



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 27, 2024 – 09:16 PM EDT

PDB ID : 1NJK  
Title : Crystal Structure of YbaW Probable Thioesterase from Escherichia coli  
Authors : Kim, Y.; Joachimiak, A.; Edwards, A.; Xu, X.; Savchenko, A.; Midwest Center  
for Structural Genomics (MCSG)  
Deposited on : 2002-12-31  
Resolution : 1.90 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	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

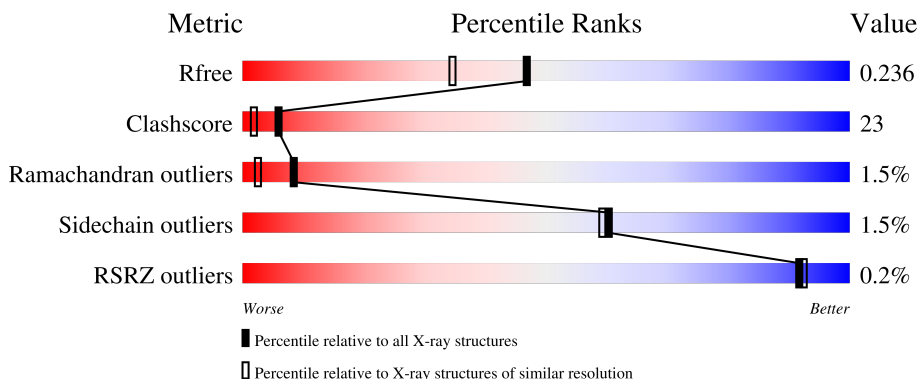
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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	156	<div> <div style="display: flex; justify-content: space-between;"> <span>51%</span> <span>32%</span> <span>•</span> <span>15%</span> </div> </div>
1	B	156	<div> <div style="display: flex; justify-content: space-between;"> <span>55%</span> <span>29%</span> <span>•</span> <span>15%</span> </div> </div>
1	C	156	<div> <div style="display: flex; justify-content: space-between;"> <span>54%</span> <span>29%</span> <span>•</span> <span>15%</span> </div> </div>
1	D	156	<div> <div style="display: flex; justify-content: space-between;"> <span>58%</span> <span>26%</span> <span>•</span> <span>15%</span> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4643 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 Hypothetical protein ybaW.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	133	Total	C	N	O	S	Se	0	0	0
			1071	680	186	201	1	3			
1	B	133	Total	C	N	O	S	Se	0	0	0
			1071	680	186	201	1	3			
1	C	133	Total	C	N	O	S	Se	0	0	0
			1071	680	186	201	1	3			
1	D	133	Total	C	N	O	S	Se	0	0	0
			1071	680	186	201	1	3			

There are 104 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	GLY	-	cloning artifact	UNP P77712
A	-19	SER	-	cloning artifact	UNP P77712
A	-18	SER	-	cloning artifact	UNP P77712
A	-17	HIS	-	cloning artifact	UNP P77712
A	-16	HIS	-	cloning artifact	UNP P77712
A	-15	HIS	-	cloning artifact	UNP P77712
A	-14	HIS	-	cloning artifact	UNP P77712
A	-13	HIS	-	cloning artifact	UNP P77712
A	-12	HIS	-	cloning artifact	UNP P77712
A	-11	SER	-	cloning artifact	UNP P77712
A	-10	SER	-	cloning artifact	UNP P77712
A	-9	GLY	-	cloning artifact	UNP P77712
A	-8	ARG	-	cloning artifact	UNP P77712
A	-7	GLU	-	cloning artifact	UNP P77712
A	-6	ASN	-	cloning artifact	UNP P77712
A	-5	LEU	-	cloning artifact	UNP P77712
A	-4	TYR	-	cloning artifact	UNP P77712
A	-3	PHE	-	cloning artifact	UNP P77712
A	-2	GLN	-	cloning artifact	UNP P77712
A	-1	GLY	-	cloning artifact	UNP P77712
A	0	HIS	-	cloning artifact	UNP P77712

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP P77712
A	44	MSE	MET	modified residue	UNP P77712
A	130	MSE	MET	modified residue	UNP P77712
A	133	GLY	-	cloning artifact	UNP P77712
A	134	HIS	-	cloning artifact	UNP P77712
B	-20	GLY	-	cloning artifact	UNP P77712
B	-19	SER	-	cloning artifact	UNP P77712
B	-18	SER	-	cloning artifact	UNP P77712
B	-17	HIS	-	cloning artifact	UNP P77712
B	-16	HIS	-	cloning artifact	UNP P77712
B	-15	HIS	-	cloning artifact	UNP P77712
B	-14	HIS	-	cloning artifact	UNP P77712
B	-13	HIS	-	cloning artifact	UNP P77712
B	-12	HIS	-	cloning artifact	UNP P77712
B	-11	SER	-	cloning artifact	UNP P77712
B	-10	SER	-	cloning artifact	UNP P77712
B	-9	GLY	-	cloning artifact	UNP P77712
B	-8	ARG	-	cloning artifact	UNP P77712
B	-7	GLU	-	cloning artifact	UNP P77712
B	-6	ASN	-	cloning artifact	UNP P77712
B	-5	LEU	-	cloning artifact	UNP P77712
B	-4	TYR	-	cloning artifact	UNP P77712
B	-3	PHE	-	cloning artifact	UNP P77712
B	-2	GLN	-	cloning artifact	UNP P77712
B	-1	GLY	-	cloning artifact	UNP P77712
B	0	HIS	-	cloning artifact	UNP P77712
B	1	MSE	MET	modified residue	UNP P77712
B	44	MSE	MET	modified residue	UNP P77712
B	130	MSE	MET	modified residue	UNP P77712
B	133	GLY	-	cloning artifact	UNP P77712
B	134	HIS	-	cloning artifact	UNP P77712
C	-20	GLY	-	cloning artifact	UNP P77712
C	-19	SER	-	cloning artifact	UNP P77712
C	-18	SER	-	cloning artifact	UNP P77712
C	-17	HIS	-	cloning artifact	UNP P77712
C	-16	HIS	-	cloning artifact	UNP P77712
C	-15	HIS	-	cloning artifact	UNP P77712
C	-14	HIS	-	cloning artifact	UNP P77712
C	-13	HIS	-	cloning artifact	UNP P77712
C	-12	HIS	-	cloning artifact	UNP P77712
C	-11	SER	-	cloning artifact	UNP P77712
C	-10	SER	-	cloning artifact	UNP P77712

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-9	GLY	-	cloning artifact	UNP P77712
C	-8	ARG	-	cloning artifact	UNP P77712
C	-7	GLU	-	cloning artifact	UNP P77712
C	-6	ASN	-	cloning artifact	UNP P77712
C	-5	LEU	-	cloning artifact	UNP P77712
C	-4	TYR	-	cloning artifact	UNP P77712
C	-3	PHE	-	cloning artifact	UNP P77712
C	-2	GLN	-	cloning artifact	UNP P77712
C	-1	GLY	-	cloning artifact	UNP P77712
C	0	HIS	-	cloning artifact	UNP P77712
C	1	MSE	MET	modified residue	UNP P77712
C	44	MSE	MET	modified residue	UNP P77712
C	130	MSE	MET	modified residue	UNP P77712
C	133	GLY	-	cloning artifact	UNP P77712
C	134	HIS	-	cloning artifact	UNP P77712
D	-20	GLY	-	cloning artifact	UNP P77712
D	-19	SER	-	cloning artifact	UNP P77712
D	-18	SER	-	cloning artifact	UNP P77712
D	-17	HIS	-	cloning artifact	UNP P77712
D	-16	HIS	-	cloning artifact	UNP P77712
D	-15	HIS	-	cloning artifact	UNP P77712
D	-14	HIS	-	cloning artifact	UNP P77712
D	-13	HIS	-	cloning artifact	UNP P77712
D	-12	HIS	-	cloning artifact	UNP P77712
D	-11	SER	-	cloning artifact	UNP P77712
D	-10	SER	-	cloning artifact	UNP P77712
D	-9	GLY	-	cloning artifact	UNP P77712
D	-8	ARG	-	cloning artifact	UNP P77712
D	-7	GLU	-	cloning artifact	UNP P77712
D	-6	ASN	-	cloning artifact	UNP P77712
D	-5	LEU	-	cloning artifact	UNP P77712
D	-4	TYR	-	cloning artifact	UNP P77712
D	-3	PHE	-	cloning artifact	UNP P77712
D	-2	GLN	-	cloning artifact	UNP P77712
D	-1	GLY	-	cloning artifact	UNP P77712
D	0	HIS	-	cloning artifact	UNP P77712
D	1	MSE	MET	modified residue	UNP P77712
D	44	MSE	MET	modified residue	UNP P77712
D	130	MSE	MET	modified residue	UNP P77712
D	133	GLY	-	cloning artifact	UNP P77712
D	134	HIS	-	cloning artifact	UNP P77712

- Molecule 2 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total I 1 1	0	0
2	C	2	Total I 2 2	0	0
2	D	1	Total I 1 1	0	0

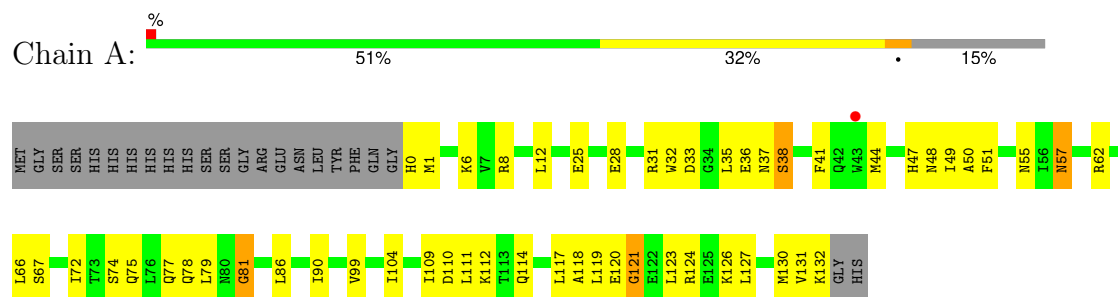
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	89	Total O 89 89	0	0
3	B	80	Total O 80 80	0	0
3	C	92	Total O 92 92	0	0
3	D	94	Total O 94 94	0	0

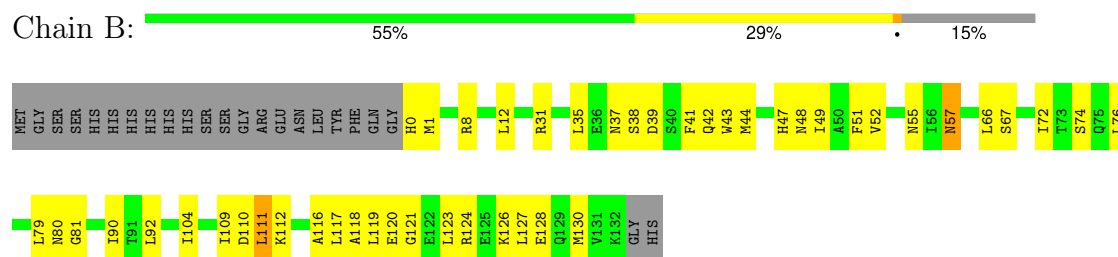
### 3 Residue-property plots

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.

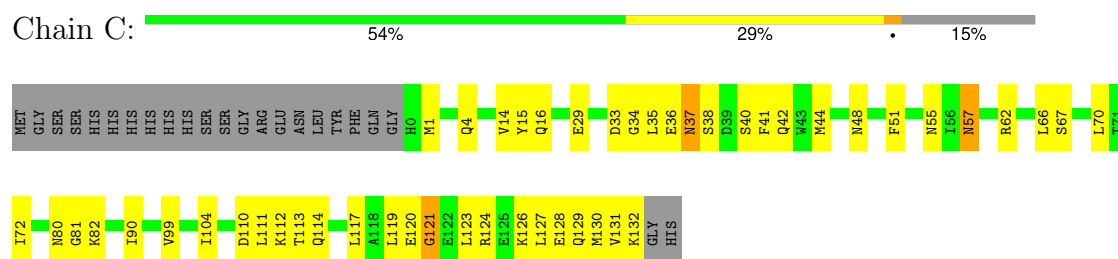
#### • Molecule 1: Hypothetical protein ybaW



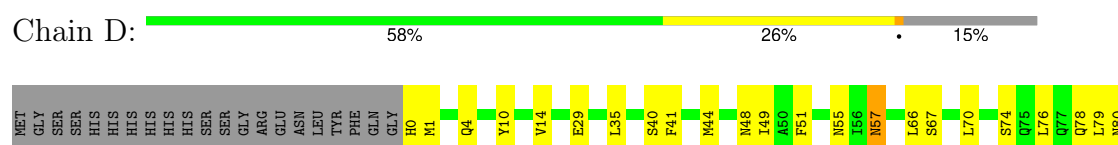
#### • Molecule 1: Hypothetical protein ybaW



#### • Molecule 1: Hypothetical protein ybaW



#### • Molecule 1: Hypothetical protein ybaW



G81	K82	I104	V107	C108	I109	D110	L111	K112	T113	Q114	K115	A116	L117	A118	L119	L123	R124	E125	K126	L127	M130	V131	K132	GLY	HIS
-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.13Å 90.13Å 83.37Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.35 – 1.90 35.35 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.6 (35.35-1.90) 96.7 (35.35-1.90)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.10 (at 1.89Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.214 , 0.240 0.210 , 0.236	Depositor DCC
$R_{free}$ test set	5659 reflections (9.71%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.4	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 49.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.045 for -h,-k,l 0.046 for h,-h-k,-l 0.488 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4643	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.74% 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: IOD

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	0.31	0/1085	0.55	0/1465
1	B	0.32	0/1085	0.56	0/1465
1	C	0.32	0/1085	0.58	0/1465
1	D	0.33	0/1085	0.57	0/1465
All	All	0.32	0/4340	0.56	0/5860

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	1071	0	1080	53	0
1	B	1071	0	1080	58	0
1	C	1071	0	1080	48	0
1	D	1071	0	1080	49	0
2	B	1	0	0	0	0
2	C	2	0	0	0	0
2	D	1	0	0	0	0
3	A	89	0	0	0	0
3	B	80	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	92	0	0	0	0
3	D	94	0	0	0	0
All	All	4643	0	4320	194	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 23.

The worst 5 of 194 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:79:LEU:HD22	1:B:127:LEU:HD13	1.50	0.93
1:B:52:VAL:HG21	1:B:109:ILE:HD12	1.54	0.86
1:A:120:GLU:HG3	1:A:121:GLY:H	1.46	0.80
1:B:123:LEU:O	1:B:126:LYS:HG2	1.83	0.78
1:C:123:LEU:HA	1:C:126:LYS:HE2	1.65	0.78

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	131/156 (84%)	115 (88%)	13 (10%)	3 (2%)	5	1
1	B	131/156 (84%)	116 (88%)	13 (10%)	2 (2%)	8	2
1	C	131/156 (84%)	120 (92%)	9 (7%)	2 (2%)	8	2
1	D	131/156 (84%)	117 (89%)	13 (10%)	1 (1%)	16	8
All	All	524/624 (84%)	468 (89%)	48 (9%)	8 (2%)	8	2

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	81	GLY
1	A	81	GLY
1	A	121	GLY
1	B	121	GLY
1	C	121	GLY

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	118/134 (88%)	117 (99%)	1 (1%)	79	80
1	B	118/134 (88%)	115 (98%)	3 (2%)	42	37
1	C	118/134 (88%)	116 (98%)	2 (2%)	56	54
1	D	118/134 (88%)	117 (99%)	1 (1%)	79	80
All	All	472/536 (88%)	465 (98%)	7 (2%)	60	59

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	92	LEU
1	C	37	ASN
1	D	57	ASN
1	C	57	ASN
1	B	57	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	2	GLN
1	D	48	ASN
1	D	57	ASN
1	D	37	ASN
1	B	57	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

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 [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	130/156 (83%)	-0.97	1 (0%) 82 84	13, 24, 120, 137	0
1	B	130/156 (83%)	-0.99	0 100 100	14, 24, 120, 133	0
1	C	130/156 (83%)	-1.11	0 100 100	12, 24, 105, 120	0
1	D	130/156 (83%)	-1.11	0 100 100	13, 25, 107, 121	0
All	All	520/624 (83%)	-1.04	1 (0%) 92 92	12, 24, 116, 137	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	43	TRP	2.4

### 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
2	IOD	D	403	1/1	0.99	0.10	59,59,59,59	1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	IOD	C	401	1/1	1.00	0.01	25,25,25,25	0
2	IOD	C	402	1/1	1.00	0.10	61,61,61,61	1
2	IOD	B	404	1/1	1.00	0.05	86,86,86,86	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.