



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2021 – 04:38 PM EDT

PDB ID : 3E4Z  
Title : Crystal structure of human insulin degrading enzyme in complex with insulin-like growth factor II  
Authors : Guo, Q.; Manolopoulou, M.; Tang, W.-J.  
Deposited on : 2008-08-12  
Resolution : 2.28 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  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
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) : 2.23.2

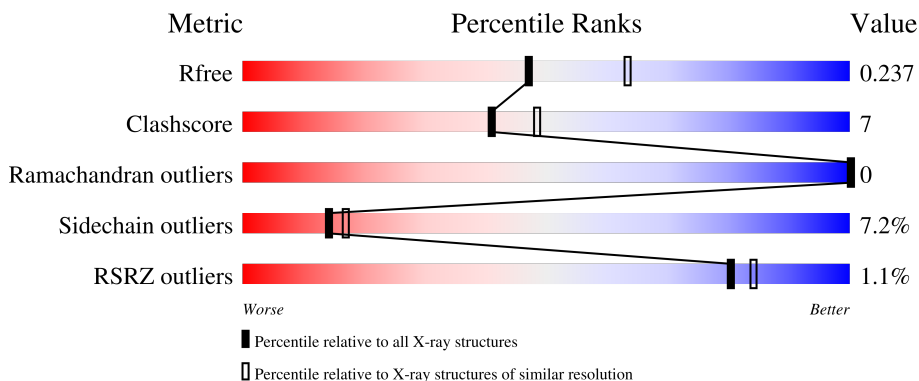
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.28 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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	990	 79% 15% . .
1	B	990	 78% 16% . .
2	C	67	 10% . . 87%
2	D	67	 9% . 87%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 16411 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 Insulin-degrading enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	953	7783	5014	1306	1441	22	0	0	0
1	B	952	7778	5011	1305	1440	22	0	0	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	MET	-	expression tag	UNP Q5T5N2
A	31	HIS	-	expression tag	UNP Q5T5N2
A	32	HIS	-	expression tag	UNP Q5T5N2
A	33	HIS	-	expression tag	UNP Q5T5N2
A	34	HIS	-	expression tag	UNP Q5T5N2
A	35	HIS	-	expression tag	UNP Q5T5N2
A	36	HIS	-	expression tag	UNP Q5T5N2
A	37	ALA	-	expression tag	UNP Q5T5N2
A	38	ALA	-	expression tag	UNP Q5T5N2
A	39	GLY	-	expression tag	UNP Q5T5N2
A	40	ILE	-	expression tag	UNP Q5T5N2
A	41	PRO	-	expression tag	UNP Q5T5N2
A	110	LEU	CYS	engineered mutation	UNP Q5T5N2
A	111	GLN	GLU	engineered mutation	UNP Q5T5N2
A	171	SER	CYS	engineered mutation	UNP Q5T5N2
A	178	ALA	CYS	engineered mutation	UNP Q5T5N2
A	257	VAL	CYS	engineered mutation	UNP Q5T5N2
A	414	LEU	CYS	engineered mutation	UNP Q5T5N2
A	573	ASN	CYS	engineered mutation	UNP Q5T5N2
A	590	SER	CYS	engineered mutation	UNP Q5T5N2
A	789	SER	CYS	engineered mutation	UNP Q5T5N2
A	812	ALA	CYS	engineered mutation	UNP Q5T5N2
A	819	ALA	CYS	engineered mutation	UNP Q5T5N2
A	904	SER	CYS	engineered mutation	UNP Q5T5N2
A	908	TYR	TRP	engineered mutation	UNP Q5T5N2

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	30	MET	-	expression tag	UNP Q5T5N2
B	31	HIS	-	expression tag	UNP Q5T5N2
B	32	HIS	-	expression tag	UNP Q5T5N2
B	33	HIS	-	expression tag	UNP Q5T5N2
B	34	HIS	-	expression tag	UNP Q5T5N2
B	35	HIS	-	expression tag	UNP Q5T5N2
B	36	HIS	-	expression tag	UNP Q5T5N2
B	37	ALA	-	expression tag	UNP Q5T5N2
B	38	ALA	-	expression tag	UNP Q5T5N2
B	39	GLY	-	expression tag	UNP Q5T5N2
B	40	ILE	-	expression tag	UNP Q5T5N2
B	41	PRO	-	expression tag	UNP Q5T5N2
B	110	LEU	CYS	engineered mutation	UNP Q5T5N2
B	111	GLN	GLU	engineered mutation	UNP Q5T5N2
B	171	SER	CYS	engineered mutation	UNP Q5T5N2
B	178	ALA	CYS	engineered mutation	UNP Q5T5N2
B	257	VAL	CYS	engineered mutation	UNP Q5T5N2
B	414	LEU	CYS	engineered mutation	UNP Q5T5N2
B	573	ASN	CYS	engineered mutation	UNP Q5T5N2
B	590	SER	CYS	engineered mutation	UNP Q5T5N2
B	789	SER	CYS	engineered mutation	UNP Q5T5N2
B	812	ALA	CYS	engineered mutation	UNP Q5T5N2
B	819	ALA	CYS	engineered mutation	UNP Q5T5N2
B	904	SER	CYS	engineered mutation	UNP Q5T5N2
B	908	TYR	TRP	engineered mutation	UNP Q5T5N2

- Molecule 2 is a protein called Insulin-like growth factor II.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace	
2	C	9	Total	C	N	O	0	0	0
			62	39	10	13			
2	D	9	Total	C	N	O	0	0	0
			62	39	10	13			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	C	1	Total	Zn	0	0
			1	1		

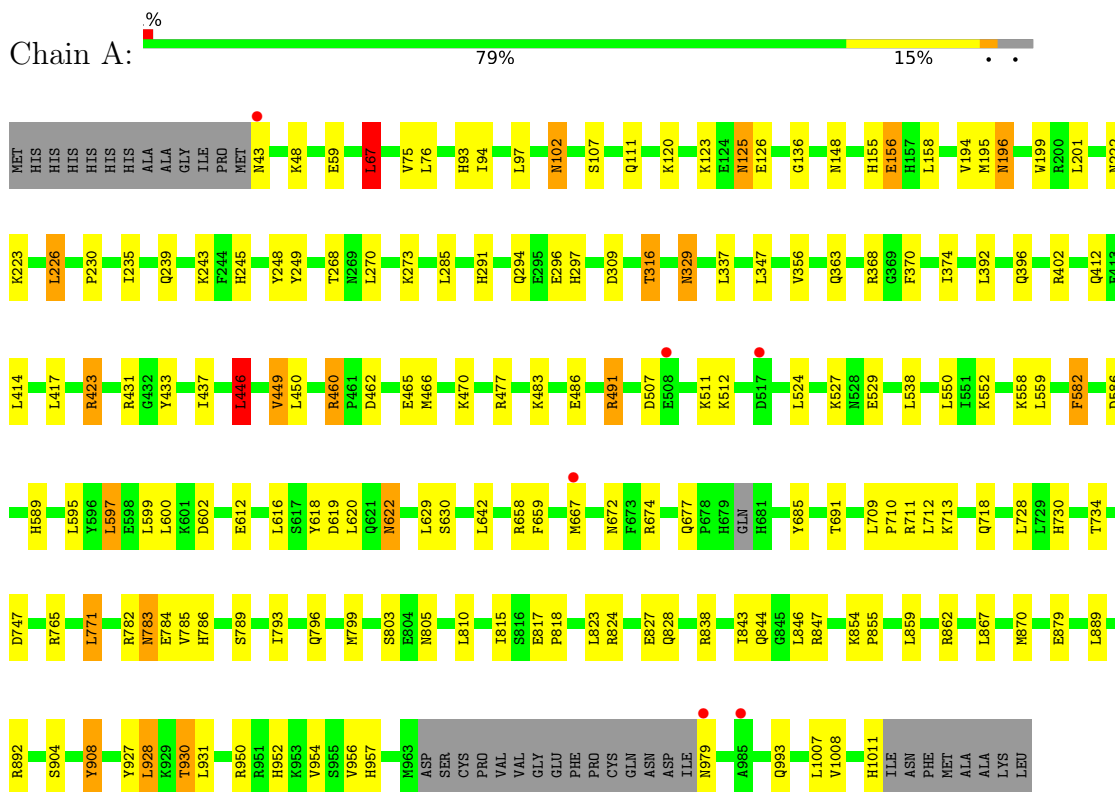
- Molecule 4 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	384	Total 384	O 384	0	0
4	B	338	Total 338	O 338	0	0
4	C	2	Total 2	O 2	0	0

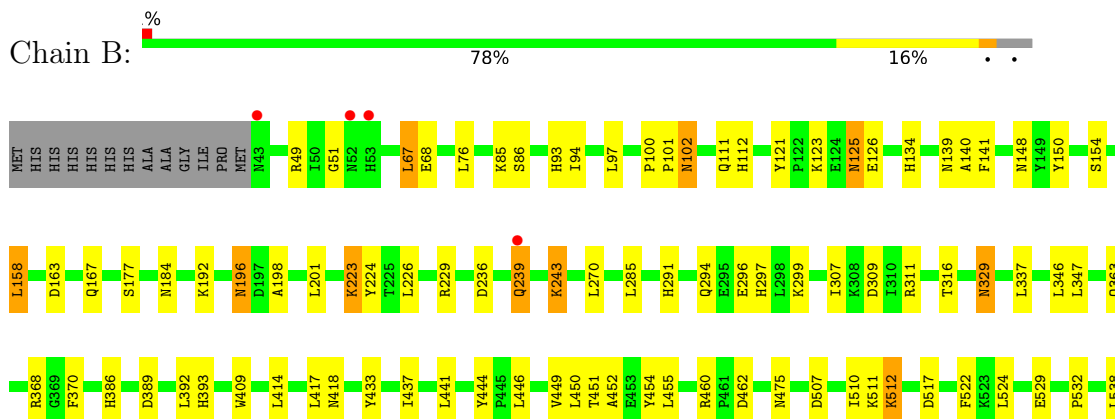
### 3 Residue-property plots

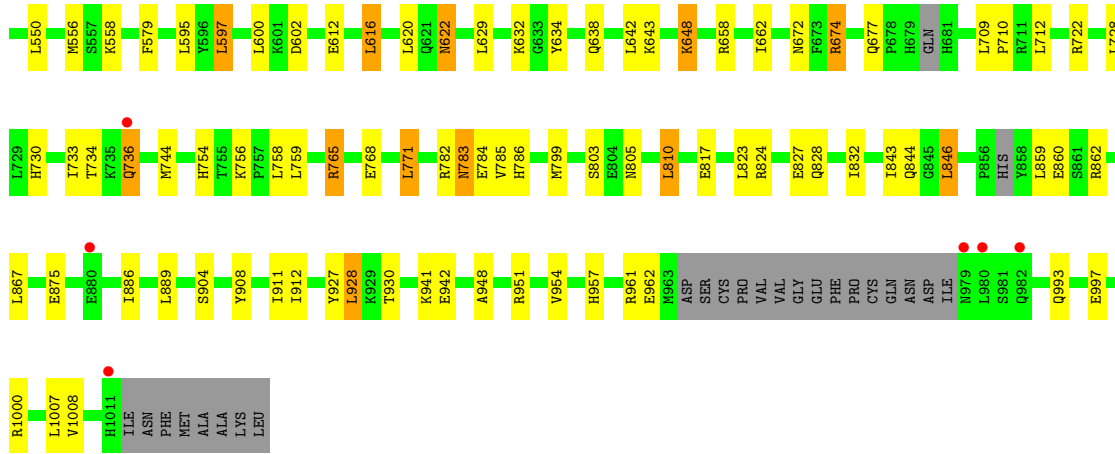
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

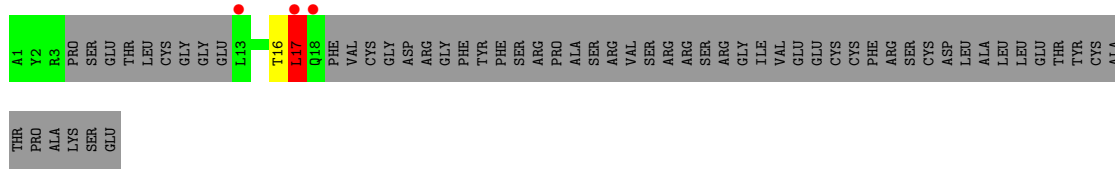


- Molecule 1: Insulin-degrading enzyme

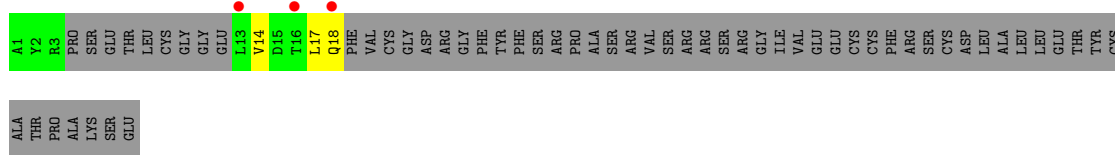




● Molecule 2: Insulin-like growth factor II



● Molecule 2: Insulin-like growth factor II



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	263.03Å 263.03Å 90.82Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.28 49.71 – 2.28	Depositor EDS
% Data completeness (in resolution range)	99.7 (50.00-2.28) 99.7 (49.71-2.28)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.196 , 0.230 0.206 , 0.237	Depositor DCC
$R_{free}$ test set	8139 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.4	Xtrriage
Anisotropy	0.051	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 52.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.018 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	16411	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.86% 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:  
ZN

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	0.59	1/7975 (0.0%)	0.68	6/10787 (0.1%)
1	B	0.57	0/7969	0.66	4/10777 (0.0%)
2	C	0.68	0/60	1.15	1/80 (1.2%)
2	D	0.61	0/60	0.79	0/80
All	All	0.58	1/16064 (0.0%)	0.67	11/21724 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	908	TYR	CD1-CE1	-5.36	1.31	1.39

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	67	LEU	CA-CB-CG	6.59	130.46	115.30
1	B	67	LEU	CA-CB-CG	6.34	129.88	115.30
1	B	311	ARG	NE-CZ-NH2	-5.75	117.42	120.30
2	C	17	LEU	CA-CB-CG	5.64	128.27	115.30
1	B	311	ARG	NE-CZ-NH1	5.51	123.06	120.30

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	7783	0	7721	106	1
1	B	7778	0	7718	111	0
2	C	62	0	60	4	0
2	D	62	0	60	4	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	384	0	0	30	0
4	B	338	0	0	23	0
4	C	2	0	0	0	0
All	All	16411	0	15559	221	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

The worst 5 of 221 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:A:908:TYR:HE1	4:A:1368:HOH:O	1.18	1.23
1:A:667:MET:HG3	4:A:1031:HOH:O	1.39	1.22
1:A:908:TYR:CE1	4:A:1368:HOH:O	1.94	1.10
4:A:1311:HOH:O	2:C:17:LEU:HB3	1.53	1.08
1:B:154:SER:HB3	4:B:1351:HOH:O	1.53	1.08

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:460:ARG:NH2	1:A:747:ASP:OD1[3_655]	2.05	0.15

## 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	947/990 (96%)	924 (98%)	23 (2%)	0	100	100
1	B	944/990 (95%)	925 (98%)	19 (2%)	0	100	100
2	C	5/67 (8%)	3 (60%)	2 (40%)	0	100	100
2	D	5/67 (8%)	5 (100%)	0	0	100	100
All	All	1901/2114 (90%)	1857 (98%)	44 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	846/880 (96%)	783 (93%)	63 (7%)	13	16
1	B	846/880 (96%)	787 (93%)	59 (7%)	15	18
2	C	6/57 (10%)	5 (83%)	1 (17%)	2	1
2	D	6/57 (10%)	6 (100%)	0	100	100
All	All	1704/1874 (91%)	1581 (93%)	123 (7%)	14	17

5 of 123 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	928	LEU
1	B	783	ASN
1	B	177	SER
1	B	771	LEU
1	B	928	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 58 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	93	HIS
1	B	922	ASN
1	B	294	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	883	GLN
1	B	770	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](#)

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](#)

Of 2 ligands modelled in this entry, 2 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	953/990 (96%)	-0.27	6 (0%) 89 91	20, 32, 47, 62	0
1	B	952/990 (96%)	-0.22	10 (1%) 80 84	25, 37, 50, 71	0
2	C	9/67 (13%)	1.58	3 (33%) 0 0	37, 66, 67, 67	0
2	D	9/67 (13%)	1.27	3 (33%) 0 0	36, 73, 75, 75	0
All	All	1923/2114 (90%)	-0.23	22 (1%) 80 84	20, 35, 49, 75	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	979	ASN	4.0
1	B	43	ASN	3.6
2	C	17	LEU	3.1
2	C	18	GLN	3.0
1	B	52	ASN	2.9

### 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	ZN	B	2	1/1	0.99	0.28	2,2,2,2	0
3	ZN	C	68	1/1	0.99	0.26	2,2,2,2	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.