



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2023 – 03:02 AM EDT

PDB ID : 1YR3  
Title : Escherichia coli purine nucleoside phosphorylase II, the product of the xapA gene  
Authors : Dandanell, G.; Szczepanowski, R.H.; Kierdaszuk, B.; Shugar, D.; Bochtler, M.  
Deposited on : 2005-02-03  
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

<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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35

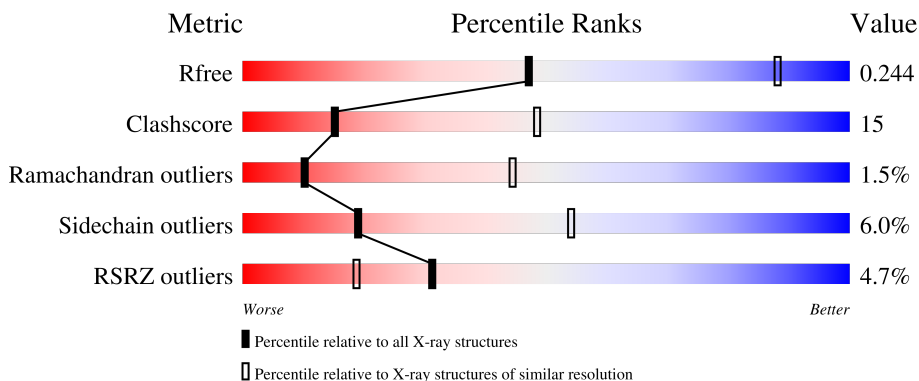
# 1 Overall quality at a glance

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.



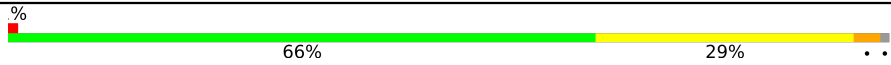
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 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	277	 8% 67% 29% ..
1	B	277	 8% 68% 28% ..
1	C	277	 8% 66% 30% ..
1	D	277	 8% 67% 29% ..
1	E	277	 2% 67% 28% ..

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	277	 <p>A horizontal bar chart representing the quality of the chain. The bar is divided into two segments: a green segment on the left labeled '66%' and a yellow segment on the right labeled '29%'. A small red square is at the beginning of the bar, and two small black dots are at the end. A '%' symbol is positioned above the start of the bar.</p>

## 2 Entry composition [i](#)

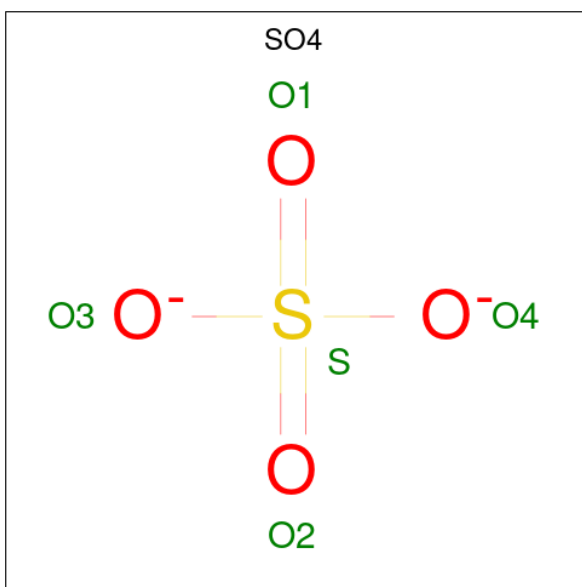
There are 3 unique types of molecules in this entry. The entry contains 12474 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Xanthosine phosphorylase.

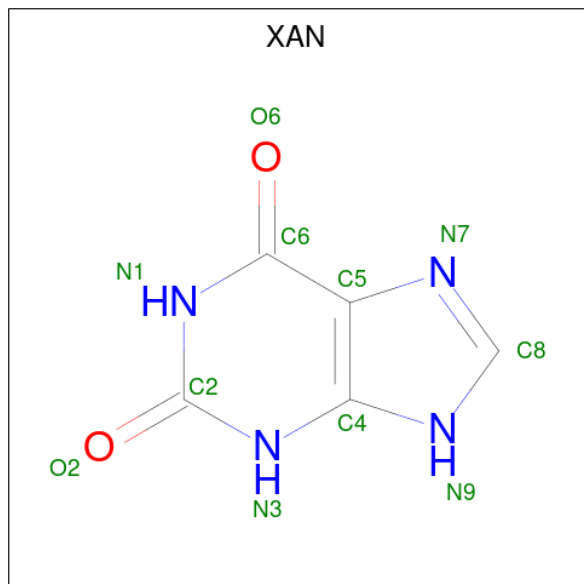
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	273	Total 2063	C 1319	N 351	O 378	S 15	0	0	0
1	B	273	Total 2063	C 1319	N 351	O 378	S 15	0	0	0
1	C	273	Total 2063	C 1319	N 351	O 378	S 15	0	0	0
1	D	273	Total 2063	C 1319	N 351	O 378	S 15	0	0	0
1	E	273	Total 2063	C 1319	N 351	O 378	S 15	0	0	0
1	F	273	Total 2063	C 1319	N 351	O 378	S 15	0	0	0

- 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	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	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is XANTHINE (three-letter code: XAN) (formula:  $C_5H_4N_4O_2$ ).

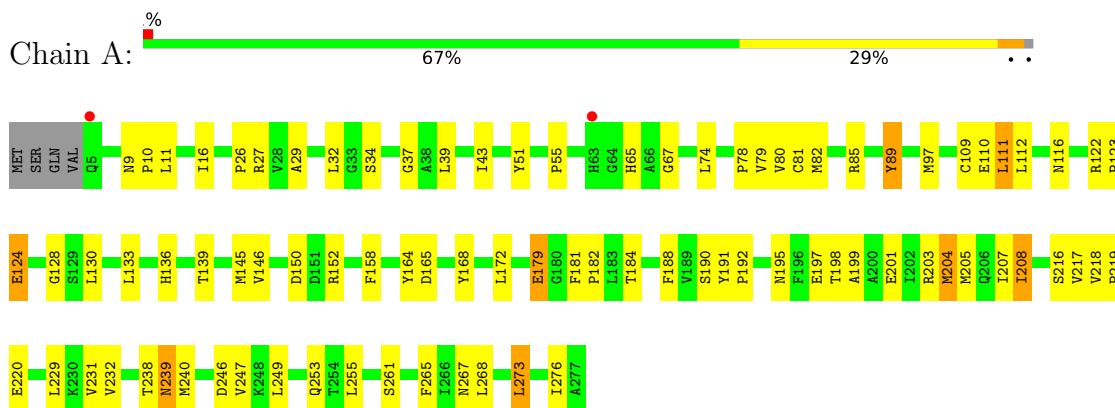


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			11	5	4	2		
3	B	1	Total	C	N	O	0	0
			11	5	4	2		
3	C	1	Total	C	N	O	0	0
			11	5	4	2		
3	D	1	Total	C	N	O	0	0
			11	5	4	2		
3	E	1	Total	C	N	O	0	0
			11	5	4	2		
3	F	1	Total	C	N	O	0	0
			11	5	4	2		

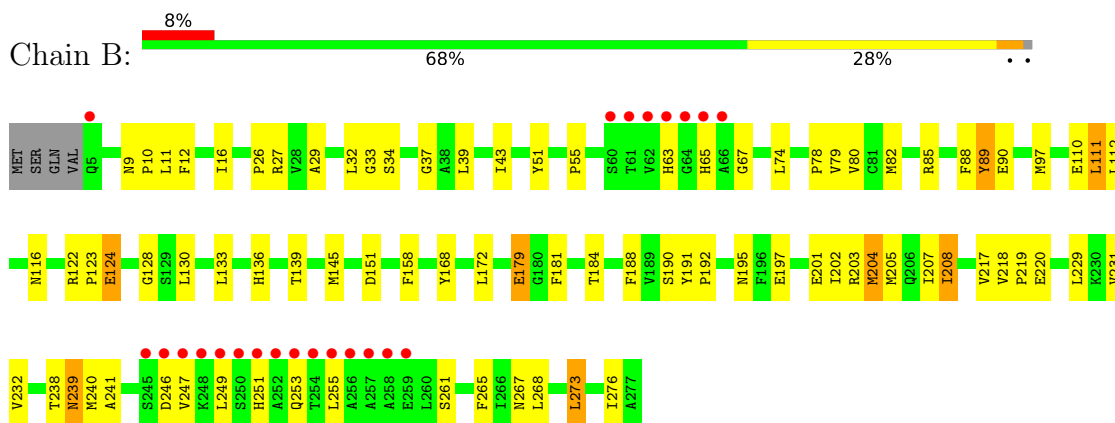
### 3 Residue-property plots [i](#)

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.

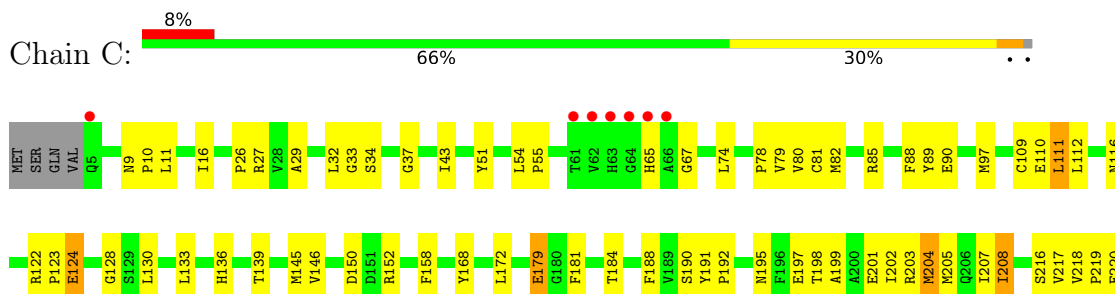
- Molecule 1: Xanthosine phosphorylase

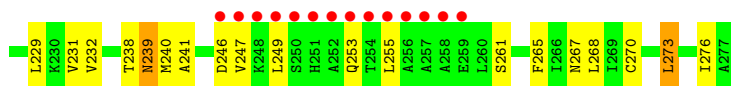


- Molecule 1: Xanthosine phosphorylase

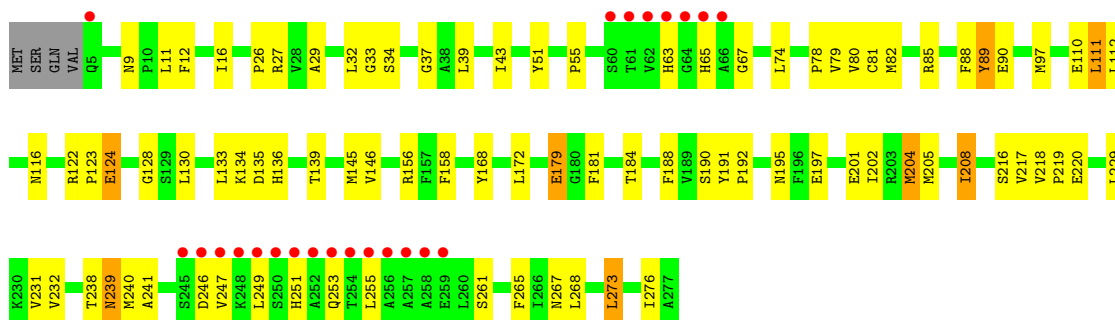


- Molecule 1: Xanthosine phosphorylase

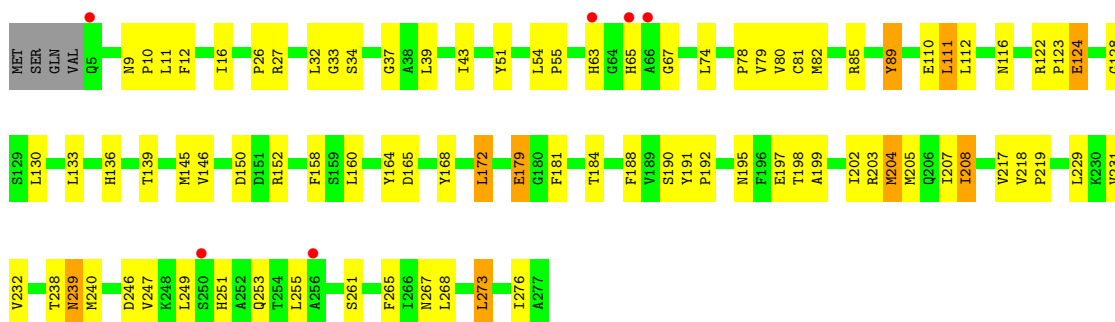




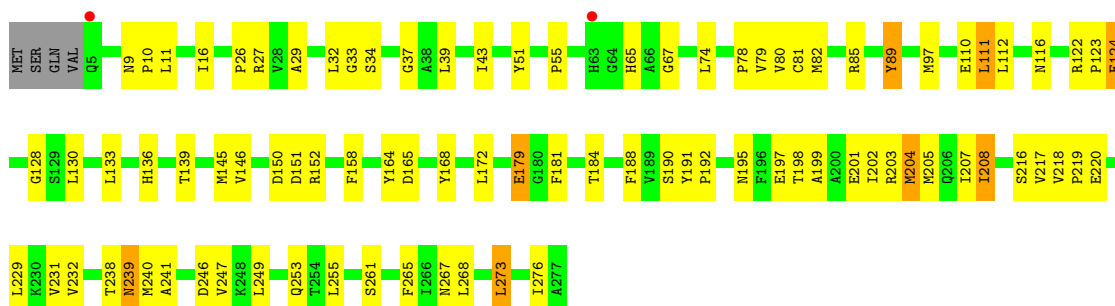
- Molecule 1: Xanthosine phosphorylase



- Molecule 1: Xanthosine phosphorylase



- Molecule 1: Xanthosine phosphorylase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	150.44Å 155.23Å 93.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 3.20 29.93 – 3.20	Depositor EDS
% Data completeness (in resolution range)	97.9 (25.00-3.20) 97.9 (29.93-3.20)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.58 (at 3.18Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.249 , 0.272 0.231 , 0.244	Depositor DCC
$R_{free}$ test set	1787 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.6	Xtrriage
Anisotropy	0.527	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 24.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.000 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	12474	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.99 % 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 1.2203e-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, 0.333 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: XAN, 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	0.43	0/2105	0.64	0/2848
1	B	0.44	0/2105	0.64	0/2848
1	C	0.43	0/2105	0.64	0/2848
1	D	0.45	0/2105	0.64	0/2848
1	E	0.43	0/2105	0.64	0/2848
1	F	0.43	0/2105	0.64	0/2848
All	All	0.44	0/12630	0.64	0/17088

There are no bond length outliers.

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	2063	0	2078	69	0
1	B	2063	0	2078	68	0
1	C	2063	0	2078	70	0
1	D	2063	0	2078	71	0
1	E	2063	0	2078	70	0
1	F	2063	0	2078	69	0
2	A	5	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	5	0	0	1	0
2	C	5	0	0	1	0
2	D	5	0	0	1	0
2	E	5	0	0	1	0
2	F	5	0	0	1	0
3	A	11	0	4	0	0
3	B	11	0	4	0	0
3	C	11	0	4	0	0
3	D	11	0	4	0	0
3	E	11	0	4	0	0
3	F	11	0	4	0	0
All	All	12474	0	12492	375	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 15.

The worst 5 of 375 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:65:HIS:HD2	1:B:85:ARG:H	1.12	0.96
1:C:65:HIS:HD2	1:C:85:ARG:H	1.13	0.96
1:E:65:HIS:HD2	1:E:85:ARG:H	1.13	0.95
1:A:65:HIS:HD2	1:A:85:ARG:H	1.14	0.94
1:F:65:HIS:HD2	1:F:85:ARG:H	1.15	0.94

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	271/277 (98%)	250 (92%)	18 (7%)	3 (1%)	<b>14</b> 51

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	271/277 (98%)	250 (92%)	16 (6%)	5 (2%)	8	41
1	C	271/277 (98%)	250 (92%)	17 (6%)	4 (2%)	10	44
1	D	271/277 (98%)	250 (92%)	17 (6%)	4 (2%)	10	44
1	E	271/277 (98%)	251 (93%)	16 (6%)	4 (2%)	10	44
1	F	271/277 (98%)	251 (93%)	15 (6%)	5 (2%)	8	41
All	All	1626/1662 (98%)	1502 (92%)	99 (6%)	25 (2%)	10	44

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	89	TYR
1	D	89	TYR
1	A	89	TYR
1	A	239	ASN
1	B	239	ASN

### 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	219/223 (98%)	206 (94%)	13 (6%)	19	54
1	B	219/223 (98%)	206 (94%)	13 (6%)	19	54
1	C	219/223 (98%)	205 (94%)	14 (6%)	17	52
1	D	219/223 (98%)	206 (94%)	13 (6%)	19	54
1	E	219/223 (98%)	206 (94%)	13 (6%)	19	54
1	F	219/223 (98%)	206 (94%)	13 (6%)	19	54
All	All	1314/1338 (98%)	1235 (94%)	79 (6%)	19	54

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	172	LEU

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	179	GLU
1	E	184	THR
1	E	273	LEU
1	F	229	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	251	HIS
1	F	253	GLN
1	E	253	GLN
1	F	87	HIS
1	B	267	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

12 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
3	XAN	C	402	-	8,12,12	2.06	2 (25%)	4,17,17	8.36	2 (50%)
2	SO4	F	505	-	4,4,4	1.85	2 (50%)	6,6,6	0.91	0
2	SO4	D	503	-	4,4,4	1.90	2 (50%)	6,6,6	0.89	0
2	SO4	C	502	-	4,4,4	1.88	2 (50%)	6,6,6	0.93	0
3	XAN	E	404	-	8,12,12	1.92	2 (25%)	4,17,17	8.39	2 (50%)
2	SO4	E	504	-	4,4,4	1.95	2 (50%)	6,6,6	0.92	0
2	SO4	A	500	-	4,4,4	1.85	2 (50%)	6,6,6	0.88	0
3	XAN	A	400	-	8,12,12	1.93	2 (25%)	4,17,17	8.35	2 (50%)
2	SO4	B	501	-	4,4,4	1.93	2 (50%)	6,6,6	0.90	0
3	XAN	D	403	-	8,12,12	1.94	1 (12%)	4,17,17	8.21	2 (50%)
3	XAN	F	405	-	8,12,12	1.86	2 (25%)	4,17,17	8.35	2 (50%)
3	XAN	B	401	-	8,12,12	2.09	1 (12%)	4,17,17	8.28	2 (50%)

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	XAN	C	402	-	-	-	0/2/2/2
3	XAN	E	404	-	-	-	0/2/2/2
3	XAN	A	400	-	-	-	0/2/2/2
3	XAN	D	403	-	-	-	0/2/2/2
3	XAN	F	405	-	-	-	0/2/2/2
3	XAN	B	401	-	-	-	0/2/2/2

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	401	XAN	C6-N1	4.89	1.41	1.33
3	C	402	XAN	C6-N1	4.53	1.40	1.33
3	D	403	XAN	C6-N1	4.26	1.40	1.33
3	E	404	XAN	C6-N1	4.16	1.40	1.33
3	A	400	XAN	C6-N1	4.12	1.40	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	405	XAN	C2-N1-C6	14.18	127.12	115.14

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	404	XAN	C2-N1-C6	14.16	127.09	115.14
3	A	400	XAN	C2-N1-C6	14.14	127.08	115.14
3	C	402	XAN	C2-N1-C6	14.07	127.02	115.14
3	B	401	XAN	C2-N1-C6	13.96	126.93	115.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	505	SO4	1	0
2	D	503	SO4	1	0
2	C	502	SO4	1	0
2	E	504	SO4	1	0
2	A	500	SO4	1	0
2	B	501	SO4	1	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	273/277 (98%)	-0.42	2 (0%) 87 81	9, 23, 63, 72	0
1	B	273/277 (98%)	-0.04	23 (8%) 11 6	10, 22, 65, 73	0
1	C	273/277 (98%)	-0.05	21 (7%) 13 7	10, 23, 65, 72	0
1	D	273/277 (98%)	-0.00	23 (8%) 11 6	10, 22, 65, 73	0
1	E	273/277 (98%)	-0.30	6 (2%) 62 48	9, 22, 63, 72	0
1	F	273/277 (98%)	-0.37	2 (0%) 87 81	8, 22, 63, 72	0
All	All	1638/1662 (98%)	-0.20	77 (4%) 31 19	8, 22, 64, 73	0

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	250	SER	9.4
1	B	256	ALA	9.0
1	D	256	ALA	8.3
1	C	256	ALA	7.8
1	C	63	HIS	7.6

### 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(Å <sup>2</sup> )	Q<0.9
3	XAN	D	403	11/11	0.85	0.29	52,53,54,54	0
3	XAN	B	401	11/11	0.88	0.26	52,53,54,54	0
3	XAN	C	402	11/11	0.92	0.21	51,52,53,53	0
3	XAN	A	400	11/11	0.92	0.23	49,50,50,51	0
2	SO4	F	505	5/5	0.94	0.29	46,46,48,49	0
3	XAN	F	405	11/11	0.94	0.25	48,49,51,51	0
3	XAN	E	404	11/11	0.95	0.21	50,50,50,51	0
2	SO4	B	501	5/5	0.95	0.32	49,49,50,50	0
2	SO4	D	503	5/5	0.96	0.26	49,50,50,51	0
2	SO4	E	504	5/5	0.97	0.23	48,48,49,49	0
2	SO4	C	502	5/5	0.97	0.24	50,50,51,51	0
2	SO4	A	500	5/5	0.98	0.29	49,49,49,50	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.