



Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 05:09 pm BST

PDB ID : 6BX8
Title : Human Mesotrypsin (PRSS3) Complexed with Tissue Factor Pathway Inhibitor Variant (TFPI1-KD1-K15R-I17C-I34C)
Authors : Coban, M.; Sankaran, B.; Cohen, I.; Hockla, A.; Papo, N.; Radisky, E.S.
Deposited on : 2017-12-18
Resolution : 1.98 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

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:

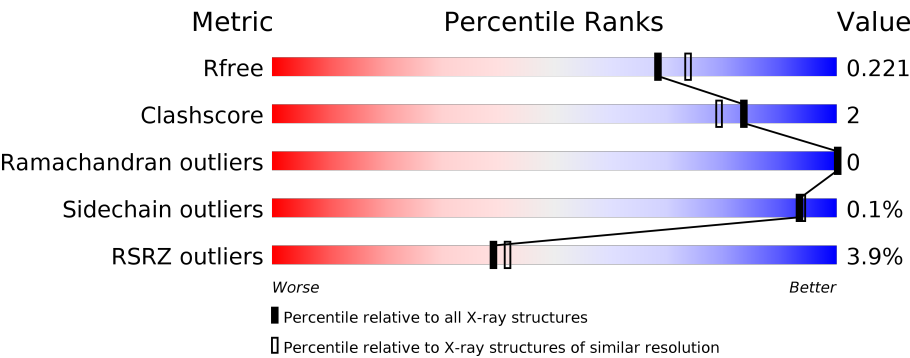
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.98 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 ≥ 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	224	<div><div></div><div>97%</div><div></div></div>
1	C	224	<div><div>%</div><div>96%</div><div></div></div>
1	E	224	<div><div>%</div><div>96%</div><div></div></div>
1	G	224	<div><div>6%</div><div>96%</div><div></div></div>
2	B	80	<div><div>9%</div><div>64%</div><div>5%</div><div>31%</div></div>
2	D	80	<div><div>3%</div><div>63%</div><div></div><div>34%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	80	
2	H	80	

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
3	SO4	B	101	-	X	-	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 16583 atoms, of which 7513 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 Trypsin-3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	224	Total	C	H	N	O	S	0	1	0
			3278	1055	1602	287	322	12			
1	C	224	Total	C	H	N	O	S	0	0	0
			3232	1046	1569	285	320	12			
1	E	224	Total	C	H	N	O	S	0	0	0
			3272	1056	1597	284	323	12			
1	G	224	Total	C	H	N	O	S	0	1	0
			3174	1036	1529	274	323	12			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	127	ALA	THR	conflict	UNP P35030
A	195	ALA	SER	engineered mutation	UNP P35030
C	127	ALA	THR	conflict	UNP P35030
C	195	ALA	SER	engineered mutation	UNP P35030
E	127	ALA	THR	conflict	UNP P35030
E	195	ALA	SER	engineered mutation	UNP P35030
G	127	ALA	THR	conflict	UNP P35030
G	195	ALA	SER	engineered mutation	UNP P35030

- Molecule 2 is a protein called Tissue factor pathway inhibitor.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	55	Total	C	H	N	O	S	0	0	0
			675	239	287	70	69	10			
2	D	53	Total	C	H	N	O	S	0	0	0
			747	254	341	69	73	10			
2	F	54	Total	C	H	N	O	S	0	0	0
			674	234	288	69	73	10			
2	H	55	Total	C	H	N	O	S	0	0	0
			690	239	300	70	72	9			

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-11	TYR	-	expression tag	UNP P10646
B	-10	VAL	-	expression tag	UNP P10646
B	-9	ASP	-	expression tag	UNP P10646
B	-8	TYR	-	expression tag	UNP P10646
B	-7	LYS	-	expression tag	UNP P10646
B	-6	ASP	-	expression tag	UNP P10646
B	-5	ASP	-	expression tag	UNP P10646
B	-4	ASP	-	expression tag	UNP P10646
B	-3	ASP	-	expression tag	UNP P10646
B	-2	LYS	-	expression tag	UNP P10646
B	-1	GLU	-	expression tag	UNP P10646
B	0	PHE	-	expression tag	UNP P10646
B	15	ARG	LYS	engineered mutation	UNP P10646
B	17	CYS	ILE	engineered mutation	UNP P10646
B	34	CYS	ILE	engineered mutation	UNP P10646
B	59	PRO	-	expression tag	UNP P10646
B	60	ARG	-	expression tag	UNP P10646
B	61	HIS	-	expression tag	UNP P10646
B	62	HIS	-	expression tag	UNP P10646
B	63	HIS	-	expression tag	UNP P10646
B	64	HIS	-	expression tag	UNP P10646
B	65	HIS	-	expression tag	UNP P10646
B	66	HIS	-	expression tag	UNP P10646
B	67	ALA	-	expression tag	UNP P10646
B	68	ASN	-	expression tag	UNP P10646
D	-11	TYR	-	expression tag	UNP P10646
D	-10	VAL	-	expression tag	UNP P10646
D	-9	ASP	-	expression tag	UNP P10646
D	-8	TYR	-	expression tag	UNP P10646
D	-7	LYS	-	expression tag	UNP P10646
D	-6	ASP	-	expression tag	UNP P10646
D	-5	ASP	-	expression tag	UNP P10646
D	-4	ASP	-	expression tag	UNP P10646
D	-3	ASP	-	expression tag	UNP P10646
D	-2	LYS	-	expression tag	UNP P10646
D	-1	GLU	-	expression tag	UNP P10646
D	0	PHE	-	expression tag	UNP P10646
D	15	ARG	LYS	engineered mutation	UNP P10646
D	17	CYS	ILE	engineered mutation	UNP P10646
D	34	CYS	ILE	engineered mutation	UNP P10646
D	59	PRO	-	expression tag	UNP P10646
D	60	ARG	-	expression tag	UNP P10646

Continued on next page...

Continued from previous page...

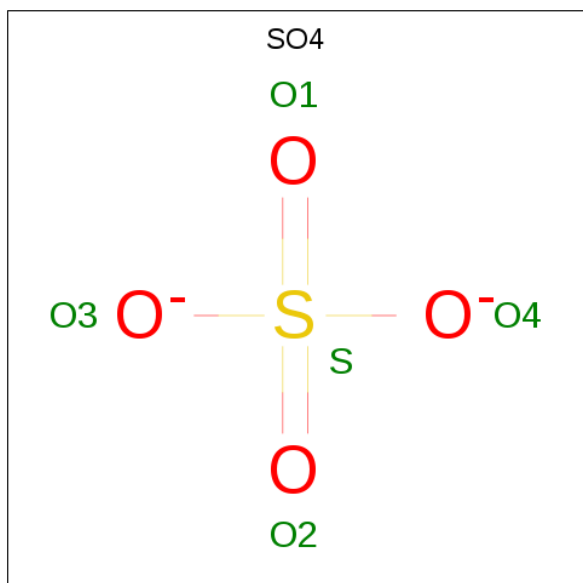
Chain	Residue	Modelled	Actual	Comment	Reference
D	61	HIS	-	expression tag	UNP P10646
D	62	HIS	-	expression tag	UNP P10646
D	63	HIS	-	expression tag	UNP P10646
D	64	HIS	-	expression tag	UNP P10646
D	65	HIS	-	expression tag	UNP P10646
D	66	HIS	-	expression tag	UNP P10646
D	67	ALA	-	expression tag	UNP P10646
D	68	ASN	-	expression tag	UNP P10646
F	-11	TYR	-	expression tag	UNP P10646
F	-10	VAL	-	expression tag	UNP P10646
F	-9	ASP	-	expression tag	UNP P10646
F	-8	TYR	-	expression tag	UNP P10646
F	-7	LYS	-	expression tag	UNP P10646
F	-6	ASP	-	expression tag	UNP P10646
F	-5	ASP	-	expression tag	UNP P10646
F	-4	ASP	-	expression tag	UNP P10646
F	-3	ASP	-	expression tag	UNP P10646
F	-2	LYS	-	expression tag	UNP P10646
F	-1	GLU	-	expression tag	UNP P10646
F	0	PHE	-	expression tag	UNP P10646
F	15	ARG	LYS	engineered mutation	UNP P10646
F	17	CYS	ILE	engineered mutation	UNP P10646
F	34	CYS	ILE	engineered mutation	UNP P10646
F	59	PRO	-	expression tag	UNP P10646
F	60	ARG	-	expression tag	UNP P10646
F	61	HIS	-	expression tag	UNP P10646
F	62	HIS	-	expression tag	UNP P10646
F	63	HIS	-	expression tag	UNP P10646
F	64	HIS	-	expression tag	UNP P10646
F	65	HIS	-	expression tag	UNP P10646
F	66	HIS	-	expression tag	UNP P10646
F	67	ALA	-	expression tag	UNP P10646
F	68	ASN	-	expression tag	UNP P10646
H	-11	TYR	-	expression tag	UNP P10646
H	-10	VAL	-	expression tag	UNP P10646
H	-9	ASP	-	expression tag	UNP P10646
H	-8	TYR	-	expression tag	UNP P10646
H	-7	LYS	-	expression tag	UNP P10646
H	-6	ASP	-	expression tag	UNP P10646
H	-5	ASP	-	expression tag	UNP P10646
H	-4	ASP	-	expression tag	UNP P10646
H	-3	ASP	-	expression tag	UNP P10646

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	-2	LYS	-	expression tag	UNP P10646
H	-1	GLU	-	expression tag	UNP P10646
H	0	PHE	-	expression tag	UNP P10646
H	15	ARG	LYS	engineered mutation	UNP P10646
H	17	CYS	ILE	engineered mutation	UNP P10646
H	34	CYS	ILE	engineered mutation	UNP P10646
H	59	PRO	-	expression tag	UNP P10646
H	60	ARG	-	expression tag	UNP P10646
H	61	HIS	-	expression tag	UNP P10646
H	62	HIS	-	expression tag	UNP P10646
H	63	HIS	-	expression tag	UNP P10646
H	64	HIS	-	expression tag	UNP P10646
H	65	HIS	-	expression tag	UNP P10646
H	66	HIS	-	expression tag	UNP P10646
H	67	ALA	-	expression tag	UNP P10646
H	68	ASN	-	expression tag	UNP P10646

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

Continued on next page...

Continued from previous page...

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

- Molecule 4 is water.

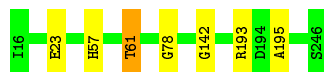
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	195	Total	O		0	0
			195	195			
4	B	35	Total	O		0	0
			35	35			
4	C	175	Total	O		0	0
			175	175			
4	D	58	Total	O		0	0
			58	58			
4	E	198	Total	O		0	0
			198	198			
4	F	39	Total	O		0	0
			39	39			
4	G	99	Total	O		0	0
			99	99			
4	H	22	Total	O		0	0
			22	22			

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: Trypsin-3

Chain A:  97%



• Molecule 1: Trypsin-3

Chain C:  96%



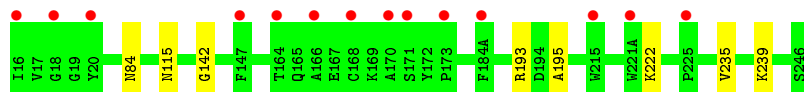
• Molecule 1: Trypsin-3

Chain E:  96%



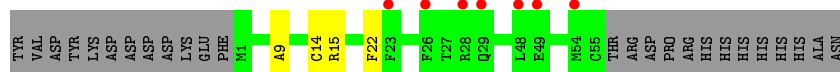
• Molecule 1: Trypsin-3

Chain G:  96%



• Molecule 2: Tissue factor pathway inhibitor

Chain B:  64% 9% 5% 31%



• Molecule 2: Tissue factor pathway inhibitor



• Molecule 2: Tissue factor pathway inhibitor



• Molecule 2: Tissue factor pathway inhibitor



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	35.14Å 87.20Å 90.84Å 94.81° 93.03° 92.46°	Depositor
Resolution (Å)	60.06 – 1.98 60.06 – 1.98	Depositor EDS
% Data completeness (in resolution range)	93.0 (60.06-1.98) 93.3 (60.06-1.98)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.85 (at 1.98Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.192 , 0.221 0.192 , 0.221	Depositor DCC
R_{free} test set	3384 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å ²)	18.5	Xtriage
Anisotropy	0.260	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.42 , 56.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.009 for -h,-l,-k	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	16583	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.98% 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	0.95	0/1713	0.90	0/2334
1	C	0.89	0/1702	0.87	0/2322
1	E	0.95	0/1714	0.88	0/2337
1	G	0.80	0/1682	0.81	0/2299
2	B	0.88	0/397	0.86	0/533
2	D	1.05	0/415	0.90	0/553
2	F	0.89	0/394	0.88	0/529
2	H	0.63	0/398	0.84	0/534
All	All	0.90	0/8415	0.87	0/11441

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	1676	1602	1602	5	1
1	C	1663	1569	1569	12	0
1	E	1675	1597	1597	13	0
1	G	1645	1529	1529	7	0
2	B	388	287	289	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	406	341	341	2	0
2	F	386	288	288	3	0
2	H	390	300	302	3	0
3	B	5	0	0	0	0
3	C	5	0	0	0	0
3	F	5	0	0	0	1
3	H	5	0	0	0	0
4	A	195	0	0	1	1
4	B	35	0	0	0	0
4	C	175	0	0	3	0
4	D	58	0	0	0	0
4	E	198	0	0	4	1
4	F	39	0	0	1	0
4	G	99	0	0	3	0
4	H	22	0	0	0	0
All	All	9070	7513	7517	37	2

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 2.

All (37) 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:C:128:PRO:HD2	1:E:128:PRO:HG3	1.49	0.94
1:C:128:PRO:CD	1:E:128:PRO:HG3	2.09	0.82
1:C:128:PRO:CG	1:E:128:PRO:HG3	2.15	0.76
1:E:169:LYS:HE2	4:E:446:HOH:O	1.91	0.70
1:G:239:LYS:HD2	4:G:374:HOH:O	1.91	0.70
1:C:128:PRO:HD2	1:E:128:PRO:CG	2.24	0.67
1:E:24:GLU:HG2	1:E:71:HIS:CD2	2.30	0.67
1:A:23:GLU:OE1	4:A:301:HOH:O	2.15	0.64
1:C:128:PRO:HG2	1:E:128:PRO:HG3	1.79	0.63
1:E:60:LYS:NZ	4:E:301:HOH:O	2.29	0.62
1:G:84:ASN:HB3	4:G:378:HOH:O	2.06	0.55
1:G:222:LYS:CB	4:G:387:HOH:O	2.55	0.54
1:G:142:GLY:HA2	1:G:193:ARG:HB3	1.90	0.54
1:E:142:GLY:HA2	1:E:193:ARG:HB3	1.89	0.53
1:G:235:VAL:O	1:G:239:LYS:HG3	2.09	0.52
1:E:156:LYS:HD2	4:E:398:HOH:O	2.09	0.52
1:C:142:GLY:HA2	1:C:193:ARG:HB3	1.92	0.51
2:H:30:CYS:SG	2:H:51:CYS:CB	2.99	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:142:GLY:HA2	1:A:193:ARG:HB3	1.92	0.50
1:E:24:GLU:HG2	1:E:71:HIS:CG	2.48	0.48
1:A:195:ALA:HB2	2:B:15:ARG:C	2.37	0.46
1:C:195:ALA:HB2	2:D:15:ARG:C	2.36	0.45
1:G:195:ALA:HB2	2:H:15:ARG:C	2.36	0.45
1:C:230:LYS:CE	4:C:410:HOH:O	2.64	0.45
1:E:195:ALA:HB2	2:F:15:ARG:C	2.37	0.44
2:H:9:ALA:HB2	2:H:22:PHE:CD2	2.53	0.43
1:C:184(A):PHE:HB2	4:C:455:HOH:O	2.19	0.42
2:F:9:ALA:HB2	2:F:22:PHE:CD2	2.55	0.42
1:C:230:LYS:HE3	4:C:410:HOH:O	2.19	0.42
1:E:135:GLU:HG3	4:E:342:HOH:O	2.18	0.42
1:A:57:HIS:CD2	2:B:14:CYS:HB3	2.53	0.42
1:C:24:GLU:HG2	1:C:25:ASN:OD1	2.19	0.42
1:A:78:GLY:HA3	1:G:115:ASN:HD22	1.86	0.41
2:B:9:ALA:HB2	2:B:22:PHE:CD2	2.55	0.41
2:D:9:ALA:HB2	2:D:22:PHE:CD2	2.56	0.41
2:F:55:CYS:C	4:F:207:HOH:O	2.59	0.40
1:C:179:SER:HB2	1:C:234:TYR:OH	2.21	0.40

All (2) 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:F:101:SO4:O1	4:E:476:HOH:O[1_655]	1.54	0.66
1:A:61:THR:H	4:A:301:HOH:O[1_455]	1.56	0.04

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	222/224 (99%)	218 (98%)	4 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	222/224 (99%)	216 (97%)	6 (3%)	0	100	100
1	E	222/224 (99%)	215 (97%)	7 (3%)	0	100	100
1	G	222/224 (99%)	218 (98%)	4 (2%)	0	100	100
2	B	53/80 (66%)	52 (98%)	1 (2%)	0	100	100
2	D	51/80 (64%)	49 (96%)	2 (4%)	0	100	100
2	F	52/80 (65%)	51 (98%)	1 (2%)	0	100	100
2	H	53/80 (66%)	53 (100%)	0	0	100	100
All	All	1097/1216 (90%)	1072 (98%)	25 (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	179/185 (97%)	178 (99%)	1 (1%)	86	85
1	C	175/185 (95%)	175 (100%)	0	100	100
1	E	179/185 (97%)	179 (100%)	0	100	100
1	G	172/185 (93%)	172 (100%)	0	100	100
2	B	32/72 (44%)	32 (100%)	0	100	100
2	D	40/72 (56%)	40 (100%)	0	100	100
2	F	34/72 (47%)	34 (100%)	0	100	100
2	H	34/72 (47%)	34 (100%)	0	100	100
All	All	845/1028 (82%)	844 (100%)	1 (0%)	93	94

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	61	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

4 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	$\# Z > 2$
3	SO4	C	301	-	4,4,4	0.79	0	6,6,6	0.73	0
3	SO4	B	101	-	4,4,4	0.97	0	6,6,6	2.98	4 (66%)
3	SO4	H	101	-	4,4,4	0.44	0	6,6,6	0.75	0
3	SO4	F	101	-	4,4,4	0.72	0	6,6,6	0.96	0

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	101	SO4	O3-S-O1	-4.65	85.06	109.31
3	B	101	SO4	O4-S-O3	4.15	126.78	109.06
3	B	101	SO4	O4-S-O2	-2.68	95.31	109.31
3	B	101	SO4	O3-S-O2	2.27	121.16	109.31

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
3	F	101	SO4	0	1

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, 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	224/224 (100%)	0.05	0 100 100	12, 21, 32, 41	0
1	C	224/224 (100%)	0.18	3 (1%) 77 78	13, 25, 40, 63	0
1	E	224/224 (100%)	0.18	3 (1%) 77 78	11, 24, 36, 55	0
1	G	224/224 (100%)	0.67	14 (6%) 20 22	27, 38, 59, 66	0
2	B	55/80 (68%)	0.85	7 (12%) 3 4	17, 35, 65, 72	0
2	D	53/80 (66%)	0.35	2 (3%) 40 43	16, 27, 46, 48	0
2	F	54/80 (67%)	0.80	6 (11%) 5 6	14, 36, 66, 72	0
2	H	55/80 (68%)	1.18	8 (14%) 2 2	34, 44, 56, 60	0
All	All	1113/1216 (91%)	0.37	43 (3%) 39 42	11, 28, 54, 72	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	28	ARG	5.4
2	F	6	ALA	4.7
2	H	6	ALA	4.5
2	F	3	SER	4.4
2	F	4	PHE	4.1
2	H	7	PHE	3.8
2	B	48	LEU	3.7
1	G	170	ALA	3.4
2	H	25	ILE	3.4
2	H	4	PHE	3.2
2	H	23	PHE	3.2
2	B	54	MET	3.1
1	E	75	VAL	3.0
1	G	215	TRP	3.0
1	G	164	THR	2.9
1	G	225	PRO	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	54	MET	2.7
1	G	147	PHE	2.7
2	B	29	GLN	2.7
1	G	16	ILE	2.7
1	G	166	ALA	2.6
1	C	78	GLY	2.6
1	G	184(A)	PHE	2.5
2	D	3	SER	2.5
2	F	26	PHE	2.5
1	G	20	TYR	2.4
2	F	55	CYS	2.4
1	E	129	PRO	2.4
2	H	26	PHE	2.4
1	C	79	ASN	2.3
2	B	23	PHE	2.3
2	D	48	LEU	2.3
2	B	49	GLU	2.3
2	H	53	LYS	2.2
2	H	55	CYS	2.2
1	E	74	LYS	2.2
1	G	171	SER	2.1
1	G	221(A)	TRP	2.1
1	G	18	GLY	2.1
1	C	77	GLU	2.0
1	G	168	CYS	2.0
1	G	173	PRO	2.0
2	B	26	PHE	2.0

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, 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
3	SO4	B	101	5/5	0.86	0.14	43,43,46,47	0
3	SO4	C	301	5/5	0.92	0.18	59,60,63,65	0
3	SO4	H	101	5/5	0.97	0.10	51,53,54,55	0
3	SO4	F	101	5/5	0.97	0.09	38,39,42,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.