



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:27 AM GMT

PDB ID : 3LZB  
Title : EGFR kinase domain complexed with an imidazo[2,1-b]thiazole inhibitor  
Authors : Swinger, K.K.  
Deposited on : 2010-03-01  
Resolution : 2.70 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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

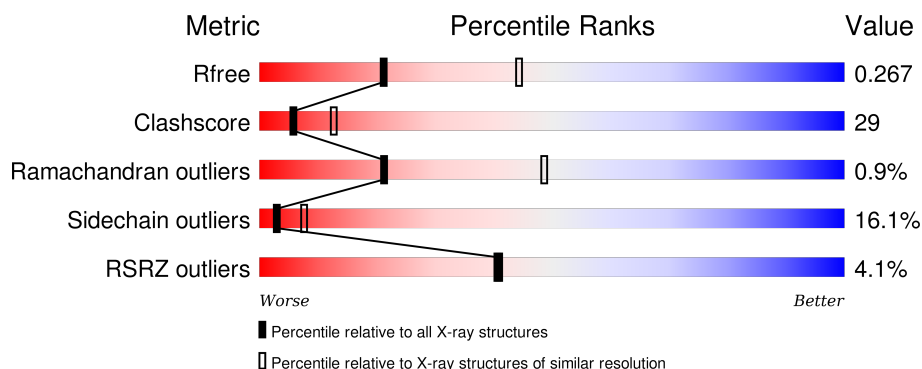
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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	327	<div> <div>2%</div> <div>46% 28% 7% 19%</div> </div>
1	B	327	<div> <div>%</div> <div>46% 28% 6% 19%</div> </div>
1	C	327	<div> <div>5%</div> <div>46% 25% 7% 21%</div> </div>
1	D	327	<div> <div>6%</div> <div>43% 28% 9% 20%</div> </div>
1	E	327	<div> <div>97%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	327	98%
1	G	327	98%
1	H	327	97%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 8885 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.

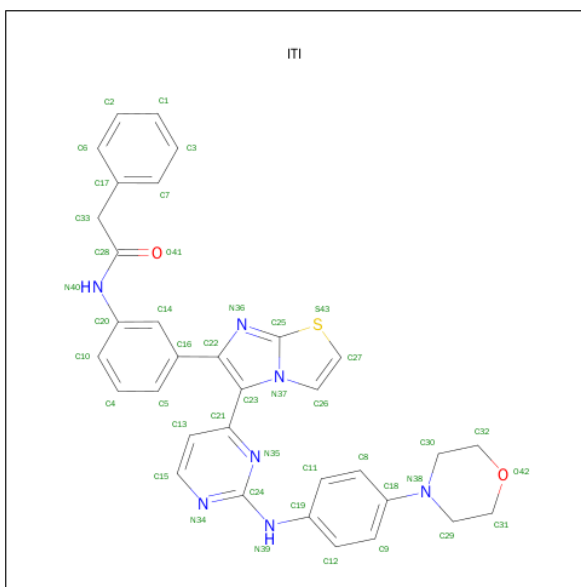
- Molecule 1 is a protein called Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	265	Total	C	N	O	S	0	0	0
			2120	1370	360	375	15			
1	B	265	Total	C	N	O	S	0	0	0
			2116	1368	360	373	15			
1	C	259	Total	C	N	O	S	0	0	0
			2078	1343	354	366	15			
1	D	261	Total	C	N	O	S	0	0	0
			2087	1350	355	367	15			
1	E	9	Total	C	N	O		0	0	0
			45	27	9	9				
1	F	7	Total	C	N	O		0	0	0
			35	21	7	7				
1	G	7	Total	C	N	O		0	0	0
			35	21	7	7				
1	H	9	Total	C	N	O		0	0	0
			45	27	9	9				

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	924	ARG	VAL	ENGINEERED	UNP P00533
B	924	ARG	VAL	ENGINEERED	UNP P00533
C	924	ARG	VAL	ENGINEERED	UNP P00533
D	924	ARG	VAL	ENGINEERED	UNP P00533
E	-29	ARG	VAL	ENGINEERED	UNP P00533
F	-28	ARG	VAL	ENGINEERED	UNP P00533
G	-30	ARG	VAL	ENGINEERED	UNP P00533
H	-31	ARG	VAL	ENGINEERED	UNP P00533

- Molecule 2 is N-[3-(5-{2-[(4-MORPHOLIN-4-YLPHENYL)AMINO]PYRIMIDIN-4-YL}IMI DAZO[2,1-B][1,3]THIAZOL-6-YL)PHENYL]-2-PHENYLACETAMIDE (three-letter code: ITI) (formula: C<sub>33</sub>H<sub>29</sub>N<sub>7</sub>O<sub>2</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			43	33	7	2	1		
2	B	1	Total	C	N	O	S	0	0
			43	33	7	2	1		
2	C	1	Total	C	N	O	S	0	0
			43	33	7	2	1		
2	D	1	Total	C	N	O	S	0	0
			43	33	7	2	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	40	Total	O	0	0
			40	40		
3	B	45	Total	O	0	0
			45	45		
3	C	35	Total	O	0	0
			35	35		
3	D	21	Total	O	0	0
			21	21		
3	E	3	Total	O	0	0
			3	3		
3	F	4	Total	O	0	0
			4	4		
3	G	2	Total	O	0	0
			2	2		
3	H	2	Total	O	0	0
			2	2		





[illegible]

- Molecule 1: Epidermal growth factor receptor

Chain F:  98%

UNK	PRO	VAL	ASN	ASN	GLY
	PRO	PRO	TRP	LYS	GLU
	ILE	ILE	CYS	GLU	ALA
	CYS	LYS	VAL	ILE	PRO
	UNK	TRP	GLN	LEU	ASN
	UNK	MET	ILE	ASP	GLN
	UNK	ALA	ALA	GLU	LEU
	UNK	LEU	LYS	ALA	ALA
	UNK	GLU	GLY	TYR	LEU
	UNK	MET	MET	VAL	ARG
	UNK	ILE	ASN	MET	ILE
	UNK	LEU	TYR	ALA	LEU
	UNK	ARG	LEU	SER	LYS
	UNK	LYS	GLU	SER	LYS
	UNK	CYS	ILE	ASP	THR
	UNK	TRP	TYR	ASN	GLU
	UNK	MET	THR	ARG	PHE
	UNK	ILE	HIS	LEU	LYS
	UNK	ASP	GLN	VAL	LYS
	UNK	ALA	SER	HIS	ILE
	UNK	ASP	ASP	ARG	LYS
	UNK	SER	VAL	ASP	VAL
	UNK	ARG	TRP	LEU	LEU
	UNK	PRO	SER	ALA	GLY
	UNK	LYS	TYR	ALA	ILE
	UNK	PHE	GLY	ARG	CYS
	UNK	ARG	VAL	ASN	LEU
	UNK	GLU	THR	VAL	THR
	UNK	LEU	VAL	LEU	GLY
	UNK	ILE	TRP	VAL	THR
	UNK	ILE	GLU	LYS	VAL
	UNK	GLU	LEU	THR	TYR
	UNK	PHE	MET	PRO	LYS
	UNK	SER	THR	GLN	GLY
	UNK	LYS	PHE	HIS	LEU
	UNK	MET	GLY	VAL	TRP
	UNK	ALA	SER	LYS	ILE
	UNK	ARG	LYS	ILE	PRO
	UNK	ASP	PRO	THR	GLU
	UNK	PRO	TYR	ASP	GLY
UNK	GLN	ASP	PHE	GLY	
UNK	ARG	GLY	PHE	LYS	
UNK	TYR	ILE	GLY	VAL	
UNK	LEU	PRO	ALA	LYS	
UNK	VAL	ALA	LYS	ILE	
UNK	ILE	SER	LEU	PRO	
UNK	GLN	GLU	LEU	VAL	
UNK	GLY	ILE	GLY	ALA	
UNK	X8	SER	ARG	ALA	
UNK	X11	SER	GLU	ILE	
UNK	X12	ILE	GLU	LYS	
UNK	X13	LEU	LYS	ASP	
UNK	X14	GLU	GLU	ARG	
UNK	UNK	GLY	TYR	ILE	
UNK	UNK	UNK	HIS	GLY	
UNK	UNK	ARG	GLN	SER	
UNK	UNK	UNK	GLN	THR	
UNK	UNK	LEU	GLY	SER	
UNK	UNK	PRO	TYR	ALA	
UNK	UNK	GLN	LYS	LYS	

- Molecule 1: Epidermal growth factor receptor

Chain G:  98%

UNK	UNK	PRO	VAL	ASN	ASN	GLY
UNK	UNK	PRO	PRO	TRP	LYS	GLU
UNK	UNK	ILE	ILE	CYS	GLU	ALA
UNK	UNK	CYS	LYS	VAL	ILE	PRO
UNK	UNK	THR	TRP	GLN	LEU	ASN
UNK	UNK	ILE	MET	ILE	ASP	GLN
UNK	UNK	ASP	ALA	ALA	GLU	ALA
UNK	UNK	VAL	LEU	LYS	ALA	LEU
UNK	UNK	TYR	GLU	GLY	TYR	LEU
UNK	UNK	MET	SER	MET	VAL	ARG
UNK	UNK	ILE	ILE	ASN	MET	ILE
UNK	UNK	MET	LEU	TYR	ALA	LEU
UNK	UNK	ARG	HIS	LEU	SER	LYS
UNK	UNK	CYS	ARG	GLU	VAL	GLU
UNK	UNK	LYS	ILE	ASP	ASP	THR
UNK	UNK	TRP	TYR	ARG	ASN	GLU
UNK	UNK	MET	THR	ARG	PRO	PHE
UNK	UNK	ILE	HIS	LEU	HIS	LYS
UNK	UNK	ASP	GLN	VAL	VAL	LYS
UNK	UNK	ALA	SER	HIS	CYS	ILE
UNK	UNK	ASP	ASP	ARG	ARG	LYS
UNK	UNK	SER	VAL	ASP	LEU	VAL
UNK	UNK	ARG	TRP	LEU	LEU	LEU
UNK	UNK	PRO	SER	ALA	GLY	GLY
UNK	UNK	LYS	TYR	ALA	ILE	SER
UNK	UNK	PHE	GLY	ARG	CYS	GLY
UNK	UNK	ARG	VAL	ASN	LEU	ALA
UNK	UNK	GLU	THR	VAL	THR	PHE
UNK	UNK	LEU	VAL	LEU	SER	GLY
UNK	UNK	ILE	TRP	VAL	THR	THR
UNK	UNK	ILE	GLU	LYS	VAL	VAL
UNK	UNK	GLU	LEU	THR	GLN	TYR
UNK	UNK	PHE	MET	PRO	LEU	LYS
UNK	UNK	SER	THR	THR	ILE	GLY
UNK	UNK	LYS	PHE	HIS	ILE	LEU
UNK	UNK	MET	GLY	VAL	GLN	TRP
UNK	UNK	ALA	SER	LYS	LEU	ILE
UNK	UNK	ARG	LYS	ILE	MET	ILE
UNK	UNK	ASP	PRO	THR	PRO	PRO
UNK	UNK	PRO	TYR	ASP	PHE	GLY
UNK	UNK	GLN	ASP	PHE	GLY	GLU
UNK	UNK	ARG	GLY	GLY	CYS	LYS
UNK	UNK	TYR	ILE	LEU	VAL	VAL
UNK	UNK	LEU	PRO	ALA	LEU	LYS
UNK	UNK	VAL	ALA	LYS	ASP	ILE
UNK	UNK	ILE	SER	LEU	TYR	PRO
UNK	UNK	GLY	GLU	LEU	VAL	VAL
UNK	UNK	GLN	ILE	GLY	ARG	ALA
UNK	UNK	X6	SER	ALA	GLU	ILE
UNK	UNK	X7	SER	GLU	HIS	LYS
UNK	UNK	X8	ILE	GLU	LYS	LYS
UNK	UNK	X9	LEU	LYS	ASN	LEU
UNK	UNK	X10	GLU	GLU	ASP	ARG
UNK	UNK	X11	LYS	TYR	ILE	GLU
UNK	UNK	X12	GLY	HIS	GLY	ALA
UNK	UNK	UNK	GLU	ALA	SER	THR
UNK	UNK	UNK	ARG	GLY	GLN	SER
UNK	UNK	UNK	LEU	GLY	TYR	PRO
UNK	UNK	UNK	PRO	GLY	LEU	LYS
UNK	UNK	UNK	GLN	LYS	LEU	ALA

- Molecule 1: Epidermal growth factor receptor

Chain H:  97%



GLY	ASN	ASN	VAL	PRO	PRO	UNK
GLU	LYS	TRP	PRO	ILE	ILE	UNK
ALA	ILE	CYS	VAL	LYS	CYS	UNK
PRO	LEU	GLN	TRP	ILE	THR	UNK
ASN	ASP	ILE	MET	ALA	ILE	UNK
GLN	ALA	ALA	LEU	ASP	VAL	UNK
ALA	GLU	LYS	GLY	TYR	UNK	UNK
LEU	TYR	MET	SER	ILE	UNK	UNK
LEU	VAL	ASN	THR	ILE	UNK	UNK
ARG	ALA	TYR	LEU	ARG	UNK	UNK
ILE	SER	LEU	HIS	THR	UNK	UNK
ILE	VAL	GLU	ARG	ILE	UNK	UNK
LYS	VAL	ASP	ARG	THR	UNK	UNK
GLU	VAL	ASP	ARG	THR	UNK	UNK
THR	ASN	ASP	TYR	ILE	UNK	UNK
GLU	ASN	ARG	THR	THR	UNK	UNK
THR	PRO	ARG	THR	HIS	UNK	UNK
PHE	HIS	LEU	GLN	SER	UNK	UNK
LYS	VAL	VAL	GLN	ASP	UNK	UNK
ILE	CYS	HIS	SER	ALA	UNK	UNK
LYS	ARG	ARG	VAL	THR	UNK	UNK
VAL	ARG	ALA	THR	GLY	UNK	UNK
THR	CYS	ALA	LEU	TYR	UNK	UNK
LYS	ILE	ALA	LEU	LYS	UNK	UNK
GLY	ILE	ALA	GLU	SER	UNK	UNK
SER	ILE	ALA	THR	GLY	UNK	UNK
GLY	CYS	ARG	VAL	VAL	UNK	UNK
ALA	LEU	ASN	THR	GLU	UNK	UNK
PHE	THR	THR	VAL	THR	UNK	UNK
GLY	SER	LEU	VAL	ILE	UNK	UNK
THR	THR	THR	TRP	TRP	UNK	UNK
VAL	VAL	LYS	GLU	ILE	UNK	UNK
THR	GLN	THR	LEU	GLU	UNK	UNK
LYS	LEU	ILE	MET	PHE	UNK	UNK
GLY	THR	THR	THR	SER	UNK	UNK
ILE	GLN	GLN	GLY	THR	UNK	UNK
PRO	LEU	LYS	SER	ALA	UNK	UNK
ILE	MET	ILE	LYS	ARG	UNK	UNK
PRO	PRO	THR	PRO	THR	UNK	UNK
GLU	PRO	THR	TYR	ARG	UNK	UNK
GLY	PHE	ASP	ASP	GLY	UNK	UNK
GLY	GLY	PHE	GLY	ILE	UNK	UNK
LYS	CYS	LEU	PRO	VAL	UNK	UNK
VAL	LEU	ALA	ILE	ILE	UNK	UNK
LYS	LEU	LYS	ALA	SER	UNK	UNK
ILE	ASP	LEU	SER	GLY	UNK	UNK
PRO	THR	GLU	THR	THR	UNK	UNK
ILE	ASP	GLU	GLY	GLY	UNK	UNK
THR	LYS	GLY	LYS	ILE	UNK	UNK
VAL	HIS	GLY	THR	THR	UNK	UNK
LYS	LYS	THR	GLY	GLY	UNK	UNK
GLU	ASN	LYS	GLY	GLY	UNK	UNK
ARG	ILE	ASN	LYS	GLY	UNK	UNK
GLU	GLY	ILE	THR	GLY	UNK	UNK
ALA	GLY	ALA	ALA	GLY	UNK	UNK
THR	SER	ALA	GLY	GLY	UNK	UNK
SER	GLN	GLU	ARG	ARG	UNK	UNK
PRO	TYR	GLY	PRO	PRO	UNK	UNK
LYS	LEU	LYS	GLN	UNK	UNK	UNK

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.59Å 70.87Å 115.18Å 90.00° 109.36° 90.00°	Depositor
Resolution (Å)	44.17 – 2.70 44.16 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.5 (44.17-2.70) 99.7 (44.16-2.70)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.76 (at 2.69Å)	Xtriage
Refinement program	BUSTER-TNT 2.5.1	Depositor
R, $R_{free}$	0.204 , 0.257 0.208 , 0.267	Depositor DCC
$R_{free}$ test set	1801 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	49.4	Xtriage
Anisotropy	0.430	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 91.1	EDS
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 36046 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8885	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.19 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.3626e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.375 respectively for untwinned datasets, and 0.333, 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: ITI

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	0.70	0/2167	0.84	1/2935 (0.0%)
1	B	0.69	1/2163 (0.0%)	0.85	2/2930 (0.1%)
1	C	0.60	0/2123	0.75	1/2873 (0.0%)
1	D	0.55	0/2132	0.75	1/2886 (0.0%)
All	All	0.64	1/8585 (0.0%)	0.80	5/11624 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	750	VAL	CB-CG1	5.56	1.64	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	728	SER	C-N-CD	-14.71	88.23	120.60
1	A	728	SER	C-N-CD	-14.55	88.58	120.60
1	B	736	LEU	CA-CB-CG	5.62	128.21	115.30
1	C	704	LYS	N-CA-C	-5.17	97.03	111.00
1	D	704	LYS	N-CA-C	-5.09	97.26	111.00

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	2120	0	2169	113	0
1	B	2116	0	2165	119	0
1	C	2078	0	2122	115	0
1	D	2087	0	2129	136	0
1	E	45	0	11	6	0
1	F	35	0	10	4	0
1	G	35	0	11	8	0
1	H	45	0	11	6	0
2	A	43	0	29	16	0
2	B	43	0	29	11	0
2	C	43	0	29	13	0
2	D	43	0	29	19	0
3	A	40	0	0	2	0
3	B	45	0	0	8	0
3	C	35	0	0	1	0
3	D	21	0	0	0	0
3	E	3	0	0	1	0
3	F	4	0	0	0	0
3	G	2	0	0	0	0
3	H	2	0	0	0	0
All	All	8885	0	8744	505	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 29.

The worst 5 of 505 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:D:715:LYS:HE3	1:D:715:LYS:H	1.23	1.00
1:C:691:ILE:HG22	1:C:692:LYS:HG2	1.41	1.00
1:D:691:ILE:HG22	1:D:692:LYS:HG2	1.41	0.99
1:B:736:LEU:HD13	1:B:758:LEU:HD11	1.43	0.99
1:C:715:LYS:H	1:C:715:LYS:HE2	1.28	0.98

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	261/327 (80%)	245 (94%)	12 (5%)	4 (2%)	13	32
1	B	261/327 (80%)	247 (95%)	11 (4%)	3 (1%)	17	42
1	C	253/327 (77%)	243 (96%)	9 (4%)	1 (0%)	39	69
1	D	255/327 (78%)	240 (94%)	14 (6%)	1 (0%)	39	69
All	All	1030/1308 (79%)	975 (95%)	46 (4%)	9 (1%)	21	49

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	728	SER
1	A	729	PRO
1	B	728	SER
1	B	729	PRO
1	B	784	ASN

### 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	233/251 (93%)	195 (84%)	38 (16%)	3	7
1	B	232/251 (92%)	199 (86%)	33 (14%)	4	10
1	C	228/251 (91%)	191 (84%)	37 (16%)	3	7
1	D	228/251 (91%)	188 (82%)	40 (18%)	2	6
All	All	921/1004 (92%)	773 (84%)	148 (16%)	3	7

5 of 148 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	925	LYS
1	C	742	MET
1	D	901	SER
1	B	936	LYS
1	C	704	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	781	HIS
1	C	784	ASN
1	D	781	HIS
1	B	781	HIS
1	C	802	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](#)

4 ligands 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
2	ITI	A	1	-	43,49,49	1.13	4 (9%)	54,68,68	1.69	12 (22%)
2	ITI	B	1	-	43,49,49	1.40	3 (6%)	54,68,68	1.64	11 (20%)
2	ITI	C	1	-	43,49,49	1.04	2 (4%)	54,68,68	1.54	9 (16%)
2	ITI	D	1	-	43,49,49	0.96	3 (6%)	54,68,68	1.24	3 (5%)

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
2	ITI	A	1	-	-	0/21/32/32	0/7/7/7
2	ITI	B	1	-	-	0/21/32/32	0/7/7/7
2	ITI	C	1	-	-	1/21/32/32	0/7/7/7
2	ITI	D	1	-	-	0/21/32/32	0/7/7/7

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	ITI	C16-C22	-7.21	1.41	1.49
2	D	1	ITI	C16-C22	-3.99	1.44	1.49
2	A	1	ITI	C16-C22	-3.93	1.44	1.49
2	C	1	ITI	C16-C22	-3.74	1.44	1.49
2	B	1	ITI	C23-C22	-2.54	1.38	1.43

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1	ITI	C32-C30-N38	-4.68	101.77	110.02
2	C	1	ITI	C20-N40-C28	-4.52	118.90	127.47
2	A	1	ITI	C9-C18-N38	-4.30	115.60	121.38
2	C	1	ITI	N34-C24-N35	-3.24	123.21	126.67
2	B	1	ITI	C15-C13-C21	-3.22	114.17	117.26

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1	ITI	C33-C28-N40-C20

There are no ring outliers.

4 monomers are involved in 59 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	ITI	16	0
2	B	1	ITI	11	0
2	C	1	ITI	13	0
2	D	1	ITI	19	0

## 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 ⓘ

### 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	265/327 (81%)	-0.02	7 (2%) 59 59	27, 47, 80, 117	0
1	B	265/327 (81%)	-0.11	2 (0%) 87 88	27, 45, 79, 112	0
1	C	259/327 (79%)	0.24	15 (5%) 26 25	39, 60, 93, 130	0
1	D	261/327 (79%)	0.36	19 (7%) 18 16	41, 64, 97, 136	0
1	E	0/327	-	-	-	-
1	F	0/327	-	-	-	-
1	G	0/327	-	-	-	-
1	H	0/327	-	-	-	-
All	All	1050/2616 (40%)	0.12	43 (4%) 41 41	27, 56, 92, 136	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	728	SER	6.0
1	A	699	PHE	5.5
1	C	735	ILE	5.4
1	B	727	THR	5.4
1	A	726	ALA	5.1

### 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(Å <sup>2</sup> )	Q<0.9
2	ITI	D	1	43/43	0.94	0.18	0.26	14,57,195,300	0
2	ITI	A	1	43/43	0.97	0.15	-0.46	11,36,81,106	0
2	ITI	C	1	43/43	0.96	0.15	-0.83	23,56,210,300	0
2	ITI	B	1	43/43	0.98	0.14	-0.93	6,33,79,236	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.