



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 5, 2026 – 01:22 PM UTC

PDB ID : 4ZIN / pdb\_00004zin  
Title : Genetically encoded Phenyl Azide Photochemistry Drive Positive and Negative Functional Modulation of a Red Fluorescent Protein  
Authors : Reddington, S.C.; Driezis, S.; Hartley, A.M.; Watson, P.D.; Rizkallah, P.J.; Jones, D.D.  
Deposited on : 2015-04-28  
Resolution : 1.67 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 : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

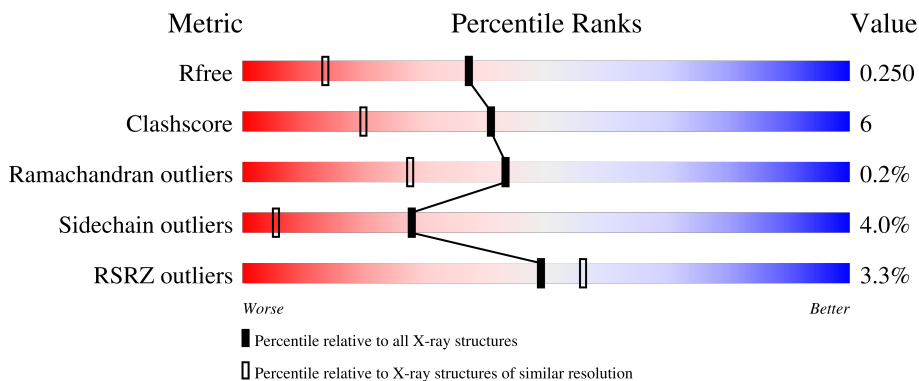
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

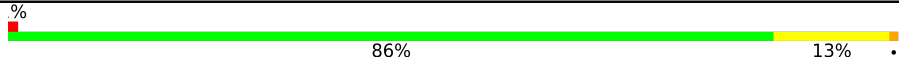
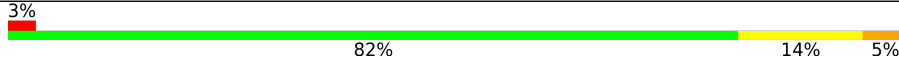

The reported resolution of this entry is 1.67 Å.

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))
$R_{free}$	180053	1054 (1.68-1.68)
Clashscore	190562	1078 (1.68-1.68)
Ramachandran outliers	187476	1068 (1.68-1.68)
Sidechain outliers	187428	1067 (1.68-1.68)
RSRZ outliers	180081	1055 (1.68-1.68)

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	217	
1	B	217	
1	C	217	

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
1	CH6	B	66	-	-	X	-
3	EDO	C	303	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5669 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 MCherry fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	217	1745	1110	296	332	7	0	0	0
1	B	217	1745	1110	296	332	7	0	0	0
1	C	217	1745	1110	296	332	7	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	66	CH6	MET	chromophore	UNP X5DSL3
A	66	CH6	TYR	chromophore	UNP X5DSL3
A	66	CH6	GLY	chromophore	UNP X5DSL3
A	143	4II	TRP	conflict	UNP X5DSL3
B	66	CH6	MET	chromophore	UNP X5DSL3
B	66	CH6	TYR	chromophore	UNP X5DSL3
B	66	CH6	GLY	chromophore	UNP X5DSL3
B	143	4II	TRP	conflict	UNP X5DSL3
C	66	CH6	MET	chromophore	UNP X5DSL3
C	66	CH6	TYR	chromophore	UNP X5DSL3
C	66	CH6	GLY	chromophore	UNP X5DSL3
C	143	4II	TRP	conflict	UNP X5DSL3

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O<sub>4</sub>S).



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

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



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

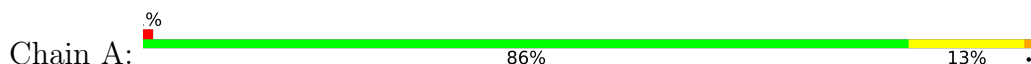
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	172	Total O 172 172	0	0
4	B	131	Total O 131 131	0	0
4	C	97	Total O 97 97	0	0

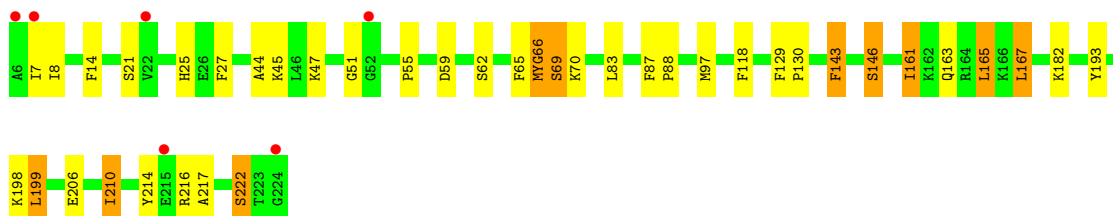
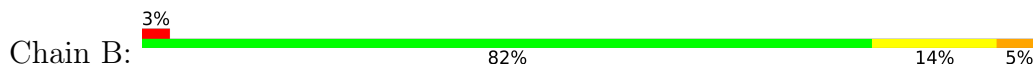
### 3 Residue-property plots [i](#)

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.

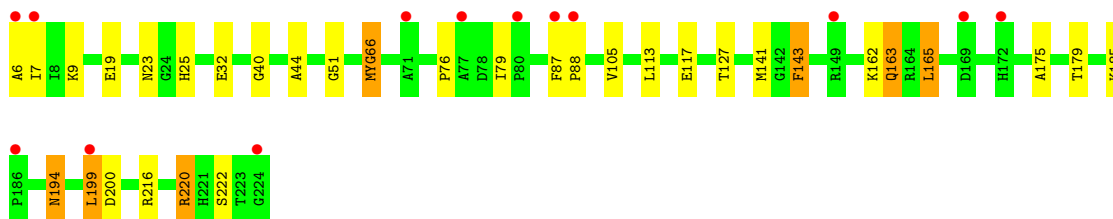
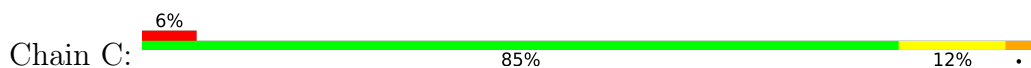
- Molecule 1: MCherry fluorescent protein



- Molecule 1: MCherry fluorescent protein



- Molecule 1: MCherry fluorescent protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.28Å 103.43Å 150.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.71 – 1.67 51.71 – 1.67	Depositor EDS
% Data completeness (in resolution range)	99.9 (51.71-1.67) 99.9 (51.71-1.67)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.21 (at 1.67Å)	Xtrriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.206 , 0.248 0.212 , 0.250	Depositor DCC
$R_{free}$ test set	3874 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.8	Xtrriage
Anisotropy	0.327	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 29.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5669	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.10% 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: CH6, 4II, SO4, EDO

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	1.27	3/1747 (0.2%)	1.15	3/2348 (0.1%)
1	B	1.31	4/1747 (0.2%)	1.14	3/2348 (0.1%)
1	C	1.13	4/1747 (0.2%)	1.08	0/2348
All	All	1.24	11/5241 (0.2%)	1.12	6/7044 (0.1%)

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	54	LEU	CA-C	-6.21	1.46	1.53
1	B	217	ALA	N-CA	6.17	1.53	1.46
1	B	27	PHE	CA-C	-5.95	1.45	1.52
1	C	179	THR	N-CA	-5.62	1.39	1.46
1	C	105	VAL	N-CA	-5.48	1.39	1.46

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	161	ILE	CB-CA-C	-6.77	98.99	110.71
1	A	125	ARG	NE-CZ-NH1	-6.28	115.22	121.50
1	A	125	ARG	CG-CD-NE	-5.38	100.16	112.00
1	B	222	SER	N-CA-C	-5.17	107.13	113.50
1	B	206	GLU	N-CA-C	5.11	117.59	111.71

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	1745	0	1686	15	0
1	B	1745	0	1686	33	0
1	C	1745	0	1686	21	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
2	C	10	0	0	1	0
3	C	4	0	6	4	0
4	A	172	0	0	2	0
4	B	131	0	0	3	0
4	C	97	0	0	2	0
All	All	5669	0	5064	67	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 6.

The worst 5 of 67 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:B:66:CH6:HA1	1:B:69:SER:OG	1.73	0.88
1:B:66:CH6:CE2	1:B:163:GLN:HE22	1.86	0.87
1:A:66:CH6:CE2	1:A:163:GLN:HE22	1.88	0.86
1:C:66:CH6:OH	4:C:401:HOH:O	2.02	0.78
1:A:13:ARG:NH1	1:A:34:GLU:OE1	2.21	0.73

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	211/217 (97%)	209 (99%)	2 (1%)	0	100	100
1	B	211/217 (97%)	206 (98%)	4 (2%)	1 (0%)	24	8
1	C	211/217 (97%)	208 (99%)	3 (1%)	0	100	100
All	All	633/651 (97%)	623 (98%)	9 (1%)	1 (0%)	43	27

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	55	PRO

### 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	182/182 (100%)	177 (97%)	5 (3%)	39	13
1	B	182/182 (100%)	174 (96%)	8 (4%)	25	4
1	C	182/182 (100%)	173 (95%)	9 (5%)	22	3
All	All	546/546 (100%)	524 (96%)	22 (4%)	28	5

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	19	GLU
1	C	163	GLN
1	C	162	LYS
1	C	165	LEU
1	B	45	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	188	GLN

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
1	C	194	ASN
1	B	42	GLN
1	B	98	ASN
1	B	163	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

6 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	CH6	C	66	1	23,24,25	3.52	5 (21%)	28,32,34	3.58	7 (25%)
1	4II	B	143	1	13,14,15	1.21	1 (7%)	11,17,19	1.70	1 (9%)
1	4II	C	143	1	13,14,15	1.52	1 (7%)	11,17,19	1.27	2 (18%)
1	4II	A	143	1	13,14,15	1.48	1 (7%)	11,17,19	0.99	0
1	CH6	A	66	1	23,24,25	3.33	5 (21%)	28,32,34	3.57	12 (42%)
1	CH6	B	66	1	23,24,25	3.52	6 (26%)	28,32,34	4.26	6 (21%)

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	CH6	C	66	1	-	3/12/31/32	0/2/2/2
1	4II	B	143	1	-	4/8/9/11	0/1/1/1
1	4II	C	143	1	-	1/8/9/11	0/1/1/1
1	4II	A	143	1	-	1/8/9/11	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CH6	A	66	1	-	5/12/31/32	0/2/2/2
1	CH6	B	66	1	-	3/12/31/32	0/2/2/2

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	66	CH6	CB2-CA2	14.75	1.49	1.35
1	B	66	CH6	CB2-CA2	14.21	1.48	1.35
1	A	66	CH6	CB2-CA2	13.59	1.48	1.35
1	A	66	CH6	C1-N2	5.67	1.40	1.32
1	B	66	CH6	C1-N2	5.63	1.40	1.32

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	66	CH6	O2-C2-CA2	-15.54	121.10	131.02
1	C	66	CH6	O2-C2-CA2	-11.93	123.41	131.02
1	A	66	CH6	O2-C2-CA2	-10.40	124.39	131.02
1	B	66	CH6	CA2-C2-N3	10.29	112.14	103.50
1	C	66	CH6	CA2-C2-N3	10.06	111.95	103.50

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	66	CH6	N1-CA1-CB1-CG1
1	A	66	CH6	C1-CA1-CB1-CG1
1	A	66	CH6	C3-CA3-N3-C1
1	A	66	CH6	C3-CA3-N3-C2
1	B	66	CH6	C3-CA3-N3-C1

There are no ring outliers.

5 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	66	CH6	2	0
1	B	143	4II	5	0
1	C	143	4II	2	0
1	A	66	CH6	1	0
1	B	66	CH6	10	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

7 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	0.44	0	6,6,6	0.75	0
2	SO4	C	301	-	4,4,4	0.59	0	6,6,6	0.61	0
2	SO4	A	302	-	4,4,4	0.59	0	6,6,6	0.84	0
2	SO4	C	302	-	4,4,4	0.52	0	6,6,6	0.35	0
2	SO4	B	301	-	4,4,4	0.52	0	6,6,6	0.34	0
2	SO4	B	302	-	4,4,4	0.53	0	6,6,6	0.51	0
3	EDO	C	303	-	3,3,3	0.57	0	2,2,2	0.55	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
3	EDO	C	303	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	303	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	301	SO4	1	0
3	C	303	EDO	4	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	215/217 (99%)	0.27	2 (0%) 81 88	16, 25, 36, 56	0
1	B	215/217 (99%)	0.56	6 (2%) 55 62	17, 27, 37, 53	0
1	C	215/217 (99%)	0.82	13 (6%) 27 36	19, 32, 46, 63	0
All	All	645/651 (99%)	0.55	21 (3%) 49 56	16, 28, 41, 63	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	7	ILE	3.9
1	C	6	ALA	3.5
1	B	224	GLY	3.4
1	B	7	ILE	3.4
1	C	186	PRO	3.0

### 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	CH6	A	66	23/24	0.81	0.20	33,40,45,45	0
1	CH6	B	66	23/24	0.81	0.23	44,57,61,63	0
1	CH6	C	66	23/24	0.84	0.18	41,44,49,50	0
1	4II	A	143	14/15	0.88	0.18	24,27,48,49	0
1	4II	C	143	14/15	0.88	0.17	25,30,47,59	0
1	4II	B	143	14/15	0.91	0.14	26,31,56,57	0

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides 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
2	SO4	B	302	5/5	0.69	0.18	61,65,71,74	0
2	SO4	C	301	5/5	0.79	0.14	45,56,68,70	0
2	SO4	B	301	5/5	0.80	0.15	59,65,73,73	0
2	SO4	A	302	5/5	0.81	0.14	36,51,61,64	0
2	SO4	C	302	5/5	0.84	0.12	61,64,67,72	0
2	SO4	A	301	5/5	0.85	0.11	40,53,57,63	0
3	EDO	C	303	4/4	0.90	0.14	24,28,31,42	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.