



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:43 PM GMT

PDB ID : 4J88  
Title : Dark-state structure of sfGFP containing the unnatural amino acid p-azido-p  
henylalanine at residue 66  
Authors : Reddington, S.C.; Jones, D.D.; Rizkallah, P.J.; Tippmann, E.M.  
Deposited on : 2013-02-14  
Resolution : 2.08 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

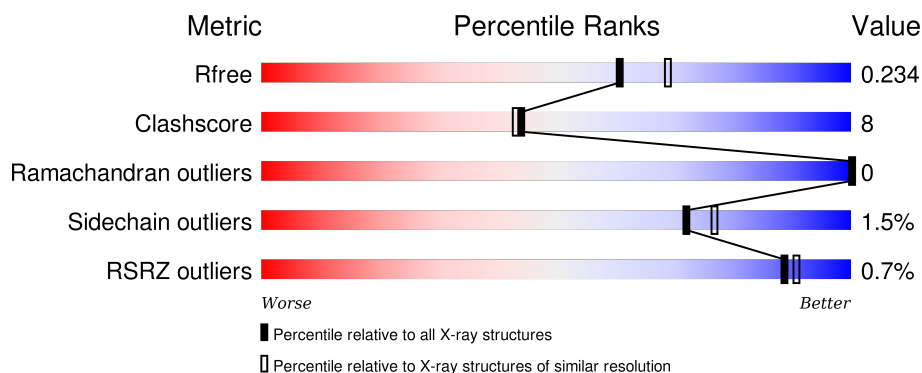
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*



The reported resolution of this entry is 2.08 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))
$R_{free}$	91344	4546 (2.10-2.06)
Clashscore	102246	5101 (2.10-2.06)
Ramachandran outliers	100387	5048 (2.10-2.06)
Sidechain outliers	100360	5049 (2.10-2.06)
RSRZ outliers	91569	4556 (2.10-2.06)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	245	 77% 15% 7%
1	B	245	 80% 12% 7%

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	303	-	-	-	X
2	SO4	A	304	-	-	-	X
2	SO4	B	303	-	-	-	X
2	SO4	B	305	-	-	-	X
3	EDO	A	306	-	-	-	X
3	EDO	A	308	-	-	-	X
3	EDO	A	312	-	-	-	X
3	EDO	A	316	-	-	-	X
3	EDO	A	321	-	-	-	X
3	EDO	A	322	-	-	-	X
3	EDO	B	306	-	-	-	X
3	EDO	B	307	-	-	-	X
4	TRS	A	324	-	-	-	X
4	TRS	A	325	-	-	-	X
4	TRS	A	326	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4141 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 fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	228	Total	C	N	O	S	0	7	0
			1862	1173	323	361	5			
1	B	227	Total	C	N	O	S	0	1	0
			1800	1140	310	345	5			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	EXPRESSION TAG	UNP P42212
A	1	VAL	-	EXPRESSION TAG	UNP P42212
A	30	ARG	SER	ENGINEERED MUTATION	UNP P42212
A	39	ASN	TYR	ENGINEERED MUTATION	UNP P42212
A	66	CQ1	PHE	SEE REMARK 999	UNP P42212
A	66	CQ1	SER	SEE REMARK 999	UNP P42212
A	64	LEU	TYR	SEE REMARK 999	UNP P42212
A	66	CQ1	GLY	SEE REMARK 999	UNP P42212
A	80	ARG	GLN	ENGINEERED MUTATION	UNP P42212
A	99	SER	PHE	ENGINEERED MUTATION	UNP P42212
A	105	THR	ASN	ENGINEERED MUTATION	UNP P42212
A	145	PHE	TYR	ENGINEERED MUTATION	UNP P42212
A	153	THR	MET	ENGINEERED MUTATION	UNP P42212
A	163	ALA	VAL	ENGINEERED MUTATION	UNP P42212
A	171	VAL	ILE	ENGINEERED MUTATION	UNP P42212
A	206	VAL	ALA	ENGINEERED MUTATION	UNP P42212
A	239	GLY	-	EXPRESSION TAG	UNP P42212
A	240	SER	-	EXPRESSION TAG	UNP P42212
A	241	HIS	-	EXPRESSION TAG	UNP P42212
A	242	HIS	-	EXPRESSION TAG	UNP P42212
A	243	HIS	-	EXPRESSION TAG	UNP P42212
A	244	HIS	-	EXPRESSION TAG	UNP P42212
A	245	HIS	-	EXPRESSION TAG	UNP P42212
A	246	HIS	-	EXPRESSION TAG	UNP P42212
B	0	MET	-	EXPRESSION TAG	UNP P42212

*Continued on next page...*

*Continued from previous page...*

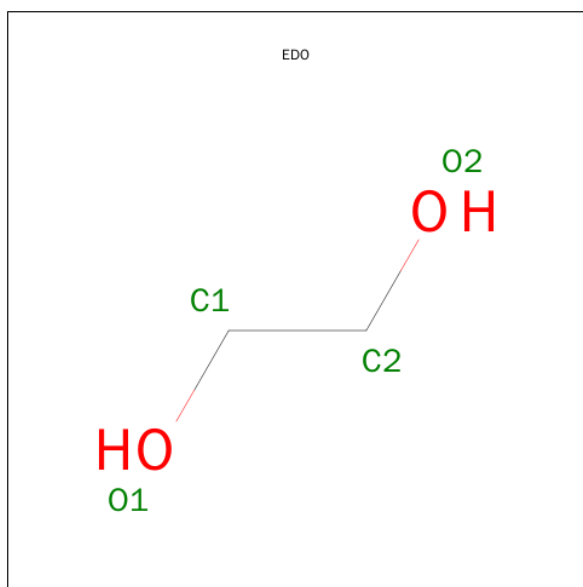
Chain	Residue	Modelled	Actual	Comment	Reference
B	1	VAL	-	EXPRESSION TAG	UNP P42212
B	30	ARG	SER	ENGINEERED MUTATION	UNP P42212
B	39	ASN	TYR	ENGINEERED MUTATION	UNP P42212
A	66	CQ1	PHE	SEE REMARK 999	UNP P42212
A	66	CQ1	SER	SEE REMARK 999	UNP P42212
A	64	LEU	TYR	SEE REMARK 999	UNP P42212
A	66	CQ1	GLY	SEE REMARK 999	UNP P42212
B	80	ARG	GLN	ENGINEERED MUTATION	UNP P42212
B	99	SER	PHE	ENGINEERED MUTATION	UNP P42212
B	105	THR	ASN	ENGINEERED MUTATION	UNP P42212
B	145	PHE	TYR	ENGINEERED MUTATION	UNP P42212
B	153	THR	MET	ENGINEERED MUTATION	UNP P42212
B	163	ALA	VAL	ENGINEERED MUTATION	UNP P42212
B	171	VAL	ILE	ENGINEERED MUTATION	UNP P42212
B	206	VAL	ALA	ENGINEERED MUTATION	UNP P42212
B	239	GLY	-	EXPRESSION TAG	UNP P42212
B	240	SER	-	EXPRESSION TAG	UNP P42212
B	241	HIS	-	EXPRESSION TAG	UNP P42212
B	242	HIS	-	EXPRESSION TAG	UNP P42212
B	243	HIS	-	EXPRESSION TAG	UNP P42212
B	244	HIS	-	EXPRESSION TAG	UNP P42212
B	245	HIS	-	EXPRESSION TAG	UNP P42212
B	246	HIS	-	EXPRESSION TAG	UNP P42212

- 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	B	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	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



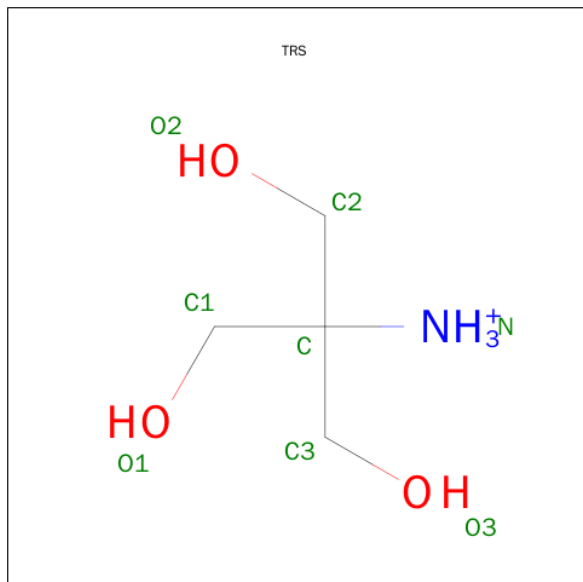
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			8	4	1	3		
4	A	1	Total	C	N	O	0	0
			8	4	1	3		
4	A	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 5 is water.

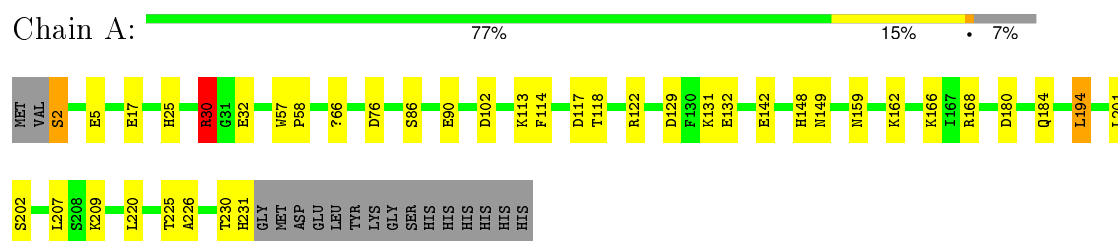
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	190	Total	O	0	0
			190	190		
5	B	124	Total	O	0	0
			124	124		



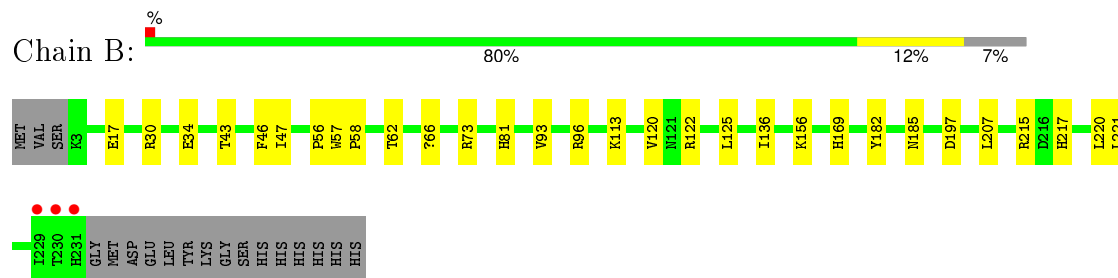
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 fluorescent protein



- Molecule 1: Green fluorescent protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.43 Å 97.56 Å 102.90 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.78 – 2.08 48.78 – 2.08	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.78-2.08) 100.0 (48.78-2.08)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.15 (at 2.08 Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.165 , 0.228 0.173 , 0.234	Depositor DCC
$R_{free}$ test set	1644 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.6	Xtriage
Anisotropy	0.572	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 47.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 32445 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4141	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: TRS, EDO, CQ1, 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	1.01	7/1876 (0.4%)	0.54	0/2538
1	B	0.89	0/1814	0.51	0/2456
All	All	0.95	7/3690 (0.2%)	0.53	0/4994

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	32	GLU	CD-OE2	-7.03	1.18	1.25
1	A	30[A]	ARG	CB-CG	6.00	1.68	1.52
1	A	30[C]	ARG	CB-CG	6.00	1.68	1.52
1	A	30[A]	ARG	CG-CD	5.34	1.65	1.51
1	A	30[C]	ARG	CG-CD	5.34	1.65	1.51
1	A	30[A]	ARG	CA-CB	5.23	1.65	1.53
1	A	30[C]	ARG	CA-CB	5.23	1.65	1.53

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	1862	0	1789	34	1
1	B	1800	0	1735	21	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	20	0	0	1	0
2	B	25	0	0	0	0
3	A	76	0	114	9	0
3	B	20	0	30	0	0
4	A	24	0	36	8	0
5	A	190	0	0	7	0
5	B	124	0	0	3	0
All	All	4141	0	3704	60	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

All (60) 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)
3:A:321:EDO:H21	5:A:550:HOH:O	1.61	1.00
4:A:326:TRS:HO1	4:A:326:TRS:HO3	0.83	0.80
1:A:5:GLU:HB2	5:A:555:HOH:O	1.87	0.73
1:A:25:HIS:HB3	3:A:318:EDO:H22	1.70	0.71
1:A:30[C]:ARG:NH1	5:A:546:HOH:O	2.23	0.69
1:A:90:GLU:OE2	5:A:564:HOH:O	2.12	0.66
1:B:43:THR:HG22	1:B:221:LEU:HD13	1.76	0.66
1:B:17:GLU:OE2	1:B:122:ARG:NH1	2.30	0.65
1:A:17:GLU:OE2	1:A:122:ARG:NH1	2.30	0.64
1:A:76:ASP:H	4:A:324:TRS:H32	1.62	0.64
1:A:2:SER:N	1:A:5:GLU:OE2	2.31	0.64
1:A:117:ASP:HB2	4:A:325:TRS:H32	1.80	0.64
1:A:225[A]:THR:HG21	5:B:522:HOH:O	1.99	0.62
1:B:17:GLU:OE1	1:B:30:ARG:NH2	2.36	0.58
1:A:25:HIS:HB3	3:A:318:EDO:C2	2.33	0.58
3:A:319:EDO:H22	5:A:580:HOH:O	2.05	0.56
2:A:302:SO4:O4	1:B:73:ARG:NH1	2.38	0.56
1:B:169:HIS:HD2	5:B:401:HOH:O	1.88	0.55
1:A:25:HIS:HE1	1:A:132:GLU:OE1	1.89	0.55
1:B:47:ILE:HD13	1:B:215:ARG:CZ	2.37	0.55
1:A:76:ASP:N	4:A:324:TRS:H32	2.22	0.54
1:A:86:SER:HB2	1:A:194:LEU:HD22	1.91	0.53
4:A:325:TRS:H22	5:A:584:HOH:O	2.09	0.53
1:B:113:LYS:O	1:B:120[B]:VAL:HG22	2.09	0.53
1:A:230:THR:O	1:A:231:HIS:C	2.48	0.51
1:B:66:CQ1:N2	1:B:66:CQ1:H3	2.26	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:102:ASP:HA	4:A:326:TRS:H21	1.92	0.50
1:A:184:GLN:HB3	3:A:322:EDO:H22	1.92	0.50
1:B:81:HIS:CD2	1:B:197:ASP:H	2.29	0.50
1:A:207:LEU:CD2	1:A:220:LEU:HD13	2.41	0.50
1:A:131:LYS:HA	3:A:323:EDO:H11	1.94	0.50
1:A:86:SER:HB2	1:A:194:LEU:CD2	2.43	0.48
1:A:86:SER:HB3	1:A:194:LEU:HD23	1.95	0.48
1:B:34:GLU:HG2	5:B:453:HOH:O	2.12	0.48
4:A:326:TRS:C1	5:A:417:HOH:O	2.62	0.48
1:A:114:PHE:O	3:A:312:EDO:C1	2.62	0.47
1:B:207:LEU:HD22	1:B:220:LEU:HD13	1.96	0.47
1:A:149:ASN:ND2	1:A:202:SER:OG	2.48	0.46
1:A:149:ASN:HA	1:A:201:LEU:O	2.16	0.46
1:B:57:TRP:N	1:B:58:PRO:CD	2.79	0.46
1:A:166:LYS:HE3	1:A:180:ASP:OD1	2.16	0.46
1:B:96:ARG:HA	1:B:182:TYR:O	2.15	0.46
1:A:57:TRP:N	1:A:58:PRO:CD	2.79	0.46
1:A:142:GLU:HB3	3:A:311:EDO:H21	1.99	0.45
1:B:62:THR:O	1:B:66:CQ1:C2	2.65	0.44
1:A:225[A]:THR:HG22	1:A:226:ALA:N	2.33	0.44
1:A:2:SER:O	1:A:5:GLU:HG2	2.17	0.44
1:B:120[B]:VAL:O	1:B:120[B]:VAL:HG23	2.18	0.43
1:A:148:HIS:CE1	1:A:168:ARG:CZ	3.02	0.43
1:B:81:HIS:HD2	1:B:197:ASP:H	1.66	0.43
1:A:209:LYS:HB2	3:A:314:EDO:H11	2.01	0.43
1:A:118:THR:N	4:A:325:TRS:O3	2.52	0.43
1:A:86:SER:CB	1:A:194:LEU:HD23	2.49	0.43
1:A:66:CQ1:H3	1:A:66:CQ1:N2	2.33	0.43
1:B:125:LEU:HD23	1:B:125:LEU:C	2.39	0.42
1:B:47:ILE:HD13	1:B:215:ARG:NH1	2.35	0.42
1:B:46:PHE:O	1:B:217:HIS:HB2	2.20	0.42
1:B:93:VAL:O	1:B:185:ASN:HA	2.19	0.41
1:A:57:TRP:HB2	1:A:58:PRO:HD3	2.03	0.41
1:B:56:PRO:HD3	1:B:136:ILE:O	2.21	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:129[B]:ASP:OD2	1:B:156:LYS:O[2_454]	2.16	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	230/245 (94%)	227 (99%)	3 (1%)	0	100	100
1	B	223/245 (91%)	221 (99%)	2 (1%)	0	100	100
All	All	453/490 (92%)	448 (99%)	5 (1%)	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	202/214 (94%)	194 (96%)	8 (4%)	38	37
1	B	194/214 (91%)	194 (100%)	0	100	100
All	All	396/428 (92%)	388 (98%)	8 (2%)	72	67

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	SER
1	A	30[A]	ARG
1	A	30[C]	ARG
1	A	113[A]	LYS
1	A	113[B]	LYS
1	A	159	ASN
1	A	162	LYS
1	A	194	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	25	HIS
1	A	135	ASN
1	A	148	HIS
1	A	149	ASN
1	A	169	HIS
1	A	177	GLN
1	B	69	GLN
1	B	81	HIS
1	B	135	ASN
1	B	149	ASN
1	B	169	HIS
1	B	177	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](#)

2 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	CQ1	A	66	1	24,25,26	3.44	9 (37%)	30,34,36	3.97	9 (30%)
1	CQ1	B	66	1	24,25,26	4.14	10 (41%)	30,34,36	4.24	6 (20%)

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	CQ1	A	66	1	-	0/16/34/35	0/2/2/2
1	CQ1	B	66	1	-	0/16/34/35	0/2/2/2

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	66	CQ1	CA2-C2	-8.16	1.39	1.48
1	A	66	CQ1	CA2-C2	-7.82	1.40	1.48
1	B	66	CQ1	C2-N3	-5.72	1.27	1.39
1	A	66	CQ1	C2-N3	-4.58	1.30	1.39
1	B	66	CQ1	O3-C3	-3.43	1.23	1.42
1	A	66	CQ1	CA1-C1	-3.30	1.46	1.51
1	A	66	CQ1	O3-C3	-2.97	1.26	1.42
1	B	66	CQ1	CA1-C1	-2.69	1.47	1.51
1	A	66	CQ1	CE2-CD2	-2.38	1.34	1.38
1	B	66	CQ1	CZ-NW	-2.13	1.37	1.43
1	B	66	CQ1	C1-N3	2.24	1.41	1.37
1	A	66	CQ1	NX-NW	2.64	1.31	1.24
1	B	66	CQ1	O2-C2	2.87	1.29	1.23
1	B	66	CQ1	NX-NW	3.40	1.33	1.24
1	A	66	CQ1	O2-C2	3.44	1.30	1.23
1	B	66	CQ1	C1-N2	4.27	1.39	1.32
1	A	66	CQ1	C1-N2	5.00	1.40	1.32
1	A	66	CQ1	CB2-CA2	10.53	1.44	1.35
1	B	66	CQ1	CB2-CA2	15.21	1.48	1.35

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	66	CQ1	O2-C2-CA2	-16.24	122.18	130.95
1	A	66	CQ1	O2-C2-CA2	-15.00	122.85	130.95
1	B	66	CQ1	CG2-CB2-CA2	-4.01	125.00	130.22
1	A	66	CQ1	CG2-CB2-CA2	-3.35	125.87	130.22
1	A	66	CQ1	CB2-CA2-N2	-2.98	123.34	128.67
1	A	66	CQ1	CG1-CB1-CA1	-2.35	109.07	112.53
1	B	66	CQ1	CA2-N2-C1	-2.33	103.59	105.71
1	B	66	CQ1	CB2-CA2-N2	-2.32	124.52	128.67
1	A	66	CQ1	CA2-N2-C1	-2.31	103.62	105.71
1	A	66	CQ1	CE1-CD1-CG2	-2.26	118.46	121.29
1	B	66	CQ1	CB2-CA2-C2	2.21	125.59	122.36
1	A	66	CQ1	CB2-CA2-C2	2.27	125.68	122.36
1	A	66	CQ1	CA3-N3-C2	2.64	128.73	122.83
1	A	66	CQ1	CA2-C2-N3	13.30	110.06	103.40

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	66	CQ1	CA2-C2-N3	14.53	110.68	103.40

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	66	CQ1	1	0
1	B	66	CQ1	2	0

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

36 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	A	301	-	4,4,4	1.26	1 (25%)	6,6,6	0.33	0
2	SO4	A	302	-	4,4,4	0.36	0	6,6,6	0.22	0
2	SO4	A	303	-	4,4,4	0.57	0	6,6,6	0.07	0
2	SO4	A	304	-	4,4,4	0.62	0	6,6,6	0.14	0
3	EDO	A	305	-	3,3,3	0.32	0	2,2,2	0.32	0
3	EDO	A	306	-	3,3,3	0.38	0	2,2,2	0.55	0
3	EDO	A	307	-	3,3,3	0.50	0	2,2,2	0.40	0
3	EDO	A	308	-	3,3,3	0.61	0	2,2,2	0.36	0
3	EDO	A	309	-	3,3,3	0.58	0	2,2,2	0.34	0
3	EDO	A	310	-	3,3,3	0.45	0	2,2,2	0.50	0
3	EDO	A	311	-	3,3,3	0.42	0	2,2,2	0.43	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	EDO	A	312	-	3,3,3	0.51	0	2,2,2	0.33	0
3	EDO	A	313	-	3,3,3	0.79	0	2,2,2	0.18	0
3	EDO	A	314	-	3,3,3	0.35	0	2,2,2	0.47	0
3	EDO	A	315	-	3,3,3	0.55	0	2,2,2	0.37	0
3	EDO	A	316	-	3,3,3	0.52	0	2,2,2	0.42	0
3	EDO	A	317	-	3,3,3	0.30	0	2,2,2	0.52	0
3	EDO	A	318	-	3,3,3	0.60	0	2,2,2	0.28	0
3	EDO	A	319	-	3,3,3	0.60	0	2,2,2	0.30	0
3	EDO	A	320	-	3,3,3	0.45	0	2,2,2	0.42	0
3	EDO	A	321	-	3,3,3	1.09	0	2,2,2	0.31	0
3	EDO	A	322	-	3,3,3	0.80	0	2,2,2	0.22	0
3	EDO	A	323	-	3,3,3	0.41	0	2,2,2	0.44	0
4	TRS	A	324	-	7,7,7	1.45	2 (28%)	9,9,9	0.77	0
4	TRS	A	325	-	7,7,7	1.04	1 (14%)	9,9,9	0.78	0
4	TRS	A	326	-	7,7,7	0.87	0	9,9,9	0.44	0
2	SO4	B	301	-	4,4,4	0.49	0	6,6,6	0.30	0
2	SO4	B	302	-	4,4,4	0.56	0	6,6,6	0.25	0
2	SO4	B	303	-	4,4,4	0.72	0	6,6,6	0.24	0
2	SO4	B	304	-	4,4,4	1.18	0	6,6,6	0.25	0
2	SO4	B	305	-	4,4,4	0.46	0	6,6,6	0.21	0
3	EDO	B	306	-	3,3,3	0.43	0	2,2,2	0.39	0
3	EDO	B	307	-	3,3,3	0.37	0	2,2,2	0.47	0
3	EDO	B	308	-	3,3,3	0.32	0	2,2,2	0.50	0
3	EDO	B	309	-	3,3,3	0.44	0	2,2,2	0.47	0
3	EDO	B	310	-	3,3,3	0.50	0	2,2,2	0.45	0

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
2	SO4	A	301	-	-	0/0/0/0	0/0/0/0
2	SO4	A	302	-	-	0/0/0/0	0/0/0/0
2	SO4	A	303	-	-	0/0/0/0	0/0/0/0
2	SO4	A	304	-	-	0/0/0/0	0/0/0/0
3	EDO	A	305	-	-	0/1/1/1	0/0/0/0
3	EDO	A	306	-	-	0/1/1/1	0/0/0/0
3	EDO	A	307	-	-	0/1/1/1	0/0/0/0
3	EDO	A	308	-	-	0/1/1/1	0/0/0/0
3	EDO	A	309	-	-	0/1/1/1	0/0/0/0
3	EDO	A	310	-	-	0/1/1/1	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	A	311	-	-	0/1/1/1	0/0/0/0
3	EDO	A	312	-	-	0/1/1/1	0/0/0/0
3	EDO	A	313	-	-	0/1/1/1	0/0/0/0
3	EDO	A	314	-	-	0/1/1/1	0/0/0/0
3	EDO	A	315	-	-	0/1/1/1	0/0/0/0
3	EDO	A	316	-	-	0/1/1/1	0/0/0/0
3	EDO	A	317	-	-	0/1/1/1	0/0/0/0
3	EDO	A	318	-	-	0/1/1/1	0/0/0/0
3	EDO	A	319	-	-	0/1/1/1	0/0/0/0
3	EDO	A	320	-	-	0/1/1/1	0/0/0/0
3	EDO	A	321	-	-	0/1/1/1	0/0/0/0
3	EDO	A	322	-	-	0/1/1/1	0/0/0/0
3	EDO	A	323	-	-	0/1/1/1	0/0/0/0
4	TRS	A	324	-	-	0/9/9/9	0/0/0/0
4	TRS	A	325	-	-	0/9/9/9	0/0/0/0
4	TRS	A	326	-	-	0/9/9/9	0/0/0/0
2	SO4	B	301	-	-	0/0/0/0	0/0/0/0
2	SO4	B	302	-	-	0/0/0/0	0/0/0/0
2	SO4	B	303	-	-	0/0/0/0	0/0/0/0
2	SO4	B	304	-	-	0/0/0/0	0/0/0/0
2	SO4	B	305	-	-	0/0/0/0	0/0/0/0
3	EDO	B	306	-	-	0/1/1/1	0/0/0/0
3	EDO	B	307	-	-	0/1/1/1	0/0/0/0
3	EDO	B	308	-	-	0/1/1/1	0/0/0/0
3	EDO	B	309	-	-	0/1/1/1	0/0/0/0
3	EDO	B	310	-	-	0/1/1/1	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	324	TRS	C-N	-2.91	1.46	1.50
4	A	324	TRS	C3-C	-2.25	1.48	1.53
4	A	325	TRS	C-N	-2.12	1.47	1.50
2	A	301	SO4	O3-S	2.28	1.55	1.47

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

12 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	302	SO4	1	0
3	A	311	EDO	1	0
3	A	312	EDO	1	0
3	A	314	EDO	1	0
3	A	318	EDO	2	0
3	A	319	EDO	1	0
3	A	321	EDO	1	0
3	A	322	EDO	1	0
3	A	323	EDO	1	0
4	A	324	TRS	2	0
4	A	325	TRS	3	0
4	A	326	TRS	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	A	227/245 (92%)	-0.30	0	100   100	15, 24, 44, 69	0
1	B	226/245 (92%)	-0.42	3 (1%)	79   83	18, 28, 43, 76	0
All	All	453/490 (92%)	-0.36	3 (0%)	89   91	15, 26, 44, 76	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	230	THR	4.1
1	B	231	HIS	2.4
1	B	229	ILE	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	CQ1	A	66	24/25	0.95	0.10	-	17,20,32,39	0
1	CQ1	B	66	24/25	0.96	0.09	-	18,21,32,38	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	A	304	5/5	0.81	0.35	10.60	71,72,82,97	0
2	SO4	B	305	5/5	0.84	0.41	7.80	73,74,83,86	0
3	EDO	A	312	4/4	0.96	0.23	7.01	28,32,33,37	0
3	EDO	A	316	4/4	0.84	0.21	5.05	42,46,51,51	0
2	SO4	A	303	5/5	0.94	0.31	4.54	72,78,85,95	0
4	TRS	A	325	8/8	0.88	0.24	4.24	42,48,58,65	0
3	EDO	A	321	4/4	0.84	0.23	3.69	31,38,40,41	0
3	EDO	A	306	4/4	0.87	0.27	3.55	58,66,67,68	0
3	EDO	B	307	4/4	0.75	0.24	3.09	53,54,56,60	0
3	EDO	A	322	4/4	0.87	0.23	2.95	52,55,55,57	0
3	EDO	A	308	4/4	0.83	0.18	2.86	46,47,48,49	0
2	SO4	B	303	5/5	0.94	0.27	2.75	57,62,64,69	0
4	TRS	A	326	8/8	0.87	0.24	2.64	49,54,62,64	0
4	TRS	A	324	8/8	0.89	0.17	2.54	38,46,57,59	0
3	EDO	B	306	4/4	0.90	0.23	2.29	43,45,49,57	0
3	EDO	A	317	4/4	0.97	0.13	1.72	31,34,37,40	0
2	SO4	B	304	5/5	0.92	0.14	1.53	47,50,55,57	0
3	EDO	A	319	4/4	0.95	0.14	1.23	37,39,41,42	0
3	EDO	B	308	4/4	0.95	0.17	0.96	45,45,47,47	0
3	EDO	A	307	4/4	0.88	0.16	0.85	45,48,50,53	0
2	SO4	B	301	5/5	0.98	0.12	0.56	36,39,39,42	0
2	SO4	A	302	5/5	0.92	0.13	0.32	30,32,36,37	5
3	EDO	B	309	4/4	0.83	0.14	0.13	52,57,62,64	0
2	SO4	A	301	5/5	0.99	0.09	-0.33	28,29,35,40	0
3	EDO	A	305	4/4	0.94	0.12	-0.48	32,33,35,36	0
3	EDO	A	315	4/4	0.72	0.19	-	49,52,52,53	0
3	EDO	A	318	4/4	0.78	0.23	-	45,47,48,49	0
3	EDO	B	310	4/4	0.70	0.21	-	53,56,57,58	0
3	EDO	A	311	4/4	0.86	0.23	-	32,39,44,49	0
3	EDO	A	310	4/4	0.89	0.10	-	43,43,44,44	0
3	EDO	A	313	4/4	0.81	0.17	-	44,46,47,48	0
2	SO4	B	302	5/5	0.92	0.29	-	64,65,69,73	0
3	EDO	A	309	4/4	0.94	0.18	-	43,54,59,64	0
3	EDO	A	323	4/4	0.75	0.31	-	58,59,61,62	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	A	314	4/4	0.93	0.17	-	41,44,47,50	0
3	EDO	A	320	4/4	0.83	0.14	-	60,60,61,63	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.