



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:57 AM GMT

PDB ID : 2EYQ  
Title : Crystal structure of Escherichia coli transcription-repair coupling factor  
Authors : Deaconescu, A.M.; Darst, S.A.  
Deposited on : 2005-11-09  
Resolution : 3.20 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  
<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

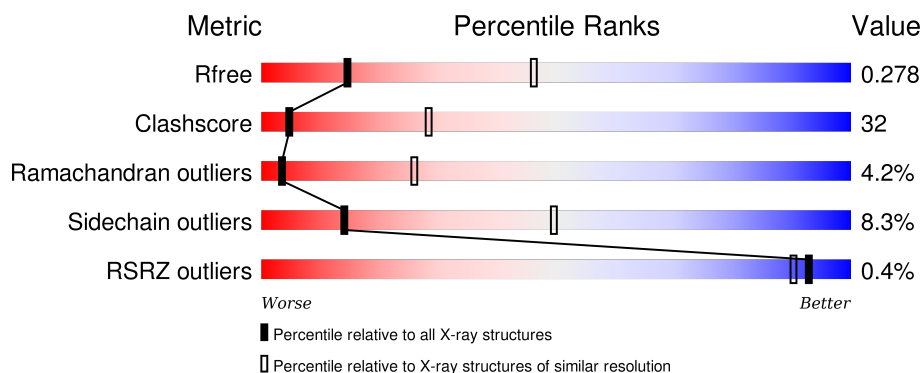
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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	1151	
1	B	1151	

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	EPE	A	1151	-	-	X	X
3	EPE	B	1150	-	-	X	X
3	EPE	B	1151	-	-	X	X
3	EPE	B	1152	-	-	-	X

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 18063 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 Transcription-repair coupling factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1146	Total	C	N	O	S	0	0	0
			8980	5683	1605	1656	36			
1	B	1143	Total	C	N	O	S	0	0	0
			8873	5611	1584	1642	36			

There are 6 discrepancies between the modelled and reference sequences:

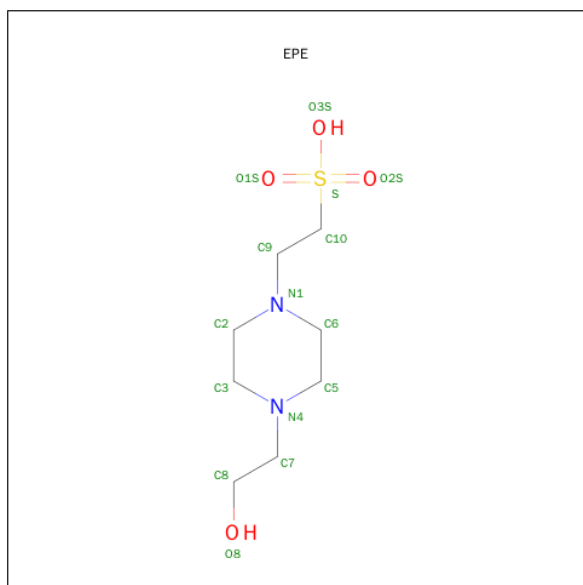
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	CLONING ARTIFACT	UNP P30958
A	-1	PRO	-	CLONING ARTIFACT	UNP P30958
A	0	HIS	-	CLONING ARTIFACT	UNP P30958
B	-2	GLY	-	CLONING ARTIFACT	UNP P30958
B	-1	PRO	-	CLONING ARTIFACT	UNP P30958
B	0	HIS	-	CLONING ARTIFACT	UNP P30958

- 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	B	1	Total O S 5 4 1	0	0

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C N O S 15 8 2 4 1	0	0
3	A	1	Total C N O S 15 8 2 4 1	0	0
3	B	1	Total C N O S 15 8 2 4 1	0	0
3	A	1	Total C N O S 15 8 2 4 1	0	0
3	B	1	Total C N O S 15 8 2 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	68	Total O 68 68	0	0

*Continued on next page...*

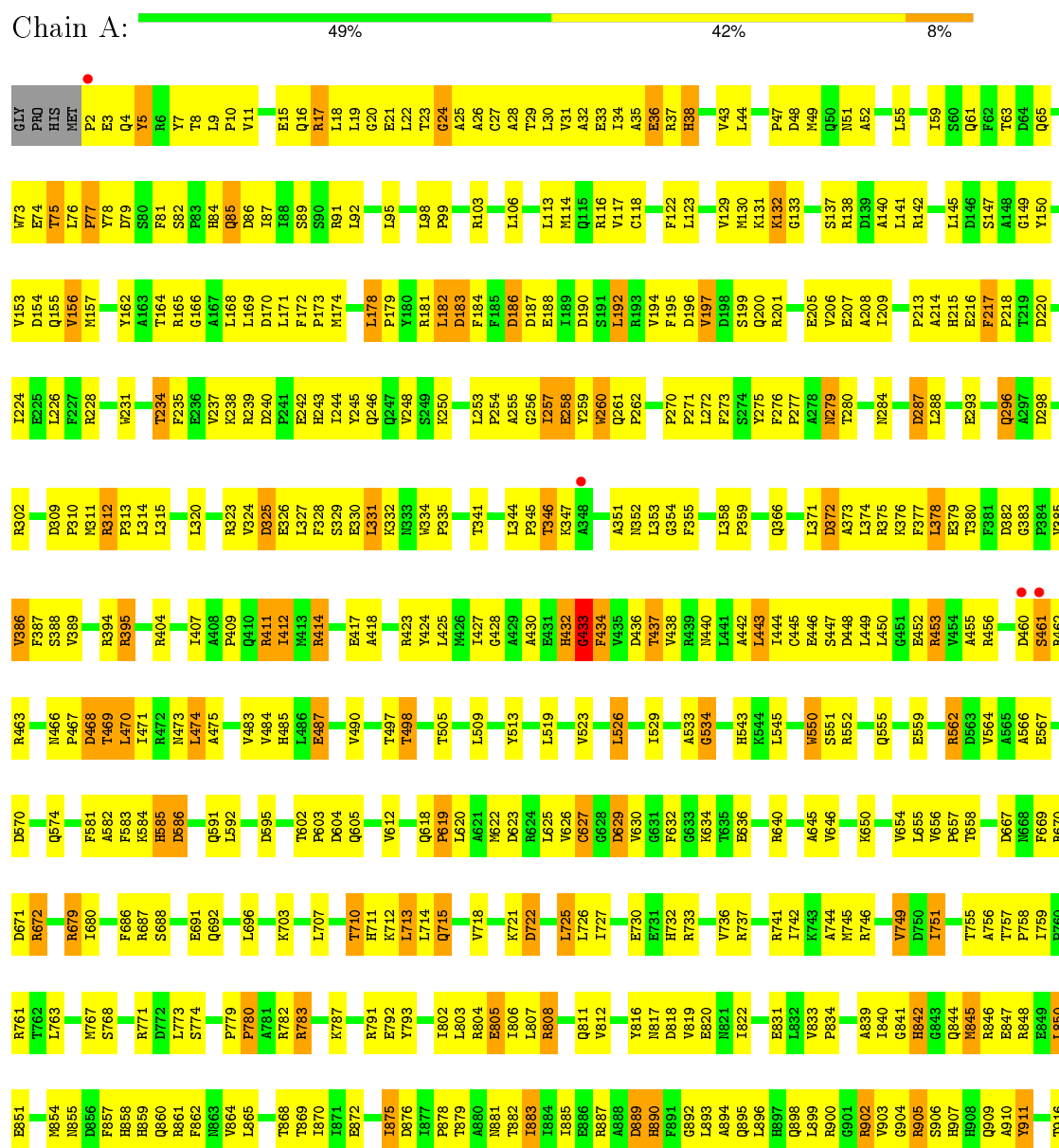
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	52	Total	O	0	0
			52	52		

### 3 Residue-property plots

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: Transcription-repair coupling factor







L979	E980	N981	A982	V983	L986	E991	P992	S993	L994	E995	D996	Q1000	Q1001	L1006	R1007	M1008	P1009	S1010	L1011	L1012	P1013	V1020	N1021	F1026	T1035	E1036	N1037	E1038	L1039	E1040	K1043	V1044	E1045	L1046	L1047	G1051	L1052	L1053	P1054	D1055	P1056	A1057	R1058	T1059	L1060	L1066	R1067	Q1068
Q1071	R1072	L1073	G1074	I1075	R1076	K1077	L1078	G1085	F1089	A1090	E1091	K1092	V1095	N1096	L1100	L1104	Q1105	K1106	Q1107	P1108	Q1109	H1110	L1113	L1119	K1120	F1121	I1122	L1125	R1128	K1129	T1130	R1131	I1132	E1133	R1136	M1139	R1140	E1141	L1142	E1143	E1144	I1147	ALA					

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.87Å 161.99Å 161.73Å 90.00° 105.09° 90.00°	Depositor
Resolution (Å)	40.00 – 3.20 29.91 – 3.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.20) 98.0 (29.91-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.59 (at 3.18Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.234 , 0.295 0.222 , 0.278	Depositor DCC
$R_{free}$ test set	3083 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	63.1	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 45.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 61409 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	18063	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% 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: EPE, 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	0.49	0/9162	0.74	3/12429 (0.0%)
1	B	0.47	0/9051	0.69	1/12284 (0.0%)
All	All	0.48	0/18213	0.72	4/24713 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	414	ARG	NE-CZ-NH2	-11.19	114.70	120.30
1	A	414	ARG	NE-CZ-NH1	9.47	125.03	120.30
1	B	422	GLY	N-CA-C	8.52	134.41	113.10
1	A	433	GLY	N-CA-C	5.70	127.34	113.10

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	8980	0	8865	583	0
1	B	8873	0	8675	551	0
2	A	10	0	0	0	0
2	B	5	0	0	0	0
3	A	30	0	36	10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	45	0	54	26	0
4	A	68	0	0	5	0
4	B	52	0	0	3	0
All	All	18063	0	17630	1130	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 32.

The worst 5 of 1130 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:254:PRO:HD2	1:B:257:ILE:HD12	1.17	1.15
1:A:474:LEU:HD21	1:A:1028:LYS:HG3	1.31	1.07
1:B:117:VAL:HG12	1:B:118:CYS:H	1.16	1.05
1:A:116:ARG:HG3	1:A:320:LEU:O	1.57	1.04
1:B:386:VAL:HB	1:B:443:LEU:HD12	1.39	1.03

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	1144/1151 (99%)	941 (82%)	153 (13%)	50 (4%)	3	24
1	B	1141/1151 (99%)	942 (83%)	153 (13%)	46 (4%)	4	27
All	All	2285/2302 (99%)	1883 (82%)	306 (13%)	96 (4%)	3	26

5 of 96 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	24	GLY

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	25	ALA
1	A	38	HIS
1	A	132	LYS
1	A	187	ASP

### 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	930/975 (95%)	843 (91%)	87 (9%)	11	41
1	B	904/975 (93%)	838 (93%)	66 (7%)	17	57
All	All	1834/1950 (94%)	1681 (92%)	153 (8%)	14	49

5 of 153 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	889	ASP
1	A	1122	ILE
1	B	890	HIS
1	A	900	ARG
1	A	945	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	84	HIS
1	B	350	ASN
1	B	948	HIS
1	B	134	GLN
1	B	230	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

8 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	1149	-	4,4,4	0.24	0	6,6,6	0.27	0
2	SO4	A	1150	-	4,4,4	0.16	0	6,6,6	0.20	0
3	EPE	A	1151	-	14,15,15	1.13	1 (7%)	18,20,20	2.05	2 (11%)
3	EPE	A	1152	-	14,15,15	1.03	1 (7%)	18,20,20	2.07	3 (16%)
2	SO4	B	1149	-	4,4,4	0.25	0	6,6,6	0.19	0
3	EPE	B	1150	-	14,15,15	0.96	1 (7%)	18,20,20	2.02	2 (11%)
3	EPE	B	1151	-	14,15,15	1.25	1 (7%)	18,20,20	1.99	2 (11%)
3	EPE	B	1152	-	14,15,15	1.09	1 (7%)	18,20,20	1.83	3 (16%)

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	1149	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1150	-	-	0/0/0/0	0/0/0/0
3	EPE	A	1151	-	-	0/9/19/19	0/1/1/1
3	EPE	A	1152	-	-	0/9/19/19	0/1/1/1
2	SO4	B	1149	-	-	0/0/0/0	0/0/0/0
3	EPE	B	1150	-	-	0/9/19/19	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EPE	B	1151	-	-	0/9/19/19	0/1/1/1
3	EPE	B	1152	-	-	0/9/19/19	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1152	EPE	O3S-S	2.07	1.51	1.46
3	A	1151	EPE	O3S-S	2.23	1.52	1.46
3	B	1150	EPE	O3S-S	2.29	1.52	1.46
3	B	1151	EPE	O3S-S	2.32	1.52	1.46
3	B	1152	EPE	O3S-S	2.60	1.53	1.46

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1151	EPE	O3S-S-O1S	-3.92	102.50	111.61
3	A	1151	EPE	O3S-S-O1S	-3.24	104.06	111.61
3	B	1150	EPE	O3S-S-O2S	-2.91	104.84	111.61
3	A	1152	EPE	O3S-S-O2S	-2.51	105.76	111.61
3	B	1152	EPE	O3S-S-O1S	-2.44	105.93	111.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 36 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1151	EPE	9	0
3	A	1152	EPE	1	0
3	B	1150	EPE	14	0
3	B	1151	EPE	12	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	1146/1151 (99%)	-0.57	4 (0%) 94 93	4, 45, 83, 109	0
1	B	1143/1151 (99%)	-0.44	6 (0%) 91 87	4, 53, 105, 123	0
All	All	2289/2302 (99%)	-0.51	10 (0%) 93 90	4, 48, 100, 123	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	PRO	5.7
1	A	461	SER	3.8
1	B	348	ALA	3.1
1	B	266	SER	2.7
1	B	459	GLN	2.4

### 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 carbohydrates 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. 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( $\text{\AA}^2$ )	Q<0.9
3	EPE	B	1151	15/15	0.75	0.46	16.59	105,108,109,109	0
3	EPE	A	1151	15/15	0.89	0.40	6.13	104,105,105,106	0
3	EPE	B	1150	15/15	0.84	0.34	3.18	81,89,91,92	0
3	EPE	B	1152	15/15	0.73	0.47	2.00	95,95,98,98	0
3	EPE	A	1152	15/15	0.87	0.38	1.30	88,89,91,92	0
2	SO4	A	1149	5/5	0.91	0.16	0.70	89,89,89,89	0
2	SO4	B	1149	5/5	0.87	0.20	-	120,120,120,121	0
2	SO4	A	1150	5/5	0.95	0.13	-	104,104,105,105	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.