Full wwPDB X-ray Structure Validation Report ⓘ

Sep 19, 2024 – 01:14 pm BST

PDB ID	:	8R9T
Title	:	Crystal structure of the C-terminal domain of Chlamydomonas reinhardtii CFAP410
Authors	:	Stadler, A.; Dong, G.
Deposited on	:	2023-11-30
Resolution	:	1.40 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:

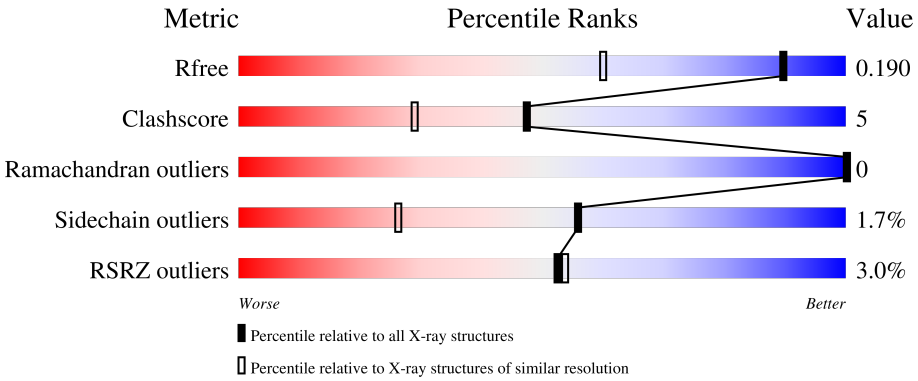
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.2

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 1.40 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}	164625	2247 (1.40-1.40)
Clashscore	180529	2446 (1.40-1.40)
Ramachandran outliers	177936	2398 (1.40-1.40)
Sidechain outliers	177891	2397 (1.40-1.40)
RSRZ outliers	164620	2246 (1.40-1.40)

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	35	91% · 6%
1	B	35	3% 80% 14% 6%
1	C	35	3% 71% 14% 6% 9%
1	D	35	6% 91% 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2279 atoms, of which 1103 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 Flagellar associated protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	33	Total	C	H	N	O	S	0	0	0
			540	169	272	45	53	1			
1	B	33	Total	C	H	N	O	S	0	0	0
			541	169	273	45	53	1			
1	C	32	Total	C	H	N	O	S	0	0	0
			533	167	269	44	52	1			
1	D	35	Total	C	H	N	O	S	0	0	0
			573	177	289	50	56	1			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	26	Total	O	0	0
			26	26		
3	B	20	Total	O	0	0
			20	20		
3	C	16	Total	O	0	0
			16	16		
3	D	29	Total	O	0	0
			29	29		

3 Residue-property plots (i)

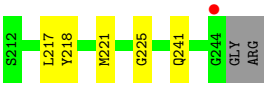
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

Chain A:

• Molecule 1: Flagellar associated protein

Chain B:

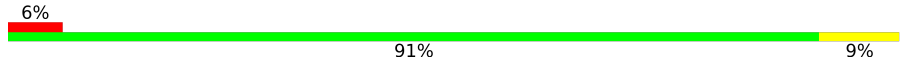


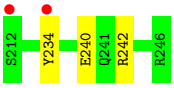
• Molecule 1: Flagellar associated protein

Chain C:

• Molecule 1: Flagellar associated protein

Chain D:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	43.44Å 48.10Å 53.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.79 – 1.40 19.79 – 1.40	Depositor EDS
% Data completeness (in resolution range)	97.8 (19.79-1.40) 98.2 (19.79-1.40)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.02 (at 1.39Å)	Xtriage
Refinement program	PHENIX 1.18_3845, PHENIX 1.18_3845	Depositor
R, R_{free}	0.163 , 0.189 0.165 , 0.190	Depositor DCC
R_{free} test set	20543 reflections (8.98%)	wwPDB-VP
Wilson B-factor (Å^2)	21.7	Xtriage
Anisotropy	0.305	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.49 , 52.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2279	wwPDB-VP
Average B, all atoms (Å^2)	31.0	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 34.15 % 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 7.1365e-04. 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

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.32	0/269	0.51	0/361
1	B	0.40	0/269	0.54	0/361
1	C	0.71	2/265 (0.8%)	1.07	3/356 (0.8%)
1	D	0.62	1/285 (0.4%)	0.59	0/380
All	All	0.54	3/1088 (0.3%)	0.71	3/1458 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	234	TYR	CB-CG	-5.79	1.43	1.51
1	C	237	ARG	CZ-NH2	5.75	1.40	1.33
1	C	237	ARG	CZ-NH1	5.54	1.40	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	237	ARG	NE-CZ-NH2	-14.59	113.00	120.30
1	C	237	ARG	CD-NE-CZ	6.71	132.99	123.60
1	C	237	ARG	NE-CZ-NH1	6.22	123.41	120.30

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	268	272	273	1	0
1	B	268	273	273	2	1
1	C	264	269	270	5	0
1	D	284	289	289	4	0
2	A	1	0	0	0	0
3	A	26	0	0	1	1
3	B	20	0	0	1	0
3	C	16	0	0	1	0
3	D	29	0	0	3	2
All	All	1176	1103	1105	12	3

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 5.

All (12) 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:240:GLU:OE1	3:D:301:HOH:O	1.88	0.90
1:D:242:ARG:O	3:D:302:HOH:O	2.04	0.75
1:C:240:GLU:OE2	3:C:301:HOH:O	2.11	0.68
1:A:212:SER:N	3:A:401:HOH:O	2.28	0.67
1:C:214:LYS:HE3	1:C:218:TYR:HE1	1.64	0.62
1:D:240:GLU:CD	3:D:303:HOH:O	2.41	0.57
1:C:217:LEU:O	1:C:221:MET:HG2	2.11	0.49
1:B:241:GLN:NE2	3:B:302:HOH:O	2.45	0.49
1:D:240:GLU:OE1	1:D:240:GLU:HA	2.14	0.48
1:B:217:LEU:O	1:B:221:MET:HG2	2.14	0.47
1:C:214:LYS:HG2	1:C:218:TYR:CE1	2.52	0.44
1:C:234:TYR:CD2	1:C:237:ARG:NH2	2.86	0.44

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:420:HOH:O	3:A:426:HOH:O[2_675]	1.69	0.51
3:D:321:HOH:O	3:D:326:HOH:O[3_745]	1.97	0.23
1:B:225:GLY:O	3:D:302:HOH:O[2_665]	2.15	0.05

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	31/35 (89%)	31 (100%)	0	0	100	100
1	B	31/35 (89%)	30 (97%)	1 (3%)	0	100	100
1	C	30/35 (86%)	30 (100%)	0	0	100	100
1	D	33/35 (94%)	32 (97%)	1 (3%)	0	100	100
All	All	125/140 (89%)	123 (98%)	2 (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	29/30 (97%)	29 (100%)	0	100	100
1	B	29/30 (97%)	28 (97%)	1 (3%)	32	6
1	C	29/30 (97%)	28 (97%)	1 (3%)	32	6
1	D	30/30 (100%)	30 (100%)	0	100	100
All	All	117/120 (98%)	115 (98%)	2 (2%)	56	26

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	218	TYR
1	C	240	GLU

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

Mol	Chain	Res	Type
1	B	241	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	33/35 (94%)	-0.38	0 100 100	19, 26, 38, 42	0
1	B	33/35 (94%)	-0.28	1 (3%) 52 54	21, 27, 44, 51	0
1	C	32/35 (91%)	0.25	1 (3%) 51 53	22, 32, 47, 51	0
1	D	35/35 (100%)	-0.03	2 (5%) 30 30	20, 28, 46, 51	0
All	All	133/140 (95%)	-0.11	4 (3%) 52 54	19, 27, 46, 51	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	234	TYR	2.4
1	C	243	ILE	2.2
1	B	244	GLY	2.2
1	D	212	SER	2.1

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CA	A	301	1/1	1.00	0.02	21,21,21,21	0

6.5 Other polymers (i)

There are no such residues in this entry.