



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 18, 2023 – 03:28 PM EDT

PDB ID : 2DXB  
Title : Recombinant thiocyanate hydrolase comprising partially-modified cobalt centers  
Authors : Arakawa, T.; Kawano, Y.; Katayama, Y.; Yohda, M.; Odaka, M.  
Deposited on : 2006-08-25  
Resolution : 2.25 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 : 2.36  
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) : 2.36

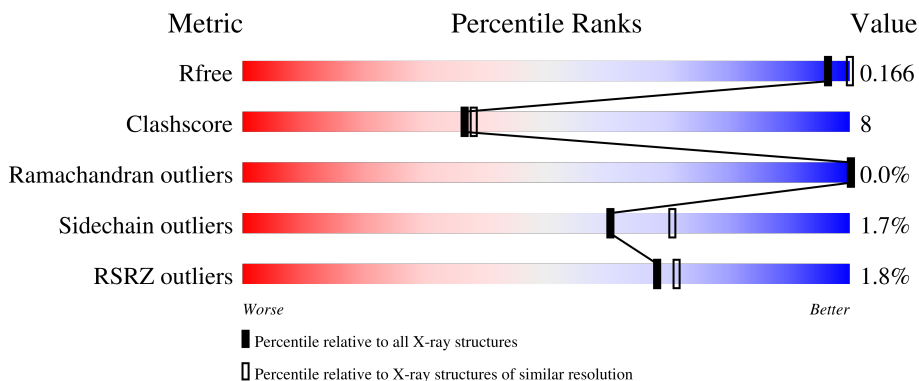
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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.



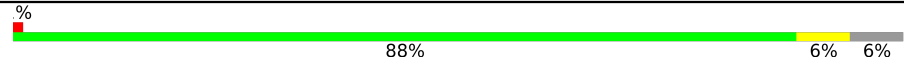

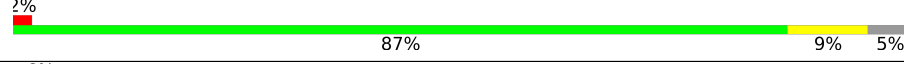

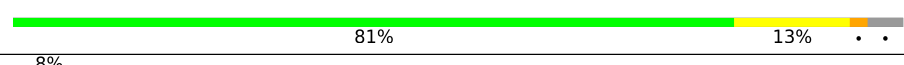
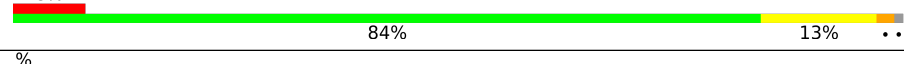
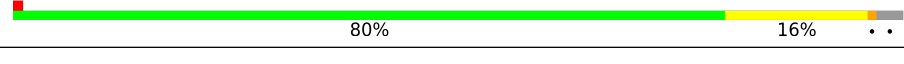

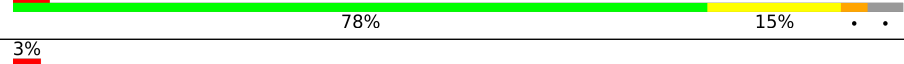

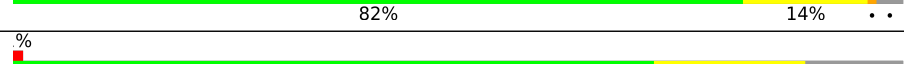
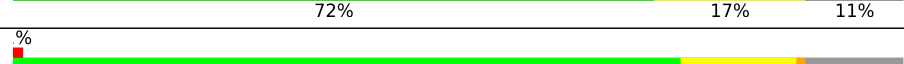
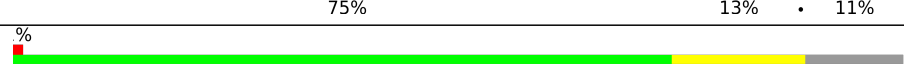
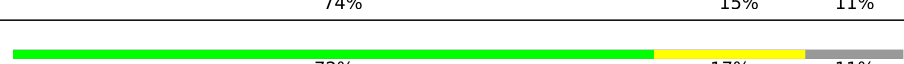

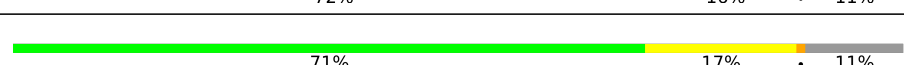
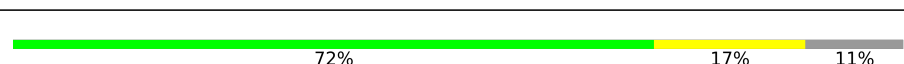
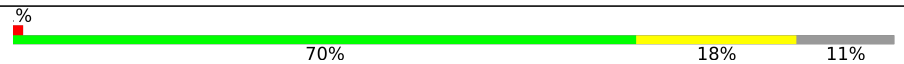

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	126	
1	D	126	
1	G	126	
1	J	126	
1	M	126	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	P	126	 % 88% 6% 6%
1	S	126	 2% 87% 8% 5%
1	V	126	 2% 87% 9% 5%
2	B	157	 6% 80% 14% ..
2	E	157	 81% 13% ..
2	H	157	 8% 84% 13% ..
2	K	157	 % 80% 16% ..
2	N	157	 79% 15% ..
2	Q	157	 4% 78% 15% ..
2	T	157	 3% 84% 13% ..
2	W	157	 6% 82% 14% ..
3	C	243	 % 72% 17% 11%
3	F	243	 % 75% 13% . 11%
3	I	243	 % 74% 15% 11%
3	L	243	 72% 17% 11%
3	O	243	 72% 16% . 11%
3	R	243	 71% 17% . 11%
3	U	243	 72% 17% 11%
3	X	243	 % 70% 18% 11%

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 34391 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 Thiocyanate hydrolase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	118	958	609	159	186	4	0	0	0
1	D	119	965	614	160	187	4	0	0	0
1	G	120	974	620	162	188	4	0	0	0
1	J	120	974	620	162	188	4	0	0	0
1	M	118	958	609	159	186	4	0	0	0
1	P	119	965	614	160	187	4	0	0	0
1	S	120	974	620	162	188	4	0	0	0
1	V	120	974	620	162	188	4	0	0	0

- Molecule 2 is a protein called Thiocyanate hydrolase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	152	1232	778	222	226	6	0	0	0
2	E	151	1226	775	221	224	6	0	0	0
2	H	156	1262	796	228	232	6	0	0	0
2	K	152	1232	778	222	226	6	0	0	0
2	N	152	1232	778	222	226	6	0	0	0
2	Q	151	1226	775	221	224	6	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	156	Total	C	N	O	S	0	0	0
			1262	796	228	232	6			
2	W	152	Total	C	N	O	S	0	0	0
			1232	778	222	226	6			

- Molecule 3 is a protein called Thiocyanate hydrolase subunit gamma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	217	Total	C	N	O	S	0	0	0
			1724	1098	304	314	8			
3	F	216	Total	C	N	O	S	0	0	0
			1715	1093	303	311	8			
3	I	217	Total	C	N	O	S	0	0	0
			1724	1098	304	314	8			
3	L	216	Total	C	N	O	S	0	0	0
			1715	1093	303	311	8			
3	O	217	Total	C	N	O	S	0	0	0
			1724	1098	304	314	8			
3	R	216	Total	C	N	O	S	0	0	0
			1715	1093	303	311	8			
3	U	217	Total	C	N	O	S	0	0	0
			1724	1098	304	314	8			
3	X	216	Total	C	N	O	S	0	0	0
			1715	1093	303	311	8			

- Molecule 4 is COBALT (III) ION (three-letter code: 3CO) (formula: Co).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	1	Total	Co	0	0
			1	1		
4	F	1	Total	Co	0	0
			1	1		
4	I	1	Total	Co	0	0
			1	1		
4	L	1	Total	Co	0	0
			1	1		
4	O	1	Total	Co	0	0
			1	1		
4	R	1	Total	Co	0	0
			1	1		
4	U	1	Total	Co	0	0
			1	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	X	1	Total	Co	0	0
			1	1		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	O	P	0	0
			5	4	1		
5	F	1	Total	O	P	0	0
			5	4	1		
5	I	1	Total	O	P	0	0
			5	4	1		
5	L	1	Total	O	P	0	0
			5	4	1		
5	O	1	Total	O	P	0	0
			5	4	1		
5	R	1	Total	O	P	0	0
			5	4	1		
5	U	1	Total	O	P	0	0
			5	4	1		
5	X	1	Total	O	P	0	0
			5	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	82	Total O 82 82	0	0
6	B	107	Total O 107 107	0	0
6	C	148	Total O 148 148	0	0
6	D	72	Total O 72 72	0	0
6	E	144	Total O 144 144	0	0
6	F	175	Total O 175 175	0	0
6	G	84	Total O 84 84	0	0
6	H	118	Total O 118 118	0	0
6	I	153	Total O 153 153	0	0
6	J	82	Total O 82 82	0	0
6	K	129	Total O 129 129	0	0
6	L	183	Total O 183 183	0	0
6	M	69	Total O 69 69	0	0
6	N	121	Total O 121 121	0	0
6	O	184	Total O 184 184	0	0
6	P	74	Total O 74 74	0	0
6	Q	108	Total O 108 108	0	0
6	R	158	Total O 158 158	0	0
6	S	75	Total O 75 75	0	0
6	T	140	Total O 140 140	0	0
6	U	191	Total O 191 191	0	0
6	V	85	Total O 85 85	0	0

*Continued on next page...*

*Continued from previous page...*


<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	W	106	Total 106	O 106	0	0
6	X	153	Total 153	O 153	0	0



### 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: Thiocyanate hydrolase subunit alpha

Chain A: 




- Molecule 1: Thiocyanate hydrolase subunit alpha

Chain D: 




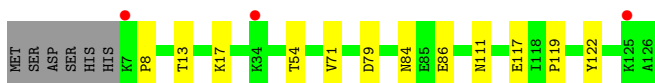
- Molecule 1: Thiocyanate hydrolase subunit alpha

Chain G: 




- Molecule 1: Thiocyanate hydrolase subunit alpha

Chain J: 

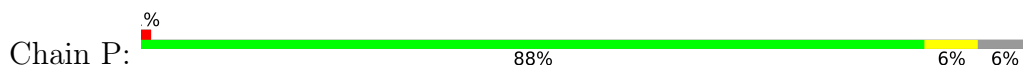


- Molecule 1: Thiocyanate hydrolase subunit alpha

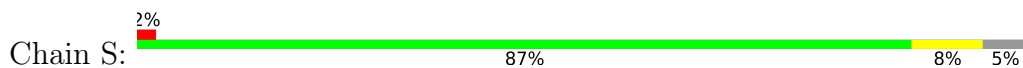
Chain M: 



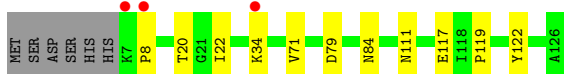
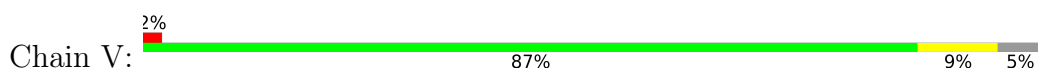
- Molecule 1: Thiocyanate hydrolase subunit alpha



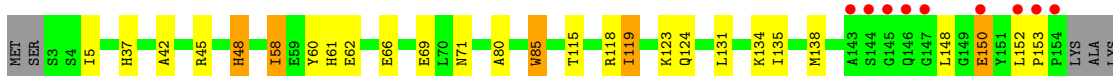
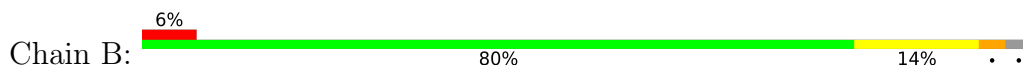
- Molecule 1: Thiocyanate hydrolase subunit alpha



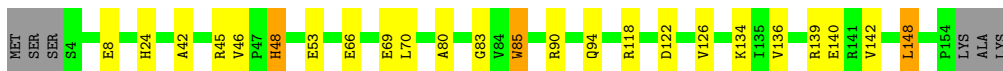
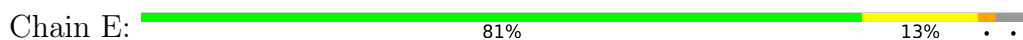
- Molecule 1: Thiocyanate hydrolase subunit alpha



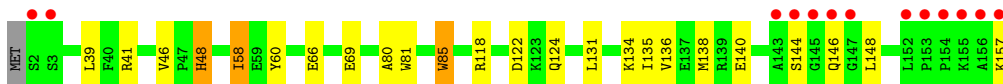
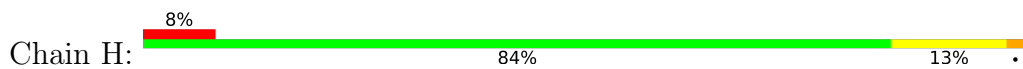
- Molecule 2: Thiocyanate hydrolase subunit beta



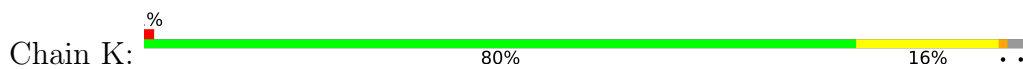
- Molecule 2: Thiocyanate hydrolase subunit beta



- Molecule 2: Thiocyanate hydrolase subunit beta

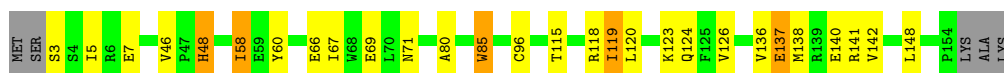


- Molecule 2: Thiocyanate hydrolase subunit beta



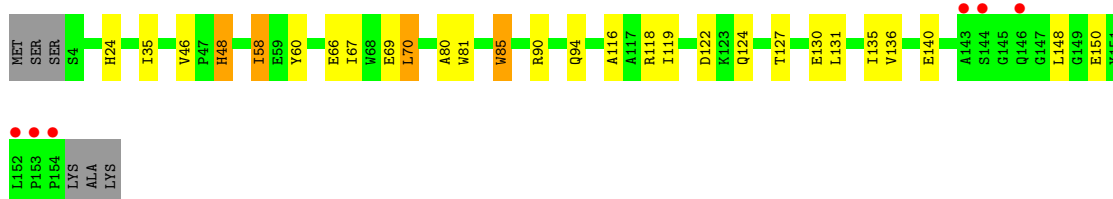
- Molecule 2: Thiocyanate hydrolase subunit beta

Chain N: 79% 15%



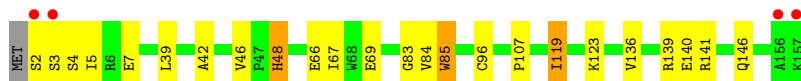
- Molecule 2: Thiocyanate hydrolase subunit beta

Chain Q: 4% 78% 15%



- Molecule 2: Thiocyanate hydrolase subunit beta

Chain T: 3% 84% 13%



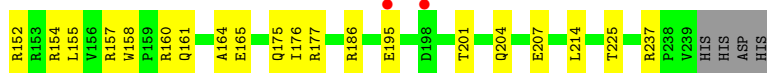
- Molecule 2: Thiocyanate hydrolase subunit beta

Chain W: 6% 82% 14%



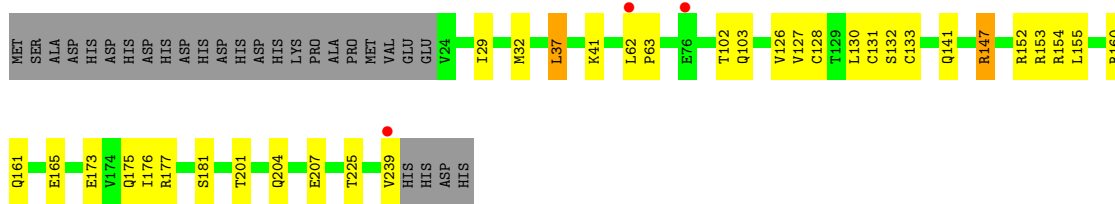
- Molecule 3: Thiocyanate hydrolase subunit gamma

Chain C: 0% 72% 17% 11%

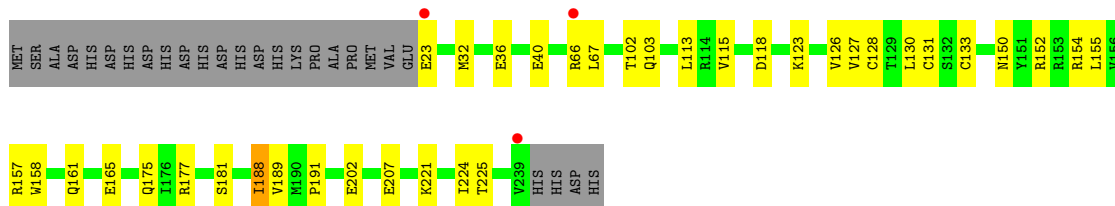


- Molecule 3: Thiocyanate hydrolase subunit gamma

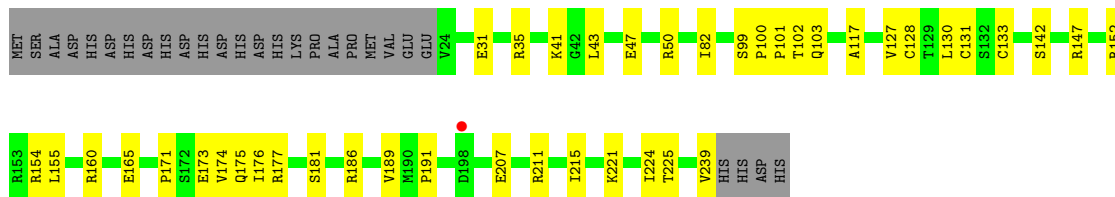
Chain F: 0% 75% 13% 11%



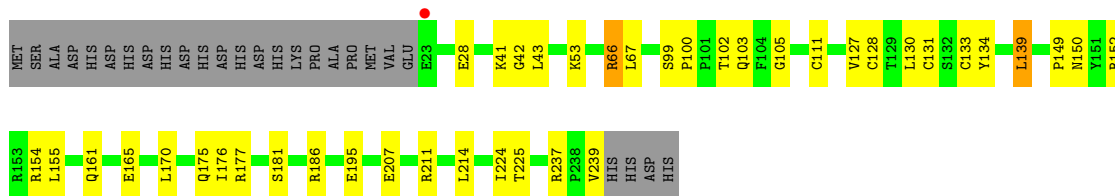
● Molecule 3: Thiocyanate hydrolase subunit gamma



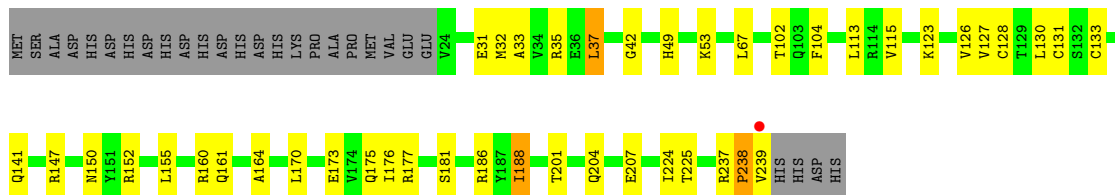
● Molecule 3: Thiocyanate hydrolase subunit gamma



● Molecule 3: Thiocyanate hydrolase subunit gamma

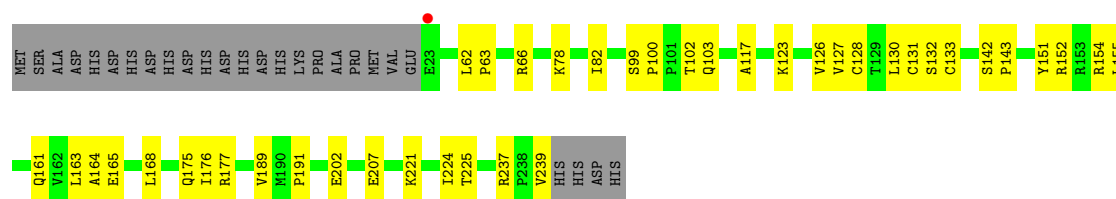


● Molecule 3: Thiocyanate hydrolase subunit gamma




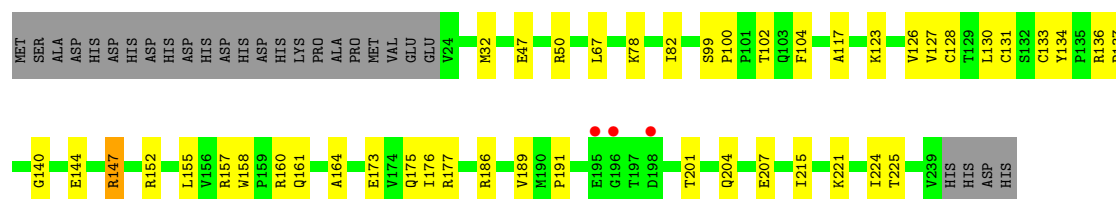
- Molecule 3: Thiocyanate hydrolase subunit gamma

Chain U:  72% 17% 11%



- Molecule 3: Thiocyanate hydrolase subunit gamma

Chain X:  70% 18% 11%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.87Å 245.14Å 244.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.41 – 2.25 49.41 – 2.25	Depositor EDS
% Data completeness (in resolution range)	97.6 (49.41-2.25) 99.5 (49.41-2.25)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.57 (at 2.25Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.171 , 0.198 0.166 , 0.166	Depositor DCC
$R_{free}$ test set	16292 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.4	Xtrriage
Anisotropy	0.517	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 52.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.016 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	34391	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.37% 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: CSO, CSD, PO4, 3CO

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	0.34	0/983	0.61	0/1331
1	D	0.34	0/991	0.62	0/1342
1	G	0.34	0/1000	0.60	0/1354
1	J	0.34	0/1000	0.60	0/1354
1	M	0.34	0/983	0.59	0/1331
1	P	0.34	0/991	0.60	0/1342
1	S	0.34	0/1000	0.60	0/1354
1	V	0.34	0/1000	0.60	0/1354
2	B	0.31	0/1264	0.57	0/1720
2	E	0.32	0/1258	0.57	0/1712
2	H	0.31	0/1294	0.56	0/1757
2	K	0.31	0/1264	0.57	0/1720
2	N	0.31	0/1264	0.57	0/1720
2	Q	0.30	0/1258	0.58	0/1712
2	T	0.31	0/1294	0.56	0/1757
2	W	0.30	0/1264	0.56	0/1720
3	C	0.30	0/1752	0.60	0/2391
3	F	0.31	0/1743	0.60	0/2379
3	I	0.31	0/1752	0.61	0/2391
3	L	0.30	0/1743	0.61	0/2379
3	O	0.31	0/1752	0.60	0/2391
3	R	0.31	0/1743	0.60	0/2379
3	U	0.30	0/1752	0.60	0/2391
3	X	0.31	0/1743	0.60	0/2379
All	All	0.32	0/32088	0.59	0/43660

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

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	958	0	906	16	0
1	D	965	0	914	8	0
1	G	974	0	926	8	0
1	J	974	0	926	11	0
1	M	958	0	906	24	0
1	P	965	0	914	5	0
1	S	974	0	926	9	0
1	V	974	0	926	11	0
2	B	1232	0	1198	27	0
2	E	1226	0	1193	22	0
2	H	1262	0	1234	19	0
2	K	1232	0	1198	23	0
2	N	1232	0	1198	34	0
2	Q	1226	0	1193	23	0
2	T	1262	0	1234	23	0
2	W	1232	0	1198	20	0
3	C	1724	0	1735	39	0
3	F	1715	0	1729	33	0
3	I	1724	0	1735	33	0
3	L	1715	0	1729	34	0
3	O	1724	0	1735	45	0
3	R	1715	0	1729	41	0
3	U	1724	0	1735	35	0
3	X	1715	0	1729	42	0
4	C	1	0	0	0	0
4	F	1	0	0	0	0
4	I	1	0	0	0	0
4	L	1	0	0	0	0
4	O	1	0	0	0	0
4	R	1	0	0	0	0
4	U	1	0	0	0	0
4	X	1	0	0	0	0
5	C	5	0	0	0	0
5	F	5	0	0	0	0
5	I	5	0	0	1	0
5	L	5	0	0	0	0
5	O	5	0	0	0	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	R	5	0	0	0	0
5	U	5	0	0	0	0
5	X	5	0	0	0	0
6	A	82	0	0	2	0
6	B	107	0	0	1	0
6	C	148	0	0	2	0
6	D	72	0	0	0	0
6	E	144	0	0	3	0
6	F	175	0	0	2	0
6	G	84	0	0	0	0
6	H	118	0	0	1	0
6	I	153	0	0	0	0
6	J	82	0	0	0	0
6	K	129	0	0	1	0
6	L	183	0	0	2	0
6	M	69	0	0	1	0
6	N	121	0	0	0	0
6	O	184	0	0	2	0
6	P	74	0	0	0	0
6	Q	108	0	0	2	0
6	R	158	0	0	2	0
6	S	75	0	0	1	0
6	T	140	0	0	4	0
6	U	191	0	0	0	0
6	V	85	0	0	1	0
6	W	106	0	0	2	0
6	X	153	0	0	2	0
All	All	34391	0	30846	493	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 8.

The worst 5 of 493 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)
3:I:115:VAL:HG22	3:I:188:ILE:HD11	1.31	1.13
3:R:115:VAL:HG22	3:R:188:ILE:HD11	1.33	1.06
3:R:113:LEU:HD11	3:R:188:ILE:HD12	1.41	1.03
1:M:26:GLN:H	1:M:26:GLN:HE21	1.04	0.96
1:M:45:ILE:HD11	1:M:96:PHE:HZ	1.41	0.85

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	116/126 (92%)	112 (97%)	4 (3%)	0	100	100
1	D	117/126 (93%)	111 (95%)	6 (5%)	0	100	100
1	G	118/126 (94%)	112 (95%)	6 (5%)	0	100	100
1	J	118/126 (94%)	113 (96%)	5 (4%)	0	100	100
1	M	116/126 (92%)	110 (95%)	6 (5%)	0	100	100
1	P	117/126 (93%)	112 (96%)	5 (4%)	0	100	100
1	S	118/126 (94%)	113 (96%)	5 (4%)	0	100	100
1	V	118/126 (94%)	114 (97%)	4 (3%)	0	100	100
2	B	150/157 (96%)	147 (98%)	3 (2%)	0	100	100
2	E	149/157 (95%)	145 (97%)	4 (3%)	0	100	100
2	H	154/157 (98%)	150 (97%)	4 (3%)	0	100	100
2	K	150/157 (96%)	146 (97%)	4 (3%)	0	100	100
2	N	150/157 (96%)	147 (98%)	3 (2%)	0	100	100
2	Q	149/157 (95%)	145 (97%)	4 (3%)	0	100	100
2	T	154/157 (98%)	150 (97%)	4 (3%)	0	100	100
2	W	150/157 (96%)	146 (97%)	4 (3%)	0	100	100
3	C	213/243 (88%)	205 (96%)	8 (4%)	0	100	100
3	F	212/243 (87%)	204 (96%)	8 (4%)	0	100	100
3	I	213/243 (88%)	205 (96%)	8 (4%)	0	100	100
3	L	212/243 (87%)	206 (97%)	6 (3%)	0	100	100
3	O	213/243 (88%)	206 (97%)	7 (3%)	0	100	100
3	R	212/243 (87%)	203 (96%)	8 (4%)	1 (0%)	29	29

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	U	213/243 (88%)	206 (97%)	7 (3%)	0	100	100
3	X	212/243 (87%)	205 (97%)	7 (3%)	0	100	100
All	All	3844/4208 (91%)	3713 (97%)	130 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	R	238	PRO

### 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	100/108 (93%)	95 (95%)	5 (5%)	24	26
1	D	101/108 (94%)	101 (100%)	0	100	100
1	G	102/108 (94%)	100 (98%)	2 (2%)	55	64
1	J	102/108 (94%)	102 (100%)	0	100	100
1	M	100/108 (93%)	98 (98%)	2 (2%)	55	64
1	P	101/108 (94%)	100 (99%)	1 (1%)	76	84
1	S	102/108 (94%)	102 (100%)	0	100	100
1	V	102/108 (94%)	102 (100%)	0	100	100
2	B	130/134 (97%)	125 (96%)	5 (4%)	33	39
2	E	129/134 (96%)	125 (97%)	4 (3%)	40	49
2	H	133/134 (99%)	129 (97%)	4 (3%)	41	50
2	K	130/134 (97%)	128 (98%)	2 (2%)	65	75
2	N	130/134 (97%)	125 (96%)	5 (4%)	33	39
2	Q	129/134 (96%)	125 (97%)	4 (3%)	40	49
2	T	133/134 (99%)	129 (97%)	4 (3%)	41	50
2	W	130/134 (97%)	128 (98%)	2 (2%)	65	75

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	188/212 (89%)	185 (98%)	3 (2%)	62	73
3	F	187/212 (88%)	184 (98%)	3 (2%)	62	73
3	I	188/212 (89%)	186 (99%)	2 (1%)	73	82
3	L	187/212 (88%)	186 (100%)	1 (0%)	88	92
3	O	188/212 (89%)	184 (98%)	4 (2%)	53	62
3	R	187/212 (88%)	183 (98%)	4 (2%)	53	62
3	U	188/212 (89%)	188 (100%)	0	100	100
3	X	187/212 (88%)	186 (100%)	1 (0%)	88	92
All	All	3354/3632 (92%)	3296 (98%)	58 (2%)	60	71

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	K	48	HIS
2	W	48	HIS
2	N	85	TRP
2	T	119	ILE
3	R	181	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
2	K	26	GLN
2	W	100	GLN
2	N	37	HIS
2	W	24	HIS
3	U	161	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](#)

16 non-standard protein/DNA/RNA residues 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	CSO	C	133	4	3,6,7	0.67	0	0,6,8	-	-
3	CSD	O	131	4	3,7,8	0.66	0	1,8,10	1.07	0
3	CSO	R	133	4	3,6,7	0.81	0	0,6,8	-	-
3	CSD	I	131	4	3,7,8	0.67	0	1,8,10	1.38	0
3	CSO	O	133	4	3,6,7	0.77	0	0,6,8	-	-
3	CSD	F	131	4	3,7,8	0.76	0	1,8,10	1.19	0
3	CSD	R	131	4	3,7,8	0.67	0	1,8,10	1.47	0
3	CSO	U	133	4	3,6,7	0.70	0	0,6,8	-	-
3	CSO	X	133	4	3,6,7	0.79	0	0,6,8	-	-
3	CSD	U	131	4	3,7,8	0.72	0	1,8,10	1.32	0
3	CSO	F	133	4	3,6,7	0.69	0	0,6,8	-	-
3	CSO	L	133	4	3,6,7	0.71	0	0,6,8	-	-
3	CSD	C	131	4	3,7,8	0.85	0	1,8,10	1.03	0
3	CSO	I	133	4	3,6,7	0.75	0	0,6,8	-	-
3	CSD	L	131	4	3,7,8	0.81	0	1,8,10	1.46	0
3	CSD	X	131	4	3,7,8	0.65	0	1,8,10	1.08	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
3	CSO	C	133	4	-	0/1/5/7	-
3	CSD	O	131	4	-	1/2/6/8	-
3	CSO	R	133	4	-	0/1/5/7	-
3	CSD	I	131	4	-	1/2/6/8	-
3	CSO	O	133	4	-	0/1/5/7	-
3	CSD	F	131	4	-	1/2/6/8	-
3	CSD	R	131	4	-	1/2/6/8	-
3	CSO	U	133	4	-	0/1/5/7	-
3	CSO	X	133	4	-	0/1/5/7	-
3	CSD	U	131	4	-	1/2/6/8	-
3	CSO	F	133	4	-	0/1/5/7	-
3	CSO	L	133	4	-	0/1/5/7	-
3	CSD	C	131	4	-	1/2/6/8	-
3	CSO	I	133	4	-	0/1/5/7	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	CSD	L	131	4	-	1/2/6/8	-
3	CSD	X	131	4	-	1/2/6/8	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	131	CSD	CA-CB-SG-OD1
3	F	131	CSD	CA-CB-SG-OD1
3	I	131	CSD	CA-CB-SG-OD1
3	L	131	CSD	CA-CB-SG-OD1
3	O	131	CSD	CA-CB-SG-OD1

There are no ring outliers.

16 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	133	CSO	2	0
3	O	131	CSD	1	0
3	R	133	CSO	1	0
3	I	131	CSD	1	0
3	O	133	CSO	1	0
3	F	131	CSD	1	0
3	R	131	CSD	1	0
3	U	133	CSO	1	0
3	X	133	CSO	1	0
3	U	131	CSD	1	0
3	F	133	CSO	1	0
3	L	133	CSO	1	0
3	C	131	CSD	1	0
3	I	133	CSO	1	0
3	L	131	CSD	1	0
3	X	131	CSD	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 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
5	PO4	R	4006	-	4,4,4	1.69	0	6,6,6	0.42	0
5	PO4	L	4004	-	4,4,4	1.79	1 (25%)	6,6,6	0.42	0
5	PO4	U	4007	-	4,4,4	1.67	0	6,6,6	0.43	0
5	PO4	O	4005	-	4,4,4	1.71	0	6,6,6	0.42	0
5	PO4	C	4001	-	4,4,4	1.69	0	6,6,6	0.42	0
5	PO4	F	4002	-	4,4,4	1.76	1 (25%)	6,6,6	0.42	0
5	PO4	X	4008	-	4,4,4	1.69	0	6,6,6	0.42	0
5	PO4	I	4003	-	4,4,4	1.71	1 (25%)	6,6,6	0.42	0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	4002	PO4	P-O4	-2.03	1.48	1.54
5	L	4004	PO4	P-O3	-2.02	1.48	1.54
5	I	4003	PO4	P-O3	-2.00	1.48	1.54

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	I	4003	PO4	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	118/126 (93%)	-0.58	0 <b>100</b> <b>100</b>	10, 19, 35, 39	0
1	D	119/126 (94%)	-0.39	2 (1%) 70 73	12, 19, 34, 43	0
1	G	120/126 (95%)	-0.59	0 <b>100</b> <b>100</b>	10, 17, 34, 56	0
1	J	120/126 (95%)	-0.40	3 (2%) 57 60	12, 20, 39, 56	0
1	M	118/126 (93%)	-0.36	2 (1%) 70 73	13, 21, 35, 44	0
1	P	119/126 (94%)	-0.58	1 (0%) 86 87	11, 18, 34, 40	0
1	S	120/126 (95%)	-0.42	2 (1%) 70 73	12, 20, 35, 52	0
1	V	120/126 (95%)	-0.51	3 (2%) 57 60	11, 19, 36, 58	0
2	B	152/157 (96%)	-0.48	9 (5%) 22 24	11, 18, 36, 49	0
2	E	151/157 (96%)	-0.61	0 <b>100</b> <b>100</b>	11, 16, 28, 32	0
2	H	156/157 (99%)	-0.25	13 (8%) 11 12	10, 17, 43, 76	0
2	K	152/157 (96%)	-0.61	1 (0%) 87 88	10, 17, 29, 37	0
2	N	152/157 (96%)	-0.49	0 <b>100</b> <b>100</b>	10, 17, 29, 38	0
2	Q	151/157 (96%)	-0.54	6 (3%) 38 40	10, 18, 38, 45	0
2	T	156/157 (99%)	-0.41	4 (2%) 56 59	10, 16, 30, 72	0
2	W	152/157 (96%)	-0.48	9 (5%) 22 24	11, 18, 38, 47	0
3	C	215/243 (88%)	-0.54	3 (1%) 75 77	11, 19, 34, 51	0
3	F	214/243 (88%)	-0.53	3 (1%) 75 77	11, 17, 30, 47	0
3	I	215/243 (88%)	-0.58	3 (1%) 75 77	10, 18, 33, 54	0
3	L	214/243 (88%)	-0.52	1 (0%) 91 91	11, 18, 28, 35	0
3	O	215/243 (88%)	-0.51	1 (0%) 91 91	11, 18, 30, 51	0
3	R	214/243 (88%)	-0.59	1 (0%) 91 91	10, 19, 33, 50	0
3	U	215/243 (88%)	-0.57	1 (0%) 91 91	11, 17, 28, 50	0
3	X	214/243 (88%)	-0.51	3 (1%) 75 77	11, 19, 32, 42	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
All	All	3892/4208 (92%)	-0.51	71 (1%) 68 71	10, 18, 33, 76	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	156	ALA	10.2
2	H	157	LYS	9.3
2	T	157	LYS	8.7
2	T	156	ALA	8.2
3	R	239	VAL	6.5

## 6.2 Non-standard residues in protein, DNA, RNA chains [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	CSO	C	133	7/8	0.97	0.08	16,17,18,19	1
3	CSO	O	133	7/8	0.97	0.10	15,15,17,19	1
3	CSO	R	133	7/8	0.97	0.11	16,16,19,22	1
3	CSO	X	133	7/8	0.97	0.10	16,16,19,21	1
3	CSD	L	131	8/9	0.98	0.10	15,19,22,22	2
3	CSO	L	133	7/8	0.98	0.13	14,15,21,24	1
3	CSD	O	131	8/9	0.98	0.14	15,17,19,20	2
3	CSD	F	131	8/9	0.98	0.11	15,18,20,20	2
3	CSO	F	133	7/8	0.98	0.11	14,15,17,21	1
3	CSO	U	133	7/8	0.98	0.13	13,14,17,21	1
3	CSD	X	131	8/9	0.98	0.12	15,18,19,21	2
3	CSO	I	133	7/8	0.98	0.11	14,15,17,19	1
3	CSD	U	131	8/9	0.99	0.13	15,16,18,20	2
3	CSD	C	131	8/9	0.99	0.08	16,17,19,19	2
3	CSD	R	131	8/9	0.99	0.08	14,18,20,21	2
3	CSD	I	131	8/9	0.99	0.10	17,19,21,22	2

## 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
5	PO4	F	4002	5/5	0.97	0.15	44,44,47,47	0
5	PO4	I	4003	5/5	0.97	0.12	41,43,44,45	0
5	PO4	R	4006	5/5	0.97	0.14	49,49,50,51	0
5	PO4	U	4007	5/5	0.97	0.15	41,42,43,44	0
5	PO4	X	4008	5/5	0.97	0.16	45,45,47,48	0
5	PO4	L	4004	5/5	0.98	0.14	44,44,45,46	0
5	PO4	O	4005	5/5	0.98	0.15	42,43,43,44	0
5	PO4	C	4001	5/5	0.98	0.10	42,43,43,43	0
4	3CO	O	301	1/1	0.98	0.10	15,15,15,15	1
4	3CO	R	301	1/1	0.98	0.07	16,16,16,16	1
4	3CO	C	301	1/1	0.99	0.06	15,15,15,15	1
4	3CO	F	301	1/1	0.99	0.09	13,13,13,13	1
4	3CO	U	301	1/1	0.99	0.10	13,13,13,13	1
4	3CO	X	301	1/1	0.99	0.07	16,16,16,16	1
4	3CO	I	301	1/1	0.99	0.09	13,13,13,13	1
4	3CO	L	301	1/1	0.99	0.10	14,14,14,14	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.