



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 2, 2017 – 02:57 AM EDT

PDB ID : 4I5L  
Title : Structural mechanism of trimeric PP2A holoenzyme involving PR70: insight for Cdc6 dephosphorylation  
Authors : Wlodarchak, N.; Satyshur, K.A.; Guo, F.; Xing, Y.  
Deposited on : unknown  
Resolution : 2.43 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

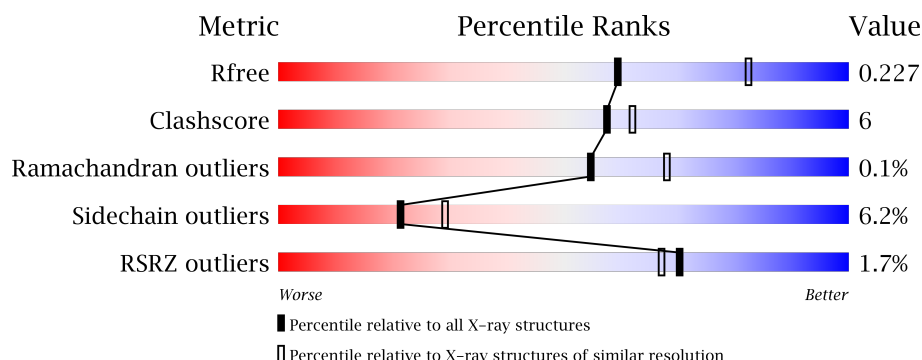
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.43 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))
$R_{free}$	100719	1152 (2.46-2.42)
Clashscore	112137	1224 (2.46-2.42)
Ramachandran outliers	110173	1217 (2.46-2.42)
Sidechain outliers	110143	1217 (2.46-2.42)
RSRZ outliers	101464	1158 (2.46-2.42)

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. 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	584	<div> <div>5%</div> <div>86% 12% .</div> </div>
1	D	584	<div> <div>5%</div> <div>81% 17% .</div> </div>
2	B	413	<div> <div>70% 13% . 14%</div> </div>
2	E	413	<div> <div>75% 9% . 14%</div> </div>
3	C	311	<div> <div>79% 13% . 5%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	311	 81% 10% 5%
4	G	7	 71% 29%
4	H	7	 71% 29%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	PEG	A	602	-	-	-	X
5	PEG	A	604	-	-	-	X
5	PEG	B	603	-	-	-	X
5	PEG	B	604	-	-	-	X
5	PEG	C	404	-	-	-	X
5	PEG	D	602	-	-	-	X
5	PEG	E	703	-	-	-	X
8	MLI	C	405	-	-	-	X
8	MLI	E	704	-	-	-	X
8	MLI	F	403	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 20426 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
1	A	582	Total	C	N	O	S	0	0	0
			4534	2882	764	860	28			
1	D	582	Total	C	N	O	S	0	0	0
			4534	2882	764	860	28			

- Molecule 2 is a protein called Serine/threonine-protein phosphatase 2A regulatory subunit B" subunit beta - Cell division control protein 6 homolog chimeric construct.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	356	Total	C	N	O	S	0	0	0
			2905	1866	486	533	20			
2	E	357	Total	C	N	O	S	0	0	0
			2912	1870	488	534	20			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	491	SER	-	LINKER	UNP Q9Y5P8
B	492	THR	-	LINKER	UNP Q9Y5P8
B	493	GLY	-	LINKER	UNP Q9Y5P8
B	494	ASN	-	LINKER	UNP Q9Y5P8
B	495	ALA	-	LINKER	UNP Q9Y5P8
B	496	SER	-	LINKER	UNP Q9Y5P8
B	497	ASP	-	LINKER	UNP Q9Y5P8
B	498	SER	-	LINKER	UNP Q9Y5P8
B	499	SER	-	LINKER	UNP Q9Y5P8
B	500	SER	-	LINKER	UNP Q9Y5P8
B	501	ASP	-	LINKER	UNP Q9Y5P8
B	502	SER	-	LINKER	UNP Q9Y5P8
B	503	SER	-	LINKER	UNP Q9Y5P8
B	504	SER	-	LINKER	UNP Q9Y5P8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	505	SER	-	LINKER	UNP Q9Y5P8
B	506	GLU	-	LINKER	UNP Q9Y5P8
B	507	GLY	-	LINKER	UNP Q9Y5P8
B	508	ASP	-	LINKER	UNP Q9Y5P8
B	509	GLY	-	LINKER	UNP Q9Y5P8
B	510	THR	-	LINKER	UNP Q9Y5P8
B	511	VAL	-	LINKER	UNP Q9Y5P8
E	491	SER	-	LINKER	UNP Q99741
E	492	THR	-	LINKER	UNP Q99741
E	493	GLY	-	LINKER	UNP Q99741
E	494	ASN	-	LINKER	UNP Q99741
E	495	ALA	-	LINKER	UNP Q99741
E	496	SER	-	LINKER	UNP Q99741
E	497	ASP	-	LINKER	UNP Q99741
E	498	SER	-	LINKER	UNP Q99741
E	499	SER	-	LINKER	UNP Q99741
E	500	SER	-	LINKER	UNP Q99741
E	501	ASP	-	LINKER	UNP Q99741
E	502	SER	-	LINKER	UNP Q99741
E	503	SER	-	LINKER	UNP Q99741
E	504	SER	-	LINKER	UNP Q99741
E	505	SER	-	LINKER	UNP Q99741
E	506	GLU	-	LINKER	UNP Q99741
E	507	GLY	-	LINKER	UNP Q99741
E	508	ASP	-	LINKER	UNP Q99741
E	509	GLY	-	LINKER	UNP Q99741
E	510	THR	-	LINKER	UNP Q99741
E	511	VAL	-	LINKER	UNP Q99741

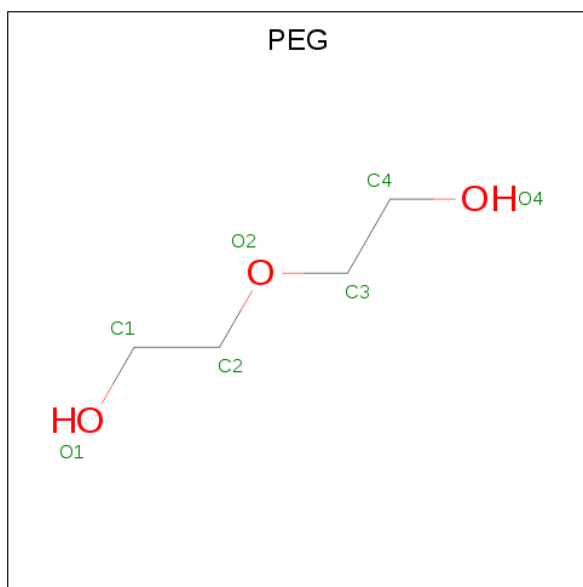
- Molecule 3 is a protein called Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform, PP2A-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	295	Total	C	N	O	S	0	0	0
			2381	1505	410	451	15			
3	F	295	Total	C	N	O	S	0	0	0
			2381	1505	410	451	15			

- Molecule 4 is a protein called Microcystin-LR (MCLR) bound form.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	G	7	Total	C	N	O	0	0	0
			71	49	10	12			
4	H	7	Total	C	N	O	0	0	0
			71	49	10	12			

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		
5	A	1	Total	C	O	0	0
			7	4	3		
5	A	1	Total	C	O	0	0
			7	4	3		
5	A	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			7	4	3		
5	C	1	Total	C	O	0	0
			7	4	3		
5	C	1	Total	C	O	0	0
			7	4	3		
5	D	1	Total	C	O	0	0
			7	4	3		
5	D	1	Total	C	O	0	0
			7	4	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total	C	O	0	0
			7	4	3		

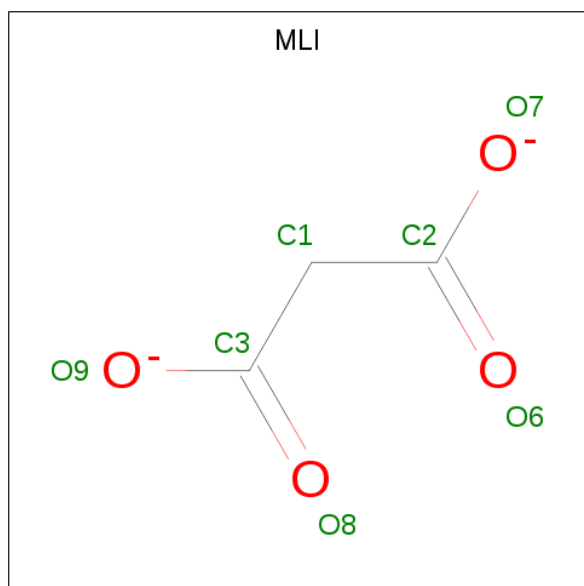
- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	2	Total	Ca	0	0
			2	2		
6	E	2	Total	Ca	0	0
			2	2		

- Molecule 7 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	2	Total	Mn	0	0
			2	2		
7	F	2	Total	Mn	0	0
			2	2		

- Molecule 8 is MALONATE ION (three-letter code: MLI) (formula: C<sub>3</sub>H<sub>2</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total	C	O	0	0
			7	3	4		
8	E	1	Total	C	O	0	0
			7	3	4		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	F	1	Total	C	O	0	0
			7	3	4		

- Molecule 9 is water.

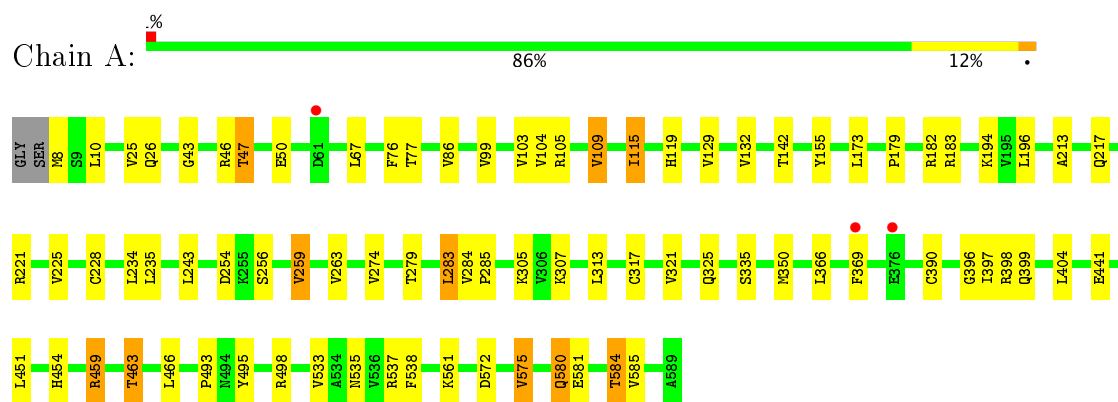
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	152	Total	O	0	0
			152	152		
9	B	82	Total	O	0	0
			82	82		
9	C	79	Total	O	0	0
			79	79		
9	D	70	Total	O	0	0
			70	70		
9	E	73	Total	O	0	0
			73	73		
9	F	71	Total	O	0	0
			71	71		
9	H	4	Total	O	0	0
			4	4		



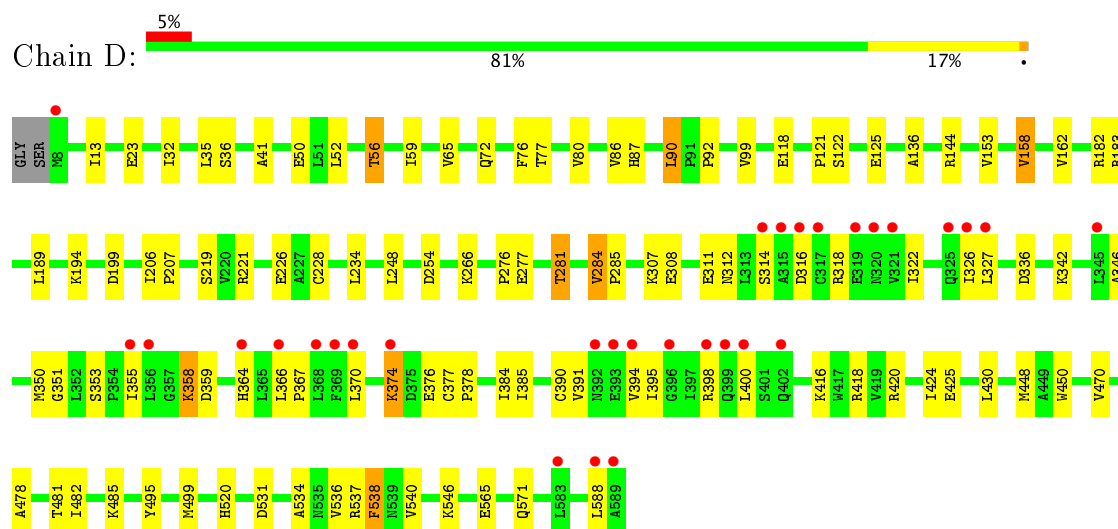
### 3 Residue-property plots [i](#)

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

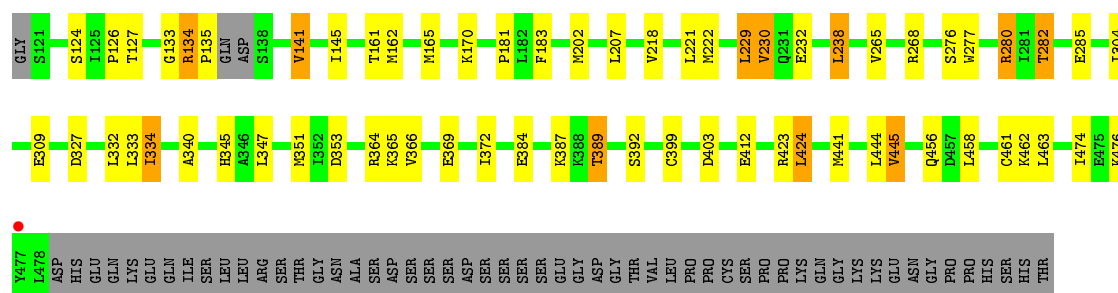


- Molecule 1: Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform



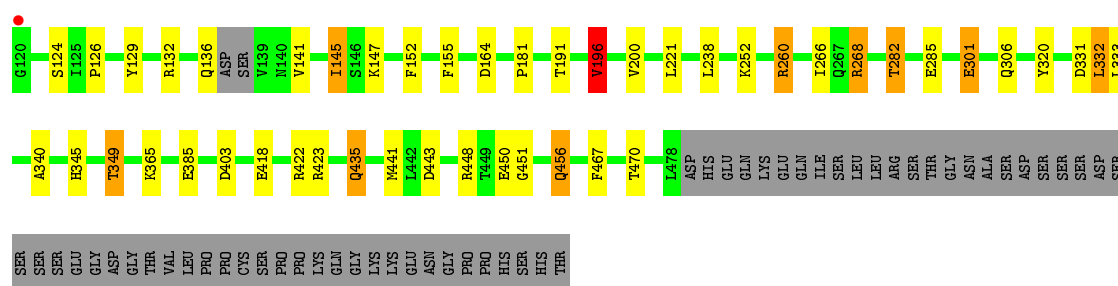
- Molecule 2: Serine/threonine-protein phosphatase 2A regulatory subunit B" subunit beta - Cell division control protein 6 homolog chimeric construct





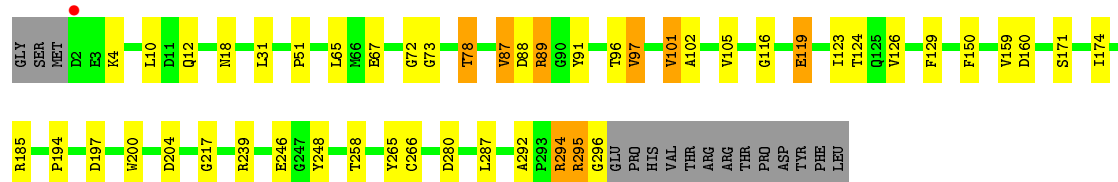
- Molecule 2: Serine/threonine-protein phosphatase 2A regulatory subunit B" subunit beta - Cell division control protein 6 homolog chimeric construct

Chain E: 75% 9% 14%



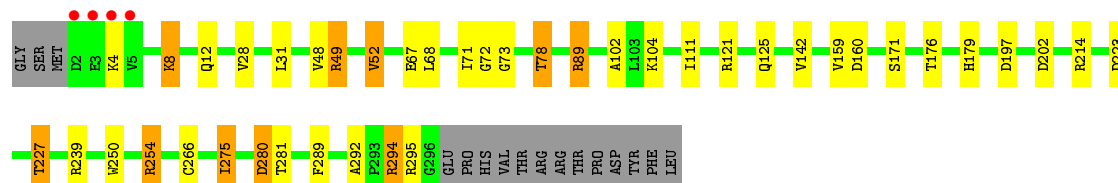
- Molecule 3: Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform, PP2A-alpha

Chain C: 79% 13% 5%



- Molecule 3: Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform, PP2A-alpha

Chain F: 81% 10% 5%



- Molecule 4: Microcystin-LR (MCLR) bound form

Chain G: 71% 29%



- Molecule 4: Microcystin-LR (MCLR) bound form

Chain H:  71% 29%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.23Å 101.07Å 343.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.85 – 2.43 49.80 – 2.43	Depositor EDS
% Data completeness (in resolution range)	98.6 (49.85-2.43) 98.6 (49.80-2.43)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.179 , 0.228 0.178 , 0.227	Depositor DCC
$R_{free}$ test set	6192 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.0	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 38.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	20426	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.17% 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: ACB, MLI, DAL, CA, MN, MAA, 1ZN, PEG, FGA

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	0.45	0/4608	0.69	4/6256 (0.1%)
1	D	0.39	0/4608	0.60	2/6256 (0.0%)
2	B	0.49	0/2979	0.69	1/4025 (0.0%)
2	E	0.44	0/2986	0.62	1/4034 (0.0%)
3	C	0.48	0/2438	0.70	1/3304 (0.0%)
3	F	0.43	0/2438	0.66	1/3304 (0.0%)
4	G	0.34	0/17	0.63	0/19
4	H	0.33	0/17	0.65	0/19
All	All	0.44	0/20091	0.66	10/27217 (0.0%)

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	459	ARG	NE-CZ-NH2	-10.75	114.93	120.30
2	B	412	GLU	OE1-CD-OE2	9.14	134.27	123.30
1	A	537	ARG	NE-CZ-NH2	-7.61	116.49	120.30
1	A	182	ARG	NE-CZ-NH2	-6.29	117.16	120.30
1	A	459	ARG	NE-CZ-NH1	6.25	123.42	120.30

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	4534	0	4642	39	0
1	D	4534	0	4642	52	0
2	B	2905	0	2816	43	0
2	E	2912	0	2822	39	0
3	C	2381	0	2284	30	0
3	F	2381	0	2284	24	0
4	G	71	0	68	2	0
4	H	71	0	69	2	0
5	A	28	0	40	0	0
5	B	14	0	20	1	0
5	C	14	0	20	3	0
5	D	14	0	20	0	0
5	E	7	0	10	0	0
6	B	2	0	0	0	0
6	E	2	0	0	0	0
7	C	2	0	0	0	0
7	F	2	0	0	0	0
8	C	7	0	2	1	0
8	E	7	0	2	0	0
8	F	7	0	2	0	0
9	A	152	0	0	1	0
9	B	82	0	0	1	0
9	C	79	0	0	0	0
9	D	70	0	0	1	0
9	E	73	0	0	1	0
9	F	71	0	0	1	0
9	H	4	0	0	0	0
All	All	20426	0	19743	220	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 6.

The worst 5 of 220 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:353:SER:HB3	1:D:394:VAL:HG11	1.52	0.90
2:E:268:ARG:HH11	2:E:268:ARG:HG2	1.36	0.90
2:B:389:THR:HG22	2:B:392:SER:H	1.37	0.88
1:D:77:THR:HG21	1:D:118:GLU:HG2	1.57	0.85
1:D:358:LYS:HD2	1:D:359:ASP:H	1.42	0.83

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	580/584 (99%)	571 (98%)	9 (2%)	0	100	100
1	D	580/584 (99%)	571 (98%)	9 (2%)	0	100	100
2	B	352/413 (85%)	344 (98%)	7 (2%)	1 (0%)	44	54
2	E	353/413 (86%)	343 (97%)	10 (3%)	0	100	100
3	C	293/311 (94%)	282 (96%)	10 (3%)	1 (0%)	44	54
3	F	293/311 (94%)	281 (96%)	12 (4%)	0	100	100
4	G	1/7 (14%)	1 (100%)	0	0	100	100
4	H	1/7 (14%)	1 (100%)	0	0	100	100
All	All	2453/2630 (93%)	2394 (98%)	57 (2%)	2 (0%)	55	68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	134	ARG
3	C	295	ARG

### 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	509/510 (100%)	479 (94%)	30 (6%)	23	31
1	D	509/510 (100%)	479 (94%)	30 (6%)	23	31
2	B	316/366 (86%)	292 (92%)	24 (8%)	15	20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	316/366 (86%)	298 (94%)	18 (6%)	24	32
3	C	260/275 (94%)	243 (94%)	17 (6%)	20	26
3	F	260/275 (94%)	245 (94%)	15 (6%)	23	32
4	G	2/2 (100%)	2 (100%)	0	100	100
4	H	2/2 (100%)	2 (100%)	0	100	100
All	All	2174/2306 (94%)	2040 (94%)	134 (6%)	21	29

5 of 134 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	101	VAL
1	D	76	PHE
3	F	52	VAL
3	C	126	VAL
3	C	280	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	GLN
3	C	122	GLN
2	E	435	GLN
3	F	141	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

10 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	DAL	G	1	4	4,4,5	1.00	0	1,4,6	0.14	0
4	ACB	G	3	4	5,8,9	2.11	1 (20%)	3,10,12	3.66	1 (33%)
4	1ZN	G	5	4	22,23,24	2.35	4 (18%)	24,29,31	2.11	4 (16%)
4	FGA	G	6	4	3,8,9	1.99	1 (33%)	2,9,11	0.93	0
4	MAA	G	7	3,4	5,5,6	0.82	0	1,5,7	0.47	0
4	DAL	H	1	4	4,4,5	1.17	1 (25%)	1,4,6	0.84	0
4	ACB	H	3	4	5,8,9	2.12	1 (20%)	3,10,12	4.16	2 (66%)
4	1ZN	H	5	4	22,23,24	2.59	5 (22%)	24,29,31	2.24	6 (25%)
4	FGA	H	6	4	3,8,9	2.10	1 (33%)	2,9,11	0.78	0
4	MAA	H	7	3,4	5,5,6	1.18	1 (20%)	1,5,7	0.91	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	DAL	G	1	4	-	0/0/2/4	0/0/0/0
4	ACB	G	3	4	-	0/6/10/12	0/0/0/0
4	1ZN	G	5	4	-	0/22/25/27	0/1/1/1
4	FGA	G	6	4	-	0/4/8/9	0/0/0/0
4	MAA	G	7	3,4	-	0/1/4/6	0/0/0/0
4	DAL	H	1	4	-	0/0/2/4	0/0/0/0
4	ACB	H	3	4	-	0/6/10/12	0/0/0/0
4	1ZN	H	5	4	-	0/22/25/27	0/1/1/1
4	FGA	H	6	4	-	0/4/8/9	0/0/0/0
4	MAA	H	7	3,4	-	0/1/4/6	0/0/0/0

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	3	ACB	OD2-CG	-4.47	1.23	1.42
4	G	3	ACB	OD2-CG	-4.43	1.23	1.42
4	G	5	1ZN	C3-C4	-4.22	1.41	1.51
4	H	5	1ZN	C3-C4	-4.16	1.41	1.51
4	H	6	FGA	OE2-CD	-3.63	1.22	1.42

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	5	1ZN	C17-C16-C15	-8.21	110.51	123.70
4	H	5	1ZN	C17-C16-C15	-8.18	110.55	123.70
4	H	5	1ZN	C19-C18-C17	-3.71	106.96	110.92
4	G	5	1ZN	C10-C12-C13	-3.29	116.15	126.68
4	G	5	1ZN	C3-C2-C10	-3.21	107.71	115.42

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	6	FGA	2	0
4	G	7	MAA	2	0
4	H	6	FGA	2	0
4	H	7	MAA	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 8 are monoatomic - leaving 14 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	PEG	A	601	-	6,6,6	0.44	0	5,5,5	0.45	0
5	PEG	A	602	-	6,6,6	0.48	0	5,5,5	0.33	0
5	PEG	A	603	-	6,6,6	0.52	0	5,5,5	0.26	0
5	PEG	A	604	-	6,6,6	0.59	0	5,5,5	0.26	0
5	PEG	B	603	-	6,6,6	0.48	0	5,5,5	0.42	0
5	PEG	B	604	-	6,6,6	0.51	0	5,5,5	0.31	0
5	PEG	C	403	-	6,6,6	0.59	0	5,5,5	0.37	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	PEG	C	404	-	6,6,6	0.47	0	5,5,5	0.33	0
8	MLI	C	405	-	0,6,6	0.00	-	0,7,7	0.00	-
5	PEG	D	601	-	6,6,6	0.48	0	5,5,5	0.30	0
5	PEG	D	602	-	6,6,6	0.44	0	5,5,5	0.43	0
5	PEG	E	703	-	6,6,6	0.54	0	5,5,5	0.32	0
8	MLI	E	704	-	0,6,6	0.00	-	0,7,7	0.00	-
8	MLI	F	403	-	0,6,6	0.00	-	0,7,7	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PEG	A	601	-	-	0/4/4/4	0/0/0/0
5	PEG	A	602	-	-	0/4/4/4	0/0/0/0
5	PEG	A	603	-	-	0/4/4/4	0/0/0/0
5	PEG	A	604	-	-	0/4/4/4	0/0/0/0
5	PEG	B	603	-	-	0/4/4/4	0/0/0/0
5	PEG	B	604	-	-	0/4/4/4	0/0/0/0
5	PEG	C	403	-	-	0/4/4/4	0/0/0/0
5	PEG	C	404	-	-	0/4/4/4	0/0/0/0
8	MLI	C	405	-	-	0/0/4/4	0/0/0/0
5	PEG	D	601	-	-	0/4/4/4	0/0/0/0
5	PEG	D	602	-	-	0/4/4/4	0/0/0/0
5	PEG	E	703	-	-	0/4/4/4	0/0/0/0
8	MLI	E	704	-	-	0/0/4/4	0/0/0/0
8	MLI	F	403	-	-	0/0/4/4	0/0/0/0

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.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	604	PEG	1	0
5	C	403	PEG	3	0
8	C	405	MLI	1	0

## 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<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	582/584 (99%)	-0.31	3 (0%) 90 91	20, 36, 59, 92	0
1	D	582/584 (99%)	0.03	31 (5%) 27 24	30, 49, 81, 107	0
2	B	356/413 (86%)	-0.25	1 (0%) 93 94	21, 37, 70, 92	0
2	E	357/413 (86%)	-0.20	1 (0%) 93 94	26, 40, 76, 92	0
3	C	295/311 (94%)	-0.39	1 (0%) 93 94	18, 34, 51, 98	0
3	F	295/311 (94%)	-0.30	4 (1%) 75 74	26, 41, 65, 115	0
4	G	2/7 (28%)	0.71	0 100 100	64, 64, 64, 75	0
4	H	2/7 (28%)	0.05	0 100 100	54, 54, 54, 64	0
All	All	2471/2630 (93%)	-0.21	41 (1%) 70 67	18, 40, 73, 115	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	589	ALA	9.8
3	F	2	ASP	4.1
1	D	398	ARG	3.5
1	D	588	LEU	3.3
1	D	319	GLU	3.3

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ACB	G	3	9/10	0.96	0.20	-	52,55,57,60	0
4	MAA	H	7	6/7	0.97	0.13	-	38,44,46,47	0
4	1ZN	H	5	23/24	0.95	0.16	-	34,37,42,44	0
4	1ZN	G	5	23/24	0.95	0.16	-	38,42,45,51	0
4	DAL	G	1	5/6	0.90	0.21	-	54,54,58,58	0
4	ACB	H	3	9/10	0.96	0.14	-	44,46,47,49	0
4	DAL	H	1	5/6	0.95	0.16	-	50,50,55,56	0
4	MAA	G	7	6/7	0.95	0.11	-	46,48,49,51	0
4	FGA	G	6	9/10	0.97	0.15	-	42,47,50,50	0
4	FGA	H	6	9/10	0.95	0.11	-	38,39,43,43	0

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	PEG	C	404	7/7	0.75	0.25	6.25	62,70,75,78	0
8	MLI	F	403	7/7	0.91	0.25	5.78	67,69,73,73	0
5	PEG	A	604	7/7	0.70	0.35	5.27	64,66,71,76	0
5	PEG	D	602	7/7	0.81	0.22	5.19	72,73,76,78	0
8	MLI	C	405	7/7	0.81	0.20	4.68	47,53,56,59	0
5	PEG	E	703	7/7	0.84	0.25	4.18	59,64,65,66	0
8	MLI	E	704	7/7	0.89	0.24	4.08	61,63,66,66	0
5	PEG	A	602	7/7	0.81	0.24	4.00	57,65,79,80	0
5	PEG	B	604	7/7	0.86	0.26	3.69	62,63,65,68	0
5	PEG	B	603	7/7	0.92	0.21	3.03	67,67,72,72	0
5	PEG	D	601	7/7	0.72	0.18	1.42	65,76,85,87	0
5	PEG	C	403	7/7	0.88	0.15	1.23	41,45,53,56	0
5	PEG	A	601	7/7	0.86	0.20	0.07	71,72,76,76	0
5	PEG	A	603	7/7	0.91	0.13	-0.22	62,65,67,68	0
6	CA	B	602	1/1	0.91	0.07	-2.64	43,43,43,43	0
6	CA	E	701	1/1	0.97	0.07	-3.08	27,27,27,27	0
7	MN	C	401	1/1	0.99	0.08	-3.47	34,34,34,34	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	MN	C	402	1/1	0.99	0.05	-5.57	40,40,40,40	0
6	CA	E	702	1/1	0.88	0.06	-6.02	48,48,48,48	0
7	MN	F	402	1/1	0.98	0.06	-6.31	43,43,43,43	0
7	MN	F	401	1/1	0.99	0.06	-6.41	34,34,34,34	0
6	CA	B	601	1/1	0.99	0.06	-9.20	22,22,22,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.