Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2023 – 04:29 PM EDT

PDB ID	:	8F0E
Title	:	N-terminal WD40 domain of beta'-COPI subunit with four chains in the asymmetric unit
Authors	:	Dey, D.; Hasan, S.S.
Deposited on	:	2022-11-02
Resolution	:	1.31 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

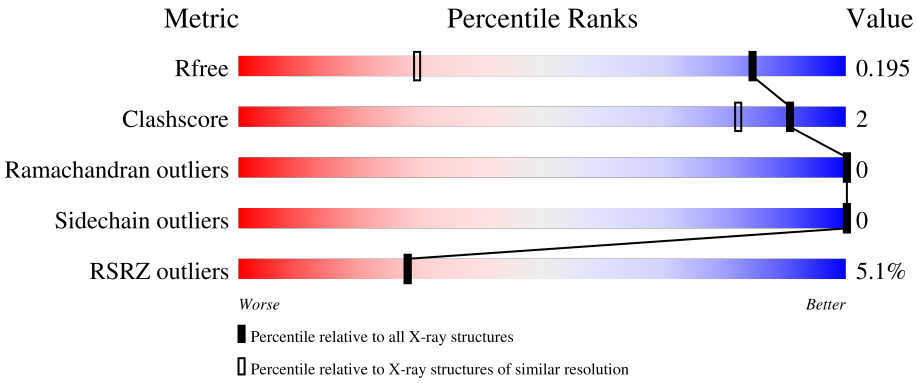
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.31 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	1611 (1.34-1.30)
Clashscore	141614	1667 (1.34-1.30)
Ramachandran outliers	138981	1615 (1.34-1.30)
Sidechain outliers	138945	1615 (1.34-1.30)
RSRZ outliers	127900	1580 (1.34-1.30)

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	302	98%
1	B	302	8%, 91%, 6%
1	C	302	2%, 96%
1	D	302	9%, 93%, 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 20182 atoms, of which 9260 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 Coatomer subunit beta'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	299	Total	C	H	N	O	S	0	0	0
			4737	1542	2332	403	453	7			
1	B	293	Total	C	H	N	O	S	0	0	0
			4638	1514	2282	392	443	7			
1	C	299	Total	C	H	N	O	S	0	0	0
			4737	1542	2332	403	453	7			
1	D	296	Total	C	H	N	O	S	0	0	0
			4698	1530	2314	400	447	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP C8Z8B3
B	0	SER	-	expression tag	UNP C8Z8B3
C	0	SER	-	expression tag	UNP C8Z8B3
D	0	SER	-	expression tag	UNP C8Z8B3

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	416	Total	O	0	0
			416	416		
2	B	277	Total	O	0	0
			277	277		
2	C	402	Total	O	0	0
			402	402		
2	D	277	Total	O	0	0
			277	277		

3 Residue-property plots (i)

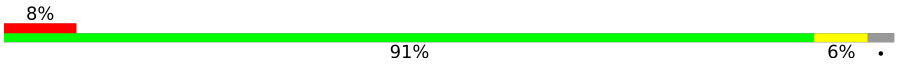
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

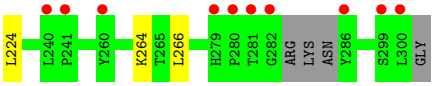
• Molecule 1: Coatomer subunit beta'

Chain A:

• Molecule 1: Coatomer subunit beta'

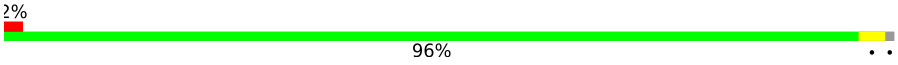
Chain B:

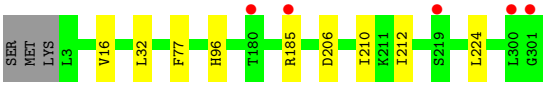




• Molecule 1: Coatomer subunit beta'

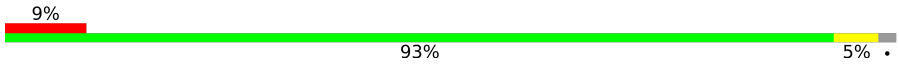
Chain C:

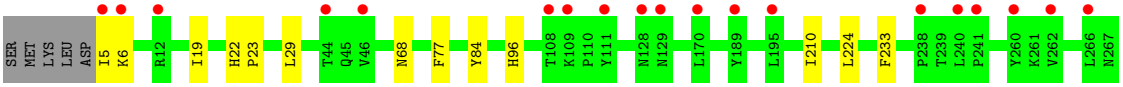


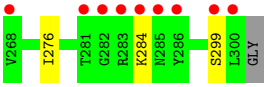


• Molecule 1: Coatomer subunit beta'

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	49.12Å 89.94Å 138.32Å 90.00° 93.76° 90.00°	Depositor
Resolution (Å)	29.13 – 1.31 29.37 – 1.31	Depositor EDS
% Data completeness (in resolution range)	99.5 (29.13-1.31) 99.7 (29.37-1.31)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 1.31Å)	Xtriage
Refinement program	REFMAC 1.20.1_4487+SVN	Depositor
R, R_{free}	0.157 , 0.191 0.163 , 0.195	Depositor DCC
R_{free} test set	14293 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	9.6	Xtriage
Anisotropy	0.106	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 50.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	20182	wwPDB-VP
Average B, all atoms (Å^2)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 92.41 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.3398e-09. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	A	0.45	0/2474	0.73	0/3373
1	B	0.40	0/2424	0.72	0/3307
1	C	0.46	0/2474	0.75	0/3373
1	D	0.40	0/2453	0.72	0/3346
All	All	0.43	0/9825	0.73	0/13399

There are no bond length outliers.

There are no bond angle outliers.

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	A	2405	2332	2332	2	0
1	B	2356	2282	2281	10	0
1	C	2405	2332	2332	10	0
1	D	2384	2314	2314	10	0
2	A	416	0	0	0	0
2	B	277	0	0	0	0
2	C	402	0	0	0	0
2	D	277	0	0	0	0
All	All	10922	9260	9259	32	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 2.

All (32) 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:D:5:ILE:O	1:D:6:LYS:HD2	1.88	0.73
1:D:284:LYS:HG3	1:D:299:SER:HB2	1.80	0.61
1:C:212:ILE:CD1	1:C:224:LEU:CD1	2.81	0.58
1:C:77:PHE:HB3	1:C:96:HIS:O	2.07	0.54
1:A:77:PHE:HB3	1:A:96:HIS:O	2.08	0.54
1:C:212:ILE:CD1	1:C:224:LEU:HD13	2.39	0.53
1:D:284:LYS:HG3	1:D:284:LYS:O	2.09	0.52
1:C:212:ILE:CD1	1:C:224:LEU:HD11	2.39	0.52
1:B:264:LYS:HE3	1:B:266:LEU:HD21	1.90	0.52
1:C:212:ILE:HD13	1:C:224:LEU:HD11	1.92	0.52
1:C:224:LEU:HD12	1:C:224:LEU:N	2.25	0.51
1:B:264:LYS:HE3	1:B:266:LEU:CD2	2.44	0.47
1:B:77:PHE:HB3	1:B:96:HIS:O	2.15	0.47
1:C:212:ILE:HD12	1:C:212:ILE:N	2.32	0.45
1:B:16:VAL:HG22	1:B:32:LEU:CD2	2.47	0.45
1:B:210:ILE:HB	1:B:224:LEU:HB2	1.99	0.45
1:A:210:ILE:HB	1:A:224:LEU:HB2	2.00	0.44
1:B:202:ILE:HG13	1:B:212:ILE:HD13	2.00	0.42
1:D:210:ILE:HB	1:D:224:LEU:HB2	2.01	0.42
1:D:68:ASN:HA	1:D:84:TYR:CZ	2.55	0.42
1:D:19:ILE:HA	1:D:29:LEU:O	2.20	0.41
1:D:77:PHE:HB3	1:D:96:HIS:O	2.20	0.41
1:C:16:VAL:HG22	1:C:32:LEU:CD2	2.50	0.41
1:C:185:ARG:HB3	1:C:206:ASP:HB2	2.02	0.41
1:B:58:VAL:HG22	1:B:74:SER:HB3	2.03	0.41
1:B:19:ILE:HA	1:B:29:LEU:O	2.21	0.41
1:B:163:ARG:HA	1:B:186:GLY:HA2	2.02	0.41
1:D:233:PHE:CE1	1:D:276:ILE:HB	2.56	0.41
1:B:68:ASN:HA	1:B:84:TYR:CZ	2.56	0.40
1:D:22:HIS:CG	1:D:23:PRO:HD2	2.56	0.40
1:D:284:LYS:CG	1:D:299:SER:HB2	2.49	0.40
1:C:210:ILE:HB	1:C:224:LEU:HB2	2.02	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	297/302 (98%)	283 (95%)	14 (5%)	0	100	100
1	B	289/302 (96%)	280 (97%)	9 (3%)	0	100	100
1	C	297/302 (98%)	285 (96%)	12 (4%)	0	100	100
1	D	294/302 (97%)	282 (96%)	12 (4%)	0	100	100
All	All	1177/1208 (97%)	1130 (96%)	47 (4%)	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	268/271 (99%)	268 (100%)	0	100	100
1	B	263/271 (97%)	263 (100%)	0	100	100
1	C	268/271 (99%)	268 (100%)	0	100	100
1	D	266/271 (98%)	266 (100%)	0	100	100
All	All	1065/1084 (98%)	1065 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	A	299/302 (99%)	-0.12	3 (1%) 82 82	5, 9, 17, 32	0
1	B	293/302 (97%)	0.39	24 (8%) 11 9	8, 15, 32, 58	0
1	C	299/302 (99%)	-0.15	5 (1%) 70 71	5, 9, 18, 26	0
1	D	296/302 (98%)	0.44	28 (9%) 8 6	8, 15, 28, 41	0
All	All	1187/1208 (98%)	0.14	60 (5%) 28 28	5, 12, 27, 58	0

All (60) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	300	LEU	9.3
1	D	5	ILE	8.7
1	D	300	LEU	6.1
1	D	46	VAL	5.4
1	D	283	ARG	5.4
1	D	281	THR	5.1
1	D	284	LYS	4.9
1	B	6	LYS	4.7
1	B	5	ILE	4.6
1	D	285	ASN	4.3
1	B	281	THR	4.1
1	B	46	VAL	3.9
1	B	286	TYR	3.9
1	D	44	THR	3.6
1	C	300	LEU	3.6
1	D	299	SER	3.6
1	D	111	TYR	3.5
1	B	129	ASN	3.3
1	B	111	TYR	3.3
1	B	183	GLN	3.3
1	D	195	LEU	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	129	ASN	3.2
1	B	42	TYR	3.2
1	D	12	ARG	3.1
1	D	6	LYS	3.1
1	B	282	GLY	3.1
1	B	12	ARG	2.9
1	D	260	TYR	2.9
1	D	282	GLY	2.9
1	D	241	PRO	2.9
1	B	240	LEU	2.9
1	D	240	LEU	2.9
1	B	108	THR	2.8
1	D	286	TYR	2.8
1	B	241	PRO	2.8
1	B	299	SER	2.8
1	C	301	GLY	2.7
1	B	170	LEU	2.7
1	A	301	GLY	2.6
1	B	172	GLN	2.6
1	B	44	THR	2.6
1	A	219	SER	2.5
1	B	279	HIS	2.5
1	B	43	GLU	2.4
1	B	260	TYR	2.4
1	D	128	ASN	2.3
1	B	189	TYR	2.3
1	D	266	LEU	2.3
1	B	280	PRO	2.2
1	D	262	VAL	2.2
1	D	189	TYR	2.2
1	D	268	VAL	2.2
1	D	238	PRO	2.2
1	C	219	SER	2.1
1	A	300	LEU	2.1
1	C	180	THR	2.1
1	D	108	THR	2.1
1	C	185	ARG	2.0
1	D	109	LYS	2.0
1	D	170	LEU	2.0

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.