



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 4, 2021 – 08:05 PM JST

PDB ID : 7COP  
Title : Hexameric Ring Complex of Engineered V1-ATPase: A3(De)3\_empty  
Authors : Kosugi, T.; Tanabe, M.; Koga, N.  
Deposited on : 2020-08-05  
Resolution : 2.77 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](mailto: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	:	2.23.1
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)	:	2.23.1

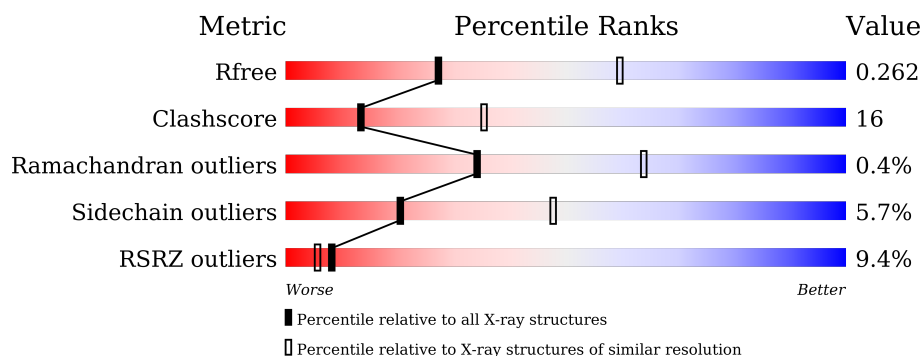
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.77 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	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	596	<div> <div>3%</div> <div>81% 17% ..</div> </div>
1	B	596	<div> <div>7%</div> <div>65% 31% ..</div> </div>
1	C	596	<div> <div>8%</div> <div>60% 34% ..</div> </div>
2	D	458	<div> <div>9%</div> <div>62% 29% . 6%</div> </div>
2	E	458	<div> <div>14%</div> <div>63% 29% 5% ..</div> </div>
2	F	458	<div> <div>17%</div> <div>57% 36% . .</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 24314 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 V-type sodium ATPase catalytic subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	586	Total	C	N	O	S	0	1	0
			4573	2872	770	905	26			
1	B	586	Total	C	N	O	S	0	0	0
			4562	2866	766	904	26			
1	C	584	Total	C	N	O	S	0	1	0
			4560	2864	768	902	26			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q08636
A	-1	SER	-	expression tag	UNP Q08636
A	0	GLY	-	expression tag	UNP Q08636
B	-2	SER	-	expression tag	UNP Q08636
B	-1	SER	-	expression tag	UNP Q08636
B	0	GLY	-	expression tag	UNP Q08636
C	-2	SER	-	expression tag	UNP Q08636
C	-1	SER	-	expression tag	UNP Q08636
C	0	GLY	-	expression tag	UNP Q08636

- Molecule 2 is a protein called V-type sodium ATPase designed non-catalytic subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	432	Total	C	N	O	S	0	0	0
			3369	2134	576	646	13			
2	E	449	Total	C	N	O	S	0	1	0
			3520	2230	605	671	14			
2	F	445	Total	C	N	O	S	0	0	0
			3477	2204	597	662	14			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	151	GLY	SER	engineered mutation	UNP Q08637
D	152	PRO	GLY	engineered mutation	UNP Q08637
D	153	PRO	SER	engineered mutation	UNP Q08637
D	155	ALA	LEU	engineered mutation	UNP Q08637
D	156	GLY	PRO	engineered mutation	UNP Q08637
D	157	LYS	HIS	engineered mutation	UNP Q08637
D	158	SER	LYS	engineered mutation	UNP Q08637
D	159	ALA	GLU	engineered mutation	UNP Q08637
D	248	GLU	THR	engineered mutation	UNP Q08637
D	339	SER	GLN	engineered mutation	UNP Q08637
E	151	GLY	SER	engineered mutation	UNP Q08637
E	152	PRO	GLY	engineered mutation	UNP Q08637
E	153	PRO	SER	engineered mutation	UNP Q08637
E	155	ALA	LEU	engineered mutation	UNP Q08637
E	156	GLY	PRO	engineered mutation	UNP Q08637
E	157	LYS	HIS	engineered mutation	UNP Q08637
E	158	SER	LYS	engineered mutation	UNP Q08637
E	159	ALA	GLU	engineered mutation	UNP Q08637
E	248	GLU	THR	engineered mutation	UNP Q08637
E	339	SER	GLN	engineered mutation	UNP Q08637
F	151	GLY	SER	engineered mutation	UNP Q08637
F	152	PRO	GLY	engineered mutation	UNP Q08637
F	153	PRO	SER	engineered mutation	UNP Q08637
F	155	ALA	LEU	engineered mutation	UNP Q08637
F	156	GLY	PRO	engineered mutation	UNP Q08637
F	157	LYS	HIS	engineered mutation	UNP Q08637
F	158	SER	LYS	engineered mutation	UNP Q08637
F	159	ALA	GLU	engineered mutation	UNP Q08637
F	248	GLU	THR	engineered mutation	UNP Q08637
F	339	SER	GLN	engineered mutation	UNP Q08637

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	100	Total O 100 100	0	0
3	B	36	Total O 36 36	0	0
3	C	45	Total O 45 45	0	0
3	D	23	Total O 23 23	0	0
3	E	29	Total O 29 29	0	0

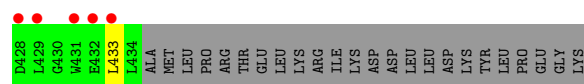
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	F	20	Total	O	0	0
			20	20		







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	122.45Å 122.65Å 128.70Å 90.00° 90.74° 90.00°	Depositor
Resolution (Å)	44.64 – 2.77 44.64 – 2.77	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.64-2.77) 100.0 (44.64-2.77)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 2.77Å)	Xtriage
Refinement program	PHENIX (1.18_3845)	Depositor
R, $R_{free}$	0.212 , 0.262 0.212 , 0.262	Depositor DCC
$R_{free}$ test set	4712 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.5	Xtriage
Anisotropy	0.147	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 59.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k 0.007 for -h,-l,-k 0.000 for -l,k,h 0.000 for -k,-h,-l 0.000 for k,h,-l 0.000 for k,l,h 0.000 for l,h,k 0.000 for l,-h,-k 0.000 for -k,-l,h 0.015 for h,-k,-l 0.011 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	24314	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.73% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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	0.30	0/4649	0.51	1/6289 (0.0%)
1	B	0.29	0/4638	0.50	0/6275
1	C	0.32	0/4636	0.54	1/6271 (0.0%)
2	D	0.37	1/3429 (0.0%)	0.60	0/4637
2	E	0.37	0/3581	0.68	6/4840 (0.1%)
2	F	0.33	0/3538	0.62	4/4782 (0.1%)
All	All	0.33	1/24471 (0.0%)	0.57	12/33094 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	D	0	2
2	F	0	1
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	401	TYR	CD2-CE2	-5.27	1.31	1.39

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	318	LEU	CB-CG-CD2	-12.12	90.39	111.00
2	E	318	LEU	CB-CG-CD1	-10.82	92.61	111.00
2	E	399	LYS	CA-CB-CG	6.55	127.80	113.40
2	E	278	TYR	CA-CB-CG	6.44	125.63	113.40
2	E	363	GLU	OE1-CD-OE2	-5.85	116.28	123.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	176	ASP	Peptide
2	D	354	LYS	Peptide
2	F	171	LEU	Peptide

## 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	4573	0	4540	68	0
1	B	4562	0	4528	142	0
1	C	4560	0	4526	166	0
2	D	3369	0	3377	122	0
2	E	3520	0	3549	166	0
2	F	3477	0	3507	132	0
3	A	100	0	0	5	0
3	B	36	0	0	4	0
3	C	45	0	0	2	0
3	D	23	0	0	1	0
3	E	29	0	0	2	0
3	F	20	0	0	1	0
All	All	24314	0	24027	776	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 16.

The worst 5 of 776 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:368:THR:O	2:D:372:LEU:HB2	1.52	1.08
2:E:403:LYS:HD2	2:E:436:MET:HG3	1.41	0.99
2:E:278:TYR:HB3	2:E:318:LEU:HB3	1.43	0.97
2:F:407:ARG:HA	2:F:410:ASN:HB2	1.50	0.91
2:E:258:ARG:HG2	2:E:274:PRO:HD3	1.52	0.89

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	585/596 (98%)	569 (97%)	16 (3%)	0	100	100
1	B	584/596 (98%)	566 (97%)	18 (3%)	0	100	100
1	C	583/596 (98%)	564 (97%)	18 (3%)	1 (0%)	47	76
2	D	430/458 (94%)	406 (94%)	22 (5%)	2 (0%)	29	58
2	E	448/458 (98%)	413 (92%)	29 (6%)	6 (1%)	12	33
2	F	443/458 (97%)	410 (93%)	30 (7%)	3 (1%)	22	50
All	All	3073/3162 (97%)	2928 (95%)	133 (4%)	12 (0%)	34	64

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	397	ILE
2	E	153	PRO
2	F	388	VAL
1	C	233	PRO
2	D	99	ASN

### 5.3.2 Protein sidechains

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	503/509 (99%)	488 (97%)	15 (3%)	41	72
1	B	502/509 (99%)	483 (96%)	19 (4%)	33	64
1	C	501/509 (98%)	467 (93%)	34 (7%)	16	39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	356/380 (94%)	333 (94%)	23 (6%)	17	41
2	E	373/380 (98%)	344 (92%)	29 (8%)	12	32
2	F	368/380 (97%)	338 (92%)	30 (8%)	11	30
All	All	2603/2667 (98%)	2453 (94%)	150 (6%)	20	47

5 of 150 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	401	TYR
2	F	397	ILE
2	E	447	ASP
2	F	276	TYR
1	C	430	TRP

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

Mol	Chain	Res	Type
2	E	137	HIS
2	E	163	GLN
2	F	326	GLN
2	F	163	GLN
2	F	167	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 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	586/596 (98%)	0.28	17 (2%) 51 46	34, 51, 83, 125	0
1	B	586/596 (98%)	0.46	40 (6%) 17 12	41, 71, 126, 147	0
1	C	584/596 (97%)	0.51	47 (8%) 12 8	43, 74, 121, 139	0
2	D	432/458 (94%)	0.60	43 (9%) 7 5	38, 68, 136, 161	0
2	E	449/458 (98%)	0.87	65 (14%) 2 1	35, 72, 143, 160	0
2	F	445/458 (97%)	0.95	77 (17%) 1 1	41, 79, 150, 170	0
All	All	3082/3162 (97%)	0.59	289 (9%) 8 6	34, 68, 133, 170	0

The worst 5 of 289 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	393	ALA	8.9
2	F	387	VAL	8.6
2	E	272	GLY	8.1
2	F	319	THR	7.6
2	F	312	THR	7.3

### 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.