



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 4, 2024 – 09:06 AM EST

PDB ID : 2RFQ  
Title : Crystal structure of 3-HSA hydroxylase from Rhodococcus sp. RHA1  
Authors : Chang, C.; Skarina, T.; Kagan, O.; Savchenko, A.; Edwards, A.M.;  
Joachimciak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2007-10-01  
Resolution : 1.65 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

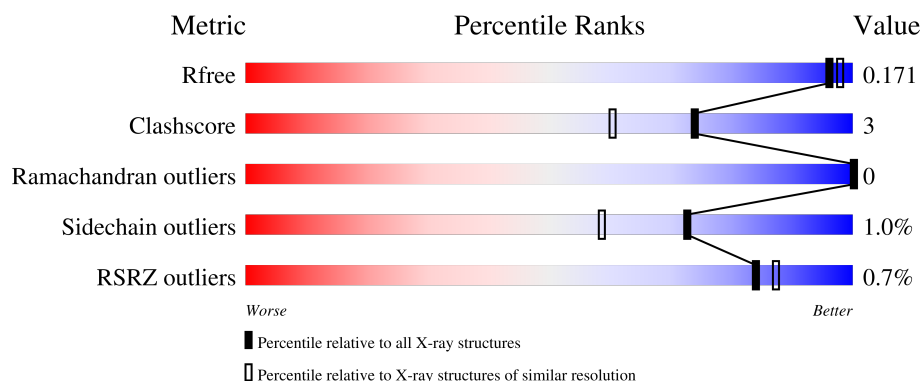
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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))
$R_{free}$	164625	2328 (1.66-1.66)
Clashscore	180529	2515 (1.66-1.66)
Ramachandran outliers	177936	2475 (1.66-1.66)
Sidechain outliers	177891	2475 (1.66-1.66)
RSRZ outliers	164620	2328 (1.66-1.66)

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 of 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	394	<div> <div>%</div> <div>92% 7% .</div> </div>
1	B	394	<div> <div>%</div> <div>91% 7% ..</div> </div>
1	C	394	<div> <div>88% 7% ..</div> </div>
1	D	394	<div> <div>%</div> <div>91% 5% .</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13979 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 3-HSA hydroxylase, oxygenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	388	Total	C	N	O	S	Se	0	8	0
			3038	1915	540	573	2	8			
1	B	387	Total	C	N	O	S	Se	0	9	0
			3039	1915	539	575	2	8			
1	C	378	Total	C	N	O	S	Se	0	6	0
			2957	1863	524	560	2	8			
1	D	379	Total	C	N	O	S	Se	0	12	0
			2991	1888	524	569	2	8			

There are 48 discrepancies between the modelled and reference sequences:

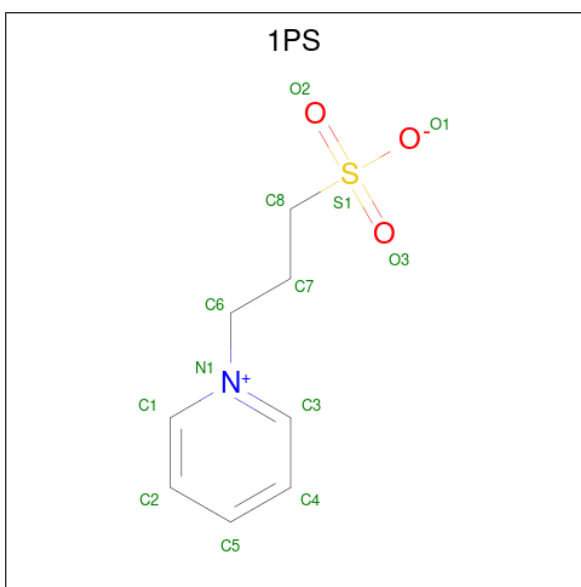
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLN	-	expression tag	UNP Q0S811
A	-1	GLY	-	expression tag	UNP Q0S811
A	0	HIS	-	expression tag	UNP Q0S811
A	1	VAL	-	expression tag	UNP Q0S811
A	2	GLY	-	expression tag	UNP Q0S811
A	3	ASP	-	expression tag	UNP Q0S811
A	4	HIS	-	expression tag	UNP Q0S811
A	5	ASP	-	expression tag	UNP Q0S811
A	6	SER	-	expression tag	UNP Q0S811
A	7	HIS	-	expression tag	UNP Q0S811
A	8	GLU	-	expression tag	UNP Q0S811
A	9	VAL	-	expression tag	UNP Q0S811
B	-2	GLN	-	expression tag	UNP Q0S811
B	-1	GLY	-	expression tag	UNP Q0S811
B	0	HIS	-	expression tag	UNP Q0S811
B	1	VAL	-	expression tag	UNP Q0S811
B	2	GLY	-	expression tag	UNP Q0S811
B	3	ASP	-	expression tag	UNP Q0S811
B	4	HIS	-	expression tag	UNP Q0S811
B	5	ASP	-	expression tag	UNP Q0S811
B	6	SER	-	expression tag	UNP Q0S811

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	7	HIS	-	expression tag	UNP Q0S811
B	8	GLU	-	expression tag	UNP Q0S811
B	9	VAL	-	expression tag	UNP Q0S811
C	-2	GLN	-	expression tag	UNP Q0S811
C	-1	GLY	-	expression tag	UNP Q0S811
C	0	HIS	-	expression tag	UNP Q0S811
C	1	VAL	-	expression tag	UNP Q0S811
C	2	GLY	-	expression tag	UNP Q0S811
C	3	ASP	-	expression tag	UNP Q0S811
C	4	HIS	-	expression tag	UNP Q0S811
C	5	ASP	-	expression tag	UNP Q0S811
C	6	SER	-	expression tag	UNP Q0S811
C	7	HIS	-	expression tag	UNP Q0S811
C	8	GLU	-	expression tag	UNP Q0S811
C	9	VAL	-	expression tag	UNP Q0S811
D	-2	GLN	-	expression tag	UNP Q0S811
D	-1	GLY	-	expression tag	UNP Q0S811
D	0	HIS	-	expression tag	UNP Q0S811
D	1	VAL	-	expression tag	UNP Q0S811
D	2	GLY	-	expression tag	UNP Q0S811
D	3	ASP	-	expression tag	UNP Q0S811
D	4	HIS	-	expression tag	UNP Q0S811
D	5	ASP	-	expression tag	UNP Q0S811
D	6	SER	-	expression tag	UNP Q0S811
D	7	HIS	-	expression tag	UNP Q0S811
D	8	GLU	-	expression tag	UNP Q0S811
D	9	VAL	-	expression tag	UNP Q0S811

- Molecule 2 is 3-PYRIDINIUM-1-YLPROPANE-1-SULFONATE (three-letter code: 1PS) (formula: C<sub>8</sub>H<sub>11</sub>NO<sub>3</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
2	A	1	Total	C	N	O	S	9	0
			13	8	1	3	1		
2	B	1	Total	C	N	O	S	5	0
			13	8	1	3	1		
2	B	1	Total	O	S			2	0
			2	1	1				
2	B	1	Total	C	N	O		5	0
			11	8	1	2			
2	C	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
2	C	1	Total	C	N	O	S	7	0
			13	8	1	3	1		
2	C	1	Total	C	N	O	S	9	0
			13	8	1	3	1		
2	C	1	Total	C	N	O	S	7	0
			13	8	1	3	1		
2	D	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
2	D	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
2	D	1	Total	C	N	O	S	9	0
			13	8	1	3	1		
2	D	1	Total	C	N	O	S	7	0
			13	8	1	3	1		

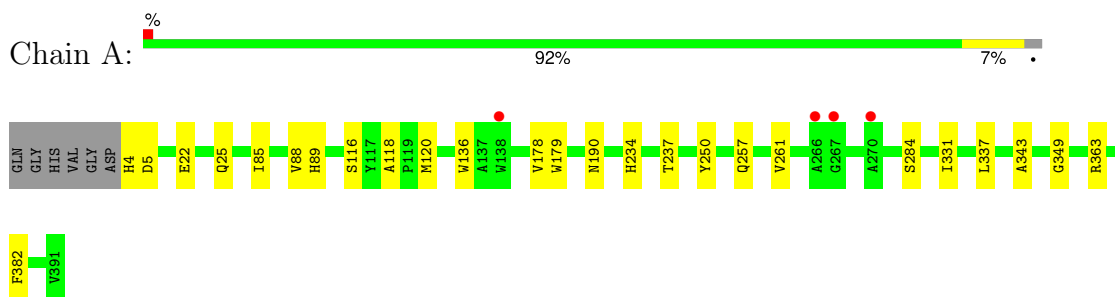
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	455	Total 455	O 455	0	0
3	B	442	Total 442	O 442	0	0
3	C	449	Total 449	O 449	0	0
3	D	452	Total 452	O 452	0	0

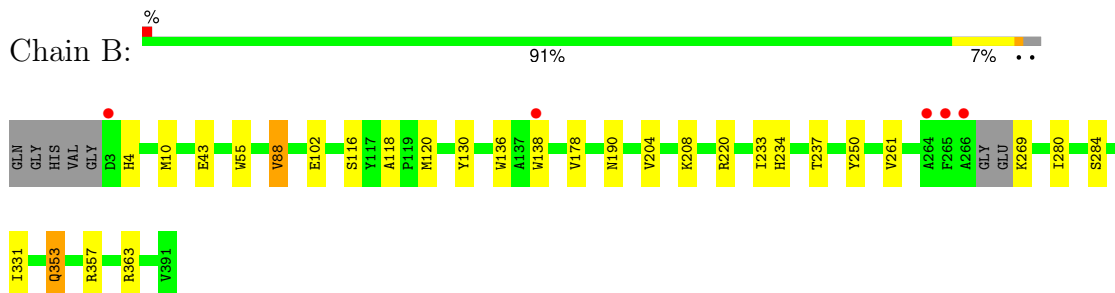
### 3 Residue-property plots

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.

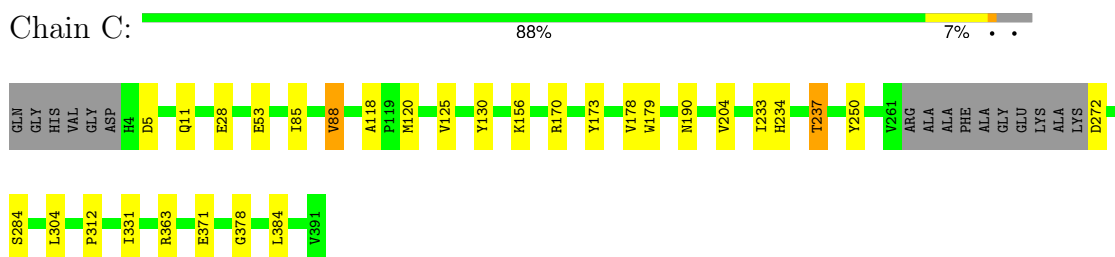
- Molecule 1: 3-HSA hydroxylase, oxygenase



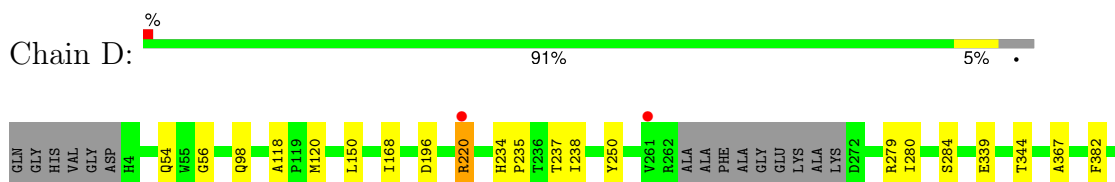
- Molecule 1: 3-HSA hydroxylase, oxygenase



- Molecule 1: 3-HSA hydroxylase, oxygenase



- Molecule 1: 3-HSA hydroxylase, oxygenase



1584



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.28Å 76.14Å 80.88Å 83.78° 75.95° 89.99°	Depositor
Resolution (Å)	50.00 – 1.65 50.00 – 1.65	Depositor EDS
% Data completeness (in resolution range)	97.2 (50.00-1.65) 97.2 (50.00-1.65)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.13 (at 1.65Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.146 , 0.171 0.145 , 0.171	Depositor DCC
$R_{free}$ test set	9922 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.0	Xtriage
Anisotropy	0.251	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 47.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13979	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.17% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 1PS

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	0.60	0/3124	0.64	0/4240
1	B	0.59	0/3133	0.65	0/4251
1	C	0.60	0/3041	0.64	0/4128
1	D	0.58	0/3087	0.66	0/4191
All	All	0.59	0/12385	0.65	0/16810

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	3038	0	2928	18	0
1	B	3039	0	2932	18	0
1	C	2957	0	2848	21	0
1	D	2991	0	2880	23	0
2	A	26	0	22	0	0
2	B	26	0	20	0	0
2	C	52	0	44	0	0
2	D	52	0	44	1	0
3	A	455	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	442	0	0	3	1
3	C	449	0	0	4	1
3	D	452	0	0	6	0
All	All	13979	0	11718	71	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.

The worst 5 of 71 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:220:ARG:HH11	1:D:220:ARG:HG3	1.10	1.06
1:D:339[B]:GLU:HG2	3:D:498:HOH:O	1.58	1.02
1:D:220:ARG:HG3	1:D:220:ARG:NH1	1.79	0.89
1:C:371:GLU:OE2	1:D:279[B]:ARG:NH2	2.05	0.88
1:B:233:ILE:O	1:B:237[A]:THR:HG23	1.77	0.84

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)
3:B:835:HOH:O	3:C:793:HOH:O[1_565]	2.03	0.17

## 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	394/394 (100%)	386 (98%)	8 (2%)	0	100	100
1	B	392/394 (100%)	387 (99%)	5 (1%)	0	100	100
1	C	380/394 (96%)	375 (99%)	5 (1%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	387/394 (98%)	382 (99%)	5 (1%)	0	100	100
All	All	1553/1576 (98%)	1530 (98%)	23 (2%)	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	313/305 (103%)	309 (99%)	4 (1%)	65	47
1	B	315/305 (103%)	309 (98%)	6 (2%)	52	30
1	C	308/305 (101%)	304 (99%)	4 (1%)	65	47
1	D	312/305 (102%)	309 (99%)	3 (1%)	73	59
All	All	1248/1220 (102%)	1231 (99%)	17 (1%)	73	43

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	196[A]	ASP
1	D	220	ARG
1	B	220[A]	ARG
1	B	220[B]	ARG
1	B	353	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	234	HIS
1	D	51	GLN
1	C	368	ASN
1	D	98	GLN
1	B	107	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 oligosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

13 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	1PS	D	393	-	13,13,13	2.22	1 (7%)	17,17,17	0.92	1 (5%)
2	1PS	D	394	-	13,13,13	2.29	1 (7%)	17,17,17	2.33	3 (17%)
2	1PS	D	392	-	13,13,13	2.06	1 (7%)	17,17,17	1.06	2 (11%)
2	1PS	B	393	-	0,1,13	-	-	-	-	-
2	1PS	C	392	-	13,13,13	2.08	1 (7%)	17,17,17	1.05	2 (11%)
2	1PS	C	395	-	13,13,13	2.16	2 (15%)	17,17,17	1.64	3 (17%)
2	1PS	B	392	-	13,13,13	1.99	1 (7%)	17,17,17	1.23	2 (11%)
2	1PS	A	393	-	13,13,13	1.08	1 (7%)	17,17,17	3.24	4 (23%)
2	1PS	A	392	-	13,13,13	1.97	1 (7%)	17,17,17	1.12	1 (5%)
2	1PS	D	395	-	13,13,13	2.32	2 (15%)	17,17,17	1.81	3 (17%)
2	1PS	C	393	-	13,13,13	2.18	1 (7%)	17,17,17	1.57	3 (17%)
2	1PS	B	394	-	9,9,13	2.64	1 (11%)	10,10,17	5.97	2 (20%)
2	1PS	C	394	-	13,13,13	1.12	1 (7%)	17,17,17	1.93	5 (29%)

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	1PS	D	393	-	-	0/7/7/7	0/1/1/1
2	1PS	D	394	-	-	2/7/7/7	0/1/1/1
2	1PS	D	392	-	-	0/7/7/7	0/1/1/1
2	1PS	C	392	-	-	0/7/7/7	0/1/1/1
2	1PS	C	395	-	-	2/7/7/7	0/1/1/1
2	1PS	B	392	-	-	1/7/7/7	0/1/1/1
2	1PS	A	393	-	-	2/7/7/7	0/1/1/1
2	1PS	A	392	-	-	0/7/7/7	0/1/1/1
2	1PS	D	395	-	-	1/7/7/7	0/1/1/1
2	1PS	C	393	-	-	2/7/7/7	0/1/1/1
2	1PS	B	394	-	-	2/3/3/7	0/1/1/1
2	1PS	C	394	-	-	2/7/7/7	0/1/1/1

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	394	1PS	C6-N1	7.70	1.67	1.49
2	D	394	1PS	C8-S1	-7.64	1.66	1.77
2	D	393	1PS	C8-S1	-7.47	1.67	1.77
2	C	393	1PS	C8-S1	-7.43	1.67	1.77
2	C	392	1PS	C8-S1	-6.97	1.67	1.77

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	394	1PS	C6-N1-C3	-13.79	90.17	119.75
2	B	394	1PS	C6-N1-C1	12.74	147.09	119.75
2	A	393	1PS	O1-S1-C8	9.25	124.09	106.00
2	A	393	1PS	O3-S1-C8	-8.50	93.88	106.73
2	D	394	1PS	O2-S1-C8	8.11	118.98	106.73

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	394	1PS	C7-C6-N1-C1
2	A	393	1PS	C7-C6-N1-C3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	C	394	1PS	C7-C6-N1-C3
2	C	393	1PS	C7-C6-N1-C1
2	D	395	1PS	C6-C7-C8-S1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	395	1PS	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	380/394 (96%)	-0.51	4 (1%) 77 81	6, 10, 20, 33	8 (2%)
1	B	379/394 (96%)	-0.36	5 (1%) 74 79	4, 10, 19, 32	9 (2%)
1	C	370/394 (93%)	-0.64	0 100 100	4, 10, 18, 33	6 (1%)
1	D	371/394 (94%)	-0.41	2 (0%) 87 90	3, 11, 19, 34	12 (3%)
All	All	1500/1576 (95%)	-0.48	11 (0%) 84 87	3, 10, 19, 34	35 (2%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	266	ALA	4.4
1	B	3	ASP	3.0
1	A	138	TRP	2.8
1	B	266	ALA	2.5
1	B	138	TRP	2.4

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

LIGAND-RSR INFOmissingINFO



## 6.5 Other polymers [i](#)

There are no such residues in this entry.