



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 6, 2021 – 11:12 am BST

PDB ID : 6YZF  
Title : Crystal structure of the M295Y variant of Ssl1  
Authors : Mielenbrink, S.; Olbrich, A.; Urlacher, V.; Span, I.  
Deposited on : 2020-05-06  
Resolution : 1.68 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 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 : 2.18  
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) : 2.18

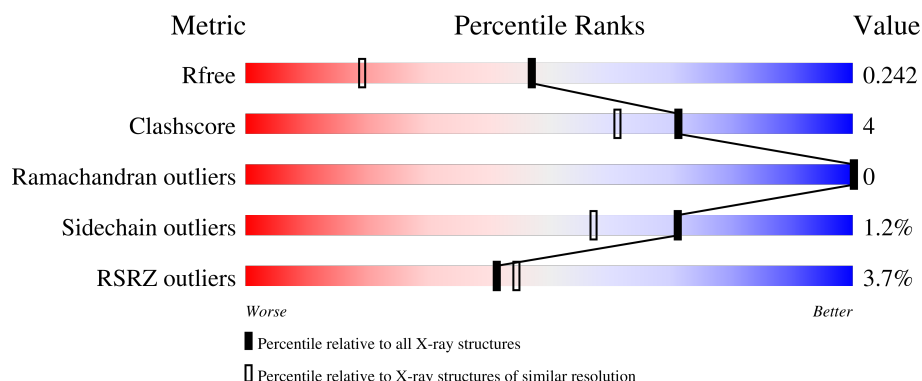
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


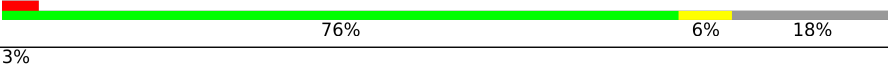
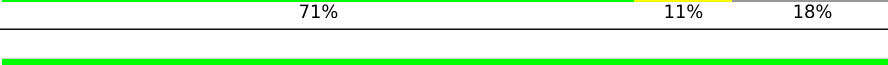
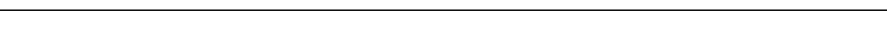
The reported resolution of this entry is 1.68 Å.

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	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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	AAA	325	
1	BBB	325	
1	CCC	325	
2	FFF	3	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12869 atoms, of which 5962 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 Copper oxidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	266	Total	C	H	N	O	S	95	0	0
			4048	1297	1977	377	388	9			
1	BBB	266	Total	C	H	N	O	S	95	0	0
			4048	1297	1977	377	388	9			
1	CCC	267	Total	C	H	N	O	S	95	0	0
			4070	1305	1987	378	391	9			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP B5HSR1
AAA	2	HIS	-	expression tag	UNP B5HSR1
AAA	3	HIS	-	expression tag	UNP B5HSR1
AAA	4	HIS	-	expression tag	UNP B5HSR1
AAA	5	HIS	-	expression tag	UNP B5HSR1
AAA	6	HIS	-	expression tag	UNP B5HSR1
AAA	7	HIS	-	expression tag	UNP B5HSR1
AAA	295	TYR	MET	engineered mutation	UNP B5HSR1
BBB	1	MET	-	initiating methionine	UNP B5HSR1
BBB	2	HIS	-	expression tag	UNP B5HSR1
BBB	3	HIS	-	expression tag	UNP B5HSR1
BBB	4	HIS	-	expression tag	UNP B5HSR1
BBB	5	HIS	-	expression tag	UNP B5HSR1
BBB	6	HIS	-	expression tag	UNP B5HSR1
BBB	7	HIS	-	expression tag	UNP B5HSR1
BBB	295	TYR	MET	engineered mutation	UNP B5HSR1
CCC	1	MET	-	initiating methionine	UNP B5HSR1
CCC	2	HIS	-	expression tag	UNP B5HSR1
CCC	3	HIS	-	expression tag	UNP B5HSR1
CCC	4	HIS	-	expression tag	UNP B5HSR1
CCC	5	HIS	-	expression tag	UNP B5HSR1
CCC	6	HIS	-	expression tag	UNP B5HSR1
CCC	7	HIS	-	expression tag	UNP B5HSR1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	295	TYR	MET	engineered mutation	UNP B5HSR1

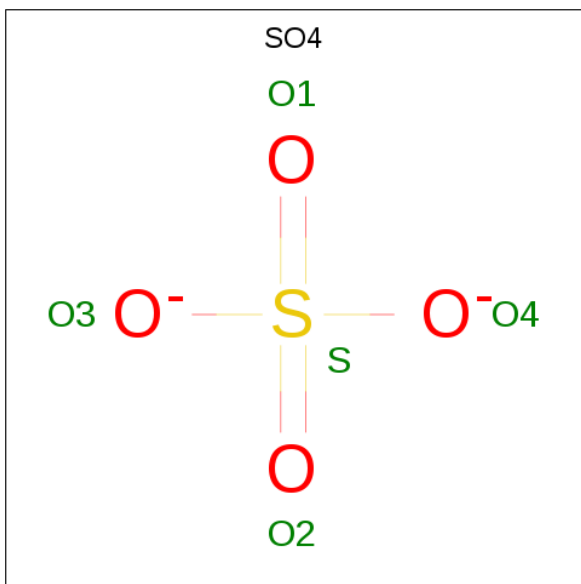
- Molecule 2 is a protein called GLU-HIS-SER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	FFF	3	Total	C	H	N	O	3	0	0
			46	14	21	5	6			

- Molecule 3 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	3	Total	Cu	0	0
			3	3		
3	BBB	3	Total	Cu	0	0
			3	3		
3	CCC	3	Total	Cu	0	0
			3	3		
3	FFF	1	Total	Cu	0	0
			1	1		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total	O	S	0	0
			5	4	1		
4	CCC	1	Total	O	S	0	0
			5	4	1		

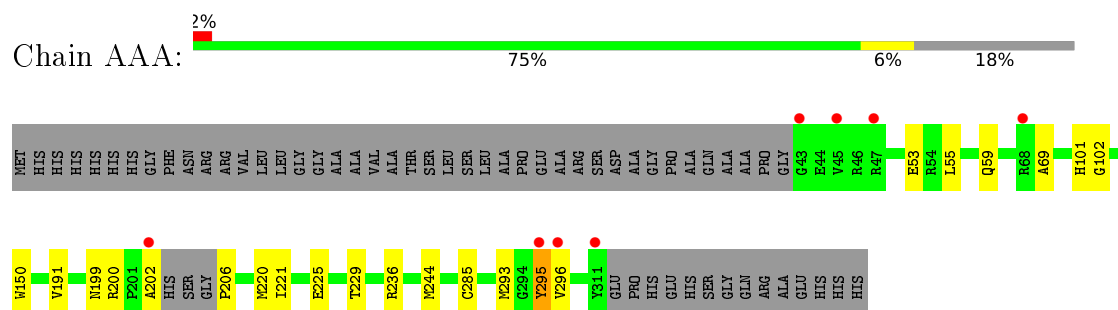
- Molecule 5 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	AAA	242	Total	O		0	0
			242	242			
5	BBB	188	Total	O		0	0
			188	188			
5	CCC	201	Total	O		0	0
			201	201			
5	FFF	1	Total	O		0	0
			1	1			

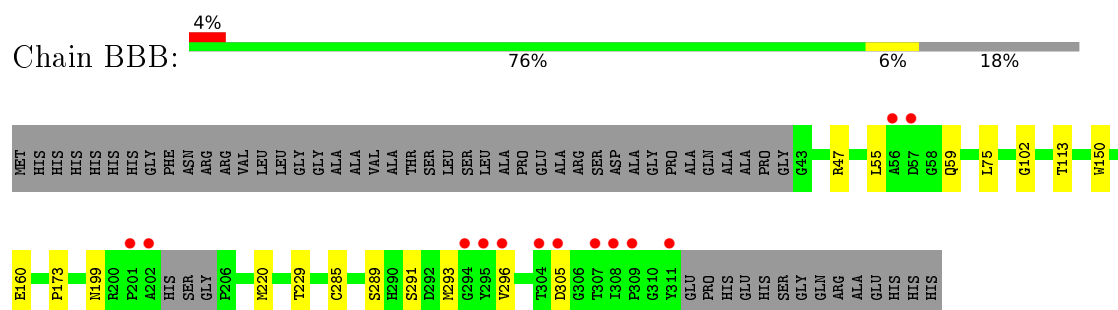
### 3 Residue-property plots [i](#)

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.

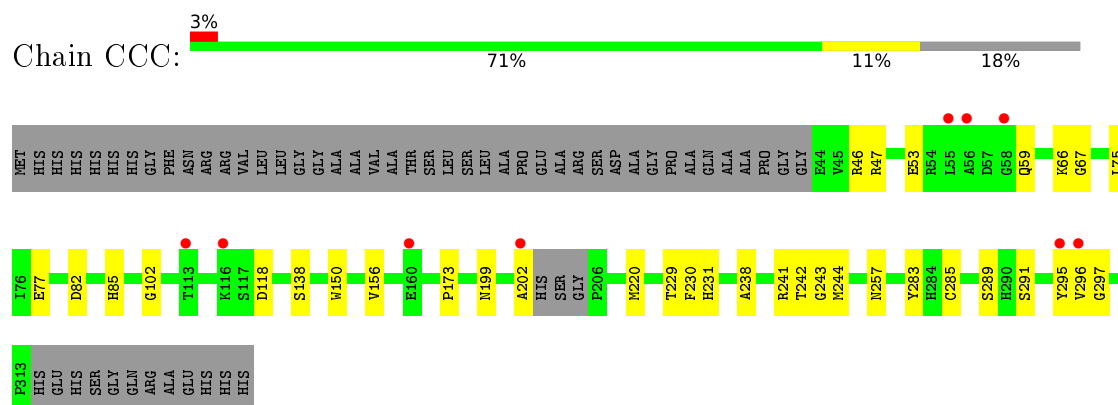
- Molecule 1: Copper oxidase



- Molecule 1: Copper oxidase



- Molecule 1: Copper oxidase



- Molecule 2: GLU-HIS-SER



There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.40Å 104.24Å 162.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.04 – 1.68 48.99 – 1.68	Depositor EDS
% Data completeness (in resolution range)	99.3 (49.04-1.68) 99.3 (48.99-1.68)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 1.68Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.191 , 0.232 0.202 , 0.242	Depositor DCC
$R_{free}$ test set	4943 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.8	Xtriage
Anisotropy	0.104	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 40.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12869	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, CU

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	AAA	0.78	0/2128	0.92	0/2886
1	BBB	0.72	0/2128	0.88	0/2886
1	CCC	0.74	0/2141	0.91	1/2905 (0.0%)
2	FFF	1.06	0/25	0.71	0/32
All	All	0.75	0/6422	0.91	1/8709 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	CCC	241	ARG	NE-CZ-NH1	-6.64	116.98	120.30

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	AAA	2071	1977	1957	13	0
1	BBB	2071	1977	1957	11	0
1	CCC	2083	1987	1967	21	0
2	FFF	25	21	19	0	0
3	AAA	3	0	0	0	0
3	BBB	3	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	CCC	3	0	0	0	0
3	FFF	1	0	0	0	0
4	AAA	10	0	0	0	0
4	CCC	5	0	0	0	0
5	AAA	242	0	0	1	0
5	BBB	188	0	0	1	0
5	CCC	201	0	0	1	0
5	FFF	1	0	0	0	0
All	All	6907	5962	5900	43	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 4.

The worst 5 of 43 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:AAA:225:GLU:OE2	1:CCC:289:SER:OG	2.17	0.63
1:CCC:59:GLN:NE2	1:CCC:199:ASN:HD22	1.98	0.61
1:CCC:202:ALA:HA	1:CCC:295:TYR:CE1	2.36	0.61
1:CCC:59:GLN:HE22	1:CCC:199:ASN:HB3	1.68	0.58
1:CCC:53:GLU:OE2	1:CCC:67:GLY:N	2.32	0.58

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	AAA	262/325 (81%)	255 (97%)	7 (3%)	0	100	100
1	BBB	262/325 (81%)	257 (98%)	5 (2%)	0	100	100
1	CCC	263/325 (81%)	255 (97%)	8 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	FFF	1/3 (33%)	1 (100%)	0	0	100	100
All	All	788/978 (81%)	768 (98%)	20 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	AAA	214/256 (84%)	210 (98%)	4 (2%)	57	38
1	BBB	214/256 (84%)	212 (99%)	2 (1%)	78	69
1	CCC	216/256 (84%)	214 (99%)	2 (1%)	78	69
2	FFF	3/3 (100%)	3 (100%)	0	100	100
All	All	647/771 (84%)	639 (99%)	8 (1%)	71	57

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	220	MET
1	CCC	138	SER
1	BBB	47	ARG
1	AAA	295	TYR
1	BBB	220	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 13 ligands modelled in this entry, 10 are monoatomic - leaving 3 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	$\# Z  > 2$
4	SO4	AAA	404	-	4,4,4	0.10	0	6,6,6	0.26	0
4	SO4	CCC	404	-	4,4,4	0.24	0	6,6,6	0.16	0
4	SO4	AAA	405	-	4,4,4	0.57	0	6,6,6	0.38	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers ⓘ

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	AAA	266/325 (81%)	-0.04	8 (3%) 50 53	15, 21, 39, 74	0
1	BBB	266/325 (81%)	0.22	13 (4%) 29 31	16, 27, 50, 96	0
1	CCC	267/325 (82%)	0.19	9 (3%) 45 48	15, 24, 45, 76	0
2	FFF	3/3 (100%)	1.12	0 100 100	34, 34, 46, 51	0
All	All	802/978 (82%)	0.13	30 (3%) 41 44	15, 24, 47, 96	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	296	VAL	10.4
1	BBB	295	TYR	9.6
1	AAA	296	VAL	7.9
1	CCC	202	ALA	7.2
1	AAA	295	TYR	6.6

### 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( $\text{\AA}^2$ )	Q<0.9
3	CU	CCC	403	1/1	0.82	0.06	46,46,46,46	0
4	SO4	AAA	405	5/5	0.85	0.15	26,30,43,46	0
3	CU	BBB	402	1/1	0.88	0.07	51,51,51,51	0
4	SO4	CCC	404	5/5	0.90	0.17	43,54,59,61	0
4	SO4	AAA	404	5/5	0.96	0.13	34,37,41,44	0
3	CU	BBB	401	1/1	0.97	0.05	29,29,29,29	0
3	CU	AAA	402	1/1	0.97	0.06	39,39,39,39	0
3	CU	AAA	403	1/1	0.98	0.14	16,16,16,16	1
3	CU	AAA	401	1/1	0.99	0.07	23,23,23,23	0
3	CU	FFF	101	1/1	0.99	0.08	18,18,18,18	1
3	CU	BBB	403	1/1	0.99	0.07	26,26,26,26	0
3	CU	CCC	401	1/1	0.99	0.09	23,23,23,23	0
3	CU	CCC	402	1/1	0.99	0.14	16,16,16,16	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.