:HG3	2:D:317:GLN:H	1.32	0.94
2:D:348:THR:HG22	2:D:349:GLN:HG3	1.57	0.87
2:D:546:THR:HG21	2:D:576:HIS:CE1	2.11	0.86
2:D:603:ALA:HB1	2:D:607:MSE:HG3	1.57	0.84

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	C	291/357 (82%)	278 (96%)	11 (4%)	2 (1%)	22	43
2	D	584/709 (82%)	548 (94%)	31 (5%)	5 (1%)	17	35
All	All	875/1066 (82%)	826 (94%)	42 (5%)	7 (1%)	19	39

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	448	ILE
1	C	450	ARG
2	D	316	GLU
2	D	31	SER
2	D	61	GLU

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	C	281/330 (85%)	274 (98%)	7 (2%)	47	73
2	D	555/641 (87%)	526 (95%)	29 (5%)	23	46
All	All	836/971 (86%)	800 (96%)	36 (4%)	29	54

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	348	THR
2	D	657	GLU
2	D	427	ASP
2	D	565	THR
2	D	122	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	531	ASN
2	D	576	HIS
2	D	665	GLN
2	D	239	GLN
2	D	358	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 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

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, 95^{th} 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	C	289/357 (80%)	-0.06	8 (2%)	53	46	30, 51, 106, 149	0
2	D	580/709 (81%)	-0.03	14 (2%)	59	53	35, 59, 102, 145	0
All	All	869/1066 (81%)	-0.04	22 (2%)	57	51	30, 56, 103, 149	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	102	HIS	5.5
2	D	38	ILE	3.9
1	C	363	VAL	3.4
1	C	425	ILE	3.2
2	D	36	LEU	2.9

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.