



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 7, 2016 – 07:59 PM EDT

PDB ID : 5A7J  
Title : Crystal structure of INPP5B in complex with benzene 1,2,4,5- tetrakisphosphate  
Authors : Tresaugues, L.; Mills, S.J.; Silvander, C.; Cozier, G.; Potter, B.V.L.; Nordlund, P.  
Deposited on : 2015-07-06  
Resolution : 2.90 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
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) : rb-20027107

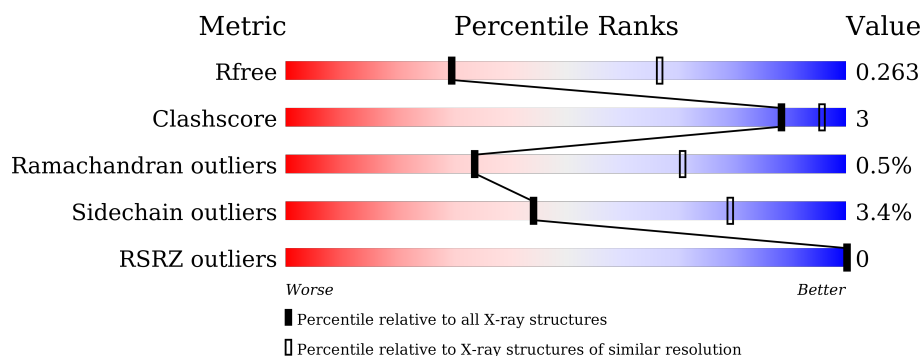
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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	313	
1	B	313	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5125 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 TYPE II INOSITOL 1,4,5-TRISPHOSPHATE 5-PHOSPHATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	309	Total	C	N	O	S	0	1	0
			2491	1599	422	456	14			
1	B	309	Total	C	N	O	S	0	1	0
			2480	1593	421	452	14			

There are 16 discrepancies between the modelled and reference sequences:

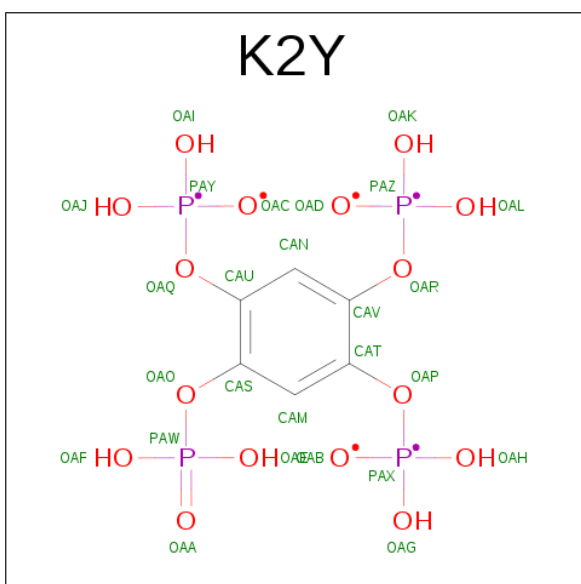
Chain	Residue	Modelled	Actual	Comment	Reference
A	258	MET	-	EXPRESSION TAG	UNP P32019
A	564	ALA	-	EXPRESSION TAG	UNP P32019
A	565	HIS	-	EXPRESSION TAG	UNP P32019
A	566	HIS	-	EXPRESSION TAG	UNP P32019
A	567	HIS	-	EXPRESSION TAG	UNP P32019
A	568	HIS	-	EXPRESSION TAG	UNP P32019
A	569	HIS	-	EXPRESSION TAG	UNP P32019
A	570	HIS	-	EXPRESSION TAG	UNP P32019
B	258	MET	-	EXPRESSION TAG	UNP P32019
B	564	ALA	-	EXPRESSION TAG	UNP P32019
B	565	HIS	-	EXPRESSION TAG	UNP P32019
B	566	HIS	-	EXPRESSION TAG	UNP P32019
B	567	HIS	-	EXPRESSION TAG	UNP P32019
B	568	HIS	-	EXPRESSION TAG	UNP P32019
B	569	HIS	-	EXPRESSION TAG	UNP P32019
B	570	HIS	-	EXPRESSION TAG	UNP P32019

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 3 is BENZENE 1,2,4,5-TETRAYL TETRAKIS[DIHYDROGEN (PHOSPHATE)] (three-letter code: K2Y) (formula: C<sub>6</sub>H<sub>10</sub>O<sub>16</sub>P<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	P	0	0
			26	6	16	4		
3	B	1	Total	C	O	P	0	0
			26	6	16	4		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

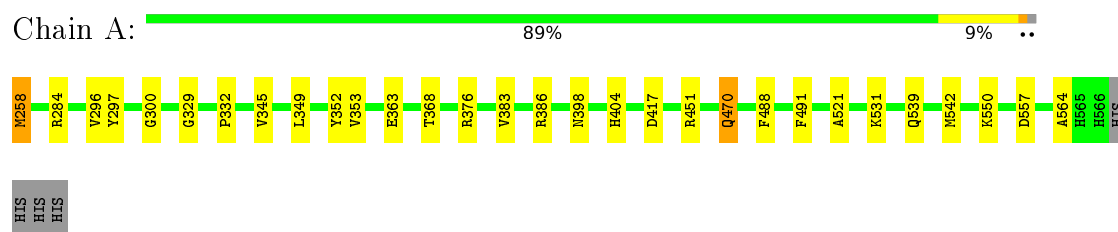
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	36	Total	O	0	0
			36	36		
5	B	40	Total	O	0	0
			40	40		

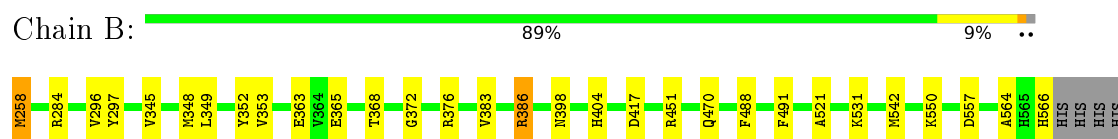
### 3 Residue-property plots

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: TYPE II INOSITOL 1,4,5-TRISPHOSPHATE 5-PHOSPHATASE



- Molecule 1: TYPE II INOSITOL 1,4,5-TRISPHOSPHATE 5-PHOSPHATASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.26Å 94.44Å 78.31Å 90.00° 106.32° 90.00°	Depositor
Resolution (Å)	48.48 – 2.90 48.48 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.48-2.90) 99.7 (48.48-2.90)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 2.91Å)	Xtriage
Refinement program	BUSTER-TNT BUSTER 2.11.1	Depositor
R, $R_{free}$	0.200 , 0.250 0.214 , 0.263	Depositor DCC
$R_{free}$ test set	813 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	34.1	Xtriage
Anisotropy	0.954	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 59.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 16223 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	5125	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.03 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.4262e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: GOL, K2Y, 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	0.43	0/2558	0.62	0/3465
1	B	0.43	0/2547	0.62	0/3452
All	All	0.43	0/5105	0.62	0/6917

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

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	2491	0	2417	13	0
1	B	2480	0	2400	12	0
2	A	5	0	0	0	0
2	B	15	0	0	1	0
3	A	26	0	0	0	0
3	B	26	0	0	0	0
4	B	6	0	8	0	0
5	A	36	0	0	0	0
5	B	40	0	0	0	0
All	All	5125	0	4825	25	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 (25) 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:284:ARG:HH22	1:A:332:PRO:HG3	1.71	0.54
1:B:348:MET:O	1:B:398:ASN:ND2	2.42	0.53
1:A:258:MET:HB2	1:A:564:ALA:O	2.11	0.51
1:B:258:MET:HB2	1:B:564:ALA:O	2.11	0.50
1:B:368:THR:HG22	1:B:383:VAL:HG12	1.94	0.50
1:A:368:THR:HG22	1:A:383:VAL:HG12	1.95	0.49
1:B:542:MET:HA	1:B:550:LYS:HD2	1.96	0.48
1:A:542:MET:HA	1:A:550:LYS:HD2	1.96	0.47
1:A:296:VAL:HG22	1:A:353:VAL:HG12	1.99	0.45
1:B:296:VAL:HG22	1:B:353:VAL:HG12	1.98	0.44
1:A:349:LEU:HD13	1:A:398:ASN:HB2	1.99	0.44
1:A:404:HIS:H	1:A:451:ARG:NH2	2.15	0.44
1:B:566:HIS:HB2	2:B:1568:SO4:S	2.58	0.44
1:A:470:GLN:H	1:A:470:GLN:HG3	1.62	0.43
1:B:404:HIS:H	1:B:451:ARG:NH2	2.16	0.43
1:B:297:TYR:HB2	1:B:352:TYR:HB2	2.00	0.43
1:A:488:PHE:HB3	1:A:491:PHE:HB2	2.00	0.43
1:A:300:GLY:HA3	1:A:398:ASN:HD22	1.83	0.43
1:B:488:PHE:HB3	1:B:491:PHE:HB2	2.00	0.43
1:A:297:TYR:HB2	1:A:352:TYR:HB2	2.01	0.42
1:B:349:LEU:HD13	1:B:398:ASN:HB2	2.00	0.42
1:B:451:ARG:HA	1:B:521:ALA:HB1	2.01	0.42
1:B:365:GLU:HB2	1:B:386:ARG:HG3	2.01	0.42
1:A:451:ARG:HA	1:A:521:ALA:HB1	2.01	0.41
1:A:284:ARG:HG3	1:A:329:GLY:HA2	2.02	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	308/313 (98%)	285 (92%)	22 (7%)	1 (0%)	46	79
1	B	308/313 (98%)	286 (93%)	20 (6%)	2 (1%)	30	67
All	All	616/626 (98%)	571 (93%)	42 (7%)	3 (0%)	34	71

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	345	VAL
1	B	345	VAL
1	B	372	GLY

### 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	266/276 (96%)	256 (96%)	10 (4%)	40	76
1	B	263/276 (95%)	253 (96%)	10 (4%)	40	76
All	All	529/552 (96%)	509 (96%)	20 (4%)	44	76

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	258	MET
1	A	363	GLU
1	A	376[A]	ARG
1	A	376[B]	ARG
1	A	386	ARG
1	A	417	ASP
1	A	470	GLN
1	A	531	LYS
1	A	539	GLN
1	A	557	ASP
1	B	258	MET
1	B	284	ARG
1	B	363	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	376[A]	ARG
1	B	376[B]	ARG
1	B	386	ARG
1	B	417	ASP
1	B	470	GLN
1	B	531	LYS
1	B	557	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	302	GLN
1	B	302	GLN
1	B	398	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

7 ligands 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
2	SO4	A	1567	-	4,4,4	0.22	0	6,6,6	0.10	0
3	K2Y	A	1568	-	25,26,26	1.95	14 (56%)	22,42,42	1.20	2 (9%)
4	GOL	B	1567	-	5,5,5	0.59	0	5,5,5	0.89	0
2	SO4	B	1568	-	4,4,4	0.26	0	6,6,6	0.17	0
2	SO4	B	1569	-	4,4,4	0.14	0	6,6,6	0.18	0
3	K2Y	B	1570	-	25,26,26	1.87	11 (44%)	22,42,42	1.15	4 (18%)
2	SO4	B	1571	-	4,4,4	0.15	0	6,6,6	0.07	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
2	SO4	A	1567	-	-	0/0/0/0	0/0/0/0
3	K2Y	A	1568	-	-	0/11/20/20	0/1/1/1
4	GOL	B	1567	-	-	0/4/4/4	0/0/0/0
2	SO4	B	1568	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1569	-	-	0/0/0/0	0/0/0/0
3	K2Y	B	1570	-	-	0/11/20/20	0/1/1/1
2	SO4	B	1571	-	-	0/0/0/0	0/0/0/0

All (25) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1568	K2Y	OAR-CAV	-2.49	1.35	1.40
3	A	1568	K2Y	PAY-OAC	-2.44	1.47	1.54
3	B	1570	K2Y	PAX-OAB	-2.35	1.48	1.54
3	B	1570	K2Y	PAY-OAC	-2.26	1.48	1.54
3	A	1568	K2Y	PAZ-OAD	-2.24	1.48	1.54
3	A	1568	K2Y	PAX-OAB	-2.19	1.48	1.54
3	B	1570	K2Y	PAZ-OAD	-2.11	1.48	1.54
3	B	1570	K2Y	OAO-CAS	-2.01	1.36	1.40
3	A	1568	K2Y	PAX-OAP	2.10	1.63	1.58
3	A	1568	K2Y	PAX-OAG	2.11	1.61	1.54
3	A	1568	K2Y	CAM-CAT	2.12	1.42	1.38
3	A	1568	K2Y	PAW-OAF	2.12	1.62	1.54
3	B	1570	K2Y	PAX-OAP	2.16	1.63	1.58
3	B	1570	K2Y	PAZ-OAK	2.20	1.61	1.54
3	B	1570	K2Y	PAY-OAI	2.26	1.61	1.54
3	B	1570	K2Y	PAW-OAF	2.29	1.62	1.54
3	A	1568	K2Y	PAY-OAJ	2.33	1.61	1.54

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1568	K2Y	PAY-OAQ	2.34	1.64	1.58
3	A	1568	K2Y	PAY-OAI	2.36	1.61	1.54
3	B	1570	K2Y	PAY-OAJ	2.37	1.61	1.54
3	B	1570	K2Y	PAX-OAH	2.46	1.62	1.54
3	B	1570	K2Y	PAZ-OAL	2.46	1.62	1.54
3	A	1568	K2Y	PAZ-OAK	2.54	1.62	1.54
3	A	1568	K2Y	PAX-OAH	2.59	1.62	1.54
3	A	1568	K2Y	PAZ-OAL	2.63	1.62	1.54

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1568	K2Y	PAZ-OAR-CAV	-3.57	114.12	124.40
3	B	1570	K2Y	OAP-CAT-CAM	-2.80	117.07	121.23
3	A	1568	K2Y	PAY-OAQ-CAU	-2.32	117.72	124.40
3	B	1570	K2Y	PAY-OAQ-CAU	-2.19	118.09	124.40
3	B	1570	K2Y	OAO-CAS-CAU	2.04	119.84	116.73
3	B	1570	K2Y	OAP-CAT-CAV	2.06	122.33	117.60

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
2	B	1568	SO4	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	309/313 (98%)	-0.20	0 100 100	17, 36, 55, 82	0
1	B	309/313 (98%)	-0.14	0 100 100	22, 38, 58, 77	0
All	All	618/626 (98%)	-0.17	0 100 100	17, 37, 57, 82	0

There are no RSRZ outliers to report.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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	GOL	B	1567	6/6	0.89	0.23	1.87	45,46,46,46	0
3	K2Y	A	1568	26/26	0.92	0.21	0.78	48,58,60,61	0
3	K2Y	B	1570	26/26	0.86	0.25	0.74	67,74,83,84	0
2	SO4	B	1569	5/5	0.96	0.15	-0.87	45,49,50,50	0
2	SO4	B	1568	5/5	0.96	0.16	-1.37	51,55,56,57	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
2	SO4	A	1567	5/5	0.94	0.12	-	72,76,77,78	0
2	SO4	B	1571	5/5	0.92	0.13	-	91,95,96,97	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.