



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 09:15 am BST

PDB ID : 5W0W  
Title : Crystal structure of Protein Phosphatase 2A bound to TIPRL  
Authors : Wu, C.; Zheng, A.; Li, J.; Satyshur, K.; Xing, Y.  
Deposited on : 2017-06-01  
Resolution : 3.80 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

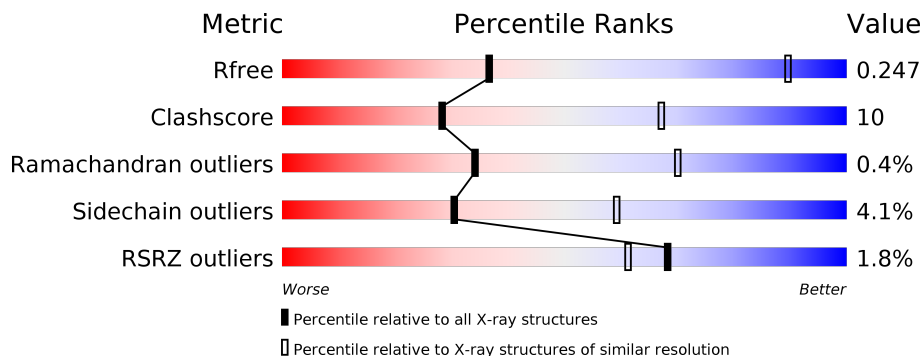
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.80 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	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)

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 on 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	585	 78% 21% ..
1	D	585	 78% 21% ..
1	G	585	 74% 25% .
1	J	585	 73% 24% ..
2	B	251	 69% 16% • 14%
2	E	251	 64% 22% 14%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	251	<p>%</p> <p>69% 16% 14%</p>
2	K	251	<p>63% 21% 14%</p>
3	C	311	<p>74% 20%</p>
3	F	311	<p>%</p> <p>65% 26% 5%</p>
3	I	311	<p>69% 25% 5%</p>
3	L	311	<p>5%</p> <p>69% 24% 5%</p>

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 34732 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 Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	582	4535	2882	764	861	14	14	0	0	0
1	D	582	4535	2882	764	861	14	14	0	0	0
1	G	584	4551	2891	768	864	14	14	0	0	0
1	J	580	4518	2870	762	858	14	14	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	GLY	-	expression tag	UNP P30153
A	6	SER	-	expression tag	UNP P30153
A	7	HIS	-	expression tag	UNP P30153
A	8	MSE	-	expression tag	UNP P30153
D	5	GLY	-	expression tag	UNP P30153
D	6	SER	-	expression tag	UNP P30153
D	7	HIS	-	expression tag	UNP P30153
D	8	MSE	-	expression tag	UNP P30153
G	5	GLY	-	expression tag	UNP P30153
G	6	SER	-	expression tag	UNP P30153
G	7	HIS	-	expression tag	UNP P30153
G	8	MSE	-	expression tag	UNP P30153
J	5	GLY	-	expression tag	UNP P30153
J	6	SER	-	expression tag	UNP P30153
J	7	HIS	-	expression tag	UNP P30153
J	8	MSE	-	expression tag	UNP P30153

- Molecule 2 is a protein called TIP41-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	216	Total 1754	C 1127	N 292	O 325	S 10	0	0	0
2	E	215	Total 1746	C 1123	N 290	O 323	S 10	0	0	0
2	H	215	Total 1743	C 1121	N 288	O 324	S 10	0	0	0
2	K	215	Total 1747	C 1123	N 288	O 326	S 10	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	9	GLY	-	expression tag	UNP Q8BH58
B	10	SER	-	expression tag	UNP Q8BH58
B	11	MET	-	expression tag	UNP Q8BH58
E	9	GLY	-	expression tag	UNP Q8BH58
E	10	SER	-	expression tag	UNP Q8BH58
E	11	MET	-	expression tag	UNP Q8BH58
H	9	GLY	-	expression tag	UNP Q8BH58
H	10	SER	-	expression tag	UNP Q8BH58
H	11	MET	-	expression tag	UNP Q8BH58
K	9	GLY	-	expression tag	UNP Q8BH58
K	10	SER	-	expression tag	UNP Q8BH58
K	11	MET	-	expression tag	UNP Q8BH58

- Molecule 3 is a protein called Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	298	Total 2424	C 1537	N 416	O 456	S 15	0	1	0
3	F	296	Total 2387	C 1513	N 407	O 452	S 15	0	0	0
3	I	296	Total 2394	C 1519	N 407	O 453	S 15	0	0	0
3	L	296	Total 2394	C 1519	N 407	O 453	S 15	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	GLY	-	expression tag	UNP P67775
C	0	SER	-	expression tag	UNP P67775

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	-1	GLY	-	expression tag	UNP P67775
F	0	SER	-	expression tag	UNP P67775
I	-1	GLY	-	expression tag	UNP P67775
I	0	SER	-	expression tag	UNP P67775
L	-1	GLY	-	expression tag	UNP P67775
L	0	SER	-	expression tag	UNP P67775

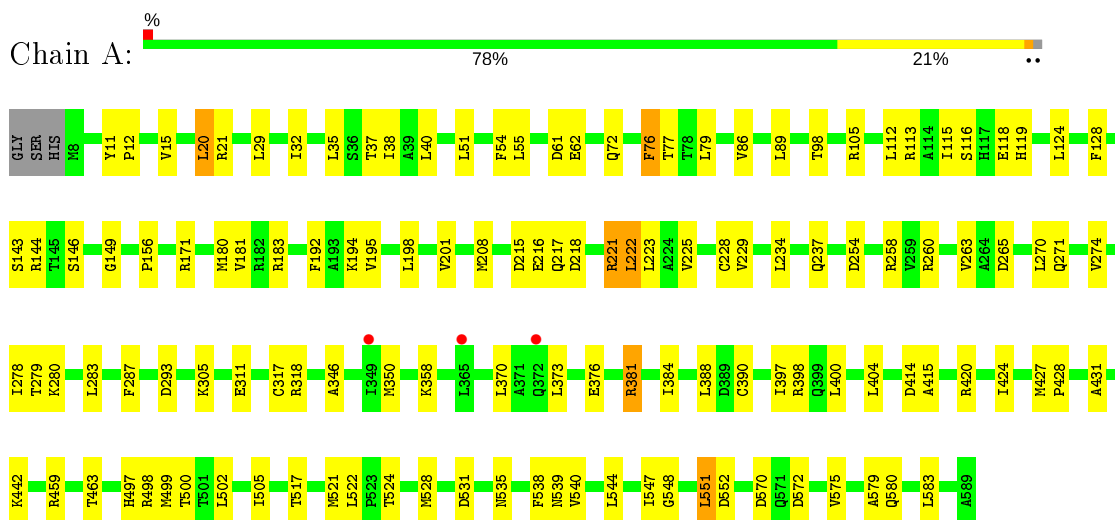
- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	I	1	Total Mn 1 1	0	0
4	L	1	Total Mn 1 1	0	0
4	C	1	Total Mn 1 1	0	0
4	F	1	Total Mn 1 1	0	0

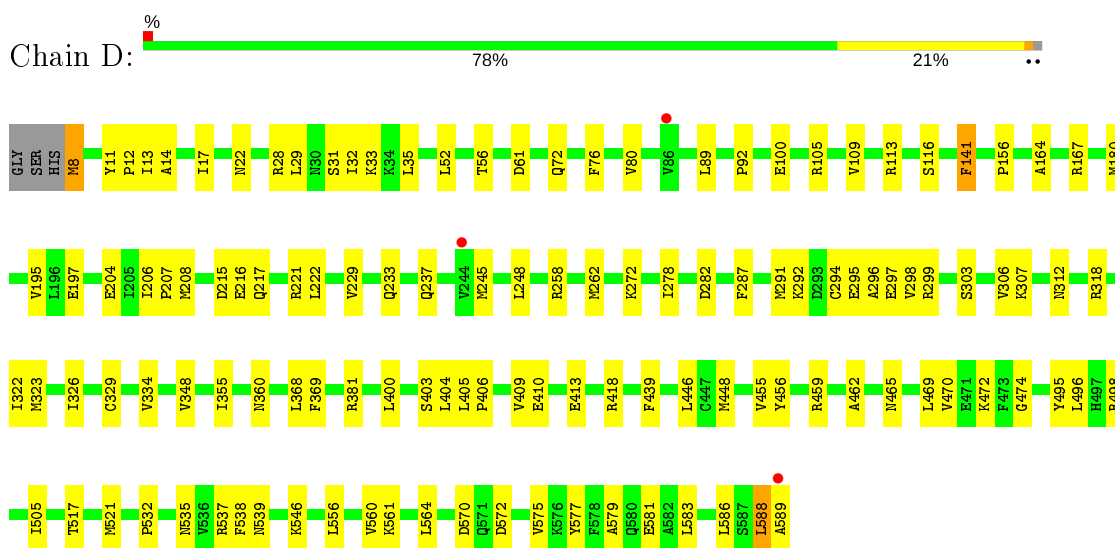
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

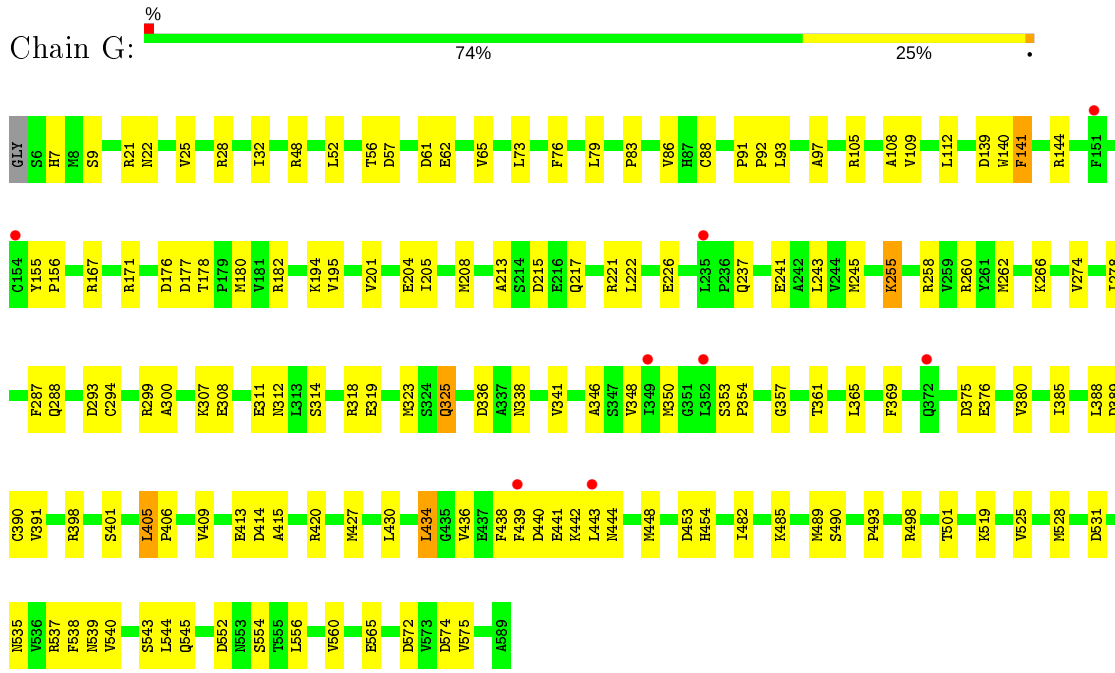
- Molecule 1: Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform



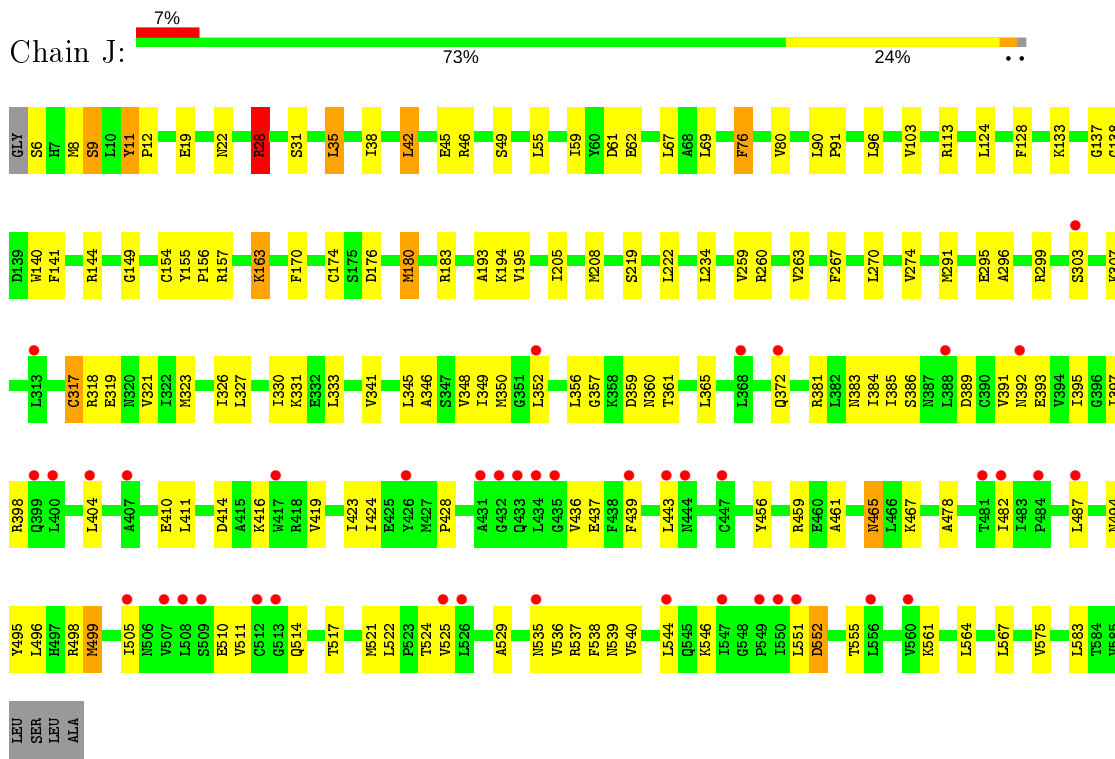
- Molecule 1: Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform



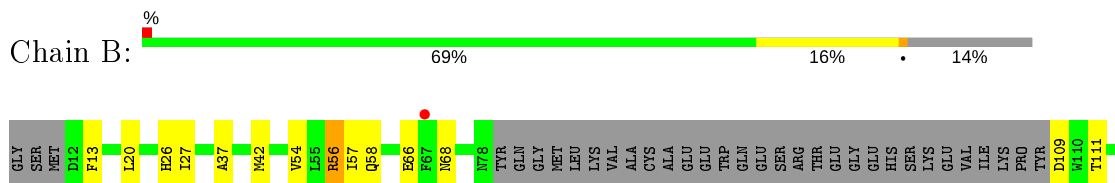
- Molecule 1: Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform



• Molecule 1: Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform



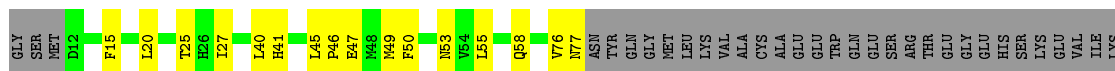
• Molecule 2: TIP41-like protein



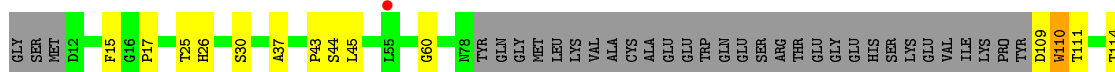




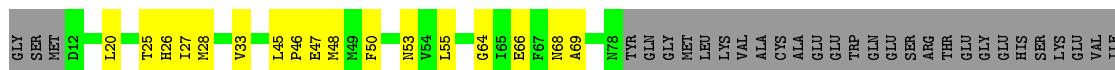
• Molecule 2: TIP41-like protein



• Molecule 2: TIP41-like protein

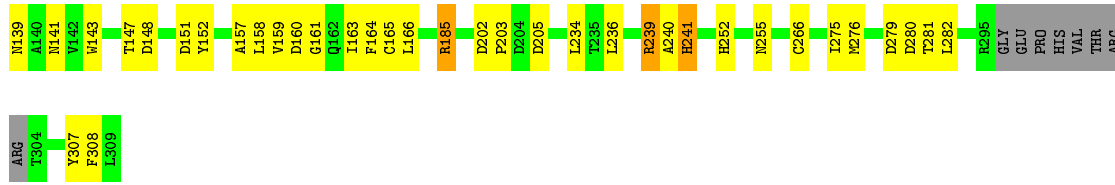


• Molecule 2: TIP41-like protein

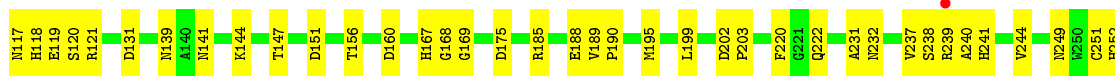
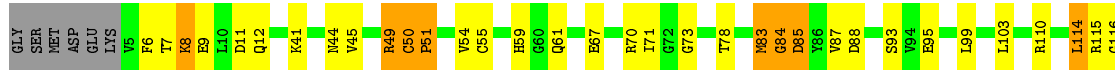


• Molecule 3: Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform

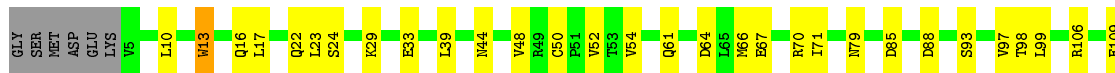




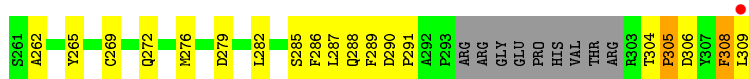
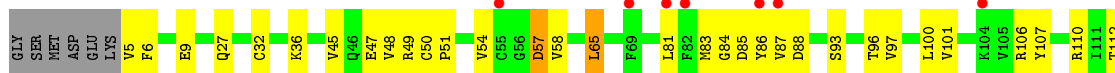
• Molecule 3: Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform



• Molecule 3: Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform



• Molecule 3: Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	150.98Å 150.98Å 285.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.42 – 3.80 49.42 – 3.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.42-3.80) 100.0 (49.42-3.60)	Depositor EDS
$R_{merge}$	0.88	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.30 (at 3.57Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, $R_{free}$	0.197 , 0.246 0.199 , 0.247	Depositor DCC
$R_{free}$ test set	5923 reflections (7.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	131.7	Xtriage
Anisotropy	0.294	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 108.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.054 for -h,-k,l 0.067 for h,-h-k,-l 0.058 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	34732	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	167.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.46% 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.25	0/4596	0.45	1/6217 (0.0%)
1	D	0.25	0/4596	0.43	0/6217
1	G	0.25	0/4612	0.44	1/6237 (0.0%)
1	J	0.27	0/4576	0.48	1/6186 (0.0%)
2	B	0.25	0/1794	0.47	0/2422
2	E	0.25	0/1786	0.51	0/2411
2	H	0.25	0/1783	0.48	0/2408
2	K	0.26	0/1787	0.50	0/2413
3	C	0.25	0/2484	0.46	0/3366
3	F	0.27	0/2445	0.49	0/3315
3	I	0.26	0/2453	0.47	0/3326
3	L	0.26	0/2453	0.50	0/3326
All	All	0.26	0/35365	0.47	3/47844 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	28	ARG	NE-CZ-NH1	5.41	123.01	120.30
1	G	405	LEU	CA-CB-CG	5.33	127.57	115.30
1	A	20	LEU	CB-CG-CD2	-5.17	102.22	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	4535	0	4642	72	0
1	D	4535	0	4642	84	0
1	G	4551	0	4652	98	0
1	J	4518	0	4615	110	0
2	B	1754	0	1745	23	0
2	E	1746	0	1739	38	0
2	H	1743	0	1732	27	0
2	K	1747	0	1736	40	0
3	C	2424	0	2324	45	0
3	F	2387	0	2285	68	0
3	I	2394	0	2292	51	0
3	L	2394	0	2292	59	0
4	C	1	0	0	0	0
4	F	1	0	0	0	0
4	I	1	0	0	0	0
4	L	1	0	0	0	0
All	All	34732	0	34696	683	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 10.

The worst 5 of 683 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:G:319:GLU:O	1:G:323:MSE:CE	1.94	1.15
3:I:54:VAL:HG12	3:I:276:MET:HG2	1.41	1.00
1:A:570:ASP:HB3	1:A:575:VAL:HG21	1.47	0.95
1:J:22:ASN:O	1:J:28:ARG:NH2	2.01	0.93
3:C:48:VAL:CG1	3:C:159:VAL:HG12	2.03	0.89

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	580/585 (99%)	555 (96%)	25 (4%)	0	100	100
1	D	580/585 (99%)	552 (95%)	27 (5%)	1 (0%)	47	79
1	G	582/585 (100%)	554 (95%)	28 (5%)	0	100	100
1	J	574/585 (98%)	532 (93%)	42 (7%)	0	100	100
2	B	212/251 (84%)	198 (93%)	13 (6%)	1 (0%)	29	66
2	E	211/251 (84%)	192 (91%)	17 (8%)	2 (1%)	17	54
2	H	211/251 (84%)	191 (90%)	18 (8%)	2 (1%)	17	54
2	K	211/251 (84%)	193 (92%)	16 (8%)	2 (1%)	17	54
3	C	295/311 (95%)	269 (91%)	25 (8%)	1 (0%)	41	74
3	F	292/311 (94%)	256 (88%)	30 (10%)	6 (2%)	7	40
3	I	292/311 (94%)	261 (89%)	29 (10%)	2 (1%)	22	60
3	L	292/311 (94%)	258 (88%)	32 (11%)	2 (1%)	22	60
All	All	4332/4588 (94%)	4011 (93%)	302 (7%)	19 (0%)	34	70

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	208	THR
2	E	208	THR
3	F	7	THR
2	H	208	THR
3	C	6	PHE

### 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	509/497 (102%)	483 (95%)	26 (5%)	24	54
1	D	509/497 (102%)	493 (97%)	16 (3%)	40	65

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	511/497 (103%)	494 (97%)	17 (3%)	38	65
1	J	507/497 (102%)	485 (96%)	22 (4%)	29	58
2	B	197/228 (86%)	192 (98%)	5 (2%)	47	70
2	E	196/228 (86%)	192 (98%)	4 (2%)	55	75
2	H	196/228 (86%)	190 (97%)	6 (3%)	40	65
2	K	197/228 (86%)	191 (97%)	6 (3%)	41	66
3	C	265/275 (96%)	252 (95%)	13 (5%)	25	55
3	F	261/275 (95%)	243 (93%)	18 (7%)	15	46
3	I	262/275 (95%)	252 (96%)	10 (4%)	33	61
3	L	262/275 (95%)	247 (94%)	15 (6%)	20	52
All	All	3872/4000 (97%)	3714 (96%)	158 (4%)	30	59

5 of 158 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	114	LEU
1	G	255	LYS
3	L	106	ARG
3	F	131	ASP
3	F	289	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	241	HIS
3	F	264	ASN
1	J	514	GLN
1	D	465	ASN
3	F	222	GLN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are 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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	J	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	J	552:ASP	C	553:ASN	N	5.66
1	J	337:ALA	C	338:ASN	N	4.02



## 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	568/585 (97%)	-0.14	3 (0%) 91 87	101, 163, 207, 237	0
1	D	568/585 (97%)	-0.25	3 (0%) 91 87	113, 161, 197, 232	0
1	G	570/585 (97%)	-0.18	8 (1%) 75 68	115, 179, 228, 250	0
1	J	566/585 (96%)	0.19	42 (7%) 14 11	104, 214, 308, 366	0
2	B	216/251 (86%)	-0.16	2 (0%) 84 79	119, 165, 210, 228	0
2	E	215/251 (85%)	-0.08	3 (1%) 75 68	108, 159, 224, 271	0
2	H	215/251 (85%)	-0.10	2 (0%) 84 79	104, 148, 208, 259	0
2	K	215/251 (85%)	-0.17	0 100 100	109, 158, 208, 226	0
3	C	298/311 (95%)	-0.20	0 100 100	87, 121, 163, 232	0
3	F	296/311 (95%)	-0.19	2 (0%) 87 83	107, 150, 192, 236	0
3	I	296/311 (95%)	-0.16	0 100 100	77, 120, 165, 203	0
3	L	296/311 (95%)	0.13	14 (4%) 31 27	157, 203, 241, 276	0
All	All	4319/4588 (94%)	-0.10	79 (1%) 68 61	77, 161, 248, 366	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	551	LEU	5.7
1	J	560	VAL	5.6
1	J	512	CYS	4.9
1	J	482	ILE	4.7
1	J	433	GLN	4.0

### 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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	MN	F	501	1/1	0.31	0.20	181,181,181,181	0
4	MN	L	501	1/1	0.31	0.27	225,225,225,225	0
4	MN	I	501	1/1	0.47	0.33	157,157,157,157	0
4	MN	C	501	1/1	0.54	0.39	204,204,204,204	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.