



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 11, 2023 – 08:10 am GMT

PDB ID : 5FVF  
Title : Room temperature structure of IrisFP determined by serial femtosecond crystallography.  
Authors : Colletier, J.P.; Gallat, F.X.; Coquelle, N.; Weik, M.  
Deposited on : 2016-02-06  
Resolution : 2.75 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

<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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

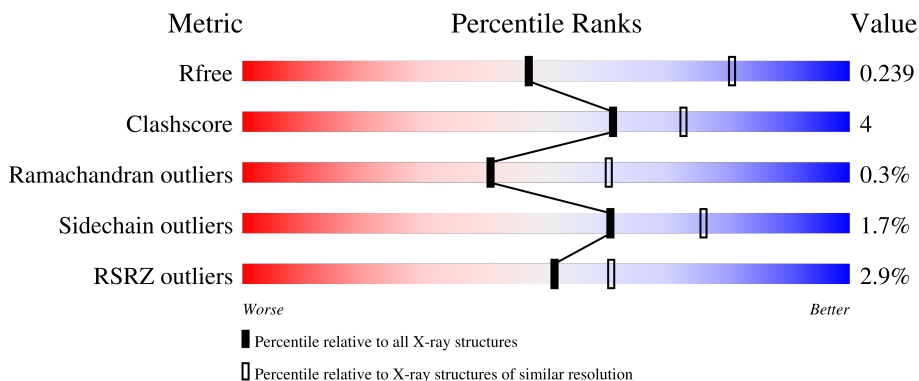
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	223	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 91%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 9%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 0%; height: 10px; background-color: grey; margin-right: 2px;"></div> </div> <p style="margin-left: 20px;">2%      91%      9%</p>
1	B	223	<div style="display: flex; align-items: center;"> <div style="width: 0%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 87%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 13%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 0%; height: 10px; background-color: grey; margin-right: 2px;"></div> </div> <p style="margin-left: 20px;">%      87%      13%</p>
1	C	223	<div style="display: flex; align-items: center;"> <div style="width: 0%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 85%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 14%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: grey; margin-right: 2px;"></div> </div> <p style="margin-left: 20px;">%      85%      14%      .</p>
1	D	223	<div style="display: flex; align-items: center;"> <div style="width: 8%; height: 10px; background-color: red; margin-right: 2px;"></div> <div style="width: 85%; height: 10px; background-color: green; margin-right: 2px;"></div> <div style="width: 14%; height: 10px; background-color: yellow; margin-right: 2px;"></div> <div style="width: 1%; height: 10px; background-color: grey; margin-right: 2px;"></div> </div> <p style="margin-left: 20px;">8%      85%      14%      .</p>

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
2	SO4	A	301	-	-	X	-
2	SO4	A	307	-	-	-	X
2	SO4	C	305	-	-	-	X
2	SO4	D	304	-	-	-	X
2	SO4	D	306	-	-	-	X
3	NH4	A	306	-	-	-	X

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7908 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 Green to red photoconvertible GFP-like protein EosFP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	223	1867	1187	328	341	11	0	7	0
1	B	223	1866	1182	326	347	11	1	8	0
1	C	221	1837	1166	319	341	11	0	6	0
1	D	221	1897	1206	333	347	11	0	13	0

There are 28 discrepancies between the modelled and reference sequences:

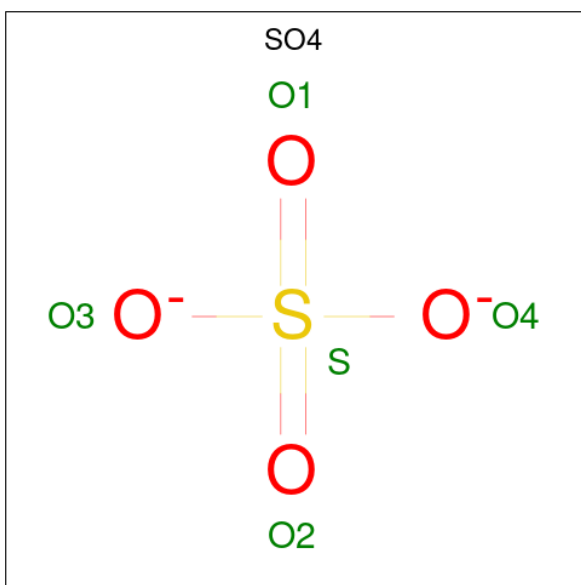
Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	HIS	-	expression tag	UNP Q5S6Z9
A	0	HIS	-	expression tag	UNP Q5S6Z9
A	64	5SQ	HIS	chromophore	UNP Q5S6Z9
A	64	5SQ	TYR	chromophore	UNP Q5S6Z9
A	64	5SQ	GLY	chromophore	UNP Q5S6Z9
A	173	SER	PHE	engineered mutation	UNP Q5S6Z9
A	191	LEU	PHE	engineered mutation	UNP Q5S6Z9
B	-1	HIS	-	expression tag	UNP Q5S6Z9
B	0	HIS	-	expression tag	UNP Q5S6Z9
B	64	5SQ	HIS	chromophore	UNP Q5S6Z9
B	64	5SQ	TYR	chromophore	UNP Q5S6Z9
B	64	5SQ	GLY	chromophore	UNP Q5S6Z9
B	173	SER	PHE	engineered mutation	UNP Q5S6Z9
B	191	LEU	PHE	engineered mutation	UNP Q5S6Z9
C	-1	HIS	-	expression tag	UNP Q5S6Z9
C	0	HIS	-	expression tag	UNP Q5S6Z9
C	64	5SQ	HIS	chromophore	UNP Q5S6Z9
C	64	5SQ	TYR	chromophore	UNP Q5S6Z9
C	64	5SQ	GLY	chromophore	UNP Q5S6Z9
C	173	SER	PHE	engineered mutation	UNP Q5S6Z9
C	191	LEU	PHE	engineered mutation	UNP Q5S6Z9

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	HIS	-	expression tag	UNP Q5S6Z9
D	0	HIS	-	expression tag	UNP Q5S6Z9
D	64	5SQ	HIS	chromophore	UNP Q5S6Z9
D	64	5SQ	TYR	chromophore	UNP Q5S6Z9
D	64	5SQ	GLY	chromophore	UNP Q5S6Z9
D	173	SER	PHE	engineered mutation	UNP Q5S6Z9
D	191	LEU	PHE	engineered mutation	UNP Q5S6Z9

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



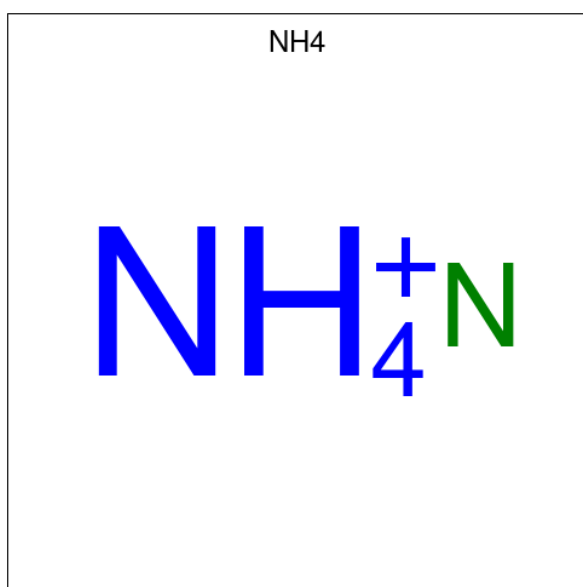
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	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	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is AMMONIUM ION (three-letter code: NH4) (formula: H<sub>4</sub>N).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total N 1 1	0	0
3	A	1	Total N 1 1	0	0

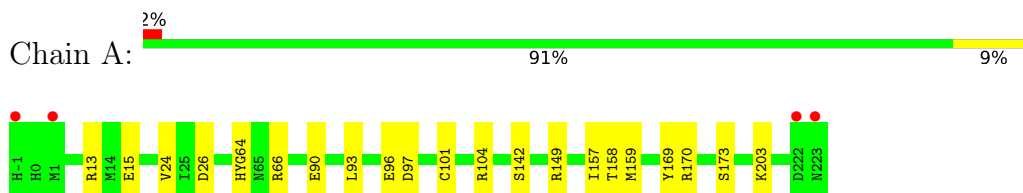
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	88	Total O 88 88	0	0
4	B	94	Total O 94 94	0	0
4	C	75	Total O 75 75	0	0
4	D	87	Total O 87 87	0	0

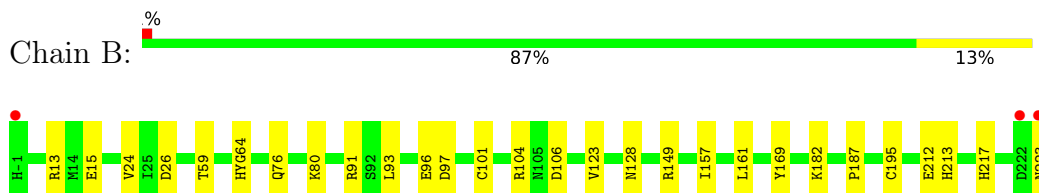
### 3 Residue-property plots [i](#)

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.

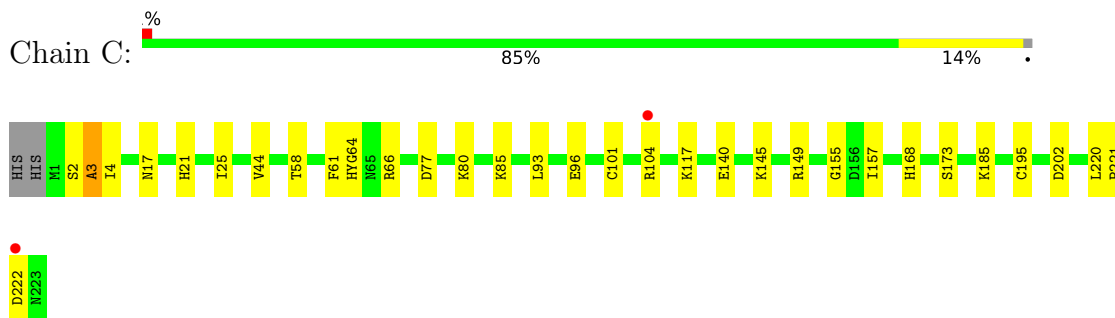
- Molecule 1: Green to red photoconvertible GFP-like protein EosFP



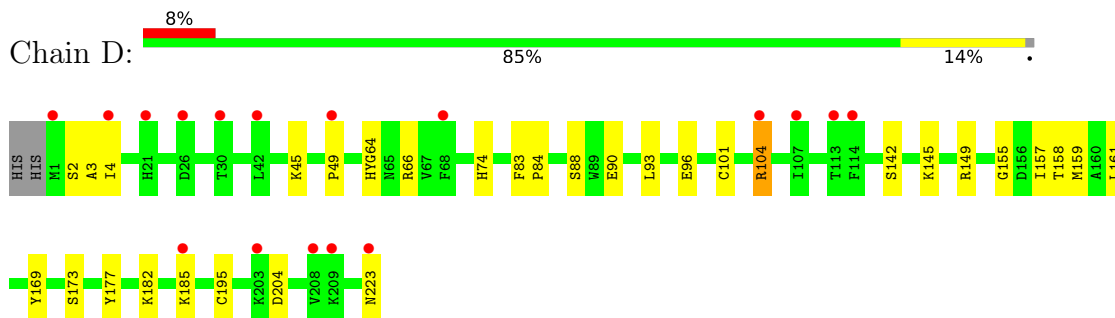
- Molecule 1: Green to red photoconvertible GFP-like protein EosFP



- Molecule 1: Green to red photoconvertible GFP-like protein EosFP



- Molecule 1: Green to red photoconvertible GFP-like protein EosFP





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	83.33Å 97.76Å 142.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.24 – 2.75 63.42 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.24-2.75) 86.4 (63.42-2.60)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.52 (at 2.61Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.176 , 0.239 0.176 , 0.239	Depositor DCC
$R_{free}$ test set	3615 reflections (9.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.0	Xtrriage
Anisotropy	0.405	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 65.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7908	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% 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: NH4, SO4, 5SQ

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	0.26	0/1892	0.46	0/2545
1	B	0.26	0/1891	0.45	0/2547
1	C	0.25	0/1857	0.46	0/2500
1	D	0.25	0/1921	0.44	0/2581
All	All	0.25	0/7561	0.45	0/10173

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	1867	0	1794	21	0
1	B	1866	0	1771	15	0
1	C	1837	0	1753	18	0
1	D	1897	0	1831	20	0
2	A	25	0	0	3	0
2	B	10	0	0	0	0
2	C	30	0	0	1	0
2	D	30	0	0	0	0
3	A	2	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	88	0	0	1	0
4	B	94	0	0	2	0
4	C	75	0	0	1	0
4	D	87	0	0	2	0
All	All	7908	0	7149	66	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 4.

All (66) 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:142:SER:HB2	1:D:157:ILE:HD11	1.63	0.81
1:A:158:THR:HB	1:A:170:ARG:HH12	1.50	0.77
1:A:149:ARG:NH2	1:D:96:GLU:OE1	2.20	0.73
1:C:2:SER:HB3	1:C:4:ILE:N	2.05	0.71
1:B:96:GLU:OE1	1:C:149:ARG:NH2	2.24	0.70
1:C:17:ASN:OD1	1:D:104[A]:ARG:NH1	2.25	0.69
1:A:158:THR:HG21	1:D:158:THR:HG21	1.76	0.68
1:A:142:SER:HB2	1:A:157:ILE:HD11	1.76	0.67
1:B:149:ARG:NH2	1:C:96:GLU:OE1	2.24	0.66
1:A:96:GLU:OE1	1:D:149:ARG:NH2	2.21	0.65
1:C:2:SER:HA	1:C:3:ALA:HB3	1.82	0.62
1:C:2:SER:HB3	1:C:4:ILE:H	1.65	0.61
1:C:104[A]:ARG:NH2	2:C:303:SO4:O3	2.36	0.58
1:A:158:THR:OG1	4:A:2067:HOH:O	2.17	0.57
1:D:157:ILE:HG23	1:D:173:SER:HB3	1.86	0.57
1:D:66[B]:ARG:NH1	1:D:177:TYR:OH	2.40	0.55
1:A:93:LEU:HB2	1:A:101:CYS:HB2	1.88	0.55
1:D:223:ASN:ND2	4:D:2068:HOH:O	2.33	0.55
1:D:2:SER:O	1:D:4:ILE:N	2.39	0.54
1:C:93:LEU:HB2	1:C:101:CYS:HB2	1.92	0.52
1:A:158:THR:HB	1:A:170:ARG:NH1	2.21	0.52
1:D:157:ILE:HG13	1:D:159:MET:HG3	1.93	0.50
1:A:158:THR:HB	1:A:170:ARG:HH22	1.76	0.50
1:B:128[B]:ASN:OD1	4:B:2060:HOH:O	2.20	0.49
1:A:203:LYS:NZ	2:A:301:SO4:O3	2.45	0.49
1:D:93:LEU:HB2	1:D:101:CYS:HB2	1.95	0.48
1:A:157:ILE:HG23	1:A:173:SER:HB3	1.94	0.47
1:B:15:GLU:HG2	1:B:24:VAL:HG22	1.95	0.47
1:C:140:GLU:OE2	1:C:168:HIS:NE2	2.34	0.47

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:15:GLU:HG2	1:A:24:VAL:HG22	1.96	0.47
1:C:202:ASP:OD2	4:C:2071:HOH:O	2.20	0.47
1:C:145:LYS:O	1:C:155:GLY:HA2	2.14	0.47
1:B:13:ARG:NE	1:B:26:ASP:OD1	2.44	0.47
1:B:217:HIS:HB2	1:B:223:ASN:HB2	1.97	0.47
1:D:90:GLU:HG2	1:D:104[B]:ARG:HD2	1.98	0.46
1:A:158:THR:HG21	1:D:158:THR:CG2	2.45	0.46
1:A:93:LEU:HD23	1:A:173:SER:HB2	1.97	0.46
1:A:104[A]:ARG:HH11	1:A:104[A]:ARG:HB3	1.80	0.46
1:D:88[B]:SER:HB3	1:D:104[B]:ARG:HH12	1.79	0.46
1:B:80:LYS:NZ	4:B:2035:HOH:O	2.49	0.46
1:C:157:ILE:HG23	1:C:173:SER:HB3	1.96	0.46
1:D:74:HIS:H	1:D:74:HIS:CD2	2.34	0.46
1:D:185[A]:LYS:NZ	4:D:2064:HOH:O	2.49	0.45
1:A:13:ARG:NE	1:A:26:ASP:OD1	2.49	0.45
1:B:76:GLN:HB3	1:B:187:PRO:HB3	1.98	0.45
1:C:25:ILE:HD12	1:C:44:VAL:HG22	1.99	0.45
1:C:58:THR:HA	1:C:61:PHE:HD2	1.82	0.44
1:A:203:LYS:NZ	2:A:301:SO4:S	2.90	0.44
1:B:161:LEU:HB2	1:B:169:TYR:HB3	1.99	0.44
1:A:97:ASP:OD1	1:A:169:TYR:OH	2.27	0.43
1:D:49:PRO:HB3	1:D:204:ASP:HB3	1.99	0.43
1:D:161:LEU:HB2	1:D:169:TYR:HB3	2.01	0.43
1:B:97:ASP:OD1	1:B:169:TYR:OH	2.19	0.43
1:A:90:GLU:HB3	1:B:123:VAL:HB	2.00	0.43
1:A:90:GLU:HG2	1:A:104[A]:ARG:HG2	2.01	0.43
1:D:145:LYS:O	1:D:155:GLY:HA2	2.19	0.42
1:C:77:ASP:OD2	1:C:80:LYS:HD2	2.20	0.42
1:B:93:LEU:HB2	1:B:101:CYS:HB2	2.02	0.42
1:C:117:LYS:HD2	1:C:117:LYS:HA	1.93	0.41
1:C:220:LEU:HD22	1:C:221:PRO:HA	2.01	0.41
1:C:17:ASN:HA	1:C:21:HIS:O	2.20	0.41
1:D:83:PHE:HB3	1:D:84:PRO:HA	2.02	0.41
1:B:212:GLU:HG2	1:B:213:HIS:N	2.35	0.41
1:A:157:ILE:HG13	1:A:159:MET:HG3	2.01	0.41
2:A:304:SO4:O3	1:B:104[A]:ARG:NH2	2.53	0.40
1:B:59:THR:HG22	1:B:91:ARG:NH1	2.36	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	226/223 (101%)	224 (99%)	2 (1%)	0	100	100
1	B	226/223 (101%)	224 (99%)	2 (1%)	0	100	100
1	C	222/223 (100%)	219 (99%)	1 (0%)	2 (1%)	17	31
1	D	229/223 (103%)	224 (98%)	4 (2%)	1 (0%)	34	53
All	All	903/892 (101%)	891 (99%)	9 (1%)	3 (0%)	41	60

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	222	ASP
1	C	3	ALA
1	D	3	ALA

### 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	198/191 (104%)	196 (99%)	2 (1%)	76	85
1	B	199/191 (104%)	194 (98%)	5 (2%)	47	67
1	C	195/191 (102%)	189 (97%)	6 (3%)	40	60
1	D	202/191 (106%)	196 (97%)	6 (3%)	41	61
All	All	794/764 (104%)	775 (98%)	19 (2%)	60	68

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66[A]	ARG
1	A	66[B]	ARG
1	B	106[A]	ASP
1	B	106[B]	ASP
1	B	157	ILE
1	B	182	LYS
1	B	195	CYS
1	C	66[A]	ARG
1	C	66[B]	ARG
1	C	85	LYS
1	C	185[A]	LYS
1	C	185[B]	LYS
1	C	195	CYS
1	D	45[A]	LYS
1	D	45[B]	LYS
1	D	104[A]	ARG
1	D	104[B]	ARG
1	D	182	LYS
1	D	195	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	-1	HIS
1	A	206	ASN
1	B	124	ASN
1	D	74	HIS
1	D	213	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](#)

4 non-standard protein/DNA/RNA residues 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
1	5SQ	B	64	1	23,27,28	5.30	4 (17%)	29,37,39	3.60	9 (31%)
1	5SQ	D	64	1	23,27,28	5.33	6 (26%)	29,37,39	3.65	9 (31%)
1	5SQ	A	64	1	23,27,28	5.30	4 (17%)	29,37,39	3.52	7 (24%)
1	5SQ	C	64	1	23,27,28	5.29	5 (21%)	29,37,39	3.75	9 (31%)

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
1	5SQ	B	64	1	-	4/12/31/32	0/3/3/3
1	5SQ	D	64	1	-	4/12/31/32	0/3/3/3
1	5SQ	A	64	1	-	4/12/31/32	0/3/3/3
1	5SQ	C	64	1	-	4/12/31/32	0/3/3/3

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	64	5SQ	CB2-CA2	22.51	1.53	1.35
1	A	64	5SQ	CB2-CA2	22.41	1.53	1.35
1	B	64	5SQ	CB2-CA2	22.38	1.53	1.35
1	C	64	5SQ	CB2-CA2	22.34	1.53	1.35
1	D	64	5SQ	O2-C2	9.62	1.43	1.23
1	B	64	5SQ	O2-C2	9.51	1.43	1.23
1	C	64	5SQ	O2-C2	9.44	1.43	1.23
1	A	64	5SQ	O2-C2	9.30	1.42	1.23
1	C	64	5SQ	CA2-C2	-4.43	1.44	1.48
1	A	64	5SQ	CA2-C2	-4.19	1.44	1.48
1	B	64	5SQ	CA2-C2	-4.03	1.44	1.48
1	D	64	5SQ	CA2-C2	-3.95	1.44	1.48
1	A	64	5SQ	CG2-CB2	3.82	1.54	1.46
1	B	64	5SQ	CG2-CB2	3.78	1.54	1.46
1	C	64	5SQ	CG2-CB2	3.68	1.53	1.46
1	D	64	5SQ	CG2-CB2	3.67	1.53	1.46
1	C	64	5SQ	C1-N2	2.02	1.35	1.32
1	D	64	5SQ	C1-N2	2.02	1.35	1.32
1	D	64	5SQ	CB1-CG1	2.00	1.58	1.51

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	64	5SQ	O2-C2-CA2	-12.60	123.88	130.96
1	B	64	5SQ	O2-C2-CA2	-11.46	124.52	130.96
1	D	64	5SQ	O2-C2-CA2	-11.37	124.58	130.96
1	A	64	5SQ	CG2-CB2-CA2	-10.08	117.60	129.94
1	A	64	5SQ	CA2-C2-N3	9.65	107.94	103.37
1	D	64	5SQ	CG2-CB2-CA2	-9.61	118.17	129.94
1	A	64	5SQ	O2-C2-CA2	-9.56	125.59	130.96
1	C	64	5SQ	CA2-C2-N3	9.55	107.89	103.37
1	B	64	5SQ	CA2-C2-N3	9.39	107.81	103.37
1	D	64	5SQ	CA2-C2-N3	9.32	107.78	103.37
1	B	64	5SQ	CG2-CB2-CA2	-9.00	118.91	129.94
1	C	64	5SQ	CG2-CB2-CA2	-8.60	119.40	129.94
1	A	64	5SQ	C2-CA2-N2	-5.04	105.40	108.93
1	D	64	5SQ	C2-CA2-N2	-4.63	105.69	108.93
1	C	64	5SQ	C2-CA2-N2	-4.36	105.88	108.93
1	B	64	5SQ	C2-CA2-N2	-4.35	105.88	108.93
1	A	64	5SQ	CA2-N2-C1	3.96	108.69	105.77
1	D	64	5SQ	CA2-N2-C1	3.80	108.58	105.77
1	C	64	5SQ	CA3-N3-C1	-3.76	122.66	127.16
1	B	64	5SQ	CA3-N3-C1	-3.31	123.19	127.16
1	C	64	5SQ	CA2-N2-C1	3.26	108.18	105.77
1	B	64	5SQ	CA2-N2-C1	3.07	108.03	105.77
1	D	64	5SQ	CA3-N3-C1	-2.97	123.60	127.16
1	C	64	5SQ	CA3-N3-C2	2.61	129.78	123.80
1	A	64	5SQ	O3-C3-CA3	-2.58	118.60	126.39
1	B	64	5SQ	O3-C3-CA3	-2.43	119.07	126.39
1	B	64	5SQ	CA3-N3-C2	2.41	129.33	123.80
1	D	64	5SQ	O3-C3-CA3	-2.40	119.14	126.39
1	C	64	5SQ	O3-C3-CA3	-2.35	119.30	126.39
1	D	64	5SQ	C2H-N2H-C1H	2.22	109.24	105.78
1	B	64	5SQ	C2H-N2H-C1H	2.19	109.20	105.78
1	A	64	5SQ	C2H-N2H-C1H	2.16	109.16	105.78
1	D	64	5SQ	CA3-N3-C2	2.15	128.72	123.80
1	C	64	5SQ	C2H-N2H-C1H	2.12	109.09	105.78

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	64	5SQ	C3-CA3-N3-C2
1	A	64	5SQ	CA1-CB1-CG1-C2H

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	A	64	5SQ	CA1-CB1-CG1-N1H
1	B	64	5SQ	C3-CA3-N3-C1
1	B	64	5SQ	C3-CA3-N3-C2
1	B	64	5SQ	CA1-CB1-CG1-C2H
1	B	64	5SQ	CA1-CB1-CG1-N1H
1	C	64	5SQ	C3-CA3-N3-C1
1	C	64	5SQ	C3-CA3-N3-C2
1	C	64	5SQ	CA1-CB1-CG1-C2H
1	C	64	5SQ	CA1-CB1-CG1-N1H
1	D	64	5SQ	C3-CA3-N3-C1
1	D	64	5SQ	C3-CA3-N3-C2
1	D	64	5SQ	CA1-CB1-CG1-C2H
1	D	64	5SQ	CA1-CB1-CG1-N1H
1	A	64	5SQ	C3-CA3-N3-C1

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 2 are modelled with single atom - leaving 19 for Mogul analysis.

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	302	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	D	301	-	4,4,4	0.15	0	6,6,6	0.05	0
2	SO4	A	303	-	4,4,4	0.12	0	6,6,6	0.05	0
2	SO4	A	304	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	B	302	-	4,4,4	0.15	0	6,6,6	0.05	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	0.05	0
2	SO4	D	306	-	4,4,4	0.13	0	6,6,6	0.05	0
2	SO4	D	304	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	C	303	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	C	301	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	D	302	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	D	305	-	4,4,4	0.13	0	6,6,6	0.06	0
2	SO4	C	302	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	A	301	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	C	306	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	C	304	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	C	305	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	D	303	-	4,4,4	0.14	0	6,6,6	0.05	0
2	SO4	A	307	-	4,4,4	0.14	0	6,6,6	0.05	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.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	304	SO4	1	0
2	C	303	SO4	1	0
2	A	301	SO4	2	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

### 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, 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	222/223 (99%)	0.28	4 (1%) 68 76	48, 61, 92, 164	0
1	B	222/223 (99%)	0.12	3 (1%) 75 82	51, 65, 90, 189	4 (1%)
1	C	220/223 (98%)	0.07	2 (0%) 84 89	51, 66, 99, 210	1 (0%)
1	D	220/223 (98%)	0.56	17 (7%) 13 16	54, 70, 111, 175	1 (0%)
All	All	884/892 (99%)	0.26	26 (2%) 51 61	48, 65, 100, 210	6 (0%)

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	223	ASN	4.6
1	B	222	ASP	4.5
1	B	223	ASN	4.5
1	A	222	ASP	4.0
1	C	222	ASP	4.0
1	B	-1	HIS	3.7
1	D	104[A]	ARG	3.5
1	D	1	MET	3.2
1	D	21[A]	HIS	3.1
1	D	209[A]	LYS	3.1
1	D	42	LEU	3.0
1	D	107	ILE	2.9
1	D	30	THR	2.8
1	A	-1	HIS	2.6
1	D	223	ASN	2.5
1	A	1	MET	2.4
1	D	4	ILE	2.4
1	D	208	VAL	2.3
1	D	49	PRO	2.2
1	D	26	ASP	2.2
1	D	203[A]	LYS	2.2

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	113	THR	2.2
1	D	68	PHE	2.2
1	D	114	PHE	2.2
1	D	185[A]	LYS	2.1
1	C	104[A]	ARG	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	5SQ	B	64	25/26	0.94	0.18	51,56,63,69	0
1	5SQ	D	64	25/26	0.94	0.21	50,62,66,68	0
1	5SQ	C	64	25/26	0.95	0.17	52,55,61,63	0
1	5SQ	A	64	25/26	0.95	0.19	45,47,56,61	0

## 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, 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
3	NH4	A	305	1/1	0.48	0.33	61,61,61,61	0
2	SO4	C	302	5/5	0.50	0.22	158,160,160,160	0
2	SO4	D	304	5/5	0.58	0.50	151,153,153,154	0
2	SO4	D	305	5/5	0.59	0.29	158,159,160,160	0
2	SO4	D	306	5/5	0.63	0.48	149,150,151,151	0
2	SO4	C	305	5/5	0.63	0.42	159,160,162,163	0
2	SO4	A	302	5/5	0.66	0.29	147,150,150,152	0
2	SO4	B	302	5/5	0.71	0.36	158,158,159,160	0
2	SO4	A	303	5/5	0.74	0.23	140,141,142,144	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	A	307	5/5	0.76	0.66	160,161,162,164	0
2	SO4	C	304	5/5	0.77	0.29	148,148,150,150	0
3	NH4	A	306	1/1	0.79	0.56	138,138,138,138	0
2	SO4	D	303	5/5	0.80	0.37	148,149,151,152	0
2	SO4	C	301	5/5	0.80	0.30	152,153,155,157	0
2	SO4	A	301	5/5	0.80	0.23	156,156,156,157	0
2	SO4	C	306	5/5	0.88	0.22	144,145,147,149	0
2	SO4	D	302	5/5	0.88	0.25	123,127,128,128	0
2	SO4	B	301	5/5	0.89	0.15	137,139,139,140	0
2	SO4	D	301	5/5	0.90	0.14	129,131,131,132	0
2	SO4	A	304	5/5	0.93	0.15	129,133,133,137	0
2	SO4	C	303	5/5	0.95	0.17	128,129,129,129	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.