



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 23, 2023 – 09:13 pm GMT

PDB ID : 8C3J  
Title : Stapled peptide SP2 in complex with humanised RadA mutant HumRadA22  
Authors : Pantelejevs, T.; Hyvonen, M.  
Deposited on : 2022-12-26  
Resolution : 3.02 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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
buster-report : 1.1.7 (2018)  
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) : 2.36

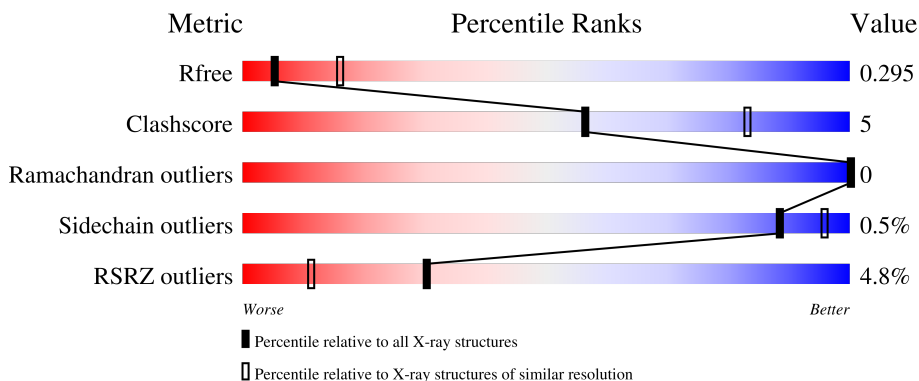
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.02 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	2399 (3.04-3.00)
Clashscore	141614	2734 (3.04-3.00)
Ramachandran outliers	138981	2640 (3.04-3.00)
Sidechain outliers	138945	2643 (3.04-3.00)
RSRZ outliers	127900	2287 (3.04-3.00)

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	231	 3% 75% 15% 10%
1	B	231	 7% 81% 10% 9%
2	C	36	 75% 8% 17%
2	J	36	 81% 17%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3714 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 DNA repair and recombination protein RadA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	207	1620	1016	293	305	6	0	0	0
1	B	210	1640	1029	296	309	6	0	0	0

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	107	MET	ARG	engineered mutation	UNP O74036
A	168	ALA	VAL	engineered mutation	UNP O74036
A	169	MET	ILE	engineered mutation	UNP O74036
A	170	TYR	TRP	engineered mutation	UNP O74036
A	182	LEU	ILE	engineered mutation	UNP O74036
A	198	ASP	LYS	engineered mutation	UNP O74036
A	199	ASN	HIS	engineered mutation	UNP O74036
A	200	VAL	ILE	engineered mutation	UNP O74036
A	201	ALA	TYR	engineered mutation	UNP O74036
A	202	TYR	VAL	engineered mutation	UNP O74036
A	213	GLN	LEU	engineered mutation	UNP O74036
A	215	LEU	VAL	engineered mutation	UNP O74036
A	216	TYR	GLN	engineered mutation	UNP O74036
A	219	SER	GLU	engineered mutation	UNP O74036
A	220	ALA	ASP	engineered mutation	UNP O74036
A	221	MET	LYS	engineered mutation	UNP O74036
A	222	MET	ILE	engineered mutation	UNP O74036
A	223	VAL	LYS	engineered mutation	UNP O74036
A	225	SER	LEU	engineered mutation	UNP O74036
A	232	TYR	VAL	engineered mutation	UNP O74036
A	263	ARG	LYS	engineered mutation	UNP O74036
A	264	PHE	HIS	engineered mutation	UNP O74036
A	266	ARG	ALA	engineered mutation	UNP O74036
A	267	MET	ASP	engineered mutation	UNP O74036
A	274	GLU	LEU	engineered mutation	UNP O74036

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	275	PHE	TYR	engineered mutation	UNP O74036
A	300	ASN	ARG	engineered mutation	UNP O74036
A	?	-	PRO	deletion	UNP O74036
A	?	-	ASP	deletion	UNP O74036
A	?	-	ALA	deletion	UNP O74036
A	?	-	PHE	deletion	UNP O74036
A	?	-	PHE	deletion	UNP O74036
A	?	-	GLY	deletion	UNP O74036
A	?	-	ASP	deletion	UNP O74036
A	?	-	PRO	deletion	UNP O74036
A	?	-	THR	deletion	UNP O74036
A	?	-	ARG	deletion	UNP O74036
A	?	-	PRO	deletion	UNP O74036
A	?	-	ILE	deletion	UNP O74036
B	107	MET	ARG	engineered mutation	UNP O74036
B	168	ALA	VAL	engineered mutation	UNP O74036
B	169	MET	ILE	engineered mutation	UNP O74036
B	170	TYR	TRP	engineered mutation	UNP O74036
B	182	LEU	ILE	engineered mutation	UNP O74036
B	198	ASP	LYS	engineered mutation	UNP O74036
B	199	ASN	HIS	engineered mutation	UNP O74036
B	200	VAL	ILE	engineered mutation	UNP O74036
B	201	ALA	TYR	engineered mutation	UNP O74036
B	202	TYR	VAL	engineered mutation	UNP O74036
B	213	GLN	LEU	engineered mutation	UNP O74036
B	215	LEU	VAL	engineered mutation	UNP O74036
B	216	TYR	GLN	engineered mutation	UNP O74036
B	219	SER	GLU	engineered mutation	UNP O74036
B	220	ALA	ASP	engineered mutation	UNP O74036
B	221	MET	LYS	engineered mutation	UNP O74036
B	222	MET	ILE	engineered mutation	UNP O74036
B	223	VAL	LYS	engineered mutation	UNP O74036
B	225	SER	LEU	engineered mutation	UNP O74036
B	232	TYR	VAL	engineered mutation	UNP O74036
B	263	ARG	LYS	engineered mutation	UNP O74036
B	264	PHE	HIS	engineered mutation	UNP O74036
B	266	ARG	ALA	engineered mutation	UNP O74036
B	267	MET	ASP	engineered mutation	UNP O74036
B	274	GLU	LEU	engineered mutation	UNP O74036
B	275	PHE	TYR	engineered mutation	UNP O74036
B	300	ASN	ARG	engineered mutation	UNP O74036
B	?	-	PRO	deletion	UNP O74036

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ASP	deletion	UNP O74036
B	?	-	ALA	deletion	UNP O74036
B	?	-	PHE	deletion	UNP O74036
B	?	-	PHE	deletion	UNP O74036
B	?	-	GLY	deletion	UNP O74036
B	?	-	ASP	deletion	UNP O74036
B	?	-	PRO	deletion	UNP O74036
B	?	-	THR	deletion	UNP O74036
B	?	-	ARG	deletion	UNP O74036
B	?	-	PRO	deletion	UNP O74036
B	?	-	ILE	deletion	UNP O74036

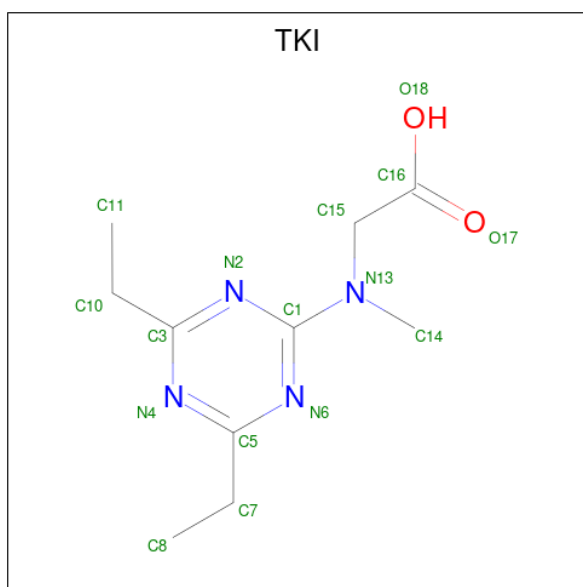
- Molecule 2 is a protein called Breast cancer type 2 susceptibility protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	30	219	139	36	42	2	0	0	0
2	J	30	219	139	36	42	2	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1231	CYS	THR	conflict	UNP P51587
C	1238	CYS	VAL	conflict	UNP P51587
C	2048	GLY	-	linker	UNP P51587
C	2049	SER	-	linker	UNP P51587
J	1231	CYS	THR	conflict	UNP P51587
J	1238	CYS	VAL	conflict	UNP P51587
J	2048	GLY	-	linker	UNP P51587
J	2049	SER	-	linker	UNP P51587

- Molecule 3 is 2-[(4,6-diethyl-1,3,5-triazin-2-yl)-methyl-amino]ethanoic acid (three-letter code: TKI) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>4</sub>O<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).

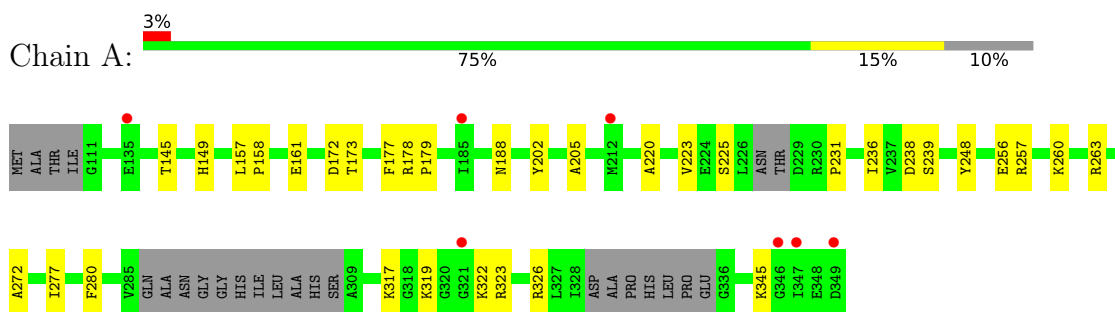


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	C	1	16	10	4	2	0	0

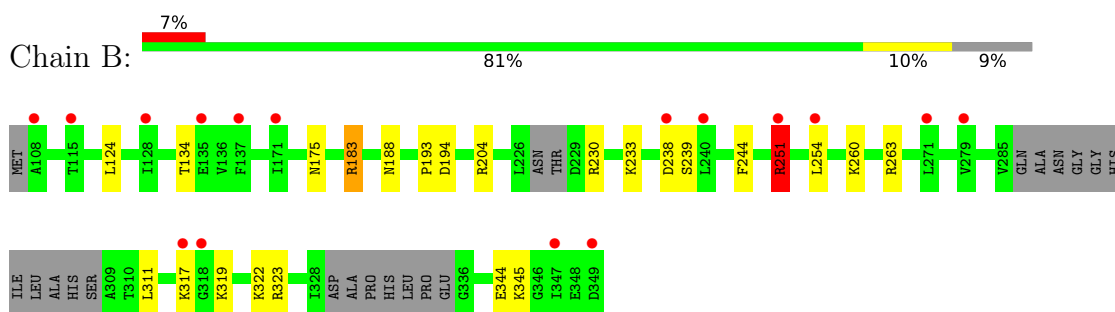
### 3 Residue-property plots [i](#)

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.

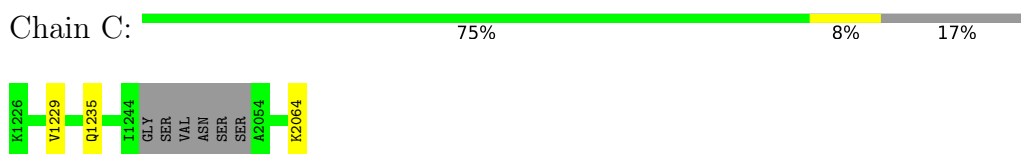
- Molecule 1: DNA repair and recombination protein RadA



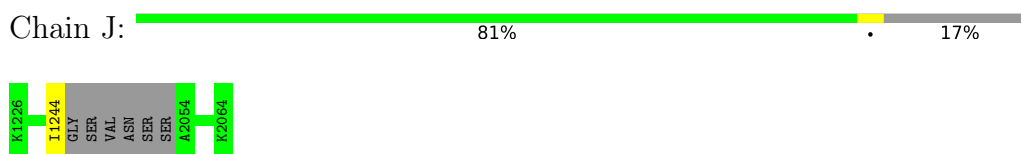
- Molecule 1: DNA repair and recombination protein RadA



- Molecule 2: Breast cancer type 2 susceptibility protein



- Molecule 2: Breast cancer type 2 susceptibility protein



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	112.56Å 112.56Å 140.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.92 – 3.02 87.92 – 3.02	Depositor EDS
% Data completeness (in resolution range)	99.9 (87.92-3.02) 100.0 (87.92-3.02)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.35 (at 3.01Å)	Xtrriage
Refinement program	BUSTER 1.19.2_4158, PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.249 , 0.281 0.251 , 0.295	Depositor DCC
$R_{free}$ test set	897 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	87.2	Xtrriage
Anisotropy	0.305	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 65.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3714	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	91.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.37% 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: TKI

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	0.48	0/1640	0.52	0/2199
1	B	0.52	0/1660	0.56	0/2227
2	C	0.36	0/220	0.57	0/290
2	J	0.37	0/220	0.58	0/290
All	All	0.49	0/3740	0.54	0/5006

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	4
All	All	0	6

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	263	ARG	Sidechain
1	A	326	ARG	Sidechain
1	B	183	ARG	Sidechain
1	B	204	ARG	Sidechain
1	B	230	ARG	Sidechain

## 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	1620	0	1634	21	0
1	B	1640	0	1657	18	2
2	C	219	0	220	3	0
2	J	219	0	222	1	0
3	C	16	0	0	1	0
All	All	3714	0	3733	37	2

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 5.

The worst 5 of 37 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:183:ARG:HG3	1:B:193:PRO:HB2	1.67	0.75
1:A:178:ARG:HH12	1:B:175:ASN:HB2	1.53	0.73
1:A:188:ASN:HB3	1:A:345:LYS:HG2	1.77	0.67
1:B:251:ARG:HD3	1:B:254:LEU:HD22	1.80	0.63
1:B:319:LYS:HB2	1:B:322:LYS:HD3	1.79	0.62

All (2) 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:B:194:ASP:OD1	1:B:251:ARG:NH2[3_555]	1.73	0.47
1:B:194:ASP:CG	1:B:251:ARG:NH2[3_555]	2.12	0.08

## 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	199/231 (86%)	196 (98%)	3 (2%)	0	100	100
1	B	202/231 (87%)	200 (99%)	2 (1%)	0	100	100
2	C	26/36 (72%)	26 (100%)	0	0	100	100
2	J	26/36 (72%)	25 (96%)	1 (4%)	0	100	100
All	All	453/534 (85%)	447 (99%)	6 (1%)	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	170/188 (90%)	170 (100%)	0	100	100
1	B	172/188 (92%)	171 (99%)	1 (1%)	86	95
2	C	24/29 (83%)	23 (96%)	1 (4%)	30	65
2	J	24/29 (83%)	24 (100%)	0	100	100
All	All	390/434 (90%)	388 (100%)	2 (0%)	88	96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	251	ARG
2	C	2064	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	199	ASN
1	B	209	ASN
2	C	1235	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	284	GLN
1	A	273	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

1 ligand is 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	$\# Z  > 2$
3	TKI	C	2101	2	16,16,16	0.70	0	17,21,21	0.79	0

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
3	TKI	C	2101	2	-	3/12/12/12	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

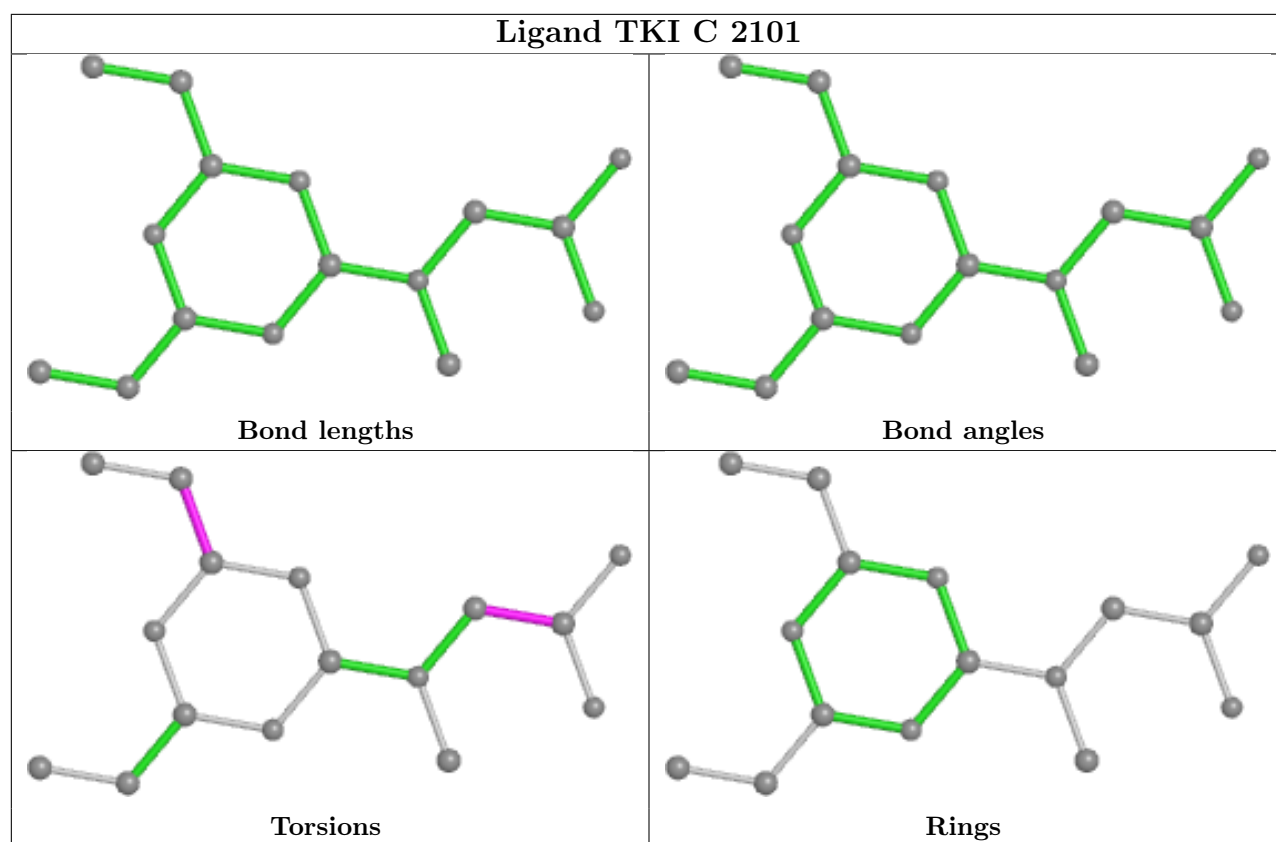
Mol	Chain	Res	Type	Atoms
3	C	2101	TKI	N13-C15-C16-O18
3	C	2101	TKI	N13-C15-C16-O17
3	C	2101	TKI	N6-C5-C7-C8

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	2101	TKI	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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	207/231 (89%)	0.64	7 (3%) 45 19	59, 88, 118, 133	0
1	B	210/231 (90%)	0.81	16 (7%) 13 4	61, 91, 122, 139	0
2	C	30/36 (83%)	0.32	0 100 100	68, 85, 107, 119	0
2	J	30/36 (83%)	0.41	0 100 100	68, 81, 100, 109	0
All	All	477/534 (89%)	0.68	23 (4%) 30 11	59, 88, 120, 139	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	347	ILE	5.8
1	A	349	ASP	4.8
1	B	317	LYS	4.6
1	B	108	ALA	4.3
1	B	137	PHE	3.4

### 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](#)

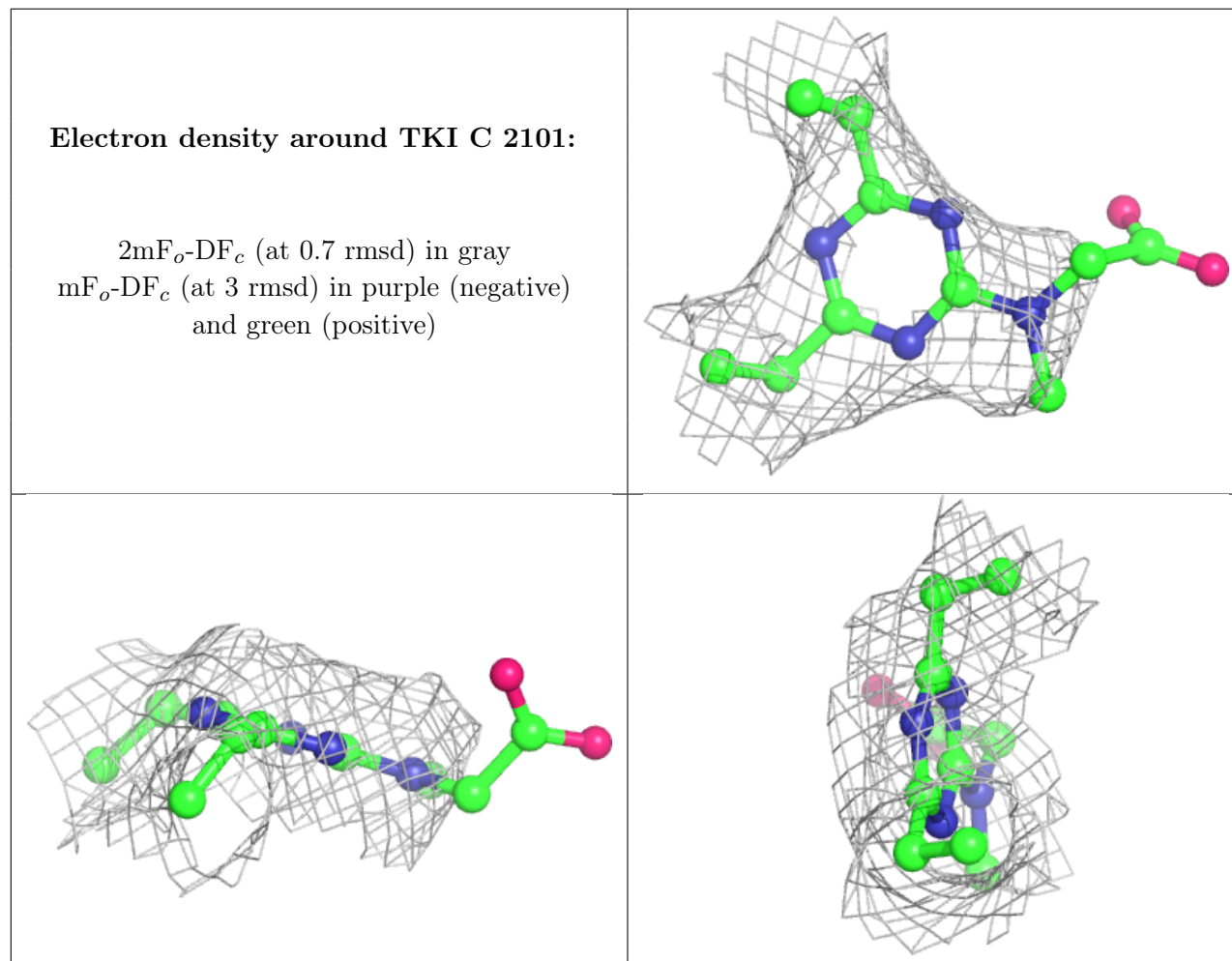
There are no monosaccharides in this entry.

### 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( $\text{\AA}^2$ )	Q<0.9
3	TKI	C	2101	16/16	0.84	0.19	112,129,150,152	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.