Full wwPDB X-ray Structure Validation Report ⓘ

Sep 23, 2020 – 10:07 PM BST

PDB ID	:	6ZEH
Title	:	Structure of PP1-spectrin alpha II chimera [PP1(7-304) + linker (G/S)x9 + spectrin alpha II (1025-1039)] bound to Phactr1 (516-580)
Authors	:	Mouilleron, S.; Treisman, R.; Fedoryshchak, R.; Lee, R.; Butler, A.M.; Prechova, M.
Deposited on	:	2020-06-16
Resolution	:	1.30 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*
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:

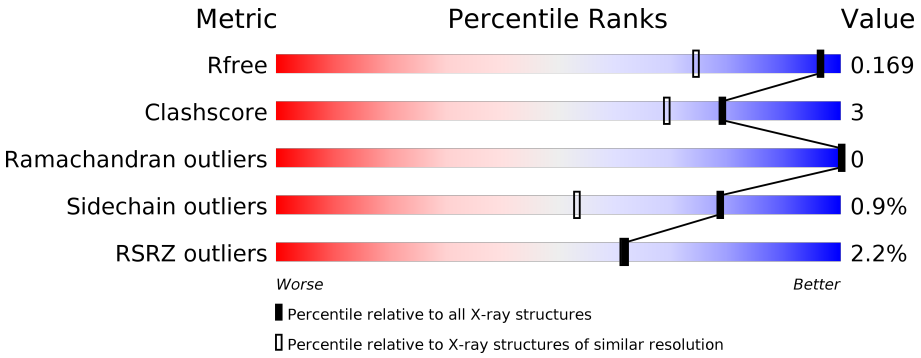
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.6

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.30 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	1058 (1.30-1.30)
Clashscore	141614	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)
RSRZ outliers	127900	1029 (1.30-1.30)

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 >=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	336	2% 83% 7% 10%
1	B	336	% 85% 7% 8%
2	C	70	4% 90% • 7%
2	D	70	% 91% • 7%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12785 atoms, of which 6007 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,Spectrin alpha chain, non-erythrocytic 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	304	Total	C	H	N	O	S	0	17	0
			4928	1595	2438	418	456	21			
1	B	308	Total	C	H	N	O	S	0	11	0
			4944	1595	2454	417	456	22			

There are 46 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
A	305	SER	-	linker	UNP P62136
A	306	GLY	-	linker	UNP P62136
A	307	SER	-	linker	UNP P62136
A	308	GLY	-	linker	UNP P62136
A	309	SER	-	linker	UNP P62136
A	310	GLY	-	linker	UNP P62136
A	311	SER	-	linker	UNP P62136
A	312	GLY	-	linker	UNP P62136
A	313	SER	-	linker	UNP P62136
A	314	GLY	-	linker	UNP P62136
A	315	SER	-	linker	UNP P62136
A	316	GLY	-	linker	UNP P62136
A	317	SER	-	linker	UNP P62136
A	318	GLY	-	linker	UNP P62136
A	319	SER	-	linker	UNP P62136
A	320	GLY	-	linker	UNP P62136
A	321	SER	-	linker	UNP P62136
A	322	GLY	-	linker	UNP P62136
B	2	GLY	-	expression tag	UNP P62136

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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
B	305	SER	-	linker	UNP P62136
B	306	GLY	-	linker	UNP P62136
B	307	SER	-	linker	UNP P62136
B	308	GLY	-	linker	UNP P62136
B	309	SER	-	linker	UNP P62136
B	310	GLY	-	linker	UNP P62136
B	311	SER	-	linker	UNP P62136
B	312	GLY	-	linker	UNP P62136
B	313	SER	-	linker	UNP P62136
B	314	GLY	-	linker	UNP P62136
B	315	SER	-	linker	UNP P62136
B	316	GLY	-	linker	UNP P62136
B	317	SER	-	linker	UNP P62136
B	318	GLY	-	linker	UNP P62136
B	319	SER	-	linker	UNP P62136
B	320	GLY	-	linker	UNP P62136
B	321	SER	-	linker	UNP P62136
B	322	GLY	-	linker	UNP P62136

- Molecule 2 is a protein called Phosphatase and actin regulator.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	65	Total	C	H	N	O	S	0	4	0
			1146	359	576	106	104	1			
2	D	65	Total	C	H	N	O	S	0	2	0
			1093	345	539	106	102	1			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	511	GLY	-	expression tag	UNP Q4VY12
C	512	PRO	-	expression tag	UNP Q4VY12
C	513	LEU	-	expression tag	UNP Q4VY12
C	514	GLY	-	expression tag	UNP Q4VY12
C	515	SER	-	expression tag	UNP Q4VY12
D	511	GLY	-	expression tag	UNP Q4VY12
D	512	PRO	-	expression tag	UNP Q4VY12
D	513	LEU	-	expression tag	UNP Q4VY12

Continued on next page...

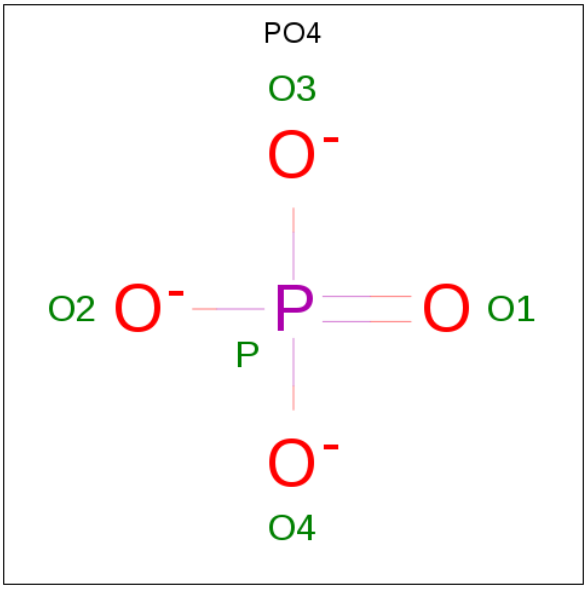
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	514	GLY	-	expression tag	UNP Q4VY12
D	515	SER	-	expression tag	UNP Q4VY12

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

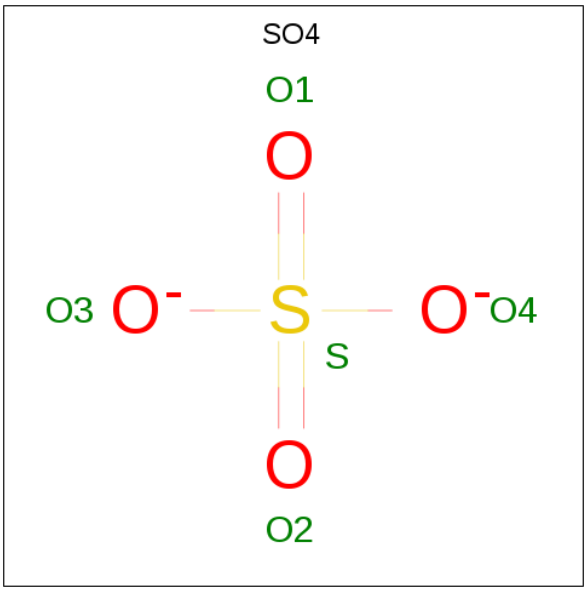
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total Mn 2 2	0	0
3	A	2	Total Mn 2 2	0	0

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 5 4 1	0	0
4	B	1	Total O P 5 4 1	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total O S 5 4 1	0	0

- Molecule 6 is water.

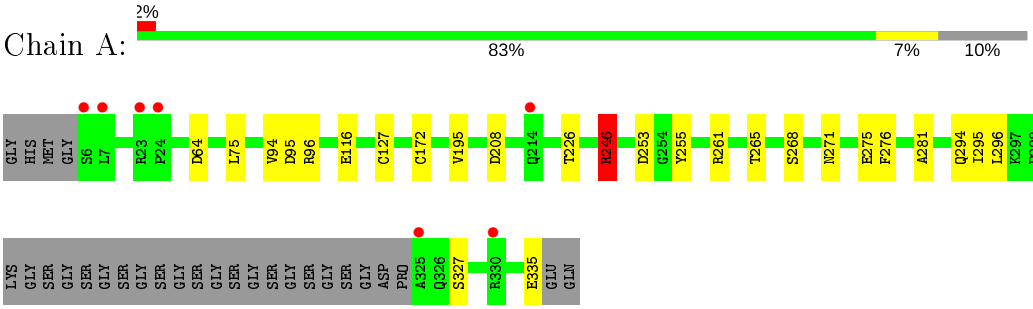
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	256	Total O 256 256	0	0
6	C	85	Total O 85 85	0	0
6	B	246	Total O 246 246	0	0
6	D	68	Total O 68 68	0	0

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.

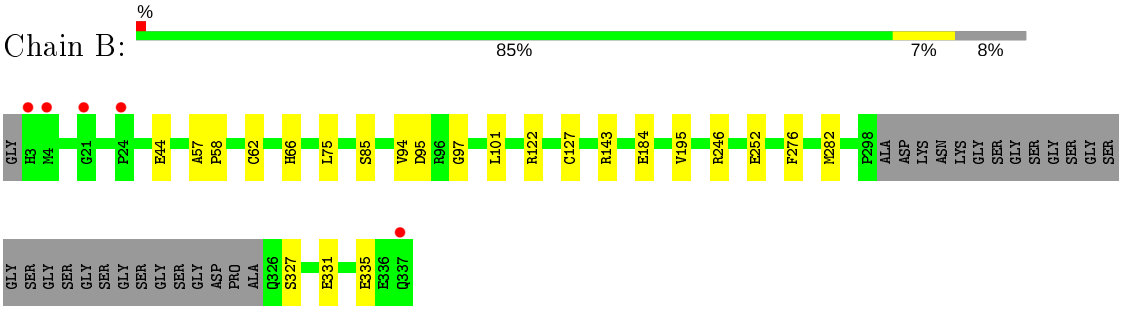
• Molecule 1: Serine/threonine-protein phosphatase PP1-alpha catalytic subunit,Spectrin alpha chain, non-erythrocytic 1

Chain A:



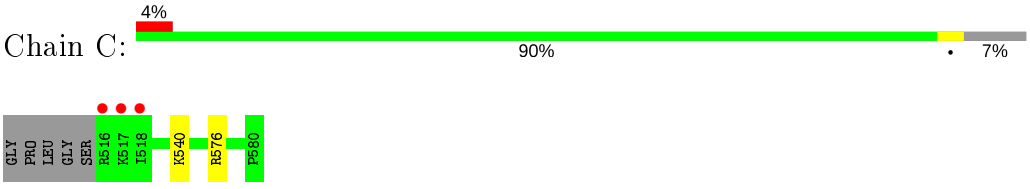
• Molecule 1: Serine/threonine-protein phosphatase PP1-alpha catalytic subunit,Spectrin alpha chain, non-erythrocytic 1

Chain B:



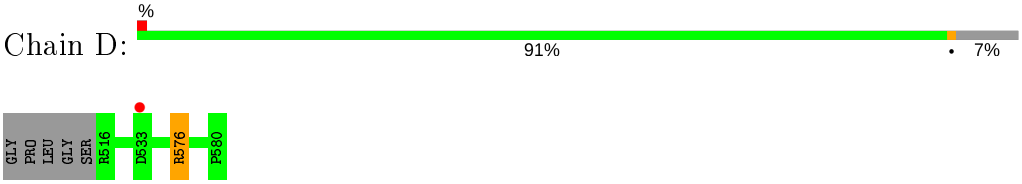
• Molecule 2: Phosphatase and actin regulator

Chain C:



• Molecule 2: Phosphatase and actin regulator

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	48.50Å 122.39Å 69.34Å 90.00° 92.18° 90.00°	Depositor
Resolution (Å)	69.29 – 1.30 69.29 – 1.30	Depositor EDS
% Data completeness (in resolution range)	98.9 (69.29-1.30) 99.3 (69.29-1.30)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 1.30Å)	Xtriage
Refinement program	PHENIX 1.18_3845, PHENIX 1.18_3845	Depositor
R, R_{free}	0.137 , 0.169 0.137 , 0.169	Depositor DCC
R_{free} test set	9653 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å^2)	10.6	Xtriage
Anisotropy	1.211	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 53.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.028 for h,-k,-l	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	12785	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.79% 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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, MN, SO4

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	0.53	0/2599	0.74	3/3512 (0.1%)
1	B	0.56	2/2588 (0.1%)	0.76	1/3494 (0.0%)
2	C	0.43	0/593	0.60	0/796
2	D	0.44	0/577	0.62	0/775
All	All	0.52	2/6357 (0.0%)	0.73	4/8577 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	252	GLU	CG-CD	5.61	1.60	1.51
1	B	252	GLU	CB-CG	5.17	1.61	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	64	ASP	CB-CG-OD1	7.34	124.90	118.30
1	B	122	ARG	NE-CZ-NH2	-5.95	117.32	120.30
1	A	246	ARG	NE-CZ-NH1	-5.23	117.69	120.30
1	A	96	ARG	NE-CZ-NH2	-5.06	117.77	120.30

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	2490	2438	2390	19	0
1	B	2490	2454	2422	15	0
2	C	570	576	576	1	0
2	D	554	539	529	1	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0
5	B	5	0	0	1	0
6	A	256	0	0	5	0
6	B	246	0	0	5	0
6	C	85	0	0	0	0
6	D	68	0	0	0	0
All	All	6778	6007	5917	34	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 3.

All (34) 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:75[B]:LEU:HD21	6:B:627:HOH:O	1.87	0.74
1:B:195:VAL:HG21	6:B:745:HOH:O	1.98	0.64
1:A:127[B]:CYS:SG	1:A:195:VAL:HG21	2.38	0.63
1:B:101[B]:LEU:CD1	1:B:143:ARG:HG3	2.31	0.61
1:B:101[B]:LEU:HD12	1:B:143:ARG:HG3	1.82	0.60
1:A:281:ALA:HB2	1:A:295:ILE:CD1	2.40	0.52
1:A:75[B]:LEU:HD21	6:A:728:HOH:O	2.10	0.50
1:B:127:CYS:HB3	6:B:745:HOH:O	2.12	0.50
1:A:275:GLU:OE2	2:C:540:LYS:NZ	2.40	0.49
1:B:85:SER:HB3	5:B:504:SO4:O2	2.12	0.49
1:B:276:PHE:CZ	1:B:327:SER:HB3	2.48	0.48
1:A:116:GLU:OE2	6:A:601:HOH:O	2.21	0.47
1:B:75[A]:LEU:HD23	1:B:75[A]:LEU:C	2.36	0.47
1:A:271:ASN:HA	1:A:276:PHE:O	2.15	0.46
1:B:66:HIS:O	1:B:97:GLY:HA3	2.16	0.45
1:A:94:VAL:O	1:A:95:ASP:HB2	2.17	0.45
1:B:184:GLU:OE2	6:B:601:HOH:O	2.20	0.45
1:A:261:ARG:NH1	6:A:604:HOH:O	2.33	0.44
1:A:75[A]:LEU:HD23	1:A:75[A]:LEU:C	2.39	0.43
1:B:331:GLU:OE1	2:D:576[A]:ARG:NH2	2.51	0.43
1:A:276:PHE:CZ	1:A:327:SER:HB3	2.53	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:172:CYS:O	1:A:246:ARG:HA	2.20	0.42
1:A:281:ALA:CB	1:A:295:ILE:HD13	2.50	0.41
1:A:253:ASP:O	1:A:295:ILE:HD12	2.20	0.41
1:A:75[B]:LEU:HD11	1:A:268:SER:HB3	2.02	0.41
1:A:208:ASP:O	1:A:226:THR:HA	2.21	0.41
1:A:294:GLN:HG3	6:A:603:HOH:O	2.21	0.41
1:B:44:GLU:OE2	6:B:602:HOH:O	2.21	0.41
1:A:296:LEU:HD13	6:A:724:HOH:O	2.20	0.40
1:A:255:TYR:HA	1:A:265:THR:O	2.21	0.40
1:B:62:CYS:SG	1:B:282[B]:MET:HG2	2.62	0.40
1:A:127[B]:CYS:SG	1:A:195:VAL:CG2	3.07	0.40
1:B:57:ALA:HB1	1:B:58:PRO:HA	2.02	0.40
1:B:94:VAL:O	1:B:95:ASP:HB2	2.22	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	317/336 (94%)	306 (96%)	11 (4%)	0	100	100
1	B	315/336 (94%)	304 (96%)	11 (4%)	0	100	100
2	C	67/70 (96%)	66 (98%)	1 (2%)	0	100	100
2	D	65/70 (93%)	64 (98%)	1 (2%)	0	100	100
All	All	764/812 (94%)	740 (97%)	24 (3%)	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	272/287 (95%)	270 (99%)	2 (1%)	84 61
1	B	276/287 (96%)	274 (99%)	2 (1%)	84 61
2	C	62/61 (102%)	61 (98%)	1 (2%)	62 28
2	D	58/61 (95%)	56 (97%)	2 (3%)	37 5
All	All	668/696 (96%)	661 (99%)	7 (1%)	78 48

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	246	ARG
1	A	335	GLU
2	C	576	ARG
1	B	246	ARG
1	B	335	GLU
2	D	576[A]	ARG
2	D	576[B]	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 4 are monoatomic - leaving 3 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	PO4	B	503	3	4,4,4	0.79	0	6,6,6	1.60	2 (33%)
5	SO4	B	504	-	4,4,4	0.40	0	6,6,6	0.46	0
4	PO4	A	503	3	4,4,4	1.26	1 (25%)	6,6,6	1.01	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	503	PO4	P-O3	-2.12	1.48	1.54

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	503	PO4	O3-P-O1	-2.68	101.11	110.89
4	B	503	PO4	O4-P-O3	2.47	115.90	107.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	504	SO4	1	0

5.7 Other polymers (i)

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^{th} 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^2)	Q<0.9
1	A	304/336 (90%)	-0.14	7 (2%)	60	59	9, 15, 29, 42	0
1	B	308/336 (91%)	-0.08	5 (1%)	72	73	10, 15, 29, 44	0
2	C	65/70 (92%)	-0.03	3 (4%)	32	30	13, 19, 30, 38	0
2	D	65/70 (92%)	-0.11	1 (1%)	73	75	14, 23, 34, 39	0
All	All	742/812 (91%)	-0.10	16 (2%)	62	61	9, 16, 30, 44	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	21	GLY	15.4
1	A	6	SER	5.5
1	B	3	HIS	4.9
1	A	325	ALA	4.7
1	A	7	LEU	3.7
2	C	518	ILE	3.7
1	A	24	PRO	3.5
2	C	517	LYS	3.3
1	B	337	GLN	3.0
2	D	533	ASP	3.0
1	B	24	PRO	2.7
1	B	4	MET	2.7
1	A	214	GLN	2.7
1	A	23	ARG	2.6
2	C	516	ARG	2.0
1	A	330	ARG	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 (i)

There are no monosaccharides 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^{th} 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^2)	Q<0.9
5	SO4	B	504	5/5	0.98	0.26	13,27,30,31	0
4	PO4	B	503	5/5	0.99	0.07	10,10,11,11	0
4	PO4	A	503	5/5	0.99	0.10	9,10,11,12	0
3	MN	B	502	1/1	1.00	0.10	10,10,10,10	0
3	MN	A	501	1/1	1.00	0.10	8,8,8,8	0
3	MN	A	502	1/1	1.00	0.10	9,9,9,9	0
3	MN	B	501	1/1	1.00	0.10	9,9,9,9	0

6.5 Other polymers (i)

There are no such residues in this entry.