



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 5, 2023 – 11:55 pm GMT

PDB ID : 2VVJ  
Title : IrisFP fluorescent protein in its red form, cis conformation  
Authors : Adam, V.; Lelimosin, M.; Boehme, S.; Desfonds, G.; Nienhaus, K.; Field, M.J.; Wiedenmann, J.; McSweeney, S.; Nienhaus, G.U.; Bourgeois, D.  
Deposited on : 2008-06-09  
Resolution : 2.00 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 : 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

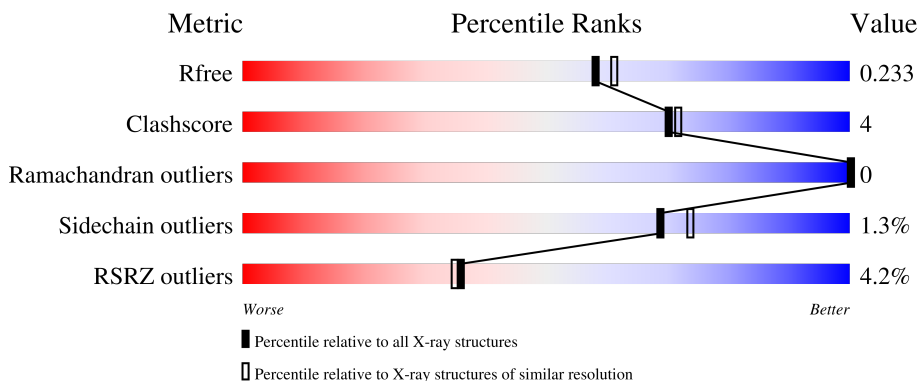
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	aa	63	94% (5% poor fit)
1	bb	63	94% (5% poor fit)
1	cc	63	98% (2% poor fit)
1	dd	63	97% (8% poor fit)
2	A	163	87% (5% poor fit, 9% outliers)

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	B	163	
2	C	163	
2	D	163	

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
4	SO3	A	1226	-	-	X	-
4	SO3	B	1228	-	-	X	X
4	SO3	C	1226	-	-	X	-
4	SO3	D	1227	-	-	X	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 8361 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	aa	60	470	299	79	89	3	0	2	0
1	bb	60	471	299	80	89	3	0	2	0
1	cc	63	499	316	87	92	4	0	2	0
1	dd	62	489	310	84	91	4	0	2	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
aa	-1	HIS	-	expression tag	UNP Q5S6Z9
aa	0	HIS	-	expression tag	UNP Q5S6Z9
bb	-1	HIS	-	expression tag	UNP Q5S6Z9
bb	0	HIS	-	expression tag	UNP Q5S6Z9
cc	-1	HIS	-	expression tag	UNP Q5S6Z9
cc	0	HIS	-	expression tag	UNP Q5S6Z9
dd	-1	HIS	-	expression tag	UNP Q5S6Z9
dd	0	HIS	-	expression tag	UNP Q5S6Z9

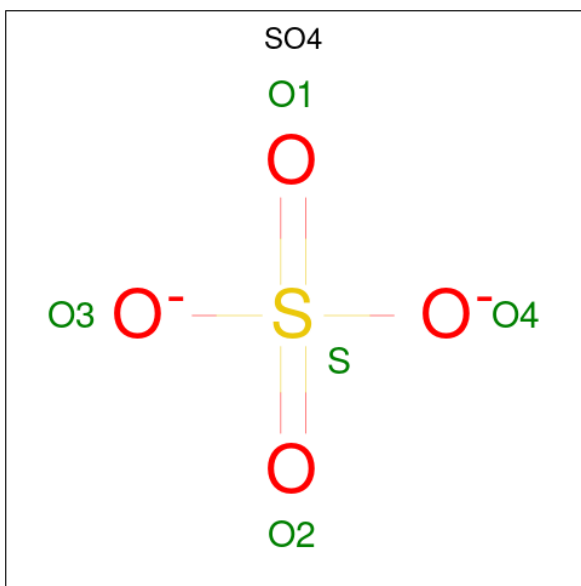
- Molecule 2 is a protein called Green to red photoconvertible GFP-like protein EosFP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	159	1333	853	225	247	8	0	5	0
2	B	160	1355	865	229	253	8	0	7	0
2	C	159	1347	861	227	251	8	0	7	0
2	D	159	1339	856	226	249	8	0	5	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	64	RC7	HIS	chromophore	UNP Q5S6Z9
A	64	RC7	TYR	chromophore	UNP Q5S6Z9
A	64	RC7	GLY	chromophore	UNP Q5S6Z9
A	173	SER	PHE	engineered mutation	UNP Q5S6Z9
A	191	LEU	PHE	engineered mutation	UNP Q5S6Z9
B	64	RC7	HIS	chromophore	UNP Q5S6Z9
B	64	RC7	TYR	chromophore	UNP Q5S6Z9
B	64	RC7	GLY	chromophore	UNP Q5S6Z9
B	173	SER	PHE	engineered mutation	UNP Q5S6Z9
B	191	LEU	PHE	engineered mutation	UNP Q5S6Z9
C	64	RC7	HIS	chromophore	UNP Q5S6Z9
C	64	RC7	TYR	chromophore	UNP Q5S6Z9
C	64	RC7	GLY	chromophore	UNP Q5S6Z9
C	173	SER	PHE	engineered mutation	UNP Q5S6Z9
C	191	LEU	PHE	engineered mutation	UNP Q5S6Z9
D	64	RC7	HIS	chromophore	UNP Q5S6Z9
D	64	RC7	TYR	chromophore	UNP Q5S6Z9
D	64	RC7	GLY	chromophore	UNP Q5S6Z9
D	173	SER	PHE	engineered mutation	UNP Q5S6Z9
D	191	LEU	PHE	engineered mutation	UNP Q5S6Z9

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



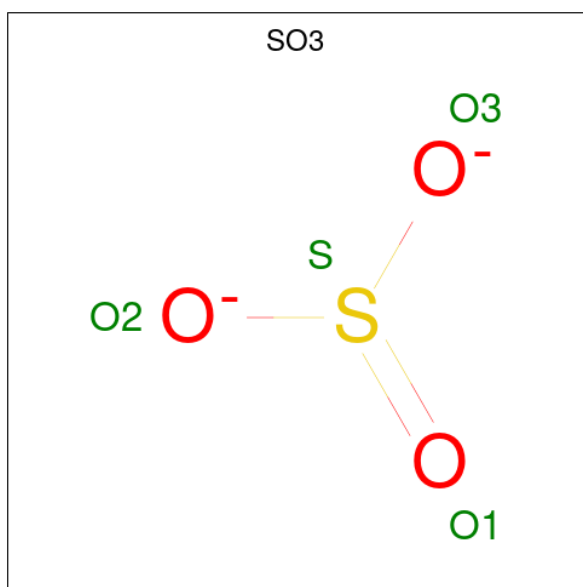
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 4 is SULFITE ION (three-letter code: SO3) (formula: O<sub>3</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 4 3 1	0	0
4	B	1	Total O S 4 3 1	0	0
4	C	1	Total O S 4 3 1	0	0
4	D	1	Total O S 4 3 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	284	Total O 284 284	0	0
5	B	225	Total O 225 225	0	0
5	C	230	Total O 230 230	0	0
5	D	228	Total O 228 228	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.

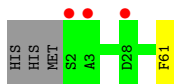
- Molecule 1: Green to red photoconvertible GFP-like protein EosFP

Chain aa: 



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

Chain bb: 



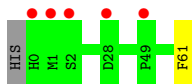
- Molecule 1: Green to red photoconvertible GFP-like protein EosFP

Chain cc: 




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

Chain dd: 



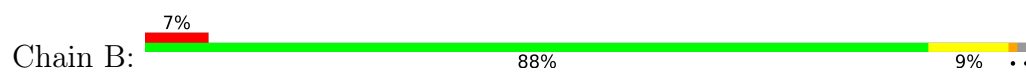
- Molecule 2: Green to red photoconvertible GFP-like protein EosFP

Chain A: 

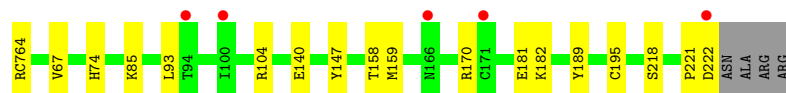
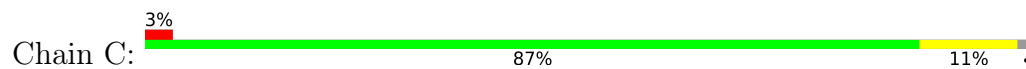


- Molecule 2: Green to red photoconvertible GFP-like protein EosFP

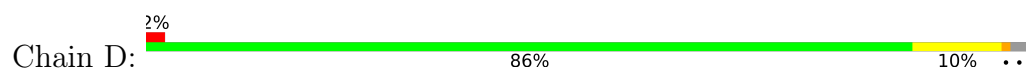




- Molecule 2: Green to red photoconvertible GFP-like protein EosFP



- Molecule 2: 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, $\alpha$ , $\beta$ , $\gamma$	85.67Å 96.62Å 140.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.69 – 2.00 42.84 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (45.69-2.00) 99.1 (42.84-2.00)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.07 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.186 , 0.232 0.188 , 0.233	Depositor DCC
$R_{free}$ test set	3927 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.1	Xtrriage
Anisotropy	0.075	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 48.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8361	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% 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: RC7, SO4, NFA, SO3

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	aa	0.34	0/473	0.50	0/634
1	bb	0.32	0/474	0.48	0/636
1	cc	0.32	0/504	0.48	0/676
1	dd	0.33	0/493	0.48	0/661
2	A	0.36	0/1360	0.52	0/1837
2	B	0.34	0/1385	0.52	0/1871
2	C	0.35	0/1377	0.51	0/1860
2	D	0.33	0/1363	0.51	0/1841
All	All	0.34	0/7429	0.51	0/10016

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	aa	470	0	463	0	0
1	bb	471	0	464	0	0
1	cc	499	0	490	0	0
1	dd	489	0	483	0	0
2	A	1333	0	1279	14	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1355	0	1296	11	0
2	C	1347	0	1290	12	0
2	D	1339	0	1279	13	0
3	A	20	0	0	0	0
3	B	20	0	0	0	0
3	C	15	0	0	0	0
3	D	20	0	0	0	0
4	A	4	0	0	3	0
4	B	4	0	0	2	0
4	C	4	0	0	3	0
4	D	4	0	0	3	0
5	A	284	0	0	1	0
5	B	225	0	0	0	0
5	C	230	0	0	0	0
5	D	228	0	0	0	0
All	All	8361	0	7044	44	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.

The worst 5 of 44 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)
2:C:195[B]:CYS:SG	4:C:1226:SO3:S	2.49	1.10
2:A:195[B]:CYS:SG	4:A:1226:SO3:S	2.51	1.07
2:A:195[B]:CYS:SG	4:A:1226:SO3:O2	2.19	1.00
2:B:195[B]:CYS:SG	4:B:1228:SO3:S	2.60	0.98
2:D:195[B]:CYS:SG	4:D:1227:SO3:O2	2.34	0.86

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	aa	60/63 (95%)	60 (100%)	0	0	100	100
1	bb	60/63 (95%)	60 (100%)	0	0	100	100
1	cc	63/63 (100%)	63 (100%)	0	0	100	100
1	dd	62/63 (98%)	62 (100%)	0	0	100	100
2	A	161/163 (99%)	161 (100%)	0	0	100	100
2	B	164/163 (101%)	162 (99%)	2 (1%)	0	100	100
2	C	163/163 (100%)	163 (100%)	0	0	100	100
2	D	161/163 (99%)	161 (100%)	0	0	100	100
All	All	894/904 (99%)	892 (100%)	2 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	aa	49/51 (96%)	49 (100%)	0	100	100
1	bb	50/51 (98%)	50 (100%)	0	100	100
1	cc	53/51 (104%)	53 (100%)	0	100	100
1	dd	52/51 (102%)	52 (100%)	0	100	100
2	A	143/141 (101%)	141 (99%)	2 (1%)	67	72
2	B	146/141 (104%)	143 (98%)	3 (2%)	53	57
2	C	145/141 (103%)	143 (99%)	2 (1%)	67	72
2	D	143/141 (101%)	140 (98%)	3 (2%)	53	57
All	All	781/768 (102%)	771 (99%)	10 (1%)	69	74

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	66	ARG
2	D	145	LYS

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
2	D	212	GLU
2	B	145	LYS
2	B	180	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	dd	38	GLN
2	D	74	HIS
2	D	81	GLN
2	B	74	HIS
1	cc	19	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

8 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	NFA	dd	61	1	12,12,12	0.47	0	15,15,15	1.06	2 (13%)
2	RC7	D	64	2	24,26,27	4.41	7 (29%)	27,35,37	4.89	12 (44%)
1	NFA	bb	61	1	12,12,12	0.50	0	15,15,15	1.01	1 (6%)
2	RC7	A	64	2	24,26,27	4.36	7 (29%)	27,35,37	4.96	13 (48%)
1	NFA	cc	61	1	12,12,12	0.50	0	15,15,15	0.96	1 (6%)
2	RC7	B	64	2	24,26,27	4.33	7 (29%)	27,35,37	4.85	12 (44%)
1	NFA	aa	61	1	12,12,12	0.52	0	15,15,15	1.09	1 (6%)
2	RC7	C	64	2	24,26,27	4.33	7 (29%)	27,35,37	4.81	12 (44%)

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	NFA	dd	61	1	-	2/8/8/8	0/1/1/1
2	RC7	D	64	2	-	3/9/28/29	0/3/3/3
1	NFA	bb	61	1	-	2/8/8/8	0/1/1/1
2	RC7	A	64	2	-	3/9/28/29	0/3/3/3
1	NFA	cc	61	1	-	2/8/8/8	0/1/1/1
2	RC7	B	64	2	-	2/9/28/29	0/3/3/3
1	NFA	aa	61	1	-	2/8/8/8	0/1/1/1
2	RC7	C	64	2	-	3/9/28/29	0/3/3/3

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	64	RC7	CB2-CA2	17.15	1.49	1.35
2	D	64	RC7	CB2-CA2	17.08	1.49	1.35
2	C	64	RC7	CB2-CA2	17.02	1.49	1.35
2	B	64	RC7	CB2-CA2	16.99	1.49	1.35
2	D	64	RC7	CA2-C2	-10.53	1.38	1.48

The worst 5 of 54 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	64	RC7	CA2-C2-N3	16.75	111.29	103.37
2	B	64	RC7	CA2-C2-N3	16.62	111.23	103.37
2	A	64	RC7	CA2-C2-N3	16.58	111.21	103.37
2	C	64	RC7	CA2-C2-N3	16.13	111.00	103.37
2	A	64	RC7	O2-C2-CA2	-14.33	122.92	130.96

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	aa	61	NFA	O-C-CA-CB
1	aa	61	NFA	NXT-C-CA-CB
2	A	64	RC7	C3-CA3-N3-C1
2	A	64	RC7	C3-CA3-N3-C2
1	bb	61	NFA	O-C-CA-CB

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	64	RC7	1	0
2	A	64	RC7	1	0
2	B	64	RC7	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

19 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	B	1224	-	4,4,4	0.12	0	6,6,6	0.07	0
3	SO4	D	1225	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	A	1224	-	4,4,4	0.14	0	6,6,6	0.07	0
4	SO3	C	1226	-	1,3,3	0.49	0	0,3,3	-	-
4	SO3	B	1228	-	1,3,3	0.54	0	0,3,3	-	-
3	SO4	A	1223	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	D	1226	-	4,4,4	0.14	0	6,6,6	0.05	0
4	SO3	D	1227	-	1,3,3	0.44	0	0,3,3	-	-
3	SO4	D	1224	-	4,4,4	0.14	0	6,6,6	0.06	0
3	SO4	D	1223	-	4,4,4	0.15	0	6,6,6	0.05	0
3	SO4	B	1226	-	4,4,4	0.13	0	6,6,6	0.06	0
3	SO4	C	1225	-	4,4,4	0.17	0	6,6,6	0.10	0
4	SO3	A	1226	-	1,3,3	0.48	0	0,3,3	-	-
3	SO4	B	1225	-	4,4,4	0.14	0	6,6,6	0.06	0
3	SO4	B	1227	-	4,4,4	0.15	0	6,6,6	0.07	0
3	SO4	A	1225	-	4,4,4	0.14	0	6,6,6	0.07	0
3	SO4	C	1224	-	4,4,4	0.15	0	6,6,6	0.05	0
3	SO4	A	1227	-	4,4,4	0.13	0	6,6,6	0.09	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SO4	C	1223	-	4,4,4	0.14	0	6,6,6	0.06	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.

4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	1226	SO3	3	0
4	B	1228	SO3	2	0
4	D	1227	SO3	3	0
4	A	1226	SO3	3	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	aa	59/63 (93%)	-0.29	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	18, 24, 30, 32	0
1	bb	59/63 (93%)	0.39	3 (5%) <span style="border: 1px solid red; padding: 2px;">28</span> <span style="border: 1px solid red; padding: 2px;">27</span>	26, 36, 50, 53	0
1	cc	62/63 (98%)	0.20	1 (1%) <span style="border: 1px solid blue; padding: 2px;">72</span> <span style="border: 1px solid blue; padding: 2px;">70</span>	24, 30, 45, 53	0
1	dd	61/63 (96%)	0.22	5 (8%) <span style="border: 1px solid red; padding: 2px;">11</span> <span style="border: 1px solid red; padding: 2px;">11</span>	25, 32, 45, 56	0
2	A	158/163 (96%)	0.13	8 (5%) <span style="border: 1px solid red; padding: 2px;">28</span> <span style="border: 1px solid red; padding: 2px;">27</span>	18, 23, 32, 43	0
2	B	159/163 (97%)	0.39	11 (6%) <span style="border: 1px solid red; padding: 2px;">16</span> <span style="border: 1px solid red; padding: 2px;">16</span>	22, 29, 40, 43	0
2	C	158/163 (96%)	0.12	5 (3%) <span style="border: 1px solid gray; padding: 2px;">47</span> <span style="border: 1px solid gray; padding: 2px;">46</span>	19, 26, 35, 44	0
2	D	158/163 (96%)	0.15	4 (2%) <span style="border: 1px solid blue; padding: 2px;">57</span> <span style="border: 1px solid blue; padding: 2px;">56</span>	21, 27, 37, 45	0
All	All	874/904 (96%)	0.18	37 (4%) <span style="border: 1px solid red; padding: 2px;">36</span> <span style="border: 1px solid red; padding: 2px;">35</span>	18, 27, 39, 56	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	bb	2	SER	5.3
1	dd	1	MET	5.3
1	dd	0	HIS	4.3
1	cc	-1	HIS	3.8
2	A	183	GLY	3.6

### 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
2	RC7	B	64	24/25	0.92	0.14	32,32,35,36	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	NFA	bb	61	12/12	0.93	0.13	30,31,31,32	0
1	NFA	cc	61	12/12	0.94	0.12	24,25,26,27	0
2	RC7	C	64	24/25	0.94	0.11	26,28,31,32	0
2	RC7	D	64	24/25	0.94	0.12	28,28,29,30	0
2	RC7	A	64	24/25	0.95	0.12	21,23,26,28	0
1	NFA	dd	61	12/12	0.95	0.10	26,26,27,27	0
1	NFA	aa	61	12/12	0.95	0.09	22,22,22,23	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( $\text{\AA}^2$ )	Q<0.9
4	SO3	B	1228	4/4	0.65	0.59	70,70,70,70	4
3	SO4	D	1226	5/5	0.71	0.22	79,79,79,79	5
3	SO4	A	1225	5/5	0.74	0.26	70,70,70,70	5
3	SO4	C	1225	5/5	0.75	0.34	38,39,39,39	5
4	SO3	A	1226	4/4	0.80	0.30	43,43,43,43	4
3	SO4	D	1225	5/5	0.82	0.26	49,50,50,50	5
4	SO3	D	1227	4/4	0.82	0.36	61,61,61,61	4
4	SO3	C	1226	4/4	0.85	0.34	55,55,55,55	4
3	SO4	B	1227	5/5	0.85	0.22	50,50,50,50	5
3	SO4	B	1226	5/5	0.88	0.19	54,54,55,55	5
3	SO4	C	1224	5/5	0.91	0.15	53,53,53,53	5
3	SO4	A	1224	5/5	0.92	0.25	43,43,43,43	5
3	SO4	A	1223	5/5	0.93	0.14	86,86,86,86	0
3	SO4	D	1224	5/5	0.93	0.11	78,78,78,78	0
3	SO4	C	1223	5/5	0.96	0.09	69,69,69,70	0
3	SO4	B	1225	5/5	0.96	0.14	76,76,76,76	0
3	SO4	D	1223	5/5	0.97	0.11	65,65,65,65	0
3	SO4	B	1224	5/5	0.97	0.08	55,55,55,55	0
3	SO4	A	1227	5/5	0.98	0.13	61,61,61,61	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.