



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 17, 2024 – 08:09 PM EDT

PDB ID : 5Y51
Title : Crystal structure of PytH_H230A
Authors : Xu, D.Q.; Ran, T.T.; He, J.; Wang, W.W.
Deposited on : 2017-08-06
Resolution : 2.30 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

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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

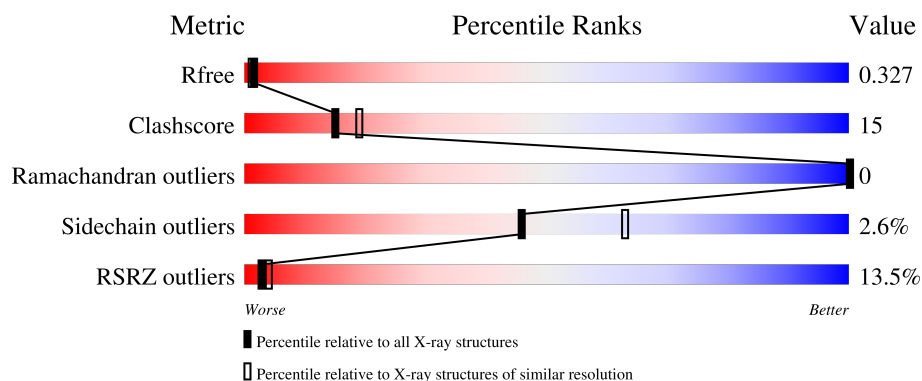
1 Overall quality at a glance

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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (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 of 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 $\leq 5\%$. 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	<div> <div>2%</div> <div>70%16%14%</div> </div>
1	B	288	<div> <div>3%</div> <div>67%17%14%</div> </div>
1	C	288	<div> <div>3%</div> <div>70%17%13%</div> </div>
1	D	288	<div> <div>3%</div> <div>72%14%13%</div> </div>
1	E	288	<div> <div>18%</div> <div>55%31%13%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	288	<div><div></div><div>41%</div><div>53%</div><div>32%</div><div>14%</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11414 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
1	A	248	Total	C	N	O	S	0	0	0
			1888	1204	325	352	7			
1	B	249	Total	C	N	O	S	0	0	0
			1892	1206	326	353	7			
1	C	251	Total	C	N	O	S	0	0	0
			1899	1210	325	357	7			
1	D	251	Total	C	N	O	S	0	0	0
			1899	1210	325	357	7			
1	E	250	Total	C	N	O	S	0	0	0
			1885	1200	324	354	7			
1	F	248	Total	C	N	O	S	0	0	0
			1888	1204	325	352	7			

There are 54 discrepancies between the modelled and reference sequences:

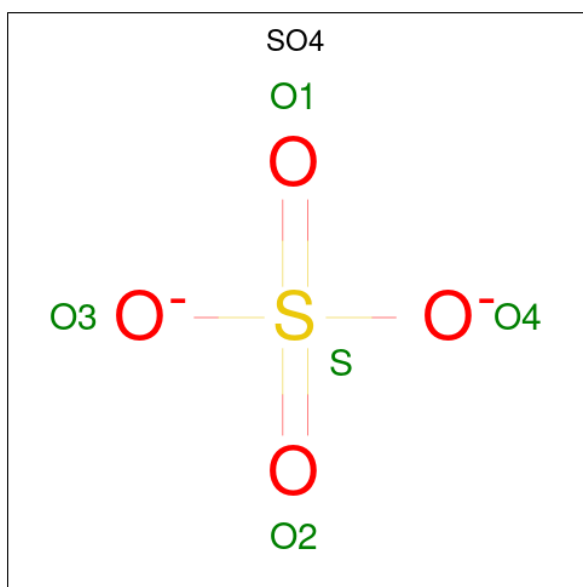
Chain	Residue	Modelled	Actual	Comment	Reference
A	230	ALA	HIS	engineered mutation	UNP D0VUS3
A	281	LEU	-	expression tag	UNP D0VUS3
A	282	GLU	-	expression tag	UNP D0VUS3
A	283	HIS	-	expression tag	UNP D0VUS3
A	284	HIS	-	expression tag	UNP D0VUS3
A	285	HIS	-	expression tag	UNP D0VUS3
A	286	HIS	-	expression tag	UNP D0VUS3
A	287	HIS	-	expression tag	UNP D0VUS3
A	288	HIS	-	expression tag	UNP D0VUS3
B	230	ALA	HIS	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	230	ALA	HIS	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	230	ALA	HIS	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	230	ALA	HIS	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	230	ALA	HIS	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: O₄S).



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		

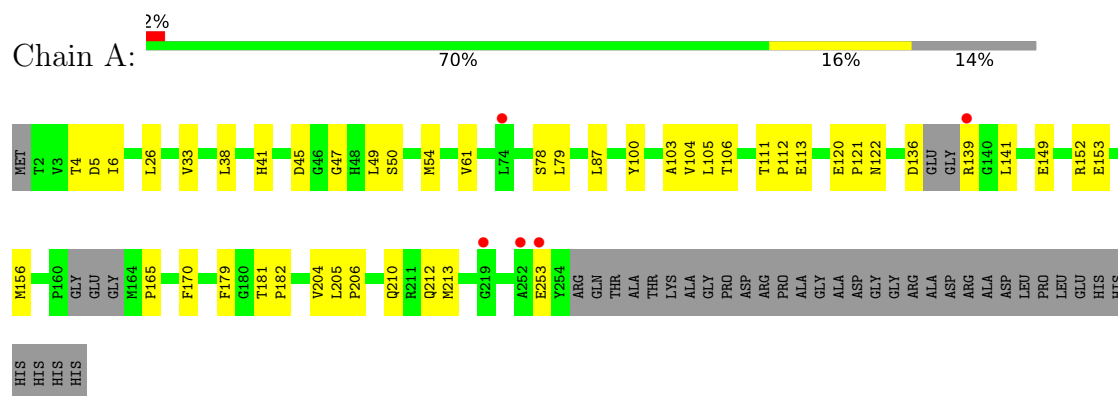
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	9	Total	O	0	0
			9	9		
3	B	7	Total	O	0	0
			7	7		
3	C	3	Total	O	0	0
			3	3		
3	D	9	Total	O	0	0
			9	9		
3	E	4	Total	O	0	0
			4	4		
3	F	1	Total	O	0	0
			1	1		

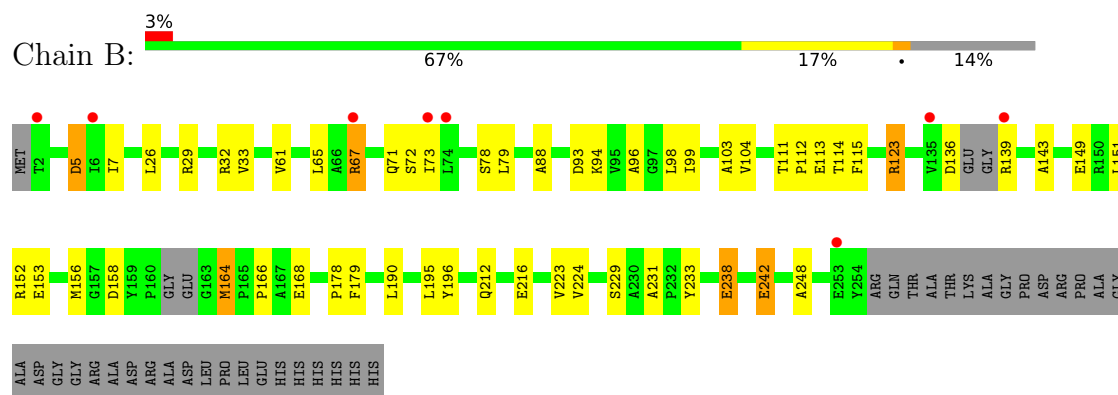
3 Residue-property plots [i](#)

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.

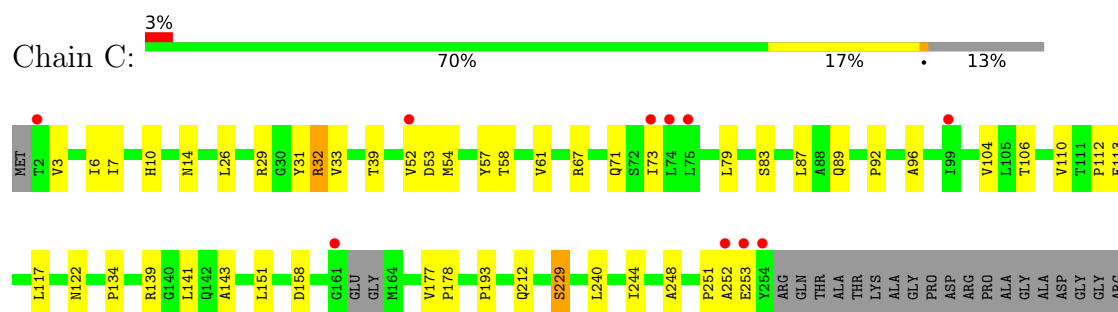
• Molecule 1: Pyrethroid hydrolase



• Molecule 1: Pyrethroid hydrolase



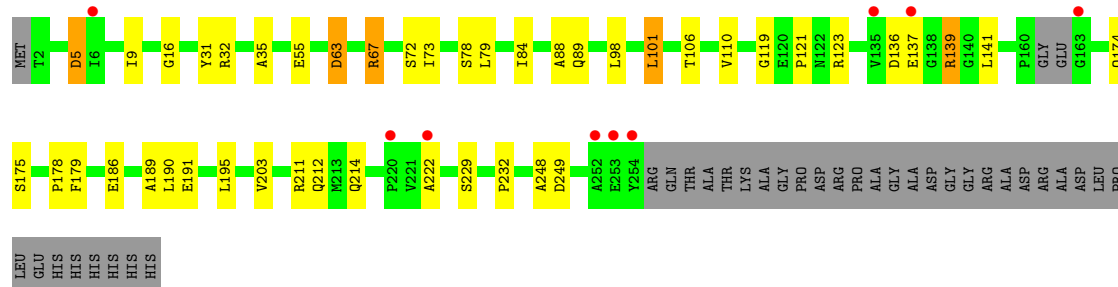
• Molecule 1: Pyrethroid hydrolase



ALA
ASP
ARG
ALA
ASP
LEU
PRO
LEU
GLU
HIS
HIS
HIS
HIS
HIS

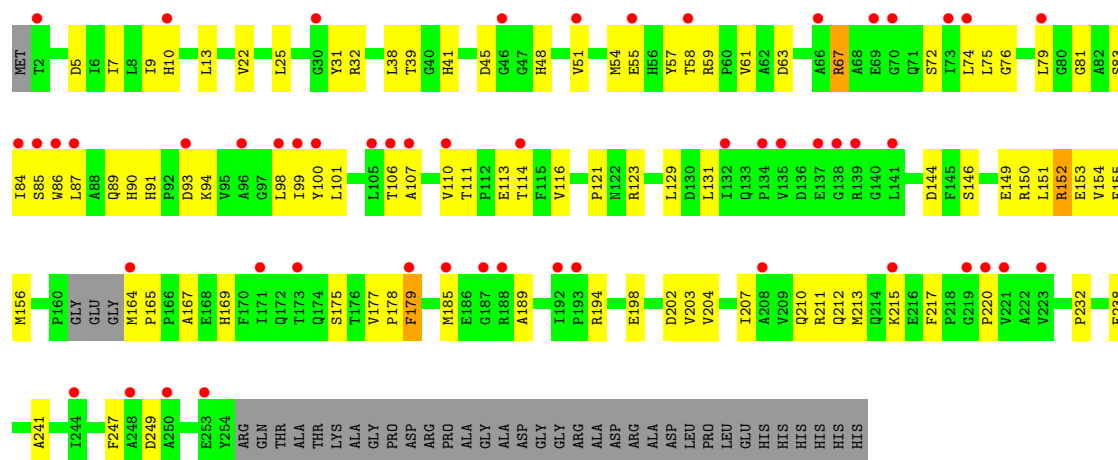
• Molecule 1: Pyrethroid hydrolase

Chain D: 3% 72% 14% 13%



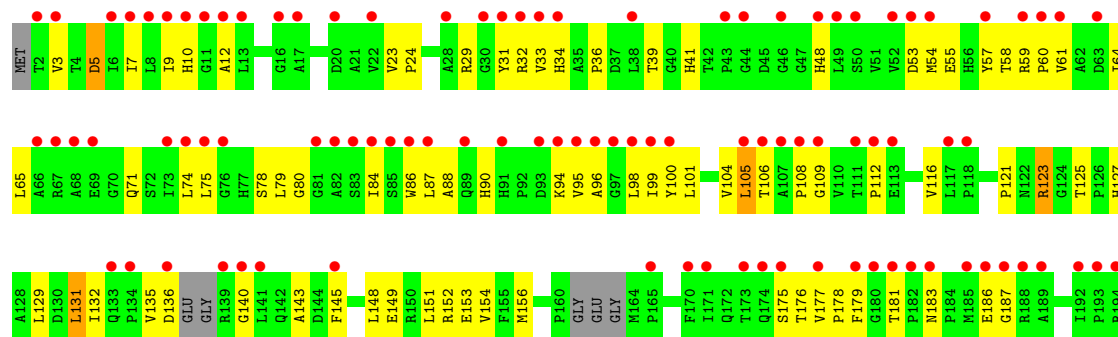
• Molecule 1: Pyrethroid hydrolase

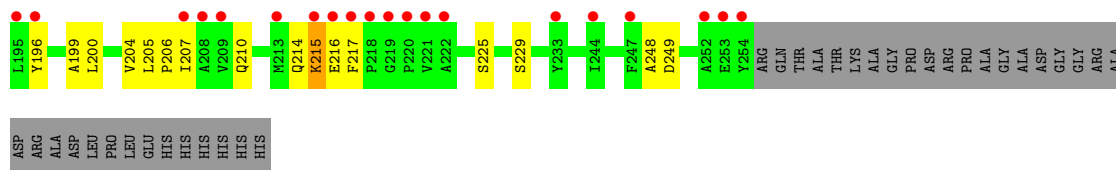
Chain E: 18% 55% 31% 13%



• Molecule 1: Pyrethroid hydrolase

Chain F: 41% 53% 32% 14%





4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α , β , γ	168.38Å 168.38Å 123.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.82 – 2.30 19.82 – 2.30	Depositor EDS
% Data completeness (in resolution range)	94.8 (19.82-2.30) 95.0 (19.82-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.04 (at 2.30Å)	Xtriage
Refinement program	PHENIX (dev_2247: ???)	Depositor
R, R_{free}	0.269 , 0.330 0.269 , 0.327	Depositor DCC
R_{free} test set	3787 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	34.5	Xtriage
Anisotropy	0.536	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 36.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	11414	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.12% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.55	0/1938	0.66	0/2652
1	B	0.56	0/1942	0.66	0/2657
1	C	0.56	0/1950	0.66	0/2670
1	D	0.57	0/1950	0.66	0/2670
1	E	0.57	0/1936	0.69	0/2652
1	F	0.51	0/1938	0.73	0/2652
All	All	0.56	0/11654	0.68	0/15953

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	1888	0	1858	33	0
1	B	1892	0	1861	33	0
1	C	1899	0	1860	39	0
1	D	1899	0	1860	33	0
1	E	1885	0	1835	100	0
1	F	1888	0	1857	105	0
2	A	5	0	0	0	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	0	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
3	A	9	0	0	0	0
3	B	7	0	0	0	0
3	C	3	0	0	0	0
3	D	9	0	0	2	0
3	E	4	0	0	3	0
3	F	1	0	0	0	0
All	All	11414	0	11131	340	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 15.

The worst 5 of 340 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:194:ARG:NH1	1:E:217:PHE:CE1	2.13	1.16
1:E:9:ILE:HD11	1:E:38:LEU:HD11	1.18	1.15
1:F:54:MET:HB2	1:F:86:TRP:CH2	1.81	1.14
1:E:9:ILE:HD11	1:E:38:LEU:CD1	1.83	1.08
1:F:39:THR:HG22	1:F:57:TYR:HA	1.28	1.08

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	242/288 (84%)	234 (97%)	8 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	243/288 (84%)	234 (96%)	9 (4%)	0	100	100
1	C	247/288 (86%)	234 (95%)	13 (5%)	0	100	100
1	D	247/288 (86%)	233 (94%)	14 (6%)	0	100	100
1	E	246/288 (85%)	226 (92%)	20 (8%)	0	100	100
1	F	242/288 (84%)	229 (95%)	13 (5%)	0	100	100
All	All	1467/1728 (85%)	1390 (95%)	77 (5%)	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	199/226 (88%)	198 (100%)	1 (0%)	88	95
1	B	199/226 (88%)	192 (96%)	7 (4%)	36	50
1	C	199/226 (88%)	195 (98%)	4 (2%)	55	72
1	D	199/226 (88%)	192 (96%)	7 (4%)	36	50
1	E	196/226 (87%)	191 (97%)	5 (3%)	46	63
1	F	199/226 (88%)	192 (96%)	7 (4%)	36	50
All	All	1191/1356 (88%)	1160 (97%)	31 (3%)	46	63

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	63	ASP
1	F	123	ARG
1	D	139	ARG
1	F	215	LYS
1	F	5	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	10	HIS
1	F	122	ASN
1	F	212	GLN
1	F	210	GLN
1	E	212	GLN

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 monosaccharides 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	D	301	-	4,4,4	0.09	0	6,6,6	0.47	0
2	SO4	B	301	-	4,4,4	0.27	0	6,6,6	0.35	0
2	SO4	C	301	-	4,4,4	0.14	0	6,6,6	0.09	0
2	SO4	A	301	-	4,4,4	0.12	0	6,6,6	0.23	0
2	SO4	E	301	-	4,4,4	0.21	0	6,6,6	0.32	0
2	SO4	F	301	-	4,4,4	0.10	0	6,6,6	0.38	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	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, 95th 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	248/288 (86%)	0.23	5 (2%) 65 71	21, 30, 43, 58	0
1	B	249/288 (86%)	0.23	8 (3%) 47 54	20, 31, 45, 66	0
1	C	251/288 (87%)	0.38	10 (3%) 38 45	22, 36, 53, 74	0
1	D	251/288 (87%)	0.34	9 (3%) 42 49	19, 32, 50, 70	0
1	E	250/288 (86%)	1.19	53 (21%) 0 1	26, 49, 66, 89	0
1	F	248/288 (86%)	2.16	117 (47%) 0 0	38, 68, 86, 108	0
All	All	1497/1728 (86%)	0.75	202 (13%) 3 4	19, 37, 75, 108	0

The worst 5 of 202 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	139	ARG	10.2
1	F	188	ARG	9.4
1	F	218	PRO	7.9
1	F	44	GLY	7.4
1	F	109	GLY	7.1

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](#)

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, 95th 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
2	SO4	F	301	5/5	0.87	0.21	30,39,51,56	0
2	SO4	D	301	5/5	0.95	0.12	41,43,47,53	0
2	SO4	E	301	5/5	0.96	0.10	39,42,46,47	0
2	SO4	A	301	5/5	0.97	0.11	38,41,43,46	0
2	SO4	B	301	5/5	0.97	0.09	31,36,42,42	0
2	SO4	C	301	5/5	0.97	0.11	31,34,38,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.