



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 12, 2024 – 07:42 AM EST

PDB ID : 2Z2S  
Title : Crystal Structure of Rhodobacter sphaeroides SigE in complex with the anti-sigma ChrR  
Authors : Darst, S.A.; Campbell, E.A.  
Deposited on : 2007-05-26  
Resolution : 2.70 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	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

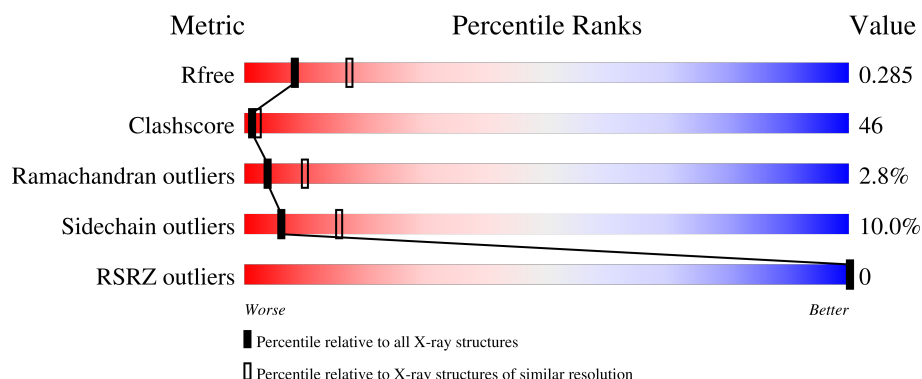
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	164625	3333 (2.70-2.70)
Clashscore	180529	3684 (2.70-2.70)
Ramachandran outliers	177936	3633 (2.70-2.70)
Sidechain outliers	177891	3633 (2.70-2.70)
RSRZ outliers	164620	3333 (2.70-2.70)

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	184	<div> <div>50%</div> <div>35%</div> <div>5%</div> <div>10%</div> </div>
1	C	184	<div> <div>45%</div> <div>44%</div> <div>5%</div> <div>7%</div> </div>
1	E	184	<div> <div>43%</div> <div>34%</div> <div>7%</div> <div>16%</div> </div>
1	G	184	<div> <div>41%</div> <div>38%</div> <div>8%</div> <div>13%</div> </div>
2	B	203	<div> <div>43%</div> <div>39%</div> <div>6%</div> <div>10%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	203	
2	F	203	
2	H	203	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ZN	D	204	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10379 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 RpoE, ECF SigE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	166	Total	C	N	O	S	Se	0	0	0
			1317	827	248	236	1	5			
1	C	172	Total	C	N	O	S	Se	0	0	0
			1333	839	253	235	1	5			
1	E	155	Total	C	N	O	S	Se	0	0	0
			1238	777	235	221	1	4			
1	G	160	Total	C	N	O	S	Se	0	0	0
			1269	799	239	226	1	4			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q3IYV6
A	-1	PRO	-	expression tag	UNP Q3IYV6
A	0	HIS	-	expression tag	UNP Q3IYV6
C	-2	GLY	-	expression tag	UNP Q3IYV6
C	-1	PRO	-	expression tag	UNP Q3IYV6
C	0	HIS	-	expression tag	UNP Q3IYV6
E	-2	GLY	-	expression tag	UNP Q3IYV6
E	-1	PRO	-	expression tag	UNP Q3IYV6
E	0	HIS	-	expression tag	UNP Q3IYV6
G	-2	GLY	-	expression tag	UNP Q3IYV6
G	-1	PRO	-	expression tag	UNP Q3IYV6
G	0	HIS	-	expression tag	UNP Q3IYV6

- Molecule 2 is a protein called Anti-Sigma factor ChrR, transcriptional activator ChrR.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	182	Total	C	N	O	S	Se	0	0	0
			1258	788	225	239	4	2			
2	D	169	Total	C	N	O	S	Se	0	0	0
			1186	746	210	224	4	2			

*Continued on next page...*

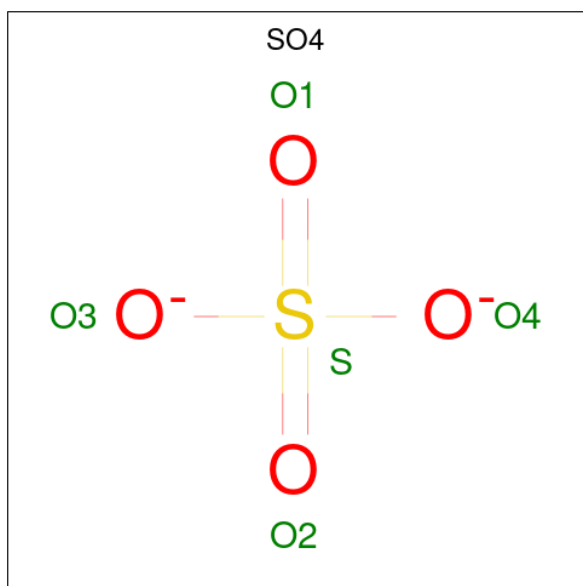
Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	F	169	Total	C	N	O	S	Se	0	0	0
			1161	732	202	221	4	2			
2	H	175	Total	C	N	O	S	Se	0	0	0
			1222	767	210	239	4	2			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	D	1	Total	Zn	0	0
			1	1		
3	F	1	Total	Zn	0	0
			1	1		
3	H	1	Total	Zn	0	0
			1	1		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	E	1	Total	O	S	0	0
			5	4	1		
4	G	1	Total	O	S	0	0
			5	4	1		

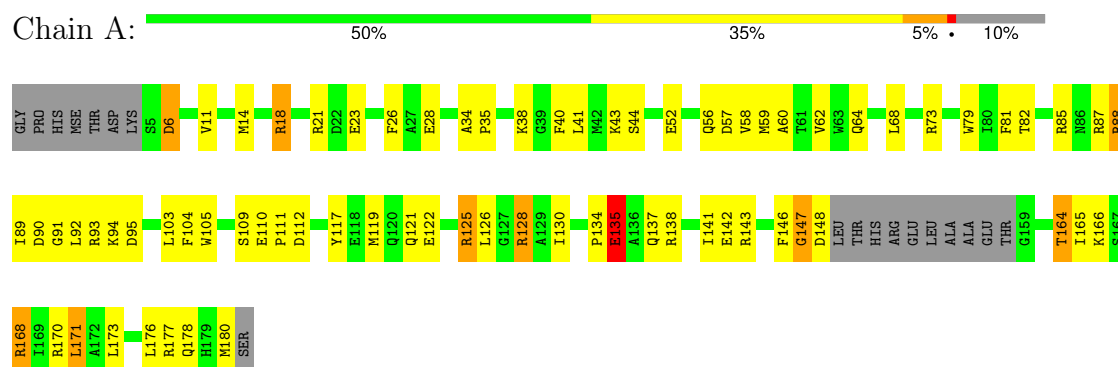
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	52	Total 52	O 52	0	0
5	B	47	Total 47	O 47	0	0
5	C	59	Total 59	O 59	0	0
5	D	38	Total 38	O 38	0	0
5	E	56	Total 56	O 56	0	0
5	F	39	Total 39	O 39	0	0
5	G	56	Total 56	O 56	0	0
5	H	34	Total 34	O 34	0	0

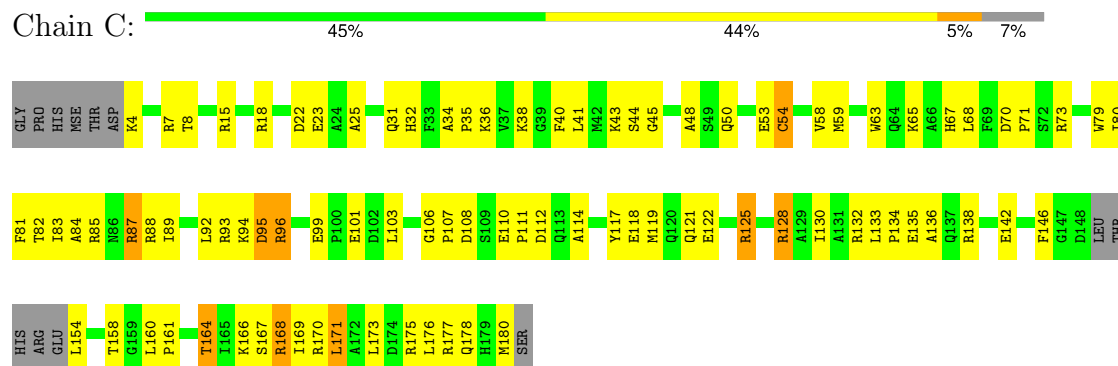
### 3 Residue-property plots

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.

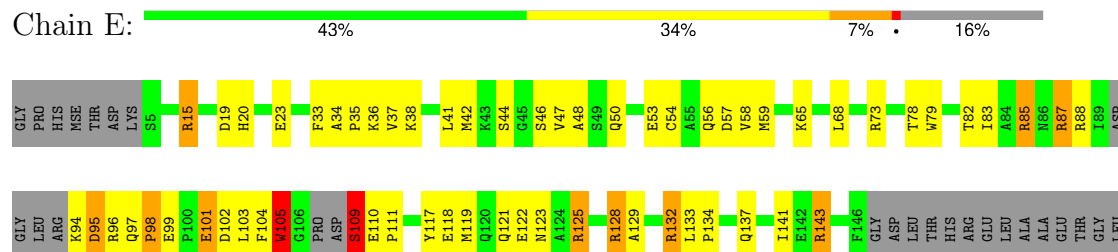
#### • Molecule 1: RpoE, ECF SigE



#### • Molecule 1: RpoE, ECF SigE



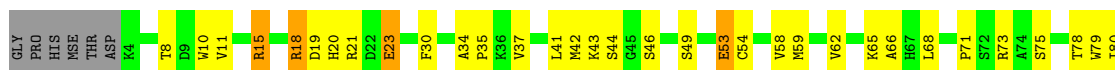
#### • Molecule 1: RpoE, ECF SigE





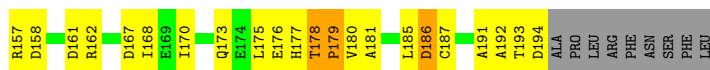
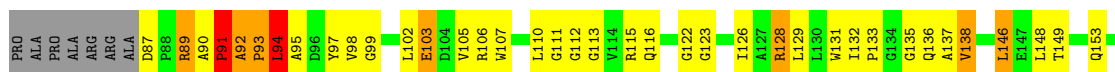
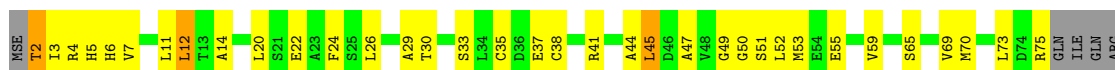
• Molecule 1: RpoE, ECF SigE

Chain G: 41% 38% 8% 13%



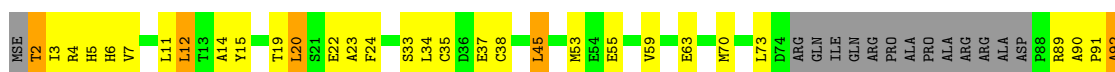
• Molecule 2: Anti-Sigma factor ChrR, transcriptional activator ChrR

Chain B: 43% 39% 6% 10%



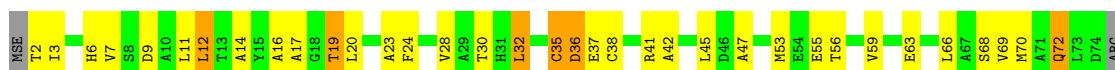
• Molecule 2: Anti-Sigma factor ChrR, transcriptional activator ChrR

Chain D: 38% 34% 8% 17%



• Molecule 2: Anti-Sigma factor ChrR, transcriptional activator ChrR

Chain F: 40% 35% 6% 17%







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	119.14Å 46.45Å 141.70Å 90.00° 91.63° 90.00°	Depositor
Resolution (Å)	25.00 – 2.70 25.00 – 2.70	Depositor EDS
% Data completeness (in resolution range)	75.2 (25.00-2.70) 69.4 (25.00-2.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.43 (at 2.60Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.250 , 0.286 0.247 , 0.285	Depositor DCC
$R_{free}$ test set	1629 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.0	Xtriage
Anisotropy	0.772	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 34.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.155 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10379	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.36% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, ZN

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	0.45	0/1340	0.66	0/1800
1	C	0.44	0/1356	0.68	1/1824 (0.1%)
1	E	0.89	3/1259 (0.2%)	0.92	5/1689 (0.3%)
1	G	0.55	0/1292	0.76	1/1738 (0.1%)
2	B	0.52	1/1276 (0.1%)	0.85	5/1741 (0.3%)
2	D	2.27	20/1202 (1.7%)	1.51	26/1635 (1.6%)
2	F	0.65	3/1173 (0.3%)	1.02	11/1598 (0.7%)
2	H	0.47	0/1237	0.75	1/1687 (0.1%)
All	All	0.95	27/10135 (0.3%)	0.92	50/13712 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	E	0	2
2	F	0	1
All	All	0	3

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	170	ILE	CA-CB	38.69	2.43	1.54
2	D	179	PRO	CA-C	32.82	2.18	1.52
2	D	170	ILE	CA-C	27.23	2.23	1.52
2	D	180	VAL	N-CA	20.70	1.87	1.46
1	E	105	TRP	C-N	-19.66	0.97	1.33

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	179	PRO	CA-CB-CG	-18.86	68.17	104.00
1	E	105	TRP	O-C-N	-16.71	94.80	123.20
2	D	180	VAL	CB-CA-C	-16.03	80.94	111.40
2	D	170	ILE	CG1-CB-CG2	-15.55	77.19	111.40
2	D	170	ILE	N-CA-CB	-14.38	77.72	110.80

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	105	TRP	Mainchain
1	E	109	SER	Mainchain
2	F	90	ALA	Mainchain

## 5.2 Too-close contacts

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	1317	0	1275	89	0
1	C	1333	0	1279	118	0
1	E	1238	0	1194	131	0
1	G	1269	0	1225	119	0
2	B	1258	0	1189	130	0
2	D	1186	0	1156	180	0
2	F	1161	0	1102	116	0
2	H	1222	0	1167	107	0
3	B	1	0	0	0	0
3	D	1	0	0	2	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0
4	E	5	0	0	1	0
4	G	5	0	0	1	0
5	A	52	0	0	19	0
5	B	47	0	0	33	0
5	C	59	0	0	32	0
5	D	38	0	0	16	0
5	E	56	0	0	26	0
5	F	39	0	0	22	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	G	56	0	0	32	0
5	H	34	0	0	18	0
All	All	10379	0	9587	899	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 46.

The worst 5 of 899 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:D:179:PRO:CD	2:D:179:PRO:CG	1.79	1.54
2:D:180:VAL:C	2:D:180:VAL:CA	1.79	1.50
2:D:180:VAL:CA	2:D:180:VAL:CB	1.90	1.48
2:D:156:PHE:HE2	2:D:179:PRO:CB	1.28	1.45
2:D:179:PRO:CA	2:D:179:PRO:N	1.80	1.42

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	162/184 (88%)	149 (92%)	11 (7%)	2 (1%)	11	28
1	C	168/184 (91%)	155 (92%)	11 (6%)	2 (1%)	11	28
1	E	147/184 (80%)	134 (91%)	10 (7%)	3 (2%)	6	16
1	G	154/184 (84%)	146 (95%)	4 (3%)	4 (3%)	4	11
2	B	178/203 (88%)	159 (89%)	13 (7%)	6 (3%)	3	7
2	D	163/203 (80%)	139 (85%)	13 (8%)	11 (7%)	1	1
2	F	159/203 (78%)	143 (90%)	13 (8%)	3 (2%)	6	17
2	H	169/203 (83%)	147 (87%)	17 (10%)	5 (3%)	3	9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1300/1548 (84%)	1172 (90%)	92 (7%)	36 (3%)	4	10

5 of 36 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	91	PRO
2	B	93	PRO
2	B	111	GLY
2	D	92	ALA
2	D	102	LEU

### 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	129/145 (89%)	116 (90%)	13 (10%)	6	15
1	C	125/145 (86%)	115 (92%)	10 (8%)	10	24
1	E	122/145 (84%)	110 (90%)	12 (10%)	6	16
1	G	124/145 (86%)	114 (92%)	10 (8%)	9	23
2	B	114/150 (76%)	103 (90%)	11 (10%)	7	17
2	D	112/150 (75%)	98 (88%)	14 (12%)	3	9
2	F	105/150 (70%)	92 (88%)	13 (12%)	4	9
2	H	115/150 (77%)	103 (90%)	12 (10%)	5	14
All	All	946/1180 (80%)	851 (90%)	95 (10%)	6	16

5 of 95 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	143	ARG
2	F	172	ASP
1	E	179	HIS
2	F	45	LEU
1	G	53	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	20	HIS
1	E	56	GLN
2	H	177	HIS
1	E	123	ASN
2	B	5	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](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	$\# Z  > 2$
4	SO4	G	182	-	4,4,4	0.28	0	6,6,6	0.10	0
4	SO4	E	182	-	4,4,4	0.27	0	6,6,6	0.32	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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	182	SO4	1	0
4	E	182	SO4	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	E	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	109:SER	C	110:GLU	N	1.08
1	E	105:TRP	C	106:GLY	N	0.97



## 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	161/184 (87%)	-1.75	0 100 100	5, 33, 65, 75	0
1	C	167/184 (90%)	-1.71	0 100 100	9, 35, 70, 89	0
1	E	151/184 (82%)	-1.71	0 100 100	11, 38, 68, 83	0
1	G	156/184 (84%)	-1.69	0 100 100	12, 39, 68, 79	0
2	B	180/203 (88%)	-1.62	0 100 100	11, 48, 84, 91	0
2	D	167/203 (82%)	-1.53	0 100 100	21, 55, 88, 94	0
2	F	167/203 (82%)	-1.61	0 100 100	21, 44, 71, 85	0
2	H	173/203 (85%)	-1.67	0 100 100	19, 44, 72, 78	0
All	All	1322/1548 (85%)	-1.66	0 100 100	5, 43, 75, 94	0

There are no RSRZ outliers to report.

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no 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	SO4	G	182	5/5	0.98	0.05	74,75,75,76	0
4	SO4	E	182	5/5	0.99	0.04	73,74,74,75	0
3	ZN	F	204	1/1	1.00	0.02	70,70,70,70	0
3	ZN	H	204	1/1	1.00	0.02	61,61,61,61	0
3	ZN	B	204	1/1	1.00	0.01	49,49,49,49	0
3	ZN	D	204	1/1	1.00	0.01	53,53,53,53	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.