



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 27, 2020 – 02:20 am BST

PDB ID : 5Y5V  
Title : Crystal structure of a novel Pyrethroid Hydrolase PytH (S78A)  
Authors : Xu, D.Q.; Ran, T.T.; He, J.; Wang, W.W.  
Deposited on : 2017-08-09  
Resolution : 2.98 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 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

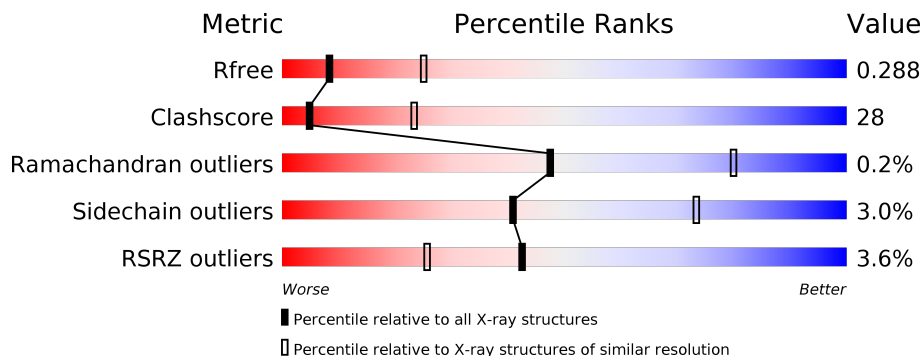
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.98 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	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

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	288	
1	B	288	
1	C	288	
1	D	288	
1	E	288	
1	F	288	

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
2	SO4	F	301	-	-	X	-

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 11233 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 Pyrethroid hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	251	1891	1207	327	350	7	0	0	0
1	B	252	1891	1207	322	355	7	0	0	0
1	C	251	1884	1204	321	352	7	0	0	0
1	D	251	1891	1206	326	352	7	0	0	0
1	E	251	1817	1159	310	342	6	0	0	0
1	F	249	1786	1141	307	332	6	0	0	0

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	78	ALA	SER	engineered mutation	UNP D0VUS3
A	281	LEU	-	expression tag	UNP D0VUS3
A	282	GLU	-	expression tag	UNP D0VUS3
A	283	HIS	-	expression tag	UNP D0VUS3
A	284	HIS	-	expression tag	UNP D0VUS3
A	285	HIS	-	expression tag	UNP D0VUS3
A	286	HIS	-	expression tag	UNP D0VUS3
A	287	HIS	-	expression tag	UNP D0VUS3
A	288	HIS	-	expression tag	UNP D0VUS3
B	78	ALA	SER	engineered mutation	UNP D0VUS3
B	281	LEU	-	expression tag	UNP D0VUS3
B	282	GLU	-	expression tag	UNP D0VUS3
B	283	HIS	-	expression tag	UNP D0VUS3
B	284	HIS	-	expression tag	UNP D0VUS3
B	285	HIS	-	expression tag	UNP D0VUS3
B	286	HIS	-	expression tag	UNP D0VUS3
B	287	HIS	-	expression tag	UNP D0VUS3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	288	HIS	-	expression tag	UNP D0VUS3
C	78	ALA	SER	engineered mutation	UNP D0VUS3
C	281	LEU	-	expression tag	UNP D0VUS3
C	282	GLU	-	expression tag	UNP D0VUS3
C	283	HIS	-	expression tag	UNP D0VUS3
C	284	HIS	-	expression tag	UNP D0VUS3
C	285	HIS	-	expression tag	UNP D0VUS3
C	286	HIS	-	expression tag	UNP D0VUS3
C	287	HIS	-	expression tag	UNP D0VUS3
C	288	HIS	-	expression tag	UNP D0VUS3
D	78	ALA	SER	engineered mutation	UNP D0VUS3
D	281	LEU	-	expression tag	UNP D0VUS3
D	282	GLU	-	expression tag	UNP D0VUS3
D	283	HIS	-	expression tag	UNP D0VUS3
D	284	HIS	-	expression tag	UNP D0VUS3
D	285	HIS	-	expression tag	UNP D0VUS3
D	286	HIS	-	expression tag	UNP D0VUS3
D	287	HIS	-	expression tag	UNP D0VUS3
D	288	HIS	-	expression tag	UNP D0VUS3
E	78	ALA	SER	engineered mutation	UNP D0VUS3
E	281	LEU	-	expression tag	UNP D0VUS3
E	282	GLU	-	expression tag	UNP D0VUS3
E	283	HIS	-	expression tag	UNP D0VUS3
E	284	HIS	-	expression tag	UNP D0VUS3
E	285	HIS	-	expression tag	UNP D0VUS3
E	286	HIS	-	expression tag	UNP D0VUS3
E	287	HIS	-	expression tag	UNP D0VUS3
E	288	HIS	-	expression tag	UNP D0VUS3
F	78	ALA	SER	engineered mutation	UNP D0VUS3
F	281	LEU	-	expression tag	UNP D0VUS3
F	282	GLU	-	expression tag	UNP D0VUS3
F	283	HIS	-	expression tag	UNP D0VUS3
F	284	HIS	-	expression tag	UNP D0VUS3
F	285	HIS	-	expression tag	UNP D0VUS3
F	286	HIS	-	expression tag	UNP D0VUS3
F	287	HIS	-	expression tag	UNP D0VUS3
F	288	HIS	-	expression tag	UNP D0VUS3

- 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	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		

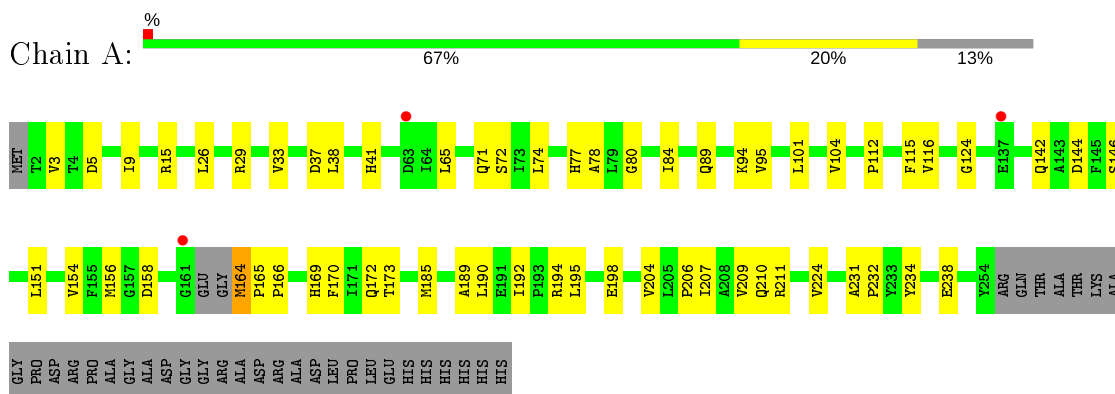
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	15	Total	O	0	0
			15	15		
3	B	8	Total	O	0	0
			8	8		
3	C	9	Total	O	0	0
			9	9		
3	D	8	Total	O	0	0
			8	8		
3	E	2	Total	O	0	0
			2	2		
3	F	1	Total	O	0	0
			1	1		

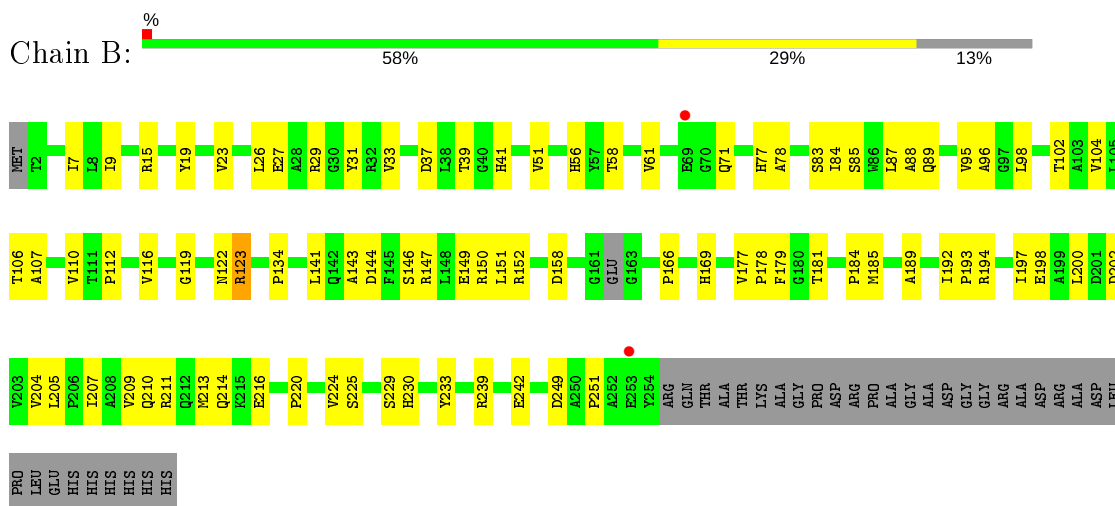
### 3 Residue-property plots [i](#)

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.

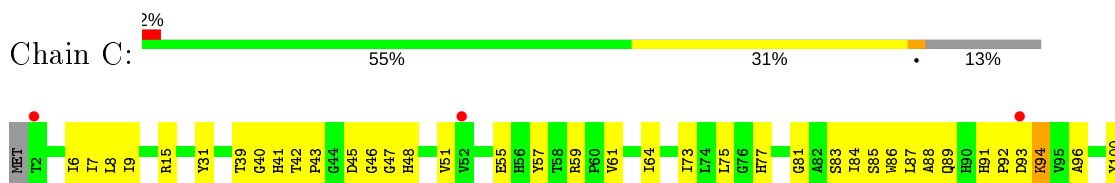
- Molecule 1: Pyrethroid hydrolase

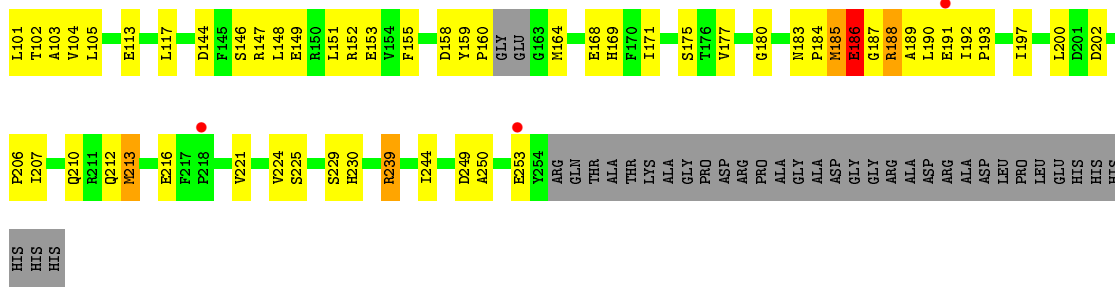


- Molecule 1: Pyrethroid hydrolase

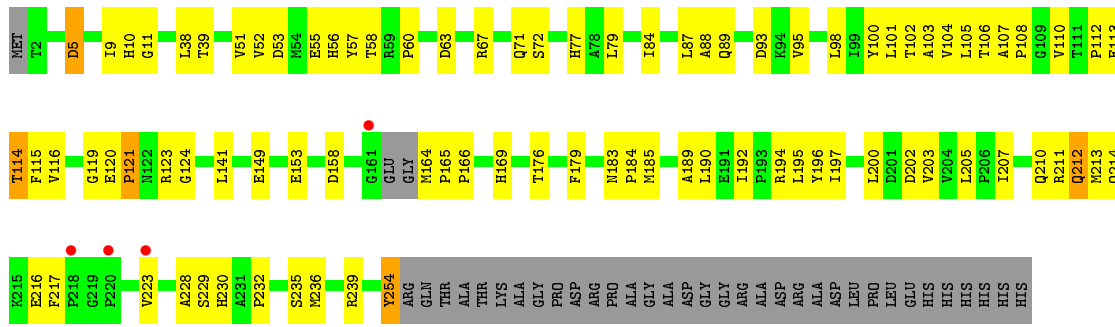


- Molecule 1: Pyrethroid hydrolase

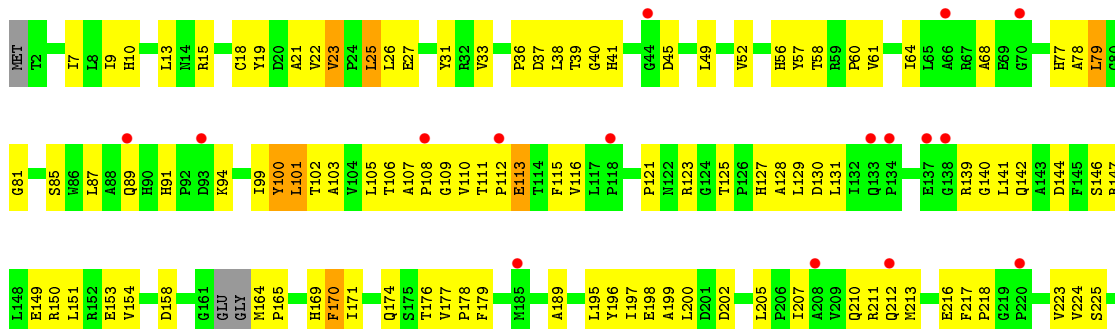




• Molecule 1: Pyrethroid hydrolase



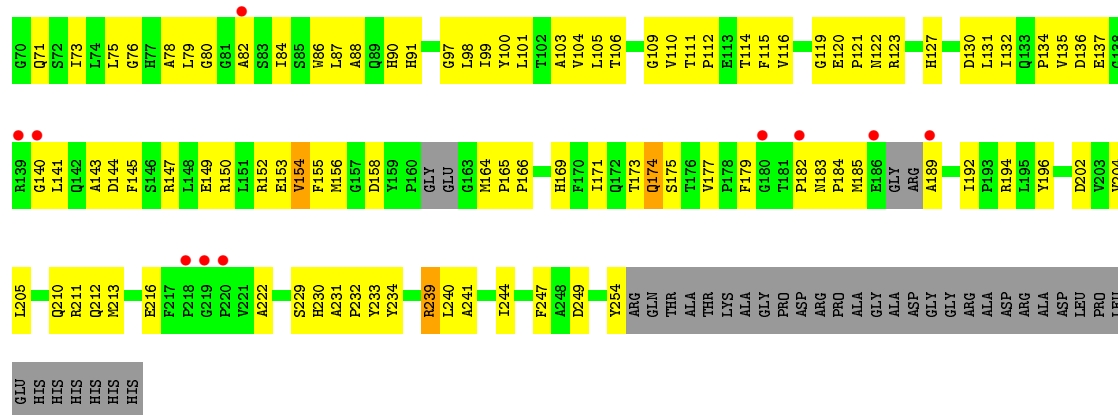
• Molecule 1: Pyrethroid hydrolase



• Molecule 1: Pyrethroid hydrolase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	168.20Å 168.20Å 123.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.82 – 2.98 19.82 – 2.98	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.82-2.98) 100.0 (19.82-2.98)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.03 (at 2.98Å)	Xtrriage
Refinement program	PHENIX (dev_2247: ???)	Depositor
R, $R_{free}$	0.216 , 0.291 0.225 , 0.288	Depositor DCC
$R_{free}$ test set	1756 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.1	Xtrriage
Anisotropy	0.932	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 48.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11233	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.10% 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: 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	0.54	0/1943	0.72	0/2662
1	B	0.50	0/1943	0.68	0/2663
1	C	0.58	0/1936	0.70	0/2654
1	D	0.58	0/1943	0.71	0/2663
1	E	0.57	0/1869	0.73	0/2571
1	F	0.59	0/1835	0.79	0/2523
All	All	0.56	0/11469	0.72	0/15736

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	1891	0	1848	45	0
1	B	1891	0	1839	65	0
1	C	1884	0	1834	93	0
1	D	1891	0	1843	87	0
1	E	1817	0	1695	162	1
1	F	1786	0	1665	155	1
2	A	5	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	5	0	0	1	0
2	C	5	0	0	1	0
2	D	5	0	0	1	0
2	E	5	0	0	0	0
2	F	5	0	0	2	0
3	A	15	0	0	1	0
3	B	8	0	0	0	0
3	C	9	0	0	0	0
3	D	8	0	0	1	0
3	E	2	0	0	0	0
3	F	1	0	0	0	0
All	All	11233	0	10724	604	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 28.

The worst 5 of 604 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:E:78:ALA:N	1:E:103:ALA:HB2	1.42	1.31
1:E:78:ALA:O	1:E:103:ALA:HB1	1.42	1.16
1:E:23:VAL:HG13	1:E:33:VAL:HG11	1.31	1.10
1:D:115:PHE:CE2	1:D:213:MET:HG2	1.88	1.08
1:F:15:ARG:NE	1:F:37:ASP:OD2	1.87	1.07

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:E:149:GLU:OE2	1:F:164:MET:N[8_554]	2.02	0.18

## 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	247/288 (86%)	235 (95%)	12 (5%)	0	100	100
1	B	248/288 (86%)	235 (95%)	13 (5%)	0	100	100
1	C	247/288 (86%)	225 (91%)	20 (8%)	2 (1%)	19	55
1	D	247/288 (86%)	230 (93%)	17 (7%)	0	100	100
1	E	247/288 (86%)	229 (93%)	18 (7%)	0	100	100
1	F	243/288 (84%)	214 (88%)	28 (12%)	1 (0%)	34	70
All	All	1479/1728 (86%)	1368 (92%)	108 (7%)	3 (0%)	47	80

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	186	GLU
1	F	36	PRO
1	C	184	PRO

### 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	196/226 (87%)	195 (100%)	1 (0%)	88	95
1	B	196/226 (87%)	194 (99%)	2 (1%)	76	91
1	C	195/226 (86%)	188 (96%)	7 (4%)	35	68
1	D	196/226 (87%)	191 (97%)	5 (3%)	46	76
1	E	177/226 (78%)	167 (94%)	10 (6%)	21	54
1	F	171/226 (76%)	162 (95%)	9 (5%)	22	56
All	All	1131/1356 (83%)	1097 (97%)	34 (3%)	41	73

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	23	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	101	LEU
1	F	211	ARG
1	E	79	LEU
1	C	186	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	77	HIS
1	F	10	HIS

### 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](#)

6 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	F	301	-	4,4,4	0.12	0	6,6,6	0.32	0
2	SO4	A	301	-	4,4,4	0.23	0	6,6,6	0.14	0
2	SO4	D	301	-	4,4,4	0.12	0	6,6,6	0.45	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SO4	B	301	-	4,4,4	0.14	0	6,6,6	0.63	0
2	SO4	C	301	-	4,4,4	0.19	0	6,6,6	0.21	0
2	SO4	E	301	-	4,4,4	0.23	0	6,6,6	0.40	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	301	SO4	2	0
2	A	301	SO4	1	0
2	D	301	SO4	1	0
2	B	301	SO4	1	0
2	C	301	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	251/288 (87%)	-0.26	3 (1%) 79 61	25, 35, 50, 60	0
1	B	252/288 (87%)	-0.34	2 (0%) 86 71	23, 34, 47, 58	0
1	C	251/288 (87%)	-0.14	6 (2%) 59 39	23, 41, 57, 68	0
1	D	251/288 (87%)	-0.18	4 (1%) 72 52	25, 36, 51, 71	0
1	E	251/288 (87%)	0.46	16 (6%) 19 10	26, 52, 70, 86	0
1	F	249/288 (86%)	0.64	23 (9%) 9 5	26, 68, 88, 100	0
All	All	1505/1728 (87%)	0.03	54 (3%) 42 26	23, 41, 78, 100	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	44	GLY	5.8
1	F	140	GLY	4.7
1	F	60	PRO	4.5
1	E	66	ALA	4.0
1	C	2	THR	4.0

### 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. 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
2	SO4	F	301	5/5	0.94	0.17	43,44,55,55	0
2	SO4	C	301	5/5	0.95	0.20	41,45,51,57	0
2	SO4	D	301	5/5	0.96	0.11	47,53,56,60	0
2	SO4	B	301	5/5	0.97	0.20	37,42,45,51	0
2	SO4	A	301	5/5	0.98	0.10	36,40,50,51	0
2	SO4	E	301	5/5	0.98	0.10	41,43,53,53	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.