Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2022 – 01:29 pm BST

PDB ID	:	7OIQ
Title	:	Crystal structure of AP2 Mu2 in complex with FCHO2 WxxPhi motif (C2 crystal form)
Authors	:	Zaccai, N.R.; Kelly, B.T.; Evans, P.R.; Owen, D.J.
Deposited on	:	2021-05-12
Resolution	:	1.85 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 following versions of software and data (see references ⓘ) were used in the production of this report:

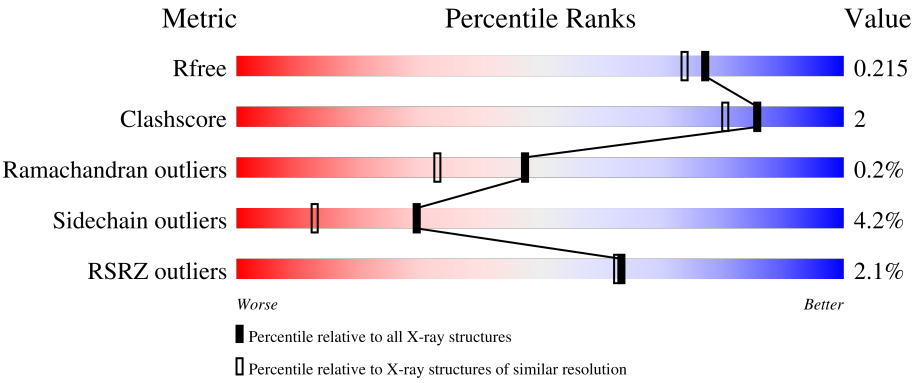
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

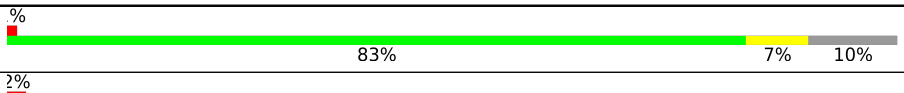
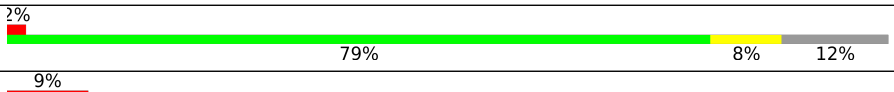
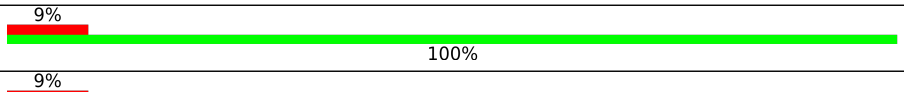
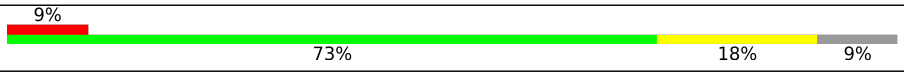
The reported resolution of this entry is 1.85 Å.

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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	AAA	285	
1	BBB	285	
2	CCC	11	
2	DDD	11	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4604 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 AP-2 complex subunit mu.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	257	Total	C	N	O	S	0	0	0
			2075	1334	364	363	14			
1	BBB	250	Total	C	N	O	S	0	0	0
			2011	1293	353	351	14			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	151	MET	-	initiating methionine	UNP P84092
AAA	152	HIS	-	expression tag	UNP P84092
AAA	153	HIS	-	expression tag	UNP P84092
AAA	154	HIS	-	expression tag	UNP P84092
AAA	155	HIS	-	expression tag	UNP P84092
AAA	156	HIS	-	expression tag	UNP P84092
AAA	157	HIS	-	expression tag	UNP P84092
BBB	151	MET	-	initiating methionine	UNP P84092
BBB	152	HIS	-	expression tag	UNP P84092
BBB	153	HIS	-	expression tag	UNP P84092
BBB	154	HIS	-	expression tag	UNP P84092
BBB	155	HIS	-	expression tag	UNP P84092
BBB	156	HIS	-	expression tag	UNP P84092
BBB	157	HIS	-	expression tag	UNP P84092

- Molecule 2 is a protein called F-BAR domain only protein 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	CCC	11	Total	C	N	O	0	0	0
			87	59	12	16			
2	DDD	10	Total	C	N	O	0	0	0
			81	56	11	14			

- Molecule 3 is water.

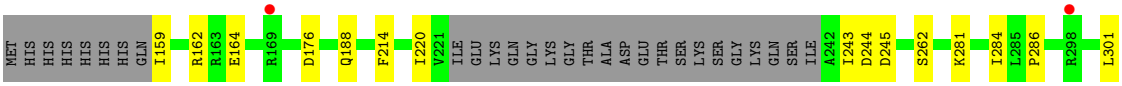
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	180	Total O 180 180	0	0
3	BBB	148	Total O 148 148	0	0
3	CCC	13	Total O 13 13	0	0
3	DDD	9	Total O 9 9	0	0

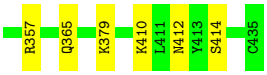
3 Residue-property plots (i)

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: AP-2 complex subunit mu

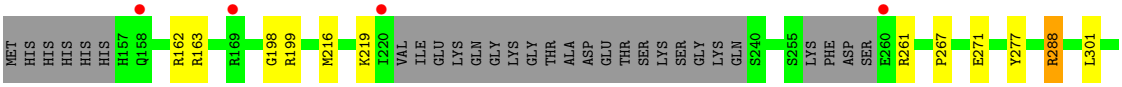
Chain AAA:

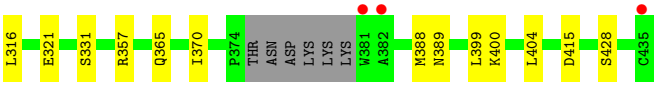




• Molecule 1: AP-2 complex subunit mu

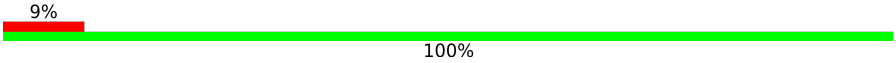
Chain BBB:





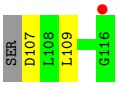
• Molecule 2: F-BAR domain only protein 2

Chain CCC:



• Molecule 2: F-BAR domain only protein 2

Chain DDD:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	118.89Å 64.54Å 108.44Å 90.00° 111.95° 90.00°	Depositor
Resolution (Å)	55.70 – 1.85 55.70 – 1.85	Depositor EDS
% Data completeness (in resolution range)	89.4 (55.70-1.85) 89.4 (55.70-1.85)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.39 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.8.0267, PHENIX 1.19rc1_4016	Depositor
R, R_{free}	0.186 , 0.209 0.195 , 0.215	Depositor DCC
R_{free} test set	2932 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	30.3	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	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	0.96	EDS
Total number of atoms	4604	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.44% 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	AAA	0.76	0/2118	0.92	0/2848
1	BBB	0.75	1/2051 (0.0%)	0.94	1/2757 (0.0%)
2	CCC	0.61	0/90	0.81	0/123
2	DDD	0.69	0/84	0.84	0/115
All	All	0.75	1/4343 (0.0%)	0.92	1/5843 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	271	GLU	CD-OE1	-6.91	1.18	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	BBB	199	ARG	NE-CZ-NH2	-6.09	117.26	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	AAA	2075	0	2163	5	0
1	BBB	2011	0	2090	13	0
2	CCC	87	0	79	0	0
2	DDD	81	0	74	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	AAA	180	0	0	0	0
3	BBB	148	0	0	0	0
3	CCC	13	0	0	0	0
3	DDD	9	0	0	0	0
All	All	4604	0	4406	19	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 (19) 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:BBB:216:MET:HE2	1:BBB:399:LEU:HD21	1.54	0.87
1:BBB:288:ARG:HH11	1:BBB:288:ARG:HG3	1.53	0.73
1:BBB:216:MET:CE	1:BBB:399:LEU:HD21	2.20	0.71
1:BBB:288:ARG:HH11	1:BBB:288:ARG:CG	2.16	0.58
1:BBB:216:MET:HE2	1:BBB:399:LEU:CD2	2.35	0.48
1:BBB:301:LEU:HB3	1:BBB:370:ILE:HB	1.98	0.46
1:BBB:162:ARG:HD2	1:BBB:267:PRO:O	2.16	0.45
1:BBB:288:ARG:CG	1:BBB:288:ARG:NH1	2.77	0.43
1:AAA:159:ILE:HD11	1:AAA:162:ARG:HB3	2.00	0.43
1:BBB:316:LEU:HD13	1:BBB:357:ARG:HD2	2.00	0.43
1:AAA:220:ILE:HG22	1:AAA:243:ILE:CD1	2.49	0.42
1:AAA:284:ILE:O	1:AAA:286:PRO:HD3	2.20	0.42
1:BBB:389:ASN:HA	1:BBB:428:SER:OG	2.19	0.42
1:BBB:216:MET:SD	1:BBB:399:LEU:HD11	2.60	0.42
2:DDD:107:ASP:OD2	2:DDD:109:LEU:HB2	2.20	0.41
1:AAA:214:PHE:O	1:AAA:262:SER:HA	2.20	0.41
1:AAA:357:ARG:NH1	1:AAA:357:ARG:HB3	2.36	0.41
1:BBB:321:GLU:O	1:BBB:388:MET:HA	2.21	0.41
1:BBB:198:GLY:HA3	1:BBB:277:TYR:CZ	2.56	0.41

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	AAA	253/285 (89%)	244 (96%)	8 (3%)	1 (0%)	34	19
1	BBB	242/285 (85%)	236 (98%)	6 (2%)	0	100	100
2	CCC	9/11 (82%)	9 (100%)	0	0	100	100
2	DDD	8/11 (73%)	8 (100%)	0	0	100	100
All	All	512/592 (86%)	497 (97%)	14 (3%)	1 (0%)	47	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	245	ASP

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	AAA	233/257 (91%)	222 (95%)	11 (5%)	26	10
1	BBB	225/257 (88%)	216 (96%)	9 (4%)	31	14
2	CCC	9/9 (100%)	9 (100%)	0	100	100
2	DDD	8/9 (89%)	8 (100%)	0	100	100
All	All	475/532 (89%)	455 (96%)	20 (4%)	30	13

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	164	GLU
1	AAA	176	ASP
1	AAA	188	GLN
1	AAA	244	ASP
1	AAA	281	LYS
1	AAA	301	LEU
1	AAA	365	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	AAA	379	LYS
1	AAA	410	LYS
1	AAA	412	ASN
1	AAA	414	SER
1	BBB	163	ARG
1	BBB	219	LYS
1	BBB	261	ARG
1	BBB	288	ARG
1	BBB	331	SER
1	BBB	365	GLN
1	BBB	400	LYS
1	BBB	404	LEU
1	BBB	415	ASP

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

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	AAA	257/285 (90%)	-0.24	2 (0%) 86 86	23, 33, 64, 84	0
1	BBB	250/285 (87%)	-0.25	7 (2%) 53 52	24, 36, 76, 108	0
2	CCC	11/11 (100%)	0.16	1 (9%) 9 8	26, 32, 66, 80	0
2	DDD	10/11 (90%)	-0.09	1 (10%) 7 6	29, 34, 53, 69	0
All	All	528/592 (89%)	-0.23	11 (2%) 63 63	23, 35, 67, 108	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	220	ILE	4.8
1	BBB	382	ALA	4.4
1	BBB	158	GLN	4.1
2	CCC	106	SER	4.0
2	DDD	116	GLY	3.0
1	AAA	298	ARG	3.0
1	BBB	260	GLU	3.0
1	BBB	381	TRP	2.9
1	AAA	169	ARG	2.7
1	BBB	435	CYS	2.3
1	BBB	169	ARG	2.3

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.