



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 06:35 PM GMT

PDB ID : 3HF3  
Title : Old Yellow Enzyme from *Thermus scotoductus* SA-01  
Authors : Opperman, D.J.; Sewell, B.T.; Litthauer, D.; Isupov, M.N.; Littlechild, J.A.;  
van Heerden, E.  
Deposited on : 2009-05-11  
Resolution : 2.20 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

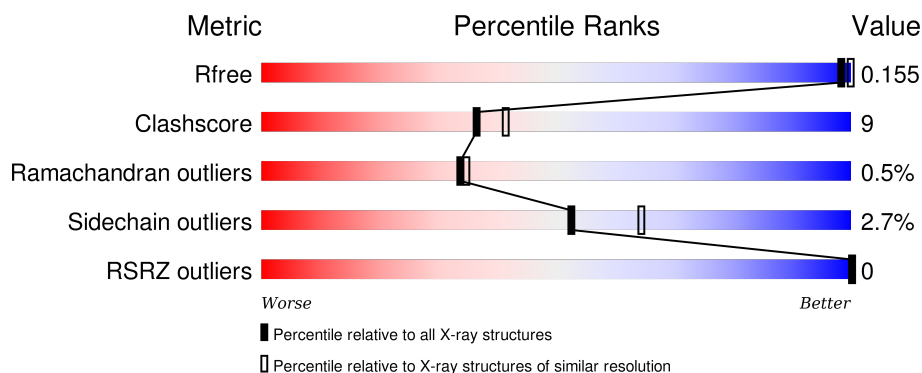
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*





The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	349	 82% 17% .
1	B	349	 82% 17%
1	C	349	 79% 19% .
1	D	349	 82% 16% .

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	SO4	A	1101	-	-	X	X
3	SO4	B	1102	-	-	X	X
3	SO4	C	1103	-	-	X	X
3	SO4	D	1104	-	-	X	X

## 2 Entry composition [i](#)

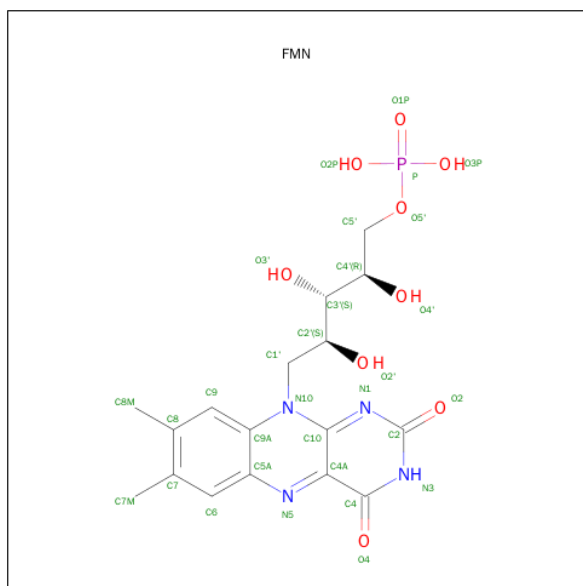
There are 4 unique types of molecules in this entry. The entry contains 11694 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Chromate reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	348	Total	C	N	O	S	0	1	0
			2680	1716	481	476	7			
1	B	348	Total	C	N	O	S	0	0	0
			2674	1712	481	474	7			
1	C	348	Total	C	N	O	S	0	0	0
			2674	1712	481	474	7			
1	D	348	Total	C	N	O	S	0	0	0
			2674	1712	481	474	7			

- Molecule 2 is FLAVIN MONONUCLEOTIDE (three-letter code: FMN) (formula: C<sub>17</sub>H<sub>21</sub>N<sub>4</sub>O<sub>9</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	17	4	9	1		
2	B	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	0	0
			31	17	4	9	1		
2	D	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

- 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		
3	B	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		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	234	Total	O	0	0
			234	234		
4	B	234	Total	O	0	0
			234	234		
4	C	190	Total	O	0	0
			190	190		

*Continued on next page...*

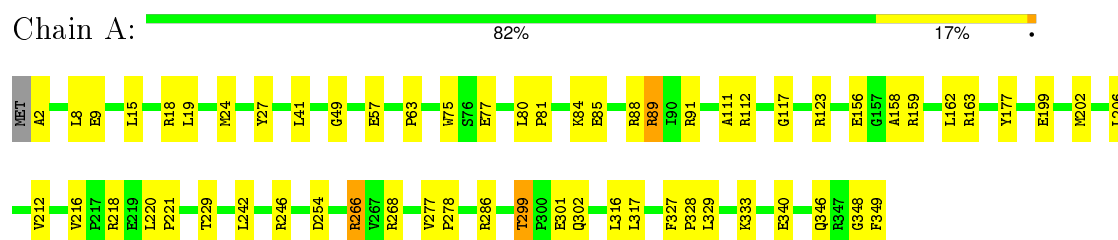
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	190	Total 190	O 190	0	0

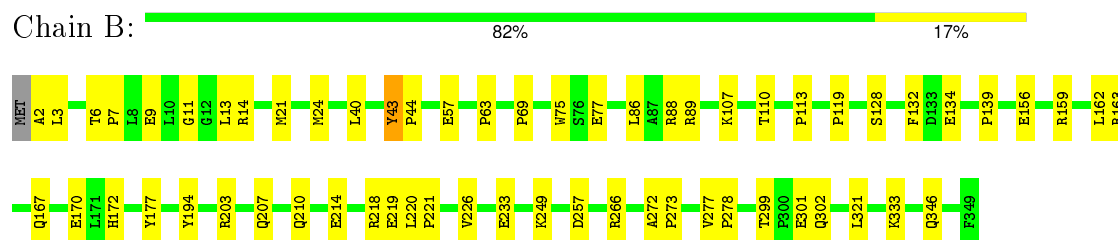
### 3 Residue-property plots

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.

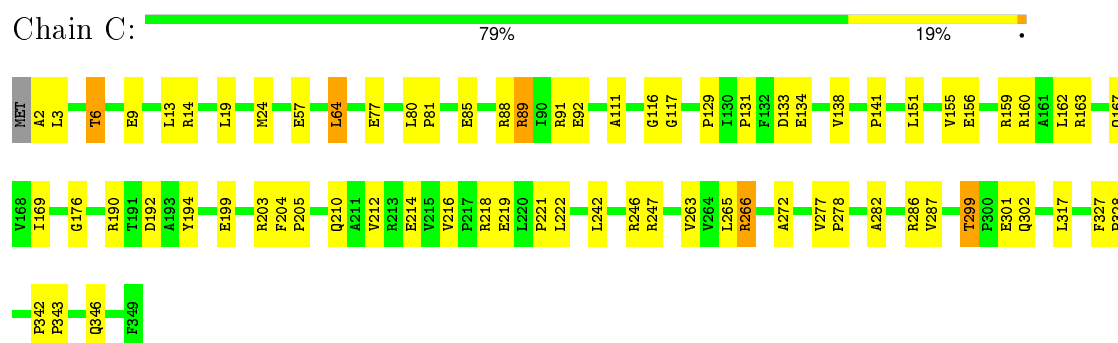
- Molecule 1: Chromate reductase



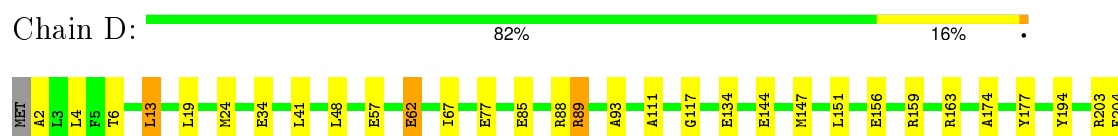
- Molecule 1: Chromate reductase



- Molecule 1: Chromate reductase



- Molecule 1: Chromate reductase



P205	L206	Q207	A211	R218	E219	L220	P221	R225	D230	W231	E232	E233	L242	R246	L256	D257	L265	R266	V267	R268	V277	P278	V283	R290	T299	Q302	R323	D324	R330	E340	V341	P342	P343	F349
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.78Å 101.88Å 180.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.28 – 2.20 38.28 – 2.20	Depositor EDS
% Data completeness (in resolution range)	94.4 (38.28-2.20) 94.2 (38.28-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.49 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, $R_{free}$	0.174 , 0.214 0.159 , 0.155	Depositor DCC
$R_{free}$ test set	4538 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.7	Xtriage
Anisotropy	0.368	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 24.9	EDS
Estimated twinning fraction	0.733 for H, K, L 0.267 for K, H, -L 0.278 for k,h,-l	Xtriage
Reported twinning fraction	0.733 for H, K, L 0.267 for K, H, -L	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 90665 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11694	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.75 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.5167e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: FMN, 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	0.60	0/2748	0.66	0/3737
1	B	0.64	2/2739 (0.1%)	0.67	0/3725
1	C	0.55	0/2739	0.64	0/3725
1	D	0.54	0/2739	0.63	0/3725
All	All	0.58	2/10965 (0.0%)	0.65	0/14912

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	43	TYR	CD1-CE1	-5.94	1.30	1.39
1	B	226	VAL	CB-CG1	5.02	1.63	1.52

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	2680	0	2725	51	0
1	B	2674	0	2719	48	0
1	C	2674	0	2719	67	0
1	D	2674	0	2719	41	0
2	A	31	0	19	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	31	0	19	2	0
2	C	31	0	19	4	0
2	D	31	0	19	3	0
3	A	5	0	0	2	0
3	B	5	0	0	4	0
3	C	5	0	0	3	0
3	D	5	0	0	3	0
4	A	234	0	0	13	0
4	B	234	0	0	12	0
4	C	190	0	0	13	0
4	D	190	0	0	7	0
All	All	11694	0	10958	207	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 9.

The worst 5 of 207 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:1003:FMN:C10	3:C:1103:SO4:O2	1.93	1.17
1:C:266:ARG:HH11	1:C:266:ARG:HG3	1.04	1.14
1:B:177:TYR:HE2	3:B:1102:SO4:O3	1.27	1.12
1:C:89:ARG:HG2	1:C:89:ARG:HH11	1.12	1.10
1:B:177:TYR:CE2	3:B:1102:SO4:O3	2.05	1.08

There are no symmetry-related clashes.

## 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	347/349 (99%)	331 (95%)	15 (4%)	1 (0%)	46 50

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	346/349 (99%)	334 (96%)	10 (3%)	2 (1%)	30	29
1	C	346/349 (99%)	334 (96%)	10 (3%)	2 (1%)	30	29
1	D	346/349 (99%)	335 (97%)	9 (3%)	2 (1%)	30	29
All	All	1385/1396 (99%)	1334 (96%)	44 (3%)	7 (0%)	34	35

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	GLU
1	C	57	GLU
1	B	11	GLY
1	B	57	GLU
1	D	57	GLU

### 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	275/275 (100%)	265 (96%)	10 (4%)	42	52
1	B	274/275 (100%)	270 (98%)	4 (2%)	72	84
1	C	274/275 (100%)	268 (98%)	6 (2%)	60	72
1	D	274/275 (100%)	264 (96%)	10 (4%)	42	52
All	All	1097/1100 (100%)	1067 (97%)	30 (3%)	52	64

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	233	GLU
1	C	89	ARG
1	D	290	ARG
1	C	64	LEU
1	C	92	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	189	GLN
1	B	167	GLN
1	B	346	GLN
1	C	167	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](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

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	FMN	A	1001	-	32,33,33	1.34	3 (9%)	34,50,50	1.82	7 (20%)
3	SO4	A	1101	-	4,4,4	0.38	0	6,6,6	0.25	0
2	FMN	B	1002	-	32,33,33	1.18	3 (9%)	34,50,50	1.76	5 (14%)
3	SO4	B	1102	-	4,4,4	0.29	0	6,6,6	0.24	0
2	FMN	C	1003	-	32,33,33	1.14	3 (9%)	34,50,50	1.69	8 (23%)
3	SO4	C	1103	-	4,4,4	0.29	0	6,6,6	0.17	0
2	FMN	D	1004	-	32,33,33	1.12	2 (6%)	34,50,50	1.58	5 (14%)
3	SO4	D	1104	-	4,4,4	0.40	0	6,6,6	0.25	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	FMN	A	1001	-	-	0/18/18/18	0/3/3/3
3	SO4	A	1101	-	-	0/0/0/0	0/0/0/0
2	FMN	B	1002	-	-	0/18/18/18	0/3/3/3
3	SO4	B	1102	-	-	0/0/0/0	0/0/0/0
2	FMN	C	1003	-	-	0/18/18/18	0/3/3/3
3	SO4	C	1103	-	-	0/0/0/0	0/0/0/0
2	FMN	D	1004	-	-	0/18/18/18	0/3/3/3
3	SO4	D	1104	-	-	0/0/0/0	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	FMN	C5A-N5	2.12	1.38	1.35
2	C	1003	FMN	C5A-N5	2.29	1.38	1.35
2	C	1003	FMN	C1'-N10	2.53	1.51	1.48
2	B	1002	FMN	C1'-N10	2.57	1.51	1.48
2	D	1004	FMN	C4-N3	2.59	1.37	1.33

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	FMN	N3-C2-N1	-4.71	119.76	127.69
2	B	1002	FMN	N3-C2-N1	-4.10	120.79	127.69
2	C	1003	FMN	N3-C2-N1	-4.00	120.95	127.69
2	D	1004	FMN	N3-C2-N1	-3.61	121.62	127.69
2	B	1002	FMN	C4A-C4-N3	-3.52	118.92	123.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	FMN	2	0
3	A	1101	SO4	2	0
2	B	1002	FMN	2	0
3	B	1102	SO4	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1003	FMN	4	0
3	C	1103	SO4	3	0
2	D	1004	FMN	3	0
3	D	1104	SO4	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	348/349 (99%)	-0.97	0 100 100	10, 19, 33, 45	0
1	B	348/349 (99%)	-0.97	0 100 100	5, 18, 33, 46	0
1	C	348/349 (99%)	-0.88	0 100 100	12, 24, 38, 47	0
1	D	348/349 (99%)	-0.89	0 100 100	12, 23, 38, 53	0
All	All	1392/1396 (99%)	-0.93	0 100 100	5, 21, 36, 53	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 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(Å <sup>2</sup> )	Q<0.9
3	SO4	A	1101	5/5	0.94	0.24	11.03	17,28,32,38	5
3	SO4	D	1104	5/5	0.94	0.28	9.17	29,36,41,44	5

*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	SO4	C	1103	5/5	0.94	0.20	8.00	44,47,48,51	5
3	SO4	B	1102	5/5	0.95	0.18	4.51	2,18,21,31	5
2	FMN	C	1003	31/31	0.98	0.09	0.24	6,16,21,21	0
2	FMN	D	1004	31/31	0.98	0.08	-0.04	9,12,19,20	0
2	FMN	A	1001	31/31	0.99	0.08	-0.20	8,12,15,16	0
2	FMN	B	1002	31/31	0.99	0.07	-0.32	6,16,19,21	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.