



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 08:15 am GMT

PDB ID : 1UPF  
Title : STRUCTURE OF THE URACIL PHOSPHORIBOSYLTRANSFERASE,  
MUTANT C128V BOUND TO THE DRUG 5-FLUOROURACIL  
Authors : Schumacher, M.A.; Carter, D.; Scott, D.; Roos, D.; Ullman, B.; Brennan, R.G.  
Deposited on : 1998-06-17  
Resolution : 2.30 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

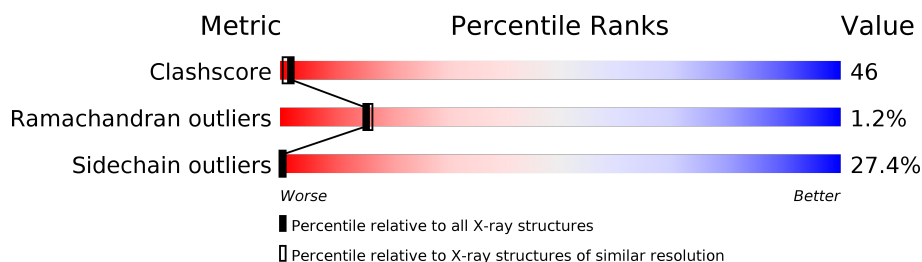
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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))
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	224	
1	B	224	
1	C	224	
1	D	224	

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	899	-	-	X	-
2	SO4	B	599	-	-	X	-
2	SO4	B	899	-	-	X	-
2	SO4	C	799	-	-	X	-
2	SO4	C	899	-	-	X	-
2	SO4	D	599	-	X	X	-
2	SO4	D	799	-	-	X	-
2	SO4	D	899	-	-	X	-
3	URF	A	999	-	X	-	-
3	URF	B	999	-	X	X	-
3	URF	C	999	-	-	X	-
3	URF	D	999	-	X	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7455 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 URACIL PHOSPHORIBOSYLTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	224	Total	C	N	O	S	0	0	0
			1788	1146	305	324	13			
1	C	224	Total	C	N	O	S	0	0	0
			1788	1146	305	324	13			
1	B	224	Total	C	N	O	S	0	0	0
			1788	1146	305	324	13			
1	A	224	Total	C	N	O	S	0	0	0
			1788	1146	305	324	13			

There are 12 discrepancies between the modelled and reference sequences:

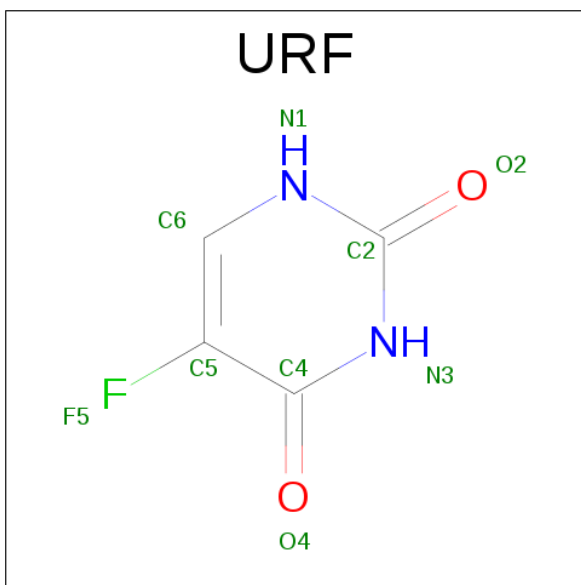
Chain	Residue	Modelled	Actual	Comment	Reference
D	84	GLN	GLU	CONFLICT	UNP Q26998
D	128	VAL	CYS	ENGINEERED	UNP Q26998
D	157	GLU	ASP	CONFLICT	UNP Q26998
C	84	GLN	GLU	CONFLICT	UNP Q26998
C	128	VAL	CYS	ENGINEERED	UNP Q26998
C	157	GLU	ASP	CONFLICT	UNP Q26998
B	84	GLN	GLU	CONFLICT	UNP Q26998
B	128	VAL	CYS	ENGINEERED	UNP Q26998
B	157	GLU	ASP	CONFLICT	UNP Q26998
A	84	GLN	GLU	CONFLICT	UNP Q26998
A	128	VAL	CYS	ENGINEERED	UNP Q26998
A	157	GLU	ASP	CONFLICT	UNP Q26998

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 3 is 5-FLUOROURACIL (three-letter code: URF) (formula: C<sub>4</sub>H<sub>3</sub>FN<sub>2</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	D	1	Total	C	F	N	O	0	0
			9	4	1	2	2		
3	C	1	Total	C	F	N	O	0	0
			9	4	1	2	2		
3	B	1	Total	C	F	N	O	0	0
			9	4	1	2	2		
3	A	1	Total	C	F	N	O	0	0
			9	4	1	2	2		

- Molecule 4 is water.

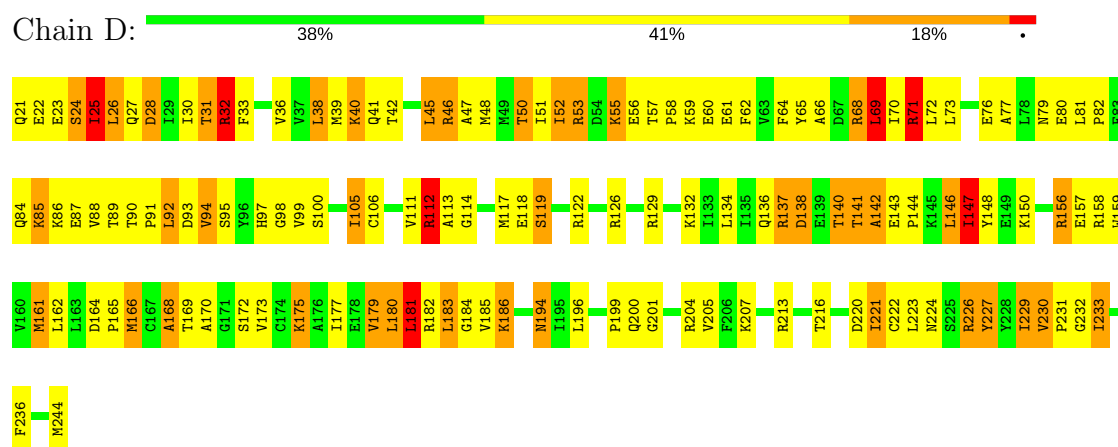
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	47	Total	O	0	0
			47	47		
4	B	47	Total	O	0	0
			47	47		
4	C	93	Total	O	0	0
			93	93		
4	D	20	Total	O	0	0
			20	20		

### 3 Residue-property plots

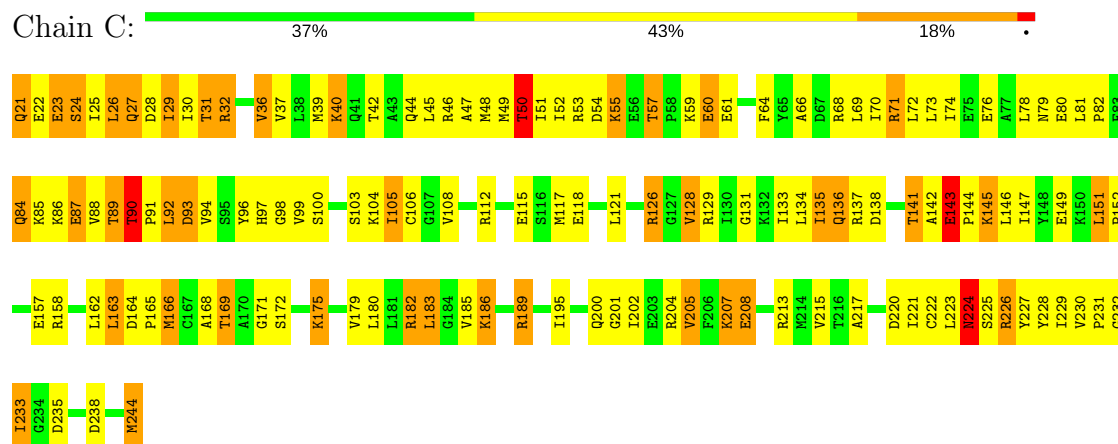
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

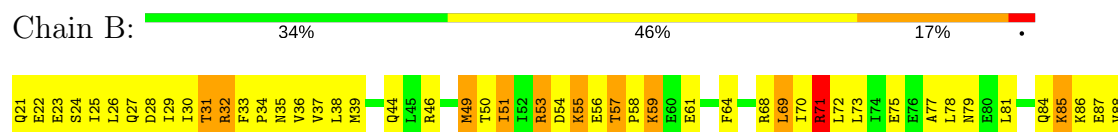
#### • Molecule 1: URACIL PHOSPHORIBOSYLTRANSFERASE

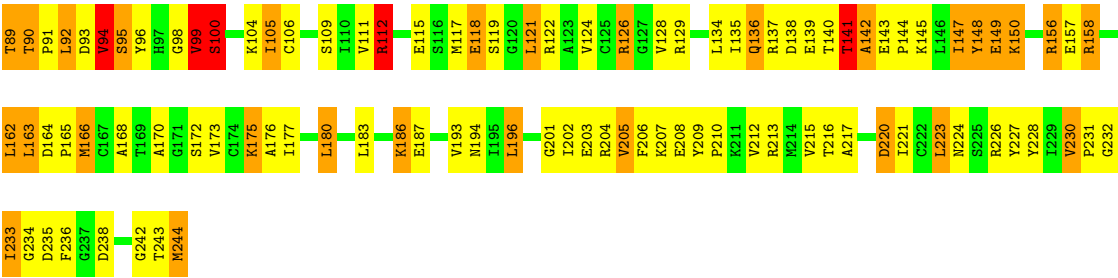


#### • Molecule 1: URACIL PHOSPHORIBOSYLTRANSFERASE

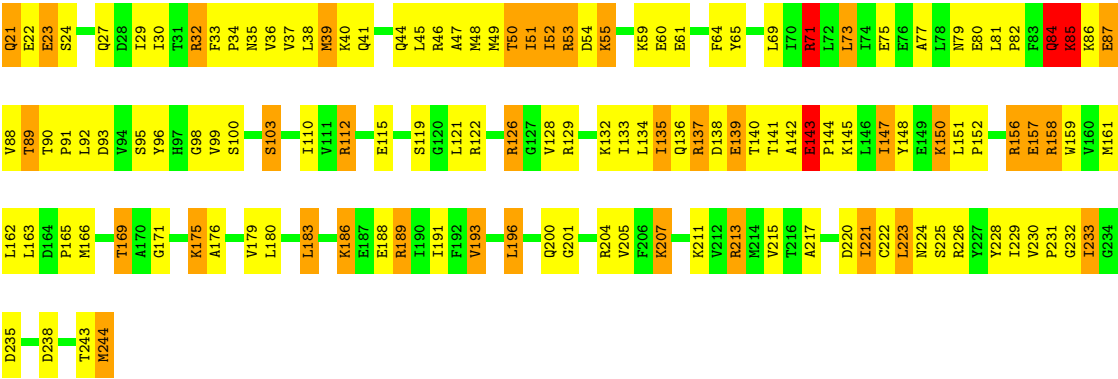
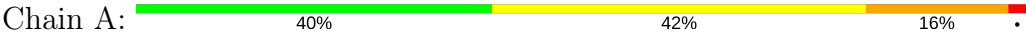


#### • Molecule 1: URACIL PHOSPHORIBOSYLTRANSFERASE





● Molecule 1: URACIL PHOSPHORIBOSYLTRANSFERASE





## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.46 Å   141.78 Å   71.46 Å 90.00°   115.03°   90.00°	Depositor
Resolution (Å)	10.00 – 2.30	Depositor
% Data completeness (in resolution range)	86.0 (10.00-2.30)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	TNT	Depositor
R, $R_{free}$	0.166 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7455	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

## 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: URF, 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	0.79	0/1817	1.32	11/2450 (0.4%)
1	B	0.80	0/1817	1.37	16/2450 (0.7%)
1	C	0.80	0/1817	1.28	10/2450 (0.4%)
1	D	0.80	1/1817 (0.1%)	1.30	9/2450 (0.4%)
All	All	0.80	1/7268 (0.0%)	1.32	46/9800 (0.5%)

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	A	1	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	229	ILE	CA-CB	-5.47	1.42	1.54

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	71	ARG	NE-CZ-NH2	-11.07	114.76	120.30
1	C	224	ASN	N-CA-CB	10.67	129.81	110.60
1	A	71	ARG	NE-CZ-NH2	-9.59	115.51	120.30
1	B	71	ARG	NE-CZ-NH1	8.82	124.71	120.30
1	B	112	ARG	NE-CZ-NH1	-7.67	116.47	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	50	THR	CA

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	143	GLU	Sidechain

## 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	1788	0	1857	166	0
1	B	1788	0	1857	187	0
1	C	1788	0	1857	195	0
1	D	1788	0	1857	192	1
2	A	15	0	0	5	0
2	B	15	0	0	5	0
2	C	15	0	0	5	0
2	D	15	0	0	7	0
3	A	9	0	3	0	0
3	B	9	0	3	4	0
3	C	9	0	3	4	0
3	D	9	0	3	4	0
4	A	47	0	0	8	0
4	B	47	0	0	8	0
4	C	93	0	0	4	1
4	D	20	0	0	9	0
All	All	7455	0	7440	670	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 46.

The worst 5 of 670 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:C:138:ASP:HB3	1:C:141:THR:HG22	1.27	1.13
1:D:90:THR:HG22	1:D:92:LEU:H	1.05	1.12

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:599:SO4:O3	2:D:599:SO4:S	2.08	1.11
1:C:141:THR:HG23	1:C:143:GLU:H	1.13	1.10
1:B:39:MET:HE2	1:B:73:LEU:HD13	1.41	1.02

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:D:223:LEU:O	4:C:1039:HOH:O[2_657]	2.01	0.19

## 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	222/224 (99%)	212 (96%)	8 (4%)	2 (1%)	20	23
1	B	222/224 (99%)	210 (95%)	6 (3%)	6 (3%)	6	4
1	C	222/224 (99%)	208 (94%)	14 (6%)	0	100	100
1	D	222/224 (99%)	207 (93%)	12 (5%)	3 (1%)	13	13
All	All	888/896 (99%)	837 (94%)	40 (4%)	11 (1%)	15	16

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	142	ALA
1	B	142	ALA
1	A	84	GLN
1	A	85	LYS
1	B	139	GLU

### 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	196/196 (100%)	146 (74%)	50 (26%)	0	0
1	B	196/196 (100%)	140 (71%)	56 (29%)	0	0
1	C	196/196 (100%)	144 (74%)	52 (26%)	0	0
1	D	196/196 (100%)	139 (71%)	57 (29%)	0	0
All	All	784/784 (100%)	569 (73%)	215 (27%)	0	0

5 of 215 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	186	LYS
1	B	69	LEU
1	A	163	LEU
1	C	205	VAL
1	B	23	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	136	GLN
1	B	35	ASN
1	A	35	ASN
1	C	97	HIS
1	A	79	ASN

### 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](#)

16 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	599	-	4,4,4	0.86	0	6,6,6	0.18	0
2	SO4	A	799	-	4,4,4	0.43	0	6,6,6	0.48	0
2	SO4	A	899	-	4,4,4	0.29	0	6,6,6	0.19	0
3	URF	A	999	-	8,9,9	2.62	5 (62%)	6,12,12	10.01	4 (66%)
2	SO4	B	599	-	4,4,4	0.46	0	6,6,6	0.33	0
2	SO4	B	799	-	4,4,4	0.89	0	6,6,6	1.15	1 (16%)
2	SO4	B	899	-	4,4,4	1.07	0	6,6,6	0.27	0
3	URF	B	999	-	8,9,9	2.31	5 (62%)	6,12,12	10.36	4 (66%)
2	SO4	C	599	-	4,4,4	0.40	0	6,6,6	0.33	0
2	SO4	C	799	-	4,4,4	0.70	0	6,6,6	0.59	0
2	SO4	C	899	-	4,4,4	1.26	0	6,6,6	0.85	0
3	URF	C	999	-	8,9,9	2.33	4 (50%)	6,12,12	9.74	4 (66%)
2	SO4	D	599	-	4,4,4	3.72	1 (25%)	6,6,6	3.29	3 (50%)
2	SO4	D	799	-	4,4,4	0.76	0	6,6,6	0.47	0
2	SO4	D	899	-	4,4,4	0.48	0	6,6,6	0.70	0
3	URF	D	999	-	8,9,9	2.56	5 (62%)	6,12,12	10.35	4 (66%)

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	599	-	-	0/0/0/0	0/0/0/0
2	SO4	A	799	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	A	899	-	-	0/0/0/0	0/0/0/0
3	URF	A	999	-	-	0/0/0/0	0/1/1/1
2	SO4	B	599	-	-	0/0/0/0	0/0/0/0
2	SO4	B	799	-	-	0/0/0/0	0/0/0/0
2	SO4	B	899	-	-	0/0/0/0	0/0/0/0
3	URF	B	999	-	-	0/0/0/0	0/1/1/1
2	SO4	C	599	-	-	0/0/0/0	0/0/0/0
2	SO4	C	799	-	-	0/0/0/0	0/0/0/0
2	SO4	C	899	-	-	0/0/0/0	0/0/0/0
3	URF	C	999	-	-	0/0/0/0	0/1/1/1
2	SO4	D	599	-	-	0/0/0/0	0/0/0/0
2	SO4	D	799	-	-	0/0/0/0	0/0/0/0
2	SO4	D	899	-	-	0/0/0/0	0/0/0/0
3	URF	D	999	-	-	0/0/0/0	0/1/1/1

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	999	URF	F5-C5	-2.63	1.31	1.35
3	B	999	URF	F5-C5	-2.62	1.31	1.35
3	C	999	URF	F5-C5	-2.52	1.31	1.35
3	A	999	URF	C6-C5	-2.38	1.34	1.37
3	B	999	URF	C6-C5	-2.36	1.34	1.37

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	999	URF	N1-C2-N3	-19.47	114.40	128.40
3	B	999	URF	N1-C2-N3	-19.45	114.41	128.40
3	A	999	URF	N1-C2-N3	-18.60	115.02	128.40
3	C	999	URF	N1-C2-N3	-18.36	115.20	128.40
3	A	999	URF	C5-C4-N3	-7.56	114.41	122.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

12 monomers are involved in 34 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	799	SO4	1	0
2	A	899	SO4	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	599	SO4	2	0
2	B	899	SO4	3	0
3	B	999	URF	4	0
2	C	799	SO4	2	0
2	C	899	SO4	3	0
3	C	999	URF	4	0
2	D	599	SO4	3	0
2	D	799	SO4	2	0
2	D	899	SO4	2	0
3	D	999	URF	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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.