



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:37 PM GMT

PDB ID : 4CXS  
Title : G4 mutant of PAS, arylsulfatase from *Pseudomonas Aeruginosa*, in complex with Phenylphosphonic acid  
Authors : Miton, C.M.; Jonas, S.; Mohammed, M.F.; Fischer, G.; Loo, B.v.; Kintses, B.; Hyvonen, M.; Tokuriki, N.; Hollfelder, F.  
Deposited on : 2014-04-08  
Resolution : 2.30 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](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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

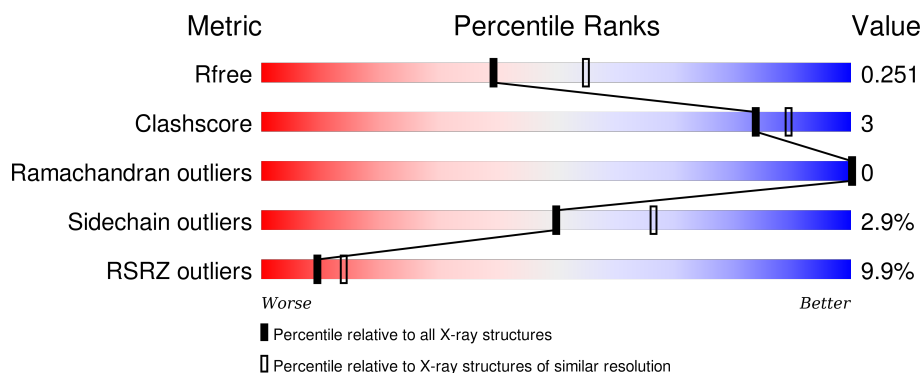
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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))
$R_{free}$	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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  $\geq 3$ , 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	536	<div> <div>10%</div> <div>86%</div> <div>10%</div> <div>••</div> </div>
1	B	536	<div> <div>10%</div> <div>87%</div> <div>9%</div> <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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	SV7	A	1532	-	-	-	X
4	SV7	B	1531	-	-	-	X
4	SV7	B	1532	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8484 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 ARYLSULFATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	522	Total	C	N	O	S	0	1	0
			4126	2626	730	761	9			
1	B	515	Total	C	N	O	S	0	1	0
			4072	2588	723	752	9			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	THR	ALA	ENGINEERED MUTATION	UNP U6AIT1
A	50	ALA	THR	ENGINEERED MUTATION	UNP U6AIT1
A	134	ALA	SER	ENGINEERED MUTATION	UNP U6AIT1
A	337	ASP	GLY	ENGINEERED MUTATION	UNP U6AIT1
A	461	GLY	GLU	ENGINEERED MUTATION	UNP U6AIT1
A	523	ASP	GLU	ENGINEERED MUTATION	UNP U6AIT1
B	22	THR	ALA	ENGINEERED MUTATION	UNP U6AIT1
B	50	ALA	THR	ENGINEERED MUTATION	UNP U6AIT1
B	134	ALA	SER	ENGINEERED MUTATION	UNP U6AIT1
B	337	ASP	GLY	ENGINEERED MUTATION	UNP U6AIT1
B	461	GLY	GLU	ENGINEERED MUTATION	UNP U6AIT1
B	523	ASP	GLU	ENGINEERED MUTATION	UNP U6AIT1

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

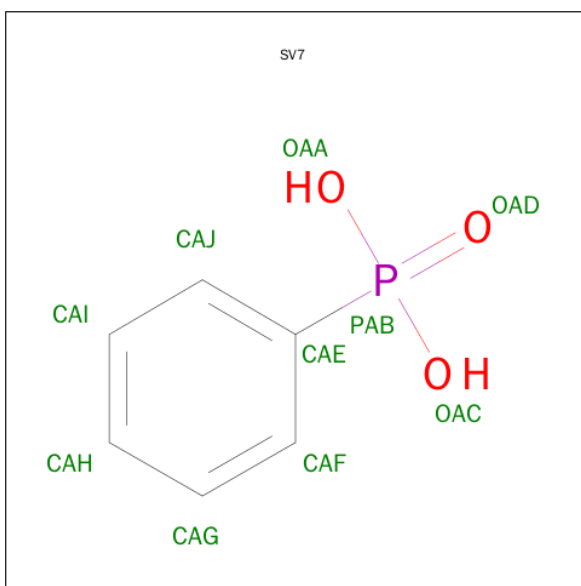
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ca	0	0
			1	1		
2	A	1	Total	Ca	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 4 is PHENYLPHOSPHONIC ACID (three-letter code: SV7) (formula:  $C_6H_7O_3P$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	O	P	0	0
			10	6	3	1		
4	A	1	Total	C	O	P	0	0
			10	6	3	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	O	P	0	0
			10	6	3	1		
4	B	1	Total	C	O	P	0	0
			10	6	3	1		
4	B	1	Total	C	O	P	0	0
			10	6	3	1		

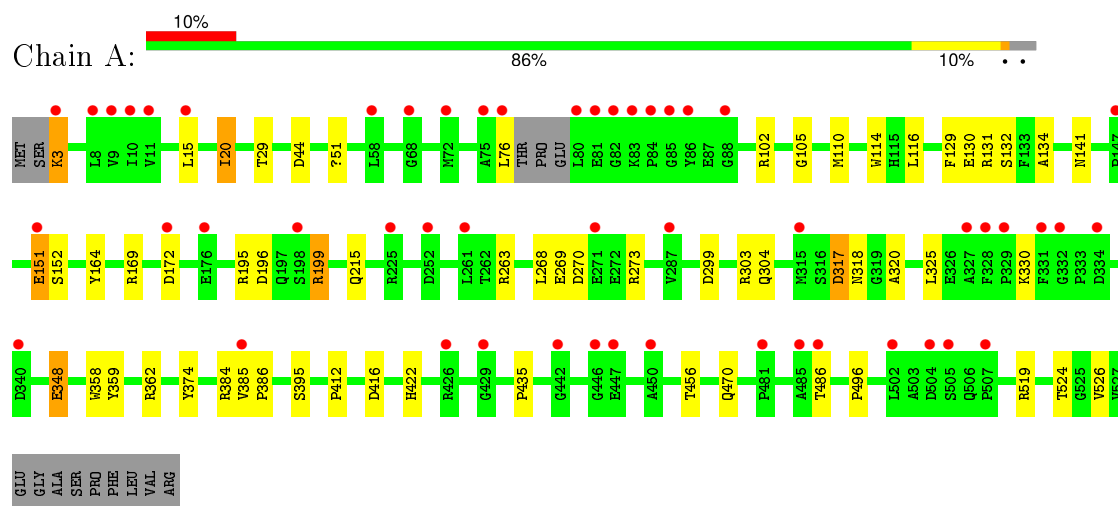
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	115	Total	O	0	0
			115	115		
5	B	109	Total	O	0	0
			109	109		

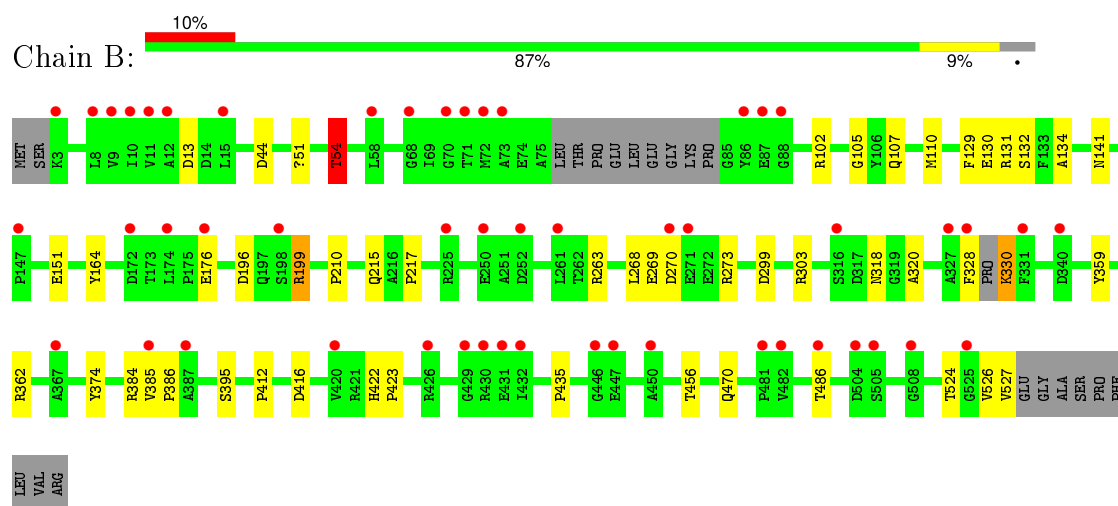
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: ARYLSULFATASE



#### • Molecule 1: ARYLSULFATASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	183.53 Å   66.18 Å   89.34 Å 90.00°   93.69°   90.00°	Depositor
Resolution (Å)	91.57 – 2.30 39.10 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.3 (91.57-2.30) 98.3 (39.10-2.30)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.37 (at 2.29 Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, $R_{free}$	0.203   ,   0.244 0.211   ,   0.251	Depositor DCC
$R_{free}$ test set	2376 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.7	Xtriage
Anisotropy	0.162	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 34.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 46965 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8484	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: DDZ, CA, SO4, SV7

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	0.74	0/4237	0.86	7/5760 (0.1%)
1	B	2.67	2/4179 (0.0%)	1.16	8/5678 (0.1%)
All	All	1.95	2/8416 (0.0%)	1.02	15/11438 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	1	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	54	THR	C-O	135.07	3.79	1.23
1	B	54	THR	CB-OG1	95.67	3.34	1.43

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	54	THR	CA-C-O	-45.77	23.98	120.10
1	B	54	THR	CA-CB-OG1	-31.63	42.57	109.00
1	B	54	THR	OG1-CB-CG2	-18.50	67.46	110.00
1	A	273	ARG	NE-CZ-NH2	-8.22	116.19	120.30
1	B	273	ARG	NE-CZ-NH2	-8.19	116.20	120.30
1	B	199	ARG	NE-CZ-NH2	7.19	123.89	120.30
1	A	519	ARG	NE-CZ-NH1	7.15	123.87	120.30
1	A	199	ARG	NE-CZ-NH2	6.72	123.66	120.30
1	A	44	ASP	CB-CG-OD1	5.80	123.52	118.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	362	ARG	NE-CZ-NH1	5.74	123.17	120.30
1	B	44	ASP	CB-CG-OD1	5.43	123.19	118.30
1	A	362	ARG	NE-CZ-NH1	5.29	122.94	120.30
1	B	151	GLU	CB-CA-C	-5.20	100.00	110.40
1	A	20	ILE	CG1-CB-CG2	-5.19	99.98	111.40
1	A	317	ASP	CB-CG-OD1	5.04	122.83	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	54	THR	CB

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	51	DDZ	Mainchain
1	B	54	THR	Mainchain

## 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	4126	0	3997	27	1
1	B	4072	0	3939	22	1
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	20	0	10	1	0
4	B	30	0	15	0	0
5	A	115	0	0	0	0
5	B	109	0	0	0	0
All	All	8484	0	7961	49	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (49) 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:A:169:ARG:NH2	1:A:172:ASP:OD1	2.21	0.74
1:A:15:LEU:HD21	1:A:20:ILE:HD13	1.72	0.71
1:A:51:DDZ:HB	1:A:317:ASP:OD2	1.92	0.70
1:A:51:DDZ:OG2	4:A:1531:SV7:OAC	2.08	0.70
1:A:15:LEU:HD21	1:A:20:ILE:CD1	2.28	0.64
1:B:328:PHE:O	1:B:330:LYS:N	2.32	0.63
1:B:299:ASP:OD2	1:B:303:ARG:NH1	2.38	0.57
1:A:299:ASP:OD2	1:A:303:ARG:NH1	2.40	0.55
1:B:141:ASN:ND2	1:B:215:GLN:HE22	2.07	0.53
1:A:195:ARG:HH22	1:A:304:GLN:HE21	1.59	0.51
1:A:141:ASN:ND2	1:A:215:GLN:HE22	2.09	0.51
1:A:412:PRO:O	1:A:416:ASP:HB2	2.12	0.49
1:B:384:ARG:C	1:B:384:ARG:HD2	2.33	0.49
1:B:456:THR:O	1:B:470:GLN:HA	2.14	0.48
1:A:105:GLY:O	1:A:199:ARG:NH1	2.47	0.48
1:A:151:GLU:HG2	1:A:152:SER:N	2.29	0.47
1:B:105:GLY:O	1:B:199:ARG:NH1	2.48	0.47
1:A:3:LYS:HD2	1:A:3:LYS:N	2.30	0.47
1:A:384:ARG:HD2	1:A:384:ARG:C	2.35	0.47
1:B:130:GLU:HG2	1:B:131:ARG:HG3	1.95	0.47
1:A:385:VAL:HB	1:A:386:PRO:HD2	1.96	0.47
1:A:422:HIS:CE1	1:A:435:PRO:HD3	2.51	0.46
1:A:130:GLU:HG2	1:A:131:ARG:HG3	1.96	0.46
1:B:328:PHE:HB3	1:B:330:LYS:HB2	1.97	0.46
1:B:412:PRO:O	1:B:416:ASP:HB2	2.16	0.46
1:B:385:VAL:HB	1:B:386:PRO:HD2	1.98	0.45
1:B:134:ALA:O	1:B:164:TYR:HA	2.17	0.45
1:A:110:MET:HB2	1:A:129:PHE:CE1	2.52	0.44
1:A:456:THR:O	1:A:470:GLN:HA	2.18	0.44
1:B:422:HIS:CE1	1:B:435:PRO:HD3	2.52	0.44
1:B:110:MET:HB2	1:B:129:PHE:CE1	2.54	0.43
1:A:263:ARG:NH1	1:A:268:LEU:HD23	2.34	0.43
1:A:320:ALA:O	1:A:374:TYR:HB2	2.19	0.43
1:A:196:ASP:OD2	1:A:199:ARG:HD3	2.19	0.43
1:A:524:THR:OG1	1:A:526:VAL:HG23	2.19	0.42
1:B:263:ARG:NH1	1:B:268:LEU:HD23	2.35	0.42
1:B:110:MET:O	1:B:132:SER:HA	2.19	0.42
1:B:107:GLN:OE1	1:B:196:ASP:N	2.42	0.42
1:B:196:ASP:OD2	1:B:199:ARG:HD3	2.20	0.41
1:A:110:MET:O	1:A:132:SER:HA	2.19	0.41
1:A:134:ALA:O	1:A:164:TYR:HA	2.21	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:114:TRP:CE2	1:A:116:LEU:HB2	2.55	0.41
1:A:325:LEU:HD12	1:A:358:TRP:CE3	2.55	0.41
1:B:422:HIS:CG	1:B:423:PRO:HD2	2.55	0.41
1:B:524:THR:OG1	1:B:526:VAL:HG23	2.21	0.41
1:B:320:ALA:O	1:B:374:TYR:HB2	2.19	0.41
1:A:20:ILE:HG13	1:A:29:THR:HG21	2.03	0.41
1:B:328:PHE:C	1:B:330:LYS:N	2.74	0.41
1:B:13:ASP:O	1:B:210:PRO:HD2	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:348:GLU:OE1	1:B:176:GLU:OE2[4_445]	2.19	0.01

## 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	518/536 (97%)	500 (96%)	18 (4%)	0	100	100
1	B	509/536 (95%)	491 (96%)	18 (4%)	0	100	100
All	All	1027/1072 (96%)	991 (96%)	36 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	415/428 (97%)	402 (97%)	13 (3%)	47	64
1	B	409/428 (96%)	398 (97%)	11 (3%)	52	70
All	All	824/856 (96%)	800 (97%)	24 (3%)	50	66

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	LYS
1	A	76	LEU
1	A	102	ARG
1	A	151	GLU
1	A	269	GLU
1	A	270	ASP
1	A	318	ASN
1	A	330	LYS
1	A	348	GLU
1	A	359	TYR
1	A	395	SER
1	A	486	THR
1	A	496	PRO
1	B	54	THR
1	B	102	ARG
1	B	217	PRO
1	B	269	GLU
1	B	270	ASP
1	B	318	ASN
1	B	330	LYS
1	B	359	TYR
1	B	395	SER
1	B	486	THR
1	B	527	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	304	GLN
1	A	402	HIS
1	A	506	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	402	HIS
1	B	506	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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
1	DDZ	A	51	1,2	2,6,7	0.33	0	1,7,9	4.12	1 (100%)
1	DDZ	B	51	1,2	2,6,7	21.60	1 (50%)	1,7,9	4.52	1 (100%)

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
1	DDZ	A	51	1,2	-	0/0/6/8	0/0/0/0
1	DDZ	B	51	1,2	-	0/0/6/8	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	51	DDZ	O-C	30.55	2.61	1.19

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	51	DDZ	C-CA-N	4.12	118.06	109.12
1	B	51	DDZ	C-CA-N	4.52	118.94	109.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	51	DDZ	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	$\# Z  > 2$
3	SO4	A	1530	-	4,4,4	0.70	0	6,6,6	0.61	0
4	SV7	A	1531	-	10,10,10	1.92	2 (20%)	14,14,14	2.01	3 (21%)
4	SV7	A	1532	-	10,10,10	2.49	3 (30%)	14,14,14	1.94	3 (21%)
3	SO4	B	1530	-	4,4,4	0.87	0	6,6,6	0.90	0
4	SV7	B	1531	-	10,10,10	2.49	1 (10%)	14,14,14	1.62	4 (28%)
4	SV7	B	1532	-	10,10,10	3.75	3 (30%)	14,14,14	1.74	3 (21%)
4	SV7	B	1533	-	10,10,10	2.84	3 (30%)	14,14,14	1.93	3 (21%)

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
3	SO4	A	1530	-	-	0/0/0/0	0/0/0/0
4	SV7	A	1531	-	-	0/6/6/6	0/1/1/1
4	SV7	A	1532	-	-	0/6/6/6	0/1/1/1
3	SO4	B	1530	-	-	0/0/0/0	0/0/0/0
4	SV7	B	1531	-	-	0/6/6/6	0/1/1/1
4	SV7	B	1532	-	-	0/6/6/6	0/1/1/1
4	SV7	B	1533	-	-	0/6/6/6	0/1/1/1

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1532	SV7	PAB-CAE	-8.18	1.65	1.79
4	B	1533	SV7	PAB-CAE	-6.90	1.67	1.79
4	A	1531	SV7	PAB-OAC	-4.74	1.44	1.54
4	B	1533	SV7	PAB-OAD	-4.17	1.41	1.50
4	A	1532	SV7	PAB-OAC	-3.37	1.47	1.54
4	B	1533	SV7	PAB-OAC	-3.13	1.47	1.54
4	B	1532	SV7	PAB-OAC	-2.71	1.48	1.54
4	A	1531	SV7	PAB-OAD	-2.68	1.44	1.50
4	A	1532	SV7	PAB-OAA	2.60	1.59	1.54
4	A	1532	SV7	PAB-CAE	5.94	1.89	1.79
4	B	1531	SV7	PAB-OAA	7.14	1.69	1.54
4	B	1532	SV7	PAB-OAA	7.89	1.70	1.54

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1533	SV7	OAA-PAB-OAD	-5.24	98.36	112.13
4	A	1531	SV7	OAA-PAB-OAD	-3.75	102.27	112.13
4	B	1532	SV7	OAA-PAB-OAD	-3.60	102.68	112.13
4	A	1532	SV7	OAA-PAB-CAE	-3.06	100.33	107.13
4	A	1531	SV7	OAC-PAB-OAA	-2.87	98.32	108.03
4	B	1531	SV7	OAA-PAB-OAD	-2.72	104.97	112.13
4	A	1532	SV7	OAC-PAB-OAA	-2.32	100.16	108.03
4	B	1531	SV7	CAI-CAJ-CAE	-2.06	117.70	120.09
4	B	1532	SV7	OAD-PAB-CAE	2.29	114.06	109.53
4	B	1531	SV7	OAD-PAB-CAE	2.39	114.26	109.53
4	B	1531	SV7	CAJ-CAE-CAF	2.86	122.88	119.25
4	B	1532	SV7	CAJ-CAE-CAF	3.03	123.09	119.25
4	B	1533	SV7	OAC-PAB-CAE	3.16	114.16	107.13
4	B	1533	SV7	OAD-PAB-CAE	3.17	115.82	109.53

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1531	SV7	OAC-PAB-CAE	4.95	118.15	107.13
4	A	1532	SV7	OAC-PAB-CAE	5.29	118.90	107.13

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
4	A	1531	SV7	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	521/536 (97%)	0.64	51 (9%)	10 14	11, 21, 47, 74	3 (0%)
1	B	514/536 (95%)	0.63	51 (9%)	9 14	11, 20, 45, 74	3 (0%)
All	All	1035/1072 (96%)	0.63	102 (9%)	9 14	11, 20, 46, 74	6 (0%)

All (102) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	84	PRO	7.3
1	A	82	GLY	7.2
1	B	328	PHE	5.7
1	A	446	GLY	4.7
1	B	72	MET	4.6
1	B	429	GLY	4.5
1	B	73	ALA	4.4
1	B	3	LYS	4.3
1	B	86	TYR	4.2
1	A	504	ASP	4.2
1	A	81	GLU	4.1
1	A	505	SER	4.0
1	B	327	ALA	3.9
1	B	447	GLU	3.9
1	A	83	LYS	3.7
1	A	85	GLY	3.7
1	B	252	ASP	3.6
1	A	172	ASP	3.5
1	B	426	ARG	3.5
1	A	3	LYS	3.5
1	A	481	PRO	3.4
1	A	86	TYR	3.3
1	A	271	GLU	3.3
1	A	252	ASP	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	10	ILE	3.3
1	A	447	GLU	3.3
1	A	429	GLY	3.2
1	B	87	GLU	3.2
1	B	481	PRO	3.1
1	B	446	GLY	3.1
1	B	10	ILE	3.0
1	A	426	ARG	3.0
1	B	71	THR	3.0
1	B	11	VAL	3.0
1	B	68	GLY	2.9
1	A	332	GLY	2.9
1	A	72	MET	2.9
1	B	172	ASP	2.9
1	B	486	THR	2.9
1	A	15	LEU	2.8
1	A	507	PRO	2.8
1	A	151	GLU	2.8
1	B	504	ASP	2.8
1	B	88	GLY	2.7
1	B	316	SER	2.7
1	B	15	LEU	2.7
1	B	450	ALA	2.7
1	B	431	GLU	2.7
1	A	198	SER	2.6
1	B	12	ALA	2.6
1	B	70	GLY	2.6
1	B	420	VAL	2.6
1	B	174	LEU	2.5
1	A	328	PHE	2.5
1	B	331	PHE	2.5
1	A	75	ALA	2.5
1	B	505	SER	2.5
1	B	340	ASP	2.5
1	A	331	PHE	2.5
1	A	80	LEU	2.5
1	A	340	ASP	2.4
1	B	508	GLY	2.4
1	B	198	SER	2.4
1	A	88	GLY	2.4
1	A	76	LEU	2.4
1	A	261	LEU	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	225	ARG	2.4
1	A	327	ALA	2.4
1	A	442	GLY	2.4
1	A	9	VAL	2.4
1	B	430	ARG	2.4
1	B	271	GLU	2.3
1	A	287	VAL	2.3
1	B	387	ALA	2.3
1	A	334	ASP	2.3
1	A	329	PRO	2.3
1	A	502	LEU	2.3
1	B	58	LEU	2.3
1	B	176	GLU	2.3
1	B	147	PRO	2.3
1	B	8	LEU	2.3
1	A	485	ALA	2.3
1	B	250	GLU	2.3
1	B	261	LEU	2.3
1	A	147	PRO	2.3
1	A	176	GLU	2.3
1	A	225	ARG	2.3
1	A	8	LEU	2.2
1	A	450	ALA	2.2
1	B	367	ALA	2.2
1	A	58	LEU	2.2
1	A	486	THR	2.2
1	A	11	VAL	2.2
1	B	525	GLY	2.2
1	A	385	VAL	2.2
1	B	270	ASP	2.1
1	B	9	VAL	2.1
1	B	385	VAL	2.1
1	B	482	VAL	2.1
1	A	315	MET	2.1
1	A	68	GLY	2.1
1	B	432	ILE	2.0

## 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(Å <sup>2</sup> )	Q<0.9
1	DDZ	B	51	7/8	0.94	0.17	-	11,12,13,14	0
1	DDZ	A	51	7/8	0.96	0.16	-	12,13,13,14	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(Å <sup>2</sup> )	Q<0.9
4	SV7	A	1532	10/10	0.74	0.23	4.23	52,56,64,66	0
4	SV7	B	1532	10/10	0.79	0.25	3.54	25,28,40,40	0
4	SV7	B	1531	10/10	0.74	0.23	3.20	24,26,39,39	0
3	SO4	B	1530	5/5	0.95	0.18	0.49	30,30,33,34	0
3	SO4	A	1530	5/5	0.97	0.17	0.37	29,30,32,33	0
4	SV7	B	1533	10/10	0.97	0.17	-0.17	29,35,42,44	0
4	SV7	A	1531	10/10	0.96	0.14	-1.31	25,31,34,34	0
2	CA	B	1529	1/1	0.95	0.08	-3.81	32,32,32,32	0
2	CA	A	1529	1/1	0.97	0.08	-4.09	31,31,31,31	0

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

There are no such residues in this entry.