



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 13, 2020 – 11:00 pm BST

PDB ID : 6GHM  
Title : Structure of PP1 alpha phosphatase bound to ASPP2  
Authors : Mouilleron, S.; Bertran, T.M.; Tapon, N.; Zhou, Y.  
Deposited on : 2018-05-08  
Resolution : 2.15 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](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

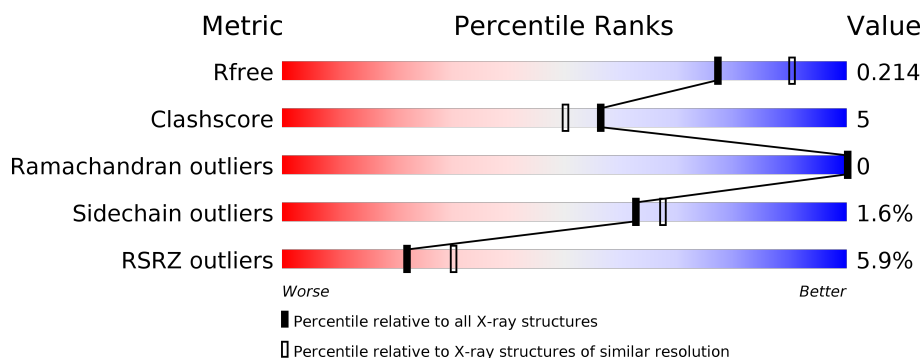
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	329	<div> <div>3%</div> <div> <div></div> <div>80%</div> <div>10%</div> <div>10%</div> </div> </div>
1	B	329	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>9%</div> <div>11%</div> </div> </div>
2	C	214	<div> <div>12%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>6%</div> </div> </div>
2	D	214	<div> <div>9%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>5%</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NHE	A	411[A]	-	-	-	X

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 8259 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 PP1-alpha catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	297	Total	C	N	O	S	0	2	0
			2376	1526	397	434	19			
1	B	294	Total	C	N	O	S	0	5	0
			2367	1526	395	427	19			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLY	-	expression tag	UNP P62136
A	3	HIS	-	expression tag	UNP P62136
A	4	MET	-	expression tag	UNP P62136
A	5	GLY	-	expression tag	UNP P62136
A	6	SER	-	expression tag	UNP P62136
B	2	GLY	-	expression tag	UNP P62136
B	3	HIS	-	expression tag	UNP P62136
B	4	MET	-	expression tag	UNP P62136
B	5	GLY	-	expression tag	UNP P62136
B	6	SER	-	expression tag	UNP P62136

- Molecule 2 is a protein called Apoptosis-stimulating of p53 protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	202	Total	C	N	O	S	0	0	0
			1554	986	254	300	14			
2	D	203	Total	C	N	O	S	0	0	0
			1563	991	254	303	15			

There are 10 discrepancies between the modelled and reference sequences:

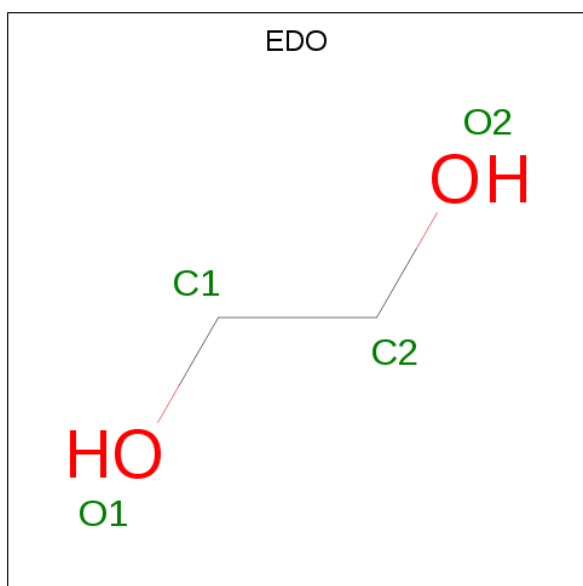
Chain	Residue	Modelled	Actual	Comment	Reference
C	915	GLY	-	expression tag	UNP Q13625

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	916	PRO	-	expression tag	UNP Q13625
C	917	LEU	-	expression tag	UNP Q13625
C	918	GLY	-	expression tag	UNP Q13625
C	919	SER	-	expression tag	UNP Q13625
D	915	GLY	-	expression tag	UNP Q13625
D	916	PRO	-	expression tag	UNP Q13625
D	917	LEU	-	expression tag	UNP Q13625
D	918	GLY	-	expression tag	UNP Q13625
D	919	SER	-	expression tag	UNP Q13625

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



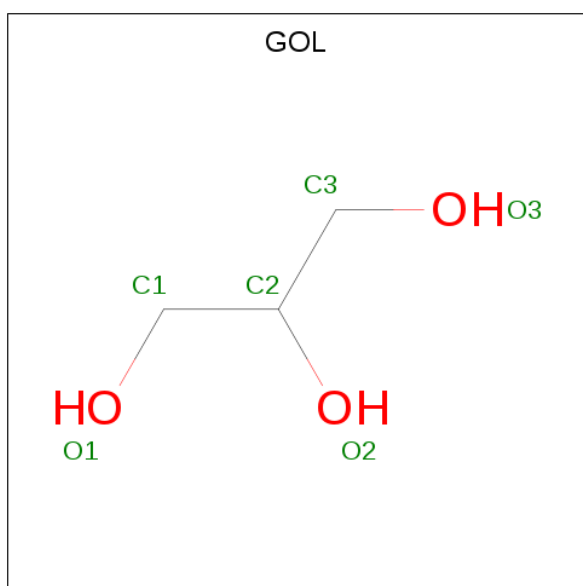
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	B	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

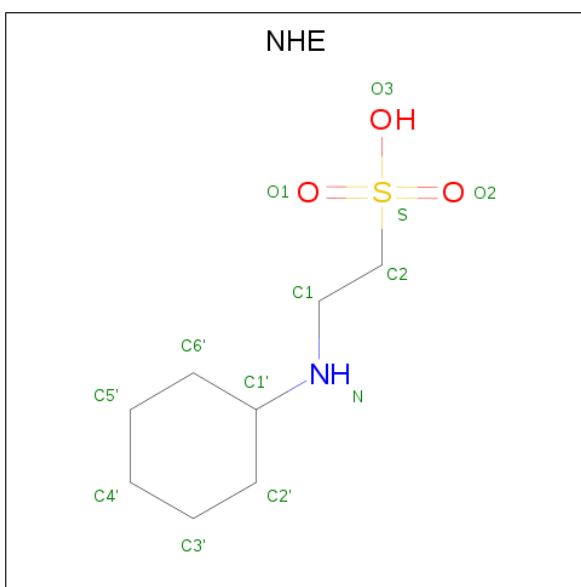
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		
3	D	1	Total	C	O	0	0
			4	2	2		
3	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is 2-[N-CYCLOHEXYLAMINO]ETHANE SULFONIC ACID (three-letter code: NHE) (formula:  $C_8H_{17}NO_3S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
5	A	1	Total	C	N	O	S	0	1
			13	8	1	3	1		
5	B	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
5	B	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
5	B	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
5	D	1	Total	C	N	O	S	0	0
			13	8	1	3	1		

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	2	Total	Mn	0	0
			2	2		
6	A	2	Total	Mn	0	0
			2	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	97	Total	O	0	0
			97	97		

*Continued on next page...*

*Continued from previous page...*

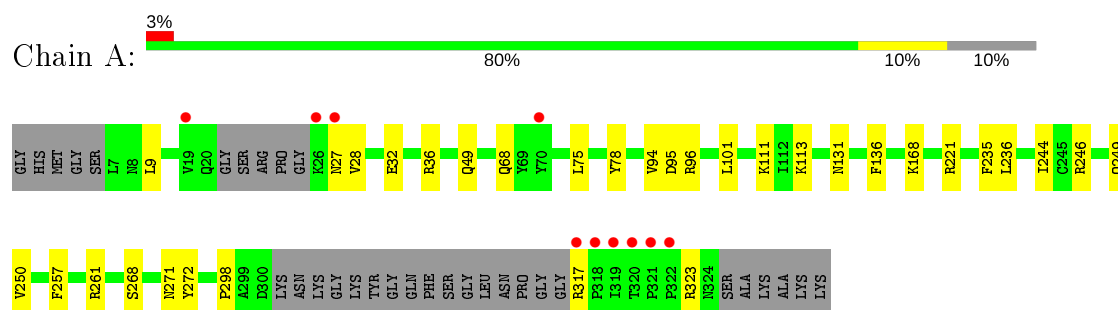
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	87	Total 87	O 87	0	1
7	C	25	Total 25	O 25	0	0
7	D	36	Total 36	O 36	0	0



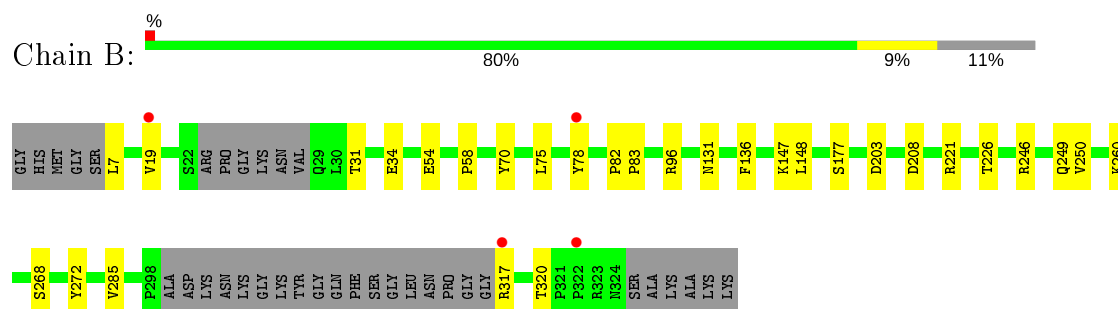
### 3 Residue-property plots [i](#)

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.

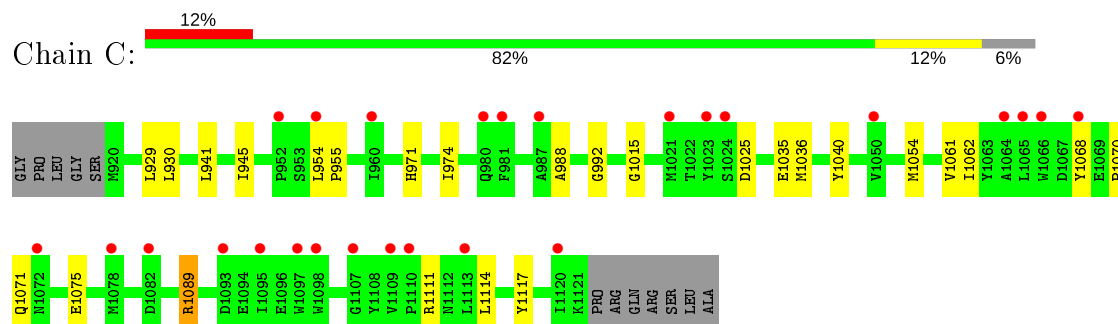
- Molecule 1: Serine/threonine-protein phosphatase PP1-alpha catalytic subunit



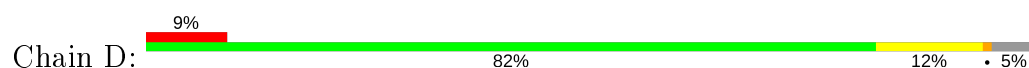
- Molecule 1: Serine/threonine-protein phosphatase PP1-alpha catalytic subunit

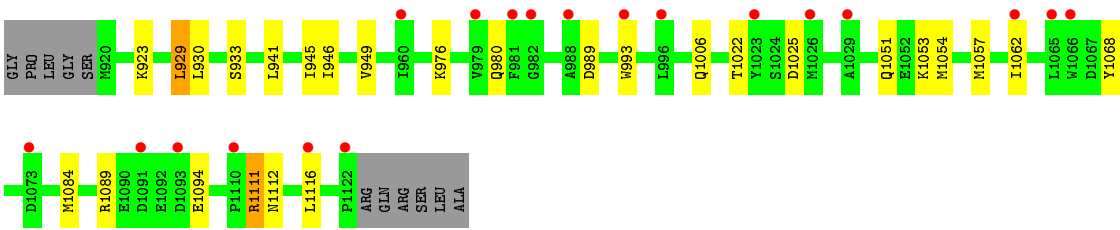


- Molecule 2: Apoptosis-stimulating of p53 protein 2



- Molecule 2: Apoptosis-stimulating of p53 protein 2





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	46.84Å 81.62Å 87.97Å 91.06° 91.84° 103.89°	Depositor
Resolution (Å)	59.65 – 2.15 59.65 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.6 (59.65-2.15) 99.6 (59.65-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.39 (at 2.14Å)	Xtriage
Refinement program	PHENIX (1.12 _2829: ???)	Depositor
R, $R_{free}$	0.177 , 0.214 0.177 , 0.214	Depositor DCC
$R_{free}$ test set	3477 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.5	Xtriage
Anisotropy	0.311	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 57.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8259	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.15% 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: NHE, GOL, MN, EDO

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	0.37	0/2436	0.52	0/3297
1	B	0.37	0/2436	0.54	0/3294
2	C	0.31	0/1590	0.45	0/2168
2	D	0.30	0/1600	0.45	0/2182
All	All	0.34	0/8062	0.50	0/10941

There are no bond length outliers.

There are no bond angle outliers.

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	2376	0	2313	25	0
1	B	2367	0	2324	23	0
2	C	1554	0	1440	15	0
2	D	1563	0	1442	18	0
3	A	24	0	36	2	0
3	B	16	0	24	0	0
3	D	8	0	12	0	0
4	A	18	0	24	1	0
4	B	6	0	8	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	26	0	33	4	0
5	B	39	0	48	7	0
5	D	13	0	17	0	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
7	A	97	0	0	5	0
7	B	87	0	0	1	0
7	C	25	0	0	2	0
7	D	36	0	0	1	0
All	All	8259	0	7721	78	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 5.

All (78) 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:B:54:GLU:OE1	7:B:501[A]:HOH:O	1.85	0.94
1:B:221:ARG:HH12	5:B:406:NHE:HC21	1.34	0.89
1:B:31:THR:HG23	1:B:34:GLU:H	1.42	0.81
1:A:271:ASN:ND2	1:A:298:PRO:HB3	1.95	0.81
1:A:271:ASN:HD21	1:A:298:PRO:HB3	1.44	0.80
1:B:250:VAL:HG12	5:B:406:NHE:H6'1	1.66	0.77
2:D:976:LYS:O	2:D:980:GLN:HG2	1.88	0.74
2:C:1061:VAL:O	2:C:1062:ILE:HD12	1.90	0.71
1:A:68:GLN:HA	3:A:406:EDO:H12	1.73	0.71
1:A:250:VAL:HG12	5:A:411[A]:NHE:H6'1	1.74	0.68
1:B:221:ARG:HH12	5:B:406:NHE:C2	2.09	0.65
5:A:411[A]:NHE:O1	7:A:501:HOH:O	2.11	0.65
1:A:96:ARG:NH2	1:A:272:TYR:OH	2.29	0.65
1:A:317:ARG:NH1	7:A:503:HOH:O	2.22	0.63
1:B:260:LYS:O	1:B:260:LYS:HD3	1.98	0.63
2:C:1070:PRO:HB3	2:C:1075:GLU:HB2	1.79	0.63
2:C:1035:GLU:HG3	2:C:1040:TYR:CZ	2.35	0.61
2:C:1062:ILE:HG21	2:C:1114:LEU:HD13	1.81	0.61
1:B:147:LYS:O	1:B:147:LYS:HD2	2.01	0.60
2:C:1071:GLN:NE2	7:C:1202:HOH:O	2.33	0.60
1:A:249:GLN:HA	5:A:411[A]:NHE:H6'2	1.83	0.60
1:A:317:ARG:NH2	7:A:503:HOH:O	2.24	0.60
2:D:933:SER:HB3	2:D:941:LEU:HD23	1.84	0.58
2:C:1036:MET:HE3	2:D:976:LYS:HD2	1.85	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:1054:MET:O	2:D:1089:ARG:NH2	2.37	0.57
1:A:317:ARG:NH1	2:D:1068:TYR:HA	2.21	0.56
1:A:9:LEU:HD13	1:A:113:LYS:HG2	1.88	0.56
2:D:1022:THR:OG1	2:D:1025:ASP:OD1	2.23	0.55
1:B:249:GLN:HA	5:B:406:NHE:H6'2	1.90	0.52
1:B:96:ARG:NH1	1:B:272:TYR:OH	2.42	0.52
2:C:1061:VAL:C	2:C:1062:ILE:HD12	2.30	0.52
2:C:929:LEU:HB3	2:C:945:ILE:HG12	1.92	0.52
1:A:257:PHE:CG	1:A:261:ARG:HG2	2.46	0.51
1:A:27:ASN:ND2	7:A:507:HOH:O	2.44	0.50
2:D:1062:ILE:HG22	2:D:1084:MET:HG3	1.94	0.50
1:A:221:ARG:HH12	5:A:411[A]:NHE:HC11	1.76	0.50
1:A:32:GLU:OE2	1:A:36:ARG:NH1	2.44	0.50
1:B:131:ASN:HB2	1:B:136:PHE:HB3	1.92	0.49
2:C:971:HIS:HB3	2:C:974:ILE:HD12	1.95	0.49
2:D:1111:ARG:HG2	2:D:1112:ASN:N	2.28	0.48
1:A:49:GLN:HE22	1:A:111:LYS:NZ	2.11	0.48
1:B:260:LYS:NZ	5:B:408:NHE:H6'1	2.28	0.47
1:A:28:VAL:O	7:A:502:HOH:O	2.20	0.47
2:D:1054:MET:HG3	2:D:1062:ILE:HD11	1.96	0.47
1:A:221:ARG:O	4:A:409:GOL:O3	2.31	0.46
1:B:19:VAL:HG23	1:B:70:TYR:HE1	1.81	0.46
2:C:1054:MET:O	2:C:1089:ARG:HD2	2.16	0.45
1:A:131:ASN:HB2	1:A:136:PHE:HB3	1.99	0.45
1:B:221:ARG:NH1	5:B:406:NHE:HC21	2.16	0.44
1:A:75:LEU:HD11	1:A:268:SER:HB3	1.99	0.44
1:B:177:SER:HB2	1:B:203:ASP:HB2	1.98	0.44
2:D:1006:GLN:H	2:D:1006:GLN:CD	2.17	0.43
2:D:1051:GLN:NE2	7:D:1303:HOH:O	2.45	0.43
1:B:317:ARG:NH1	7:C:1202:HOH:O	2.51	0.43
1:B:260:LYS:O	1:B:260:LYS:CD	2.67	0.43
2:D:946:ILE:O	2:D:949:VAL:HG22	2.19	0.43
1:B:148:LEU:HD12	1:B:148:LEU:HA	1.89	0.43
1:B:208:ASP:O	1:B:226:THR:HA	2.18	0.43
1:A:168:LYS:HA	1:A:168:LYS:HD3	1.85	0.42
1:A:94:VAL:O	1:A:95:ASP:HB2	2.19	0.42
2:C:988:ALA:HB1	2:C:992:GLY:HA2	2.00	0.42
2:D:989:ASP:OD1	2:D:993:TRP:N	2.43	0.42
1:B:82:PRO:HA	1:B:83:PRO:HA	1.85	0.42
1:A:236:LEU:HD21	1:A:244:ILE:HG13	2.00	0.42
2:D:929:LEU:HB3	2:D:945:ILE:HG12	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:250:VAL:H	5:B:406:NHE:H6'1	1.84	0.42
2:D:1057:MET:HE1	2:D:1116:LEU:HD11	2.01	0.42
2:D:1053:LYS:HB2	2:D:1057:MET:HE2	2.02	0.41
1:B:320:THR:HG21	2:C:1068:TYR:CE1	2.55	0.41
1:A:101:LEU:HA	1:A:101:LEU:HD23	1.90	0.41
1:A:323:ARG:NE	2:D:1094:GLU:OE2	2.38	0.41
2:C:1015:GLY:HA2	2:C:1117:TYR:O	2.21	0.41
1:A:235:PHE:HB2	3:A:403:EDO:H11	2.03	0.40
1:B:75:LEU:HD11	1:B:268:SER:HB3	2.02	0.40
1:B:58:PRO:HA	1:B:285:VAL:O	2.21	0.40
2:D:929:LEU:HD12	2:D:929:LEU:HA	1.96	0.40
2:C:1025:ASP:N	2:C:1025:ASP:OD1	2.51	0.40
2:C:954:LEU:HD13	2:C:955:PRO:N	2.36	0.40

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	293/329 (89%)	280 (96%)	13 (4%)	0	100	100
1	B	293/329 (89%)	280 (96%)	13 (4%)	0	100	100
2	C	200/214 (94%)	192 (96%)	8 (4%)	0	100	100
2	D	201/214 (94%)	195 (97%)	6 (3%)	0	100	100
All	All	987/1086 (91%)	947 (96%)	40 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	257/285 (90%)	255 (99%)	2 (1%)	81	86
1	B	256/285 (90%)	253 (99%)	3 (1%)	71	76
2	C	161/183 (88%)	157 (98%)	4 (2%)	47	49
2	D	162/183 (88%)	158 (98%)	4 (2%)	47	49
All	All	836/936 (89%)	823 (98%)	13 (2%)	62	67

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	78	TYR
1	A	246	ARG
1	B	7	LEU
1	B	78	TYR
1	B	246	ARG
2	C	930	LEU
2	C	941	LEU
2	C	1089	ARG
2	C	1111	ARG
2	D	923	LYS
2	D	929	LEU
2	D	930	LEU
2	D	1111	ARG

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	49	GLN
1	A	117	ASN
1	A	157	ASN
1	A	271	ASN



### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 26 ligands modelled in this entry, 4 are monoatomic - leaving 22 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$
4	GOL	A	407	-	5,5,5	0.88	0	5,5,5	0.85	0
3	EDO	A	403	-	3,3,3	0.50	0	2,2,2	0.31	0
5	NHE	A	411[A]	6	13,13,13	2.38	2 (15%)	16,17,17	1.76	4 (25%)
4	GOL	A	409	-	5,5,5	0.87	0	5,5,5	1.09	0
3	EDO	A	402	-	3,3,3	0.47	0	2,2,2	0.47	0
3	EDO	B	401	-	3,3,3	0.46	0	2,2,2	0.25	0
5	NHE	B	407	-	13,13,13	2.93	2 (15%)	16,17,17	1.16	2 (12%)
3	EDO	D	1201	-	3,3,3	0.46	0	2,2,2	0.25	0
3	EDO	A	405	-	3,3,3	0.47	0	2,2,2	0.43	0
5	NHE	A	410	-	13,13,13	2.93	2 (15%)	16,17,17	1.24	3 (18%)
3	EDO	D	1202	-	3,3,3	0.46	0	2,2,2	0.38	0
5	NHE	D	1203	-	13,13,13	2.39	2 (15%)	16,17,17	1.51	3 (18%)
5	NHE	B	408	-	13,13,13	2.70	2 (15%)	16,17,17	1.56	4 (25%)
4	GOL	A	408	-	5,5,5	1.06	0	5,5,5	0.90	0
4	GOL	B	405	-	5,5,5	1.06	0	5,5,5	0.94	0
3	EDO	B	403	-	3,3,3	0.49	0	2,2,2	0.45	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	EDO	A	406	-	3,3,3	0.49	0	2,2,2	0.28	0
3	EDO	B	402	-	3,3,3	0.48	0	2,2,2	0.24	0
5	NHE	B	406	6	13,13,13	2.74	2 (15%)	16,17,17	1.48	3 (18%)
3	EDO	A	404	-	3,3,3	0.52	0	2,2,2	0.29	0
3	EDO	B	404	-	3,3,3	0.49	0	2,2,2	0.33	0
3	EDO	A	401	-	3,3,3	0.42	0	2,2,2	0.45	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	GOL	A	407	-	-	0/4/4/4	-
3	EDO	A	403	-	-	1/1/1/1	-
5	NHE	A	411[A]	6	-	6/7/15/15	0/1/1/1
4	GOL	A	409	-	-	2/4/4/4	-
3	EDO	A	402	-	-	0/1/1/1	-
3	EDO	B	401	-	-	0/1/1/1	-
5	NHE	B	407	-	-	3/7/15/15	0/1/1/1
3	EDO	D	1201	-	-	0/1/1/1	-
3	EDO	A	405	-	-	1/1/1/1	-
5	NHE	A	410	-	-	3/7/15/15	0/1/1/1
3	EDO	D	1202	-	-	0/1/1/1	-
5	NHE	D	1203	-	-	5/7/15/15	0/1/1/1
5	NHE	B	408	-	-	0/7/15/15	0/1/1/1
4	GOL	A	408	-	-	1/4/4/4	-
4	GOL	B	405	-	-	3/4/4/4	-
3	EDO	B	403	-	-	0/1/1/1	-
3	EDO	A	406	-	-	0/1/1/1	-
3	EDO	B	402	-	-	0/1/1/1	-
5	NHE	B	406	6	-	2/7/15/15	0/1/1/1
3	EDO	A	404	-	-	0/1/1/1	-
3	EDO	B	404	-	-	0/1/1/1	-
3	EDO	A	401	-	-	0/1/1/1	-

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	410	NHE	C2-S	-8.18	1.65	1.77

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	407	NHE	C2-S	-7.97	1.66	1.77
5	B	406	NHE	C2-S	-7.22	1.67	1.77
5	A	411[A]	NHE	C2-S	-7.15	1.67	1.77
5	D	1203	NHE	C2-S	-7.10	1.67	1.77
5	B	408	NHE	C2-S	-7.01	1.67	1.77
5	B	407	NHE	O2-S	6.59	1.64	1.45
5	B	408	NHE	O1-S	6.51	1.64	1.45
5	B	406	NHE	O2-S	6.48	1.64	1.45
5	A	410	NHE	O1-S	6.34	1.63	1.45
5	D	1203	NHE	O3-S	4.52	1.63	1.47
5	A	411[A]	NHE	O3-S	4.39	1.63	1.47

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	411[A]	NHE	C1-N-C1'	4.33	122.63	114.14
5	A	411[A]	NHE	O1-S-C2	3.69	111.36	106.92
5	B	406	NHE	C1-N-C1'	3.53	121.08	114.14
5	D	1203	NHE	O1-S-C2	3.48	111.11	106.92
5	D	1203	NHE	C1-N-C1'	3.43	120.88	114.14
5	B	406	NHE	O1-S-C2	3.10	110.64	106.92
5	A	410	NHE	C1-N-C1'	3.04	120.11	114.14
5	B	408	NHE	O2-S-C2	2.87	110.36	106.92
5	B	408	NHE	C6'-C1'-C2'	-2.84	105.90	110.82
5	B	407	NHE	C1-N-C1'	2.83	119.69	114.14
5	B	408	NHE	C1-N-C1'	2.65	119.34	114.14
5	B	408	NHE	O3-S-C2	2.61	109.99	105.77
5	B	407	NHE	O3-S-C2	2.48	109.78	105.77
5	D	1203	NHE	O2-S-C2	2.26	109.63	106.92
5	A	411[A]	NHE	O3-S-C2	2.18	109.29	105.77
5	A	410	NHE	O3-S-C2	2.13	109.22	105.77
5	B	406	NHE	O3-S-C2	2.11	109.18	105.77
5	A	411[A]	NHE	O2-S-C2	2.06	109.40	106.92
5	A	410	NHE	O2-S-C2	2.02	109.35	106.92

There are no chirality outliers.

All (27) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	411[A]	NHE	C2'-C1'-N-C1
4	A	409	GOL	O1-C1-C2-C3
5	B	407	NHE	C2'-C1'-N-C1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	B	407	NHE	N-C1-C2-S
5	A	410	NHE	C1-C2-S-O1
5	A	410	NHE	C1-C2-S-O3
5	D	1203	NHE	C6'-C1'-N-C1
5	D	1203	NHE	C1-C2-S-O1
5	D	1203	NHE	C1-C2-S-O3
4	B	405	GOL	O1-C1-C2-C3
5	B	406	NHE	C2'-C1'-N-C1
5	A	411[A]	NHE	C2-C1-N-C1'
5	B	407	NHE	C2-C1-N-C1'
4	A	409	GOL	O1-C1-C2-O2
5	A	411[A]	NHE	C6'-C1'-N-C1
5	D	1203	NHE	C2'-C1'-N-C1
4	B	405	GOL	O1-C1-C2-O2
5	B	406	NHE	C2-C1-N-C1'
5	A	411[A]	NHE	C1-C2-S-O3
5	A	411[A]	NHE	C1-C2-S-O1
5	A	411[A]	NHE	C1-C2-S-O2
5	A	410	NHE	C1-C2-S-O2
5	D	1203	NHE	C1-C2-S-O2
4	A	408	GOL	C1-C2-C3-O3
4	B	405	GOL	O2-C2-C3-O3
3	A	403	EDO	O1-C1-C2-O2
3	A	405	EDO	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	403	EDO	1	0
5	A	411[A]	NHE	4	0
4	A	409	GOL	1	0
5	B	408	NHE	1	0
3	A	406	EDO	1	0
5	B	406	NHE	6	0

## 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, 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	297/329 (90%)	0.53	10 (3%) 45 53	26, 37, 76, 104	0
1	B	294/329 (89%)	0.45	4 (1%) 75 80	25, 38, 70, 101	0
2	C	202/214 (94%)	0.92	26 (12%) 3 5	44, 57, 83, 118	0
2	D	203/214 (94%)	0.80	19 (9%) 8 12	38, 56, 89, 113	0
All	All	996/1086 (91%)	0.64	59 (5%) 22 30	25, 46, 79, 118	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	1122	PRO	6.8
1	A	319	ILE	6.3
2	C	1095	ILE	5.7
1	A	317	ARG	5.6
1	A	320	THR	4.9
2	C	1066	TRP	4.8
2	C	954	LEU	4.6
1	A	26	LYS	4.0
2	C	987	ALA	3.7
2	C	1072	ASN	3.7
2	D	1023	TYR	3.7
1	B	78	TYR	3.5
2	D	1066	TRP	3.5
2	C	1023	TYR	3.3
2	C	1097	TRP	3.3
2	D	1065	LEU	3.2
2	C	960	ILE	3.2
1	A	318	PRO	3.1
2	C	980	GLN	3.1
2	C	1110	PRO	3.0
2	C	1120	ILE	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	C	1113	LEU	3.0
2	D	1091	ASP	2.8
1	B	19	VAL	2.8
2	D	1093	ASP	2.7
2	D	981	PHE	2.7
2	C	1065	LEU	2.7
1	A	27	ASN	2.7
1	A	19	VAL	2.7
2	C	1109	VAL	2.7
2	C	981	PHE	2.7
1	A	70	TYR	2.7
1	B	322	PRO	2.6
2	D	1110	PRO	2.6
2	D	1062	ILE	2.5
2	D	1029	ALA	2.5
2	D	1026	MET	2.5
1	A	322	PRO	2.5
1	A	321	PRO	2.4
2	C	1050	VAL	2.4
2	D	1116	LEU	2.3
2	D	982	GLY	2.3
2	D	996	LEU	2.3
2	D	960	ILE	2.3
2	C	1078	MET	2.3
1	B	317	ARG	2.3
2	C	1068	TYR	2.2
2	D	993	TRP	2.2
2	D	988	ALA	2.2
2	C	1098	TRP	2.2
2	C	1093	ASP	2.1
2	C	1107	GLY	2.1
2	C	1021	MET	2.1
2	C	1024	SER	2.1
2	D	979	VAL	2.1
2	C	1064	ALA	2.1
2	D	1073	ASP	2.0
2	C	1082	ASP	2.0
2	C	952	PRO	2.0

## 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 carbohydrates in this entry.

### 6.4 Ligands ⓘ

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
3	EDO	A	403	4/4	0.75	0.20	62,64,65,65	0
5	NHE	A	411[A]	13/13	0.77	0.44	32,37,44,45	13
3	EDO	A	406	4/4	0.80	0.25	66,67,68,68	0
4	GOL	B	405	6/6	0.81	0.22	62,72,72,74	0
5	NHE	B	406	13/13	0.82	0.38	31,43,49,49	13
3	EDO	D	1201	4/4	0.83	0.24	52,56,56,63	0
3	EDO	A	404	4/4	0.84	0.18	40,43,46,56	0
4	GOL	A	409	6/6	0.85	0.26	63,68,70,74	0
3	EDO	B	402	4/4	0.85	0.17	59,60,62,63	0
3	EDO	A	401	4/4	0.86	0.21	53,56,57,65	0
3	EDO	B	403	4/4	0.87	0.15	51,51,52,53	0
3	EDO	A	405	4/4	0.87	0.18	46,51,52,52	0
4	GOL	A	408	6/6	0.88	0.19	65,71,73,74	0
4	GOL	A	407	6/6	0.88	0.17	44,46,50,50	0
3	EDO	D	1202	4/4	0.89	0.11	66,68,71,72	0
5	NHE	B	408	13/13	0.90	0.17	78,81,88,88	0
3	EDO	B	401	4/4	0.91	0.20	57,63,68,72	0
5	NHE	B	407	13/13	0.91	0.17	51,57,79,81	0
5	NHE	D	1203	13/13	0.91	0.24	76,79,80,81	0
3	EDO	A	402	4/4	0.91	0.18	46,52,52,53	0
3	EDO	B	404	4/4	0.93	0.11	52,53,54,56	0
5	NHE	A	410	13/13	0.95	0.17	52,55,74,74	0
6	MN	A	412	1/1	0.99	0.16	32,32,32,32	0
6	MN	B	409	1/1	0.99	0.18	32,32,32,32	0
6	MN	B	410	1/1	1.00	0.18	27,27,27,27	0
6	MN	A	413	1/1	1.00	0.16	26,26,26,26	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.