



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 15, 2018 – 10:31 PM EST

PDB ID : 4IOF  
Title : Crystal structure analysis of Fab-bound human Insulin Degrading Enzyme (IDE)  
Authors : McCord, L.A.; Liang, W.G.; Hoey, R.; Dowdell, E.; Koide, A.; Koide, S.; Tang, W.J.  
Deposited on : 2013-01-07  
Resolution : 3.35 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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	rb-20030736
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030736

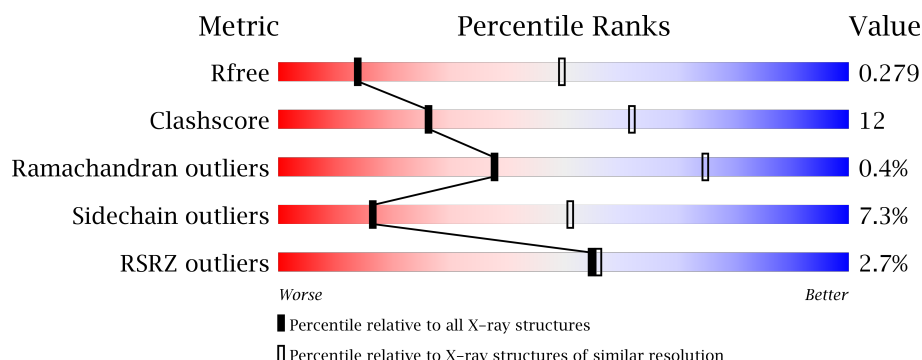
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.35 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))
$R_{free}$	100719	1156 (3.42-3.30)
Clashscore	112137	1231 (3.42-3.30)
Ramachandran outliers	110173	1212 (3.42-3.30)
Sidechain outliers	110143	1211 (3.42-3.30)
RSRZ outliers	101464	1165 (3.42-3.30)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	<div> <div>0.1%</div> <div>75%</div> <div>18%</div> <div>6%</div> </div>
1	B	990	<div> <div>2%</div> <div>63%</div> <div>22%</div> <div>13%</div> </div>
2	C	263	<div> <div>50%</div> <div>30%</div> <div>18%</div> </div>
2	E	263	<div> <div>3%</div> <div>43%</div> <div>33%</div> <div>20%</div> </div>
3	D	239	<div> <div>5%</div> <div>49%</div> <div>29%</div> <div>5%</div> <div>17%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	239	 <p>A horizontal bar chart showing the quality of chain F. The bar is divided into four segments: a small red segment at the beginning labeled '7%', a large green segment labeled '45%', a large yellow segment labeled '33%', and a small grey segment at the end labeled '18%'. There are two small black dots on the boundary between the yellow and grey segments.</p>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 20939 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
1	A	932	Total	C	N	O	S	0	0	0
			7622	4911	1274	1415	22			
1	B	861	Total	C	N	O	S	0	0	0
			7040	4550	1173	1296	21			

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	MET	-	EXPRESSION TAG	UNP P14735
A	31	HIS	-	EXPRESSION TAG	UNP P14735
A	32	HIS	-	EXPRESSION TAG	UNP P14735
A	33	HIS	-	EXPRESSION TAG	UNP P14735
A	34	HIS	-	EXPRESSION TAG	UNP P14735
A	35	HIS	-	EXPRESSION TAG	UNP P14735
A	36	HIS	-	EXPRESSION TAG	UNP P14735
A	37	ALA	-	EXPRESSION TAG	UNP P14735
A	38	ALA	-	EXPRESSION TAG	UNP P14735
A	39	GLY	-	EXPRESSION TAG	UNP P14735
A	40	ILE	-	EXPRESSION TAG	UNP P14735
A	41	PRO	-	EXPRESSION TAG	UNP P14735
A	110	LEU	CYS	CONFLICT	UNP P14735
A	171	SER	CYS	CONFLICT	UNP P14735
A	178	ALA	CYS	CONFLICT	UNP P14735
A	257	VAL	CYS	CONFLICT	UNP P14735
A	414	LEU	CYS	CONFLICT	UNP P14735
A	573	ASN	CYS	CONFLICT	UNP P14735
A	590	SER	CYS	CONFLICT	UNP P14735
A	789	SER	CYS	CONFLICT	UNP P14735
A	812	ALA	CYS	CONFLICT	UNP P14735
A	819	ALA	CYS	CONFLICT	UNP P14735
A	904	SER	CYS	CONFLICT	UNP P14735
A	966	ASN	CYS	CONFLICT	UNP P14735
A	974	ALA	CYS	CONFLICT	UNP P14735

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	30	MET	-	EXPRESSION TAG	UNP P14735
B	31	HIS	-	EXPRESSION TAG	UNP P14735
B	32	HIS	-	EXPRESSION TAG	UNP P14735
B	33	HIS	-	EXPRESSION TAG	UNP P14735
B	34	HIS	-	EXPRESSION TAG	UNP P14735
B	35	HIS	-	EXPRESSION TAG	UNP P14735
B	36	HIS	-	EXPRESSION TAG	UNP P14735
B	37	ALA	-	EXPRESSION TAG	UNP P14735
B	38	ALA	-	EXPRESSION TAG	UNP P14735
B	39	GLY	-	EXPRESSION TAG	UNP P14735
B	40	ILE	-	EXPRESSION TAG	UNP P14735
B	41	PRO	-	EXPRESSION TAG	UNP P14735
B	110	LEU	CYS	CONFLICT	UNP P14735
B	171	SER	CYS	CONFLICT	UNP P14735
B	178	ALA	CYS	CONFLICT	UNP P14735
B	257	VAL	CYS	CONFLICT	UNP P14735
B	414	LEU	CYS	CONFLICT	UNP P14735
B	573	ASN	CYS	CONFLICT	UNP P14735
B	590	SER	CYS	CONFLICT	UNP P14735
B	789	SER	CYS	CONFLICT	UNP P14735
B	812	ALA	CYS	CONFLICT	UNP P14735
B	819	ALA	CYS	CONFLICT	UNP P14735
B	904	SER	CYS	CONFLICT	UNP P14735
B	966	ASN	CYS	CONFLICT	UNP P14735
B	974	ALA	CYS	CONFLICT	UNP P14735

- Molecule 2 is a protein called Fab-bound IDE, heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	216	Total	C	N	O	S	0	0	0
			1637	1043	268	319	7			
2	E	211	Total	C	N	O	S	0	0	0
			1602	1023	263	309	7			

- Molecule 3 is a protein called Fab-bound IDE, light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	199	Total	C	N	O	S	0	0	0
			1528	962	255	306	5			
3	F	197	Total	C	N	O	S	0	0	0
			1508	947	250	306	5			

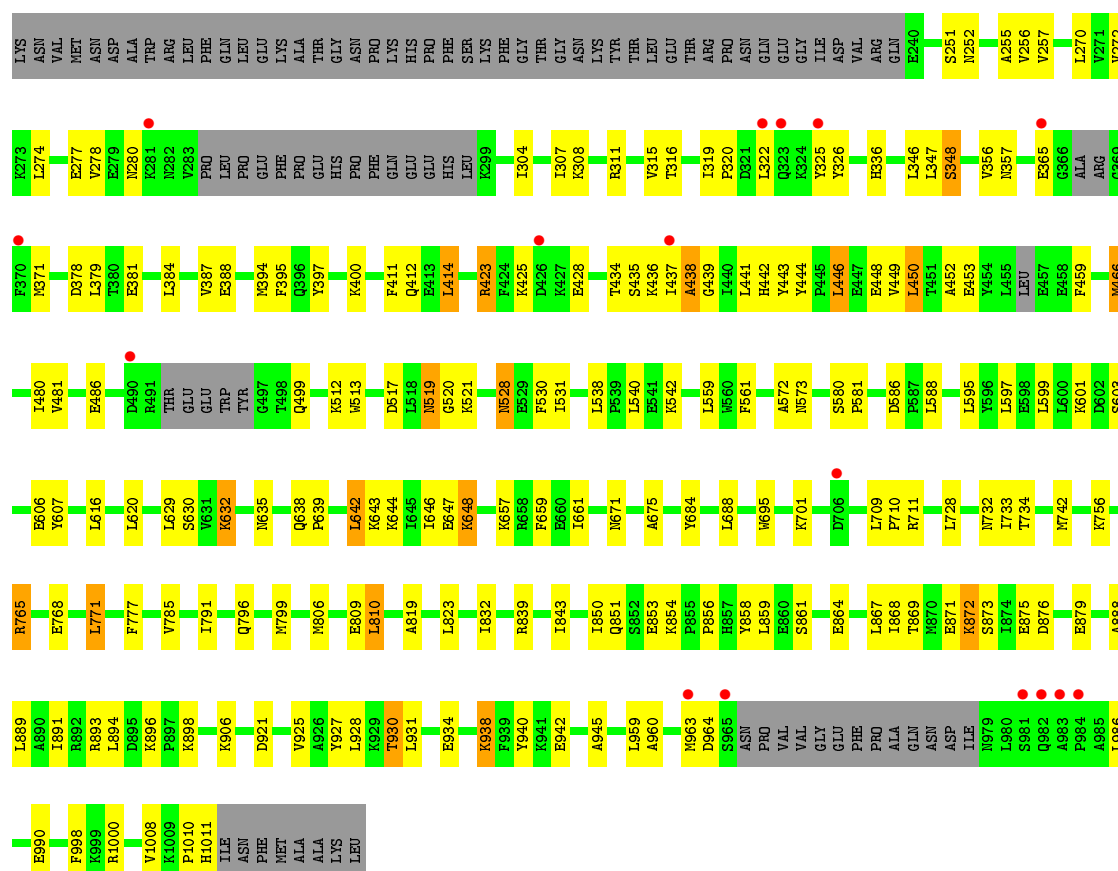
- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Zn	0	0
			1	1		

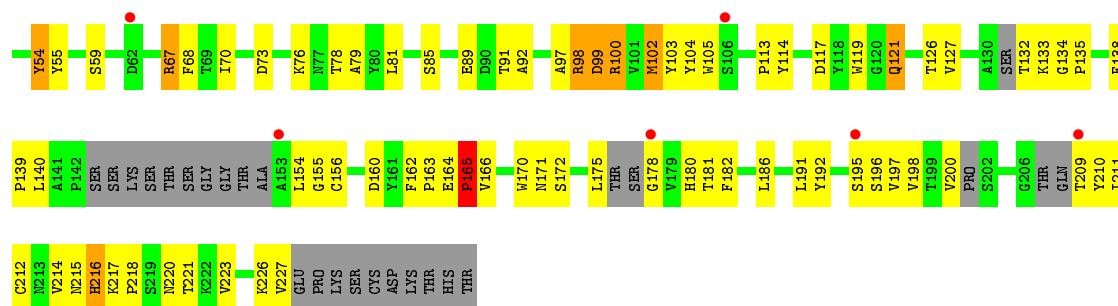
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	O	0	0
			1	1		

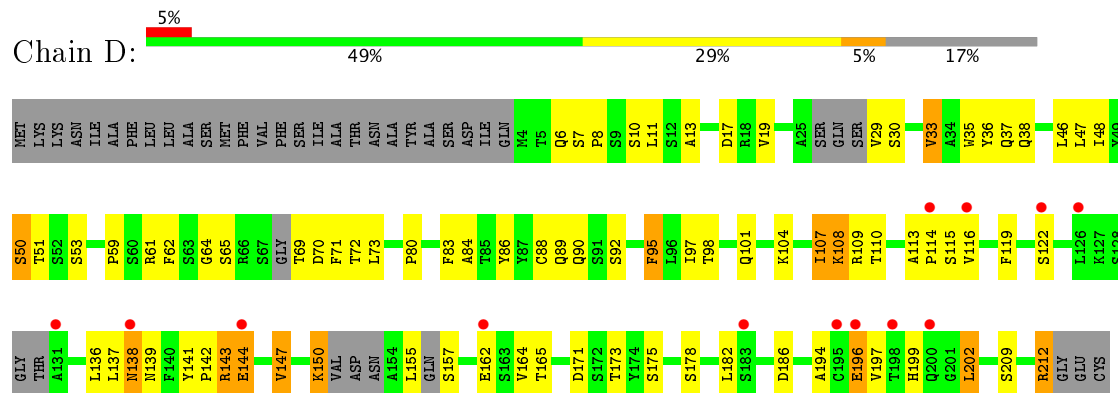




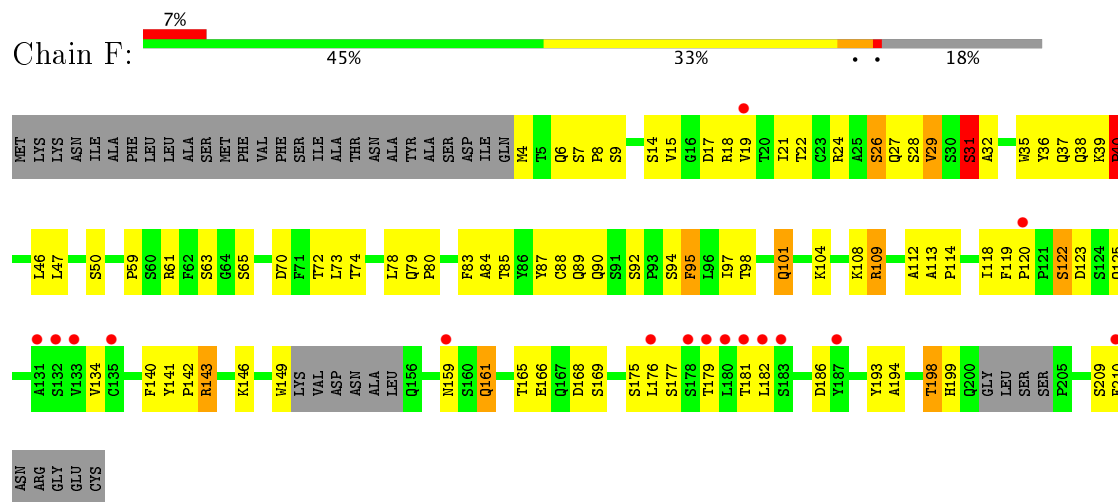




- Molecule 3: Fab-bound IDE, light chain



- Molecule 3: Fab-bound IDE, light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	56.95Å 131.66Å 377.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.85 – 3.35 49.85 – 3.35	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.85-3.35) 95.1 (49.85-3.35)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.88 (at 3.33Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.232 , 0.280 0.235 , 0.279	Depositor DCC
$R_{free}$ test set	1996 reflections (4.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.9	Xtriage
Anisotropy	0.237	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 10.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.82	EDS
Total number of atoms	20939	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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: 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	0.32	0/7803	0.68	4/10549 (0.0%)
1	B	0.34	0/7206	0.72	6/9733 (0.1%)
2	C	0.34	0/1679	0.88	1/2285 (0.0%)
2	E	0.50	0/1641	0.85	2/2229 (0.1%)
3	D	0.33	0/1557	0.83	3/2106 (0.1%)
3	F	0.40	0/1540	0.90	5/2088 (0.2%)
All	All	0.35	0/21426	0.75	21/28990 (0.1%)

There are no bond length outliers.

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	280	ASN	CB-CA-C	-15.27	79.86	110.40
1	B	171	SER	C-N-CD	-13.15	91.66	120.60
3	F	29	VAL	CB-CA-C	-11.61	89.35	111.40
2	E	220	ASN	CB-CA-C	-9.30	91.81	110.40
1	A	51	GLY	N-CA-C	-8.68	91.40	113.10

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	7622	0	7588	104	0
1	B	7040	0	7032	158	0
2	C	1637	0	1577	55	0
2	E	1602	0	1548	77	0
3	D	1528	0	1501	54	0
3	F	1508	0	1472	60	0
4	A	1	0	0	0	0
5	B	1	0	0	0	0
All	All	20939	0	20718	488	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 12.

The worst 5 of 488 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)
2:E:216:HIS:ND1	2:E:218:PRO:HD2	1.68	1.09
3:F:27:GLN:HG3	3:F:28:SER:H	1.17	1.08
1:B:118:THR:HG21	1:B:167:GLN:CB	1.87	1.04
1:B:118:THR:HG21	1:B:167:GLN:HB3	1.03	1.00
1:B:118:THR:CG2	1:B:167:GLN:HB3	1.97	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	922/990 (93%)	886 (96%)	35 (4%)	1 (0%)	55 87
1	B	845/990 (85%)	804 (95%)	38 (4%)	3 (0%)	38 75
2	C	206/263 (78%)	188 (91%)	15 (7%)	3 (2%)	12 46
2	E	199/263 (76%)	178 (89%)	19 (10%)	2 (1%)	18 56

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	D	187/239 (78%)	165 (88%)	21 (11%)	1 (0%)	32	70
3	F	191/239 (80%)	162 (85%)	28 (15%)	1 (0%)	32	70
All	All	2550/2984 (86%)	2383 (94%)	156 (6%)	11 (0%)	38	75

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	138	ASN
1	A	282	ASN
1	B	95	GLY
1	B	172	PRO
1	B	1010	PRO

### 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	832/879 (95%)	785 (94%)	47 (6%)	25	61
1	B	768/879 (87%)	719 (94%)	49 (6%)	20	57
2	C	181/220 (82%)	169 (93%)	12 (7%)	19	56
2	E	176/220 (80%)	157 (89%)	19 (11%)	7	30
3	D	178/210 (85%)	155 (87%)	23 (13%)	5	21
3	F	176/210 (84%)	158 (90%)	18 (10%)	8	33
All	All	2311/2618 (88%)	2143 (93%)	168 (7%)	16	51

5 of 168 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	648	LYS
1	B	964	ASP
3	F	63	SER
1	B	771	LEU
1	B	872	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	638	GLN
1	B	743	GLN
2	E	39	GLN
1	B	528	ASN
1	B	573	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is 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 ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	932/990 (94%)	0.01	11 (1%) 79 80	12, 39, 101, 152	0
1	B	861/990 (86%)	0.05	23 (2%) 55 56	11, 41, 100, 151	0
2	C	216/263 (82%)	0.25	1 (0%) 90 92	19, 56, 97, 112	0
2	E	211/263 (80%)	0.19	7 (3%) 47 47	18, 55, 98, 144	0
3	D	199/239 (83%)	0.47	13 (6%) 20 20	22, 55, 102, 138	0
3	F	197/239 (82%)	0.55	16 (8%) 13 13	20, 69, 116, 147	0
All	All	2616/2984 (87%)	0.13	71 (2%) 55 56	11, 46, 103, 152	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	179	THR	6.7
3	F	135	CYS	5.6
1	B	102	ASN	5.0
3	F	181	THR	4.0
1	B	370	PHE	3.8

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ZN	A	1101	1/1	0.96	0.05	-	75,75,75,75	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.