WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 6, 2023 – 11:47 AM JST

PDB ID	:	8IGU
Title	:	Hexameric Ring Complex of Engineered V1-ATPase: A3(De)3_empty
Authors	:	Kosugi, T.; Tanabe, M.; Koga, N.
Deposited on	:	2023-02-21
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*
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:

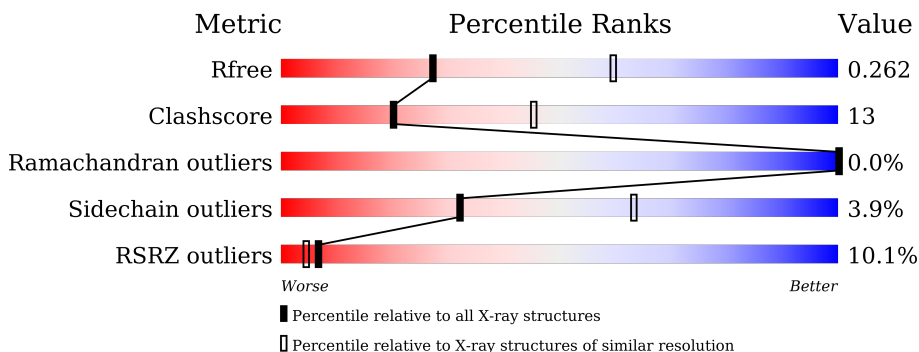|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

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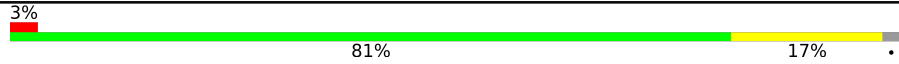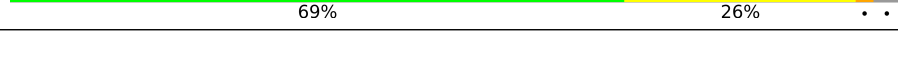*X-RAY DIFFRACTION*

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	Percentile Ranks	Value
Rfree		0.262
Clashscore		13
Ramachandran outliers		0.0%
Sidechain outliers		3.9%
RSRZ outliers		10.1%

Worse ← → *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

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 >=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	596	3% / 81% / 17% ·
1	B	596	7% / 68% / 29% ··
1	C	596	8% / 64% / 30% ··
2	D	458	11% / 66% / 27% · 6%
2	E	458	15% / 68% / 27% ··
2	F	458	18% / 69% / 26% ··

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 24292 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	0	0
			4562	2866	766	904	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 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	0	0
			3509	2224	601	670	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	99	Total 99	O 99	0	0
3	B	35	Total 35	O 35	0	0
3	C	47	Total 47	O 47	0	0
3	D	22	Total 22	O 22	0	0
3	E	30	Total 30	O 30	0	0

Continued on next page...

Continued from previous page...

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	F	20	Total 20	O 20	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.
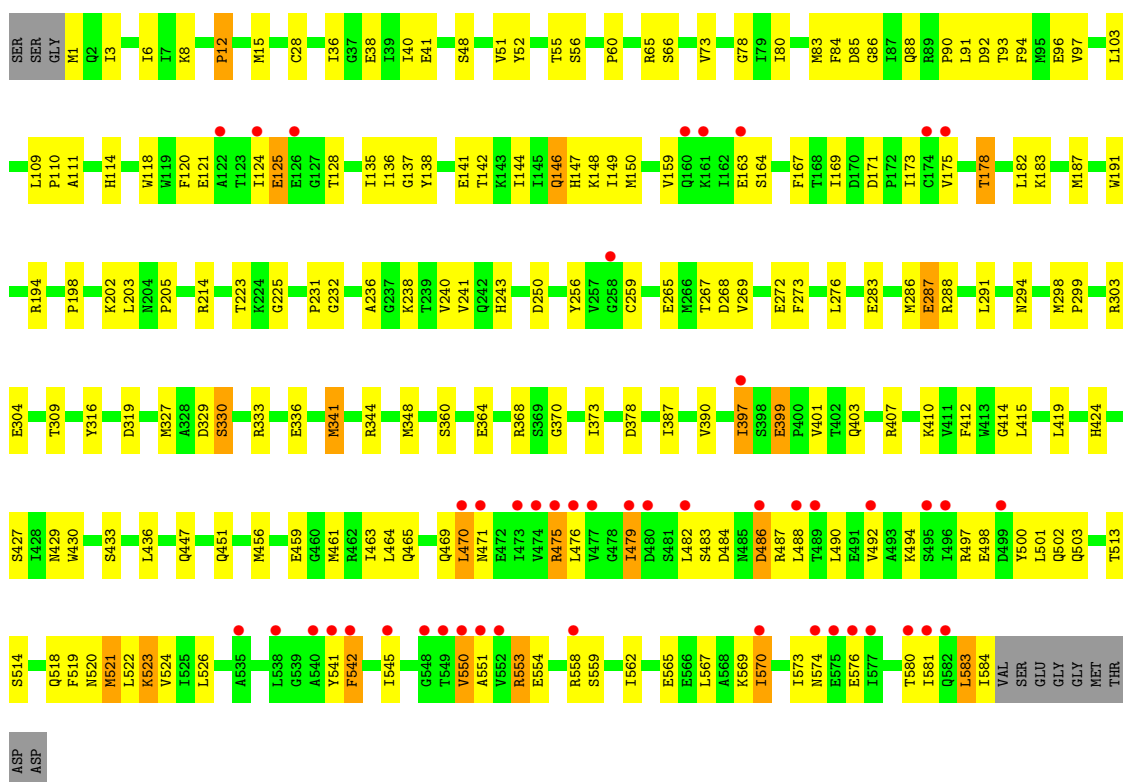
• Molecule 1: V-type sodium ATPase catalytic subunit A
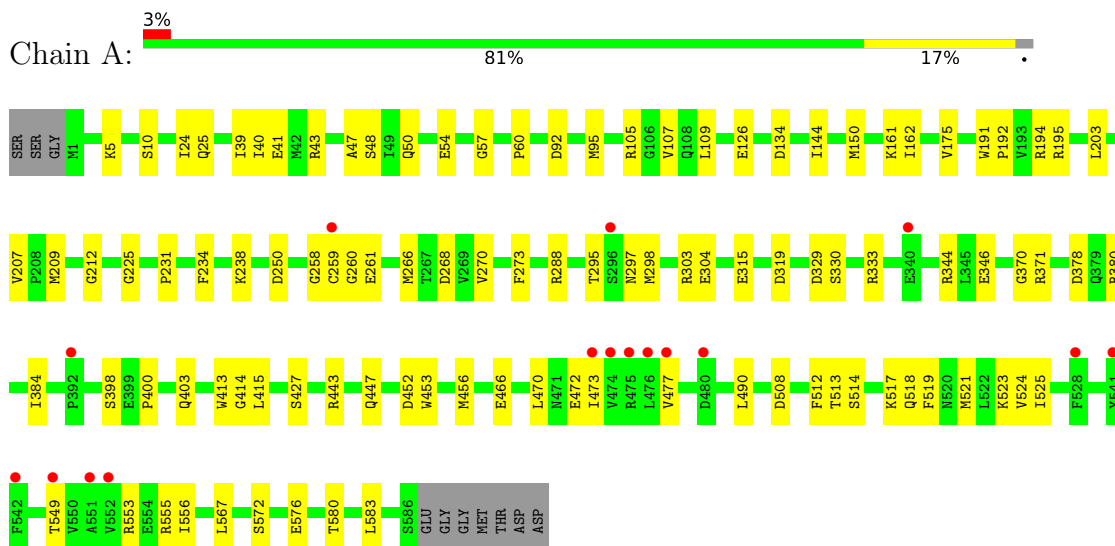
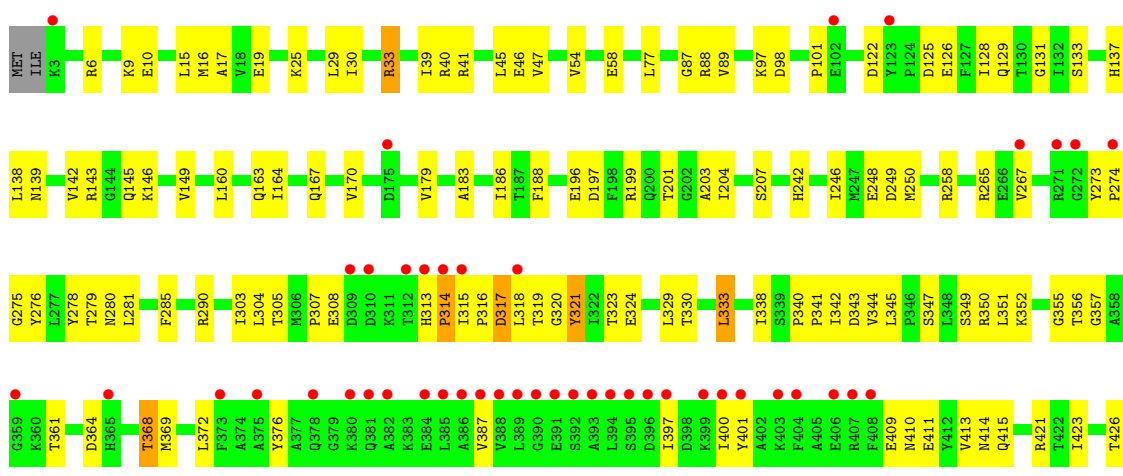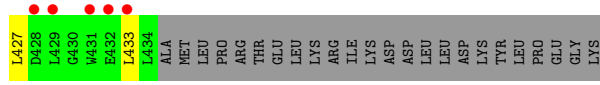• Molecule 1: V-type sodium ATPase catalytic subunit A

• Molecule 1: V-type sodium ATPase catalytic subunit A
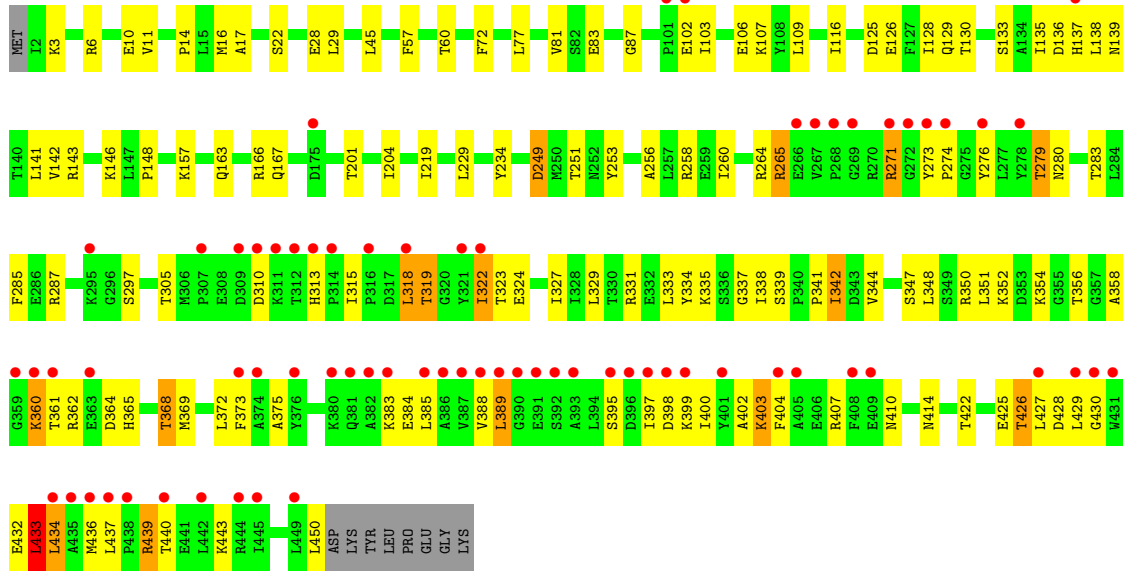
Chain C:

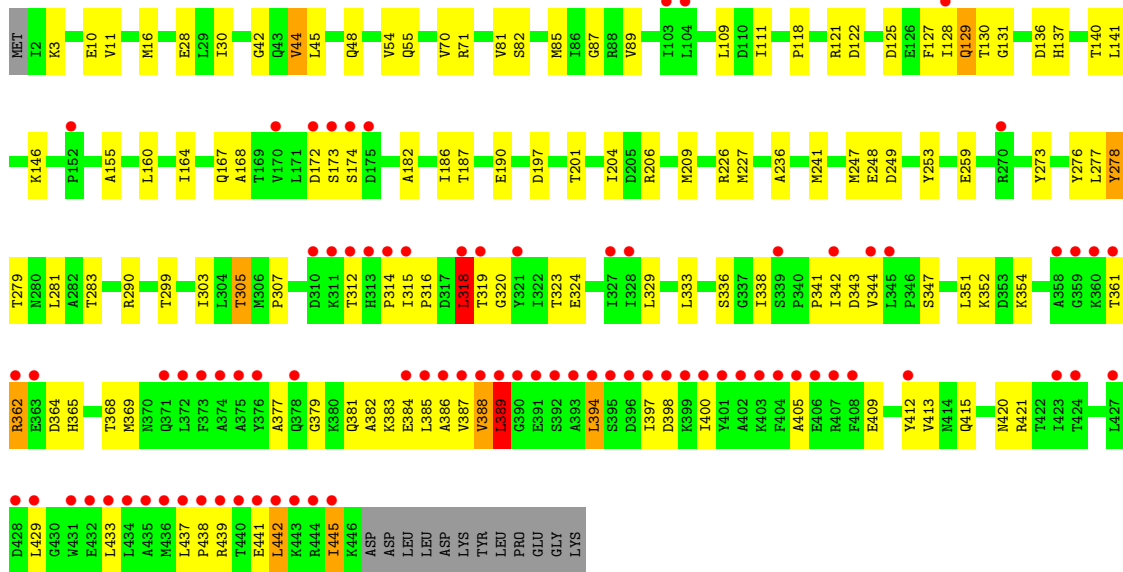• Molecule 2: V-type sodium ATPase subunit B

Chain D:
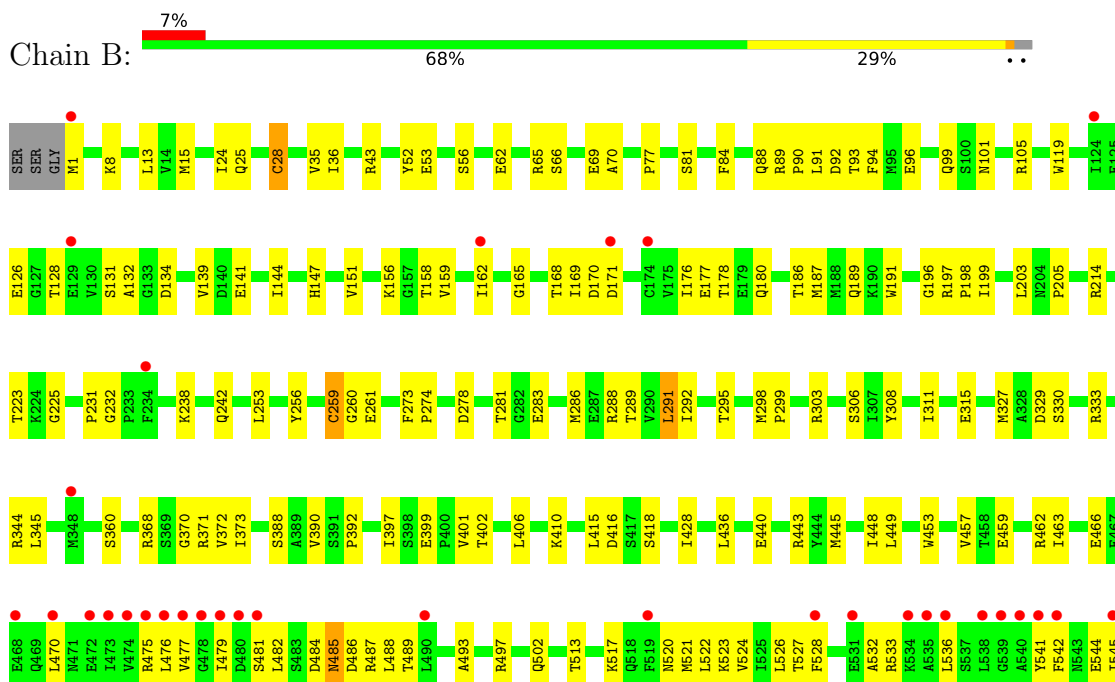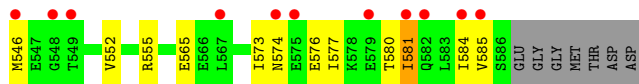
• Molecule 2: V-type sodium ATPase subunit B

• Molecule 2: V-type sodium ATPase subunit B

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	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
$< I/\sigma(I) >$ [1]	2.17 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.18_3845	Depositor
R, R_{free}	0.216 , 0.262 0.216 , 0.262	Depositor DCC
R_{free} test set	4712 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å^2)	64.5	Xtriage
Anisotropy	0.147	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 59.2	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	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	0.93	EDS
Total number of atoms	24292	wwPDB-VP
Average B, all atoms (Å^2)	74.0	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.*

---

[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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	0.33	0/4638	0.53	1/6275 (0.0%)
1	B	0.30	0/4638	0.52	0/6275
1	C	0.35	0/4636	0.59	7/6271 (0.1%)
2	D	0.31	0/3429	0.59	2/4637 (0.0%)
2	E	0.36	0/3570	0.65	4/4826 (0.1%)
2	F	0.35	0/3538	0.66	5/4782 (0.1%)
All	All	0.33	0/24449	0.59	19/33066 (0.1%)

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	388	VAL	CB-CA-C	-13.66	85.45	111.40
2	F	389	LEU	N-CA-CB	9.69	129.77	110.40
2	E	439	ARG	CB-CA-C	8.70	127.81	110.40
1	C	553	ARG	NE-CZ-NH1	-7.27	116.67	120.30
2	D	314	PRO	N-CA-C	-6.68	94.72	112.10

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	4562	0	4528	63	1
1	B	4562	0	4528	117	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	4560	0	4526	137	0
2	D	3369	0	3377	103	0
2	E	3509	0	3535	99	0
2	F	3477	0	3507	99	0
3	A	99	0	0	1	0
3	B	35	0	0	3	0
3	C	47	0	0	3	0
3	D	22	0	0	1	0
3	E	30	0	0	1	0
3	F	20	0	0	2	0
All	All	24292	0	24001	602	1

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 13.

The worst 5 of 602 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:E:333:LEU:HB3	2:E:338:ILE:HG21	1.43	0.99
2:F:388:VAL:HG12	2:F:388:VAL:O	1.63	0.94
2:E:434:LEU:HA	2:E:437:LEU:HD23	1.47	0.94
2:F:128:ILE:HD12	2:F:141:LEU:HG	1.52	0.91
1:C:570:ILE:O	1:C:573:ILE:HG12	1.71	0.90

All (1) 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)
1:A:5:LYS:NZ	1:A:508:ASP:OD2[2_656]	2.10	0.10

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	584/596 (98%)	574 (98%)	10 (2%)	0	100	100
1	B	584/596 (98%)	567 (97%)	17 (3%)	0	100	100
1	C	583/596 (98%)	567 (97%)	16 (3%)	0	100	100
2	D	430/458 (94%)	408 (95%)	22 (5%)	0	100	100
2	E	447/458 (98%)	422 (94%)	25 (6%)	0	100	100
2	F	443/458 (97%)	412 (93%)	30 (7%)	1 (0%)	47	76
All	All	3071/3162 (97%)	2950 (96%)	120 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	173	SER

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	502/509 (99%)	494 (98%)	8 (2%)	62	86
1	B	502/509 (99%)	487 (97%)	15 (3%)	41	72
1	C	501/509 (98%)	475 (95%)	26 (5%)	23	52
2	D	356/380 (94%)	343 (96%)	13 (4%)	34	65
2	E	372/380 (98%)	346 (93%)	26 (7%)	15	37
2	F	368/380 (97%)	354 (96%)	14 (4%)	33	64
All	All	2601/2667 (98%)	2499 (96%)	102 (4%)	32	63

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	333	LEU
2	E	313	HIS
2	F	433	LEU
2	D	387	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	103	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	582	GLN
2	D	280	ASN
2	D	167	GLN
2	E	163	GLN
1	B	180	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 ⓘ

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, 95th 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(Å²)	Q<0.9
1	A	586/596 (98%)	0.25	16 (2%)	54	49	32, 49, 82, 119	0
1	B	586/596 (98%)	0.43	44 (7%)	14	10	39, 69, 129, 145	0
1	C	584/596 (97%)	0.53	47 (8%)	12	8	40, 71, 122, 142	0
2	D	432/458 (94%)	0.65	50 (11%)	4	3	37, 65, 141, 166	0
2	E	449/458 (98%)	0.93	70 (15%)	2	1	33, 71, 142, 167	0
2	F	445/458 (97%)	1.00	84 (18%)	1	1	41, 79, 157, 179	0
All	All	3082/3162 (97%)	0.60	311 (10%)	7	4	32, 66, 133, 179	0

The worst 5 of 311 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	397	ILE	11.0
2	F	319	THR	9.0
2	F	387	VAL	8.2
2	E	405	ALA	7.8
2	F	404	PHE	7.7

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.