



# Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2020 – 10:29 pm BST

PDB ID : 6BDB  
Title : I-OnuI K227Y, D236A bound to A3G substrate (pre-cleavage complex)  
Authors : Brown, C.; Zhang, K.; McMurrough, T.A.; Laforet, M.; Gloor, G.B.; Edgell, D.R.; Junop, M.  
Deposited on : 2017-10-22  
Resolution : 1.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 2.11  
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) : 2.11

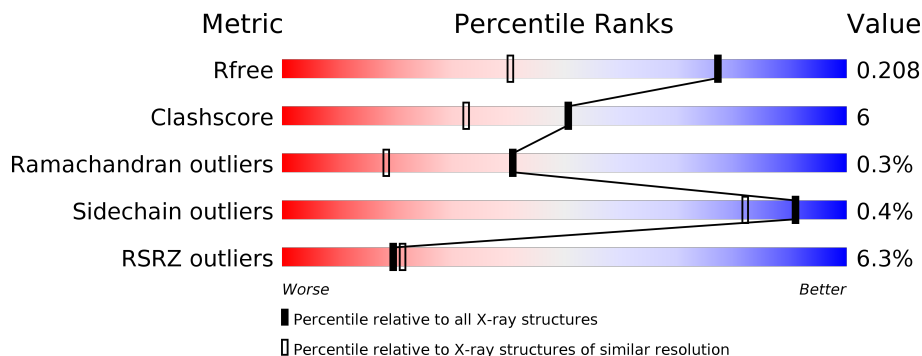
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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 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	307	
2	B	26	
3	C	26	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3983 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 Ribosomal protein 3/homing endonuclease-like protein fusion.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	295	2340	1518	388	425	9	0	2	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q4VWW5
A	-2	PRO	-	expression tag	UNP Q4VWW5
A	-1	LEU	-	expression tag	UNP Q4VWW5
A	0	GLY	-	expression tag	UNP Q4VWW5
A	7	GLU	ARG	conflict	UNP Q4VWW5
A	227	TYR	LYS	engineered mutation	UNP Q4VWW5
A	236	ALA	ASP	engineered mutation	UNP Q4VWW5

- Molecule 2 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	26	513	250	77	161	25	0	0	0

- Molecule 3 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	26	547	259	116	147	25	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	342	Total	O	0	0
			342	342		

*Continued on next page...*

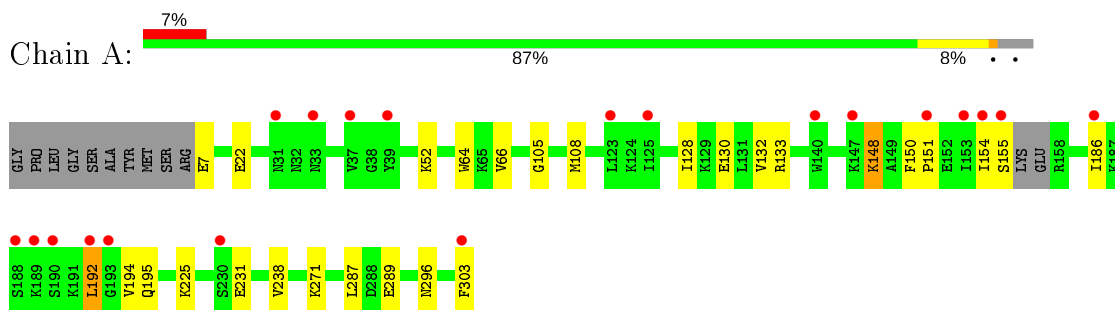
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	B	125	Total 125	O 125	0	0
4	C	116	Total 116	O 116	0	0

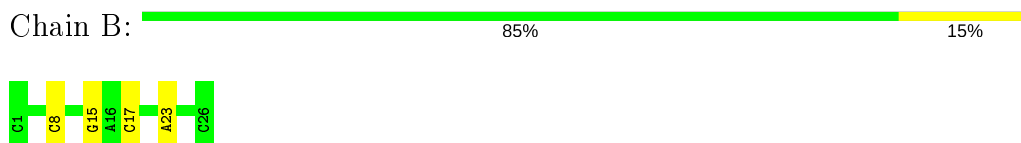
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

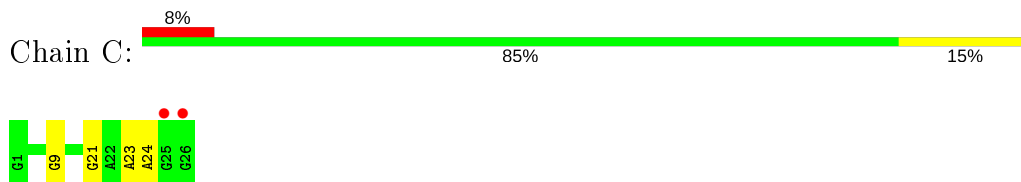
- Molecule 1: Ribosomal protein 3/homing endonuclease-like protein fusion



- Molecule 2: DNA (26-MER)



- Molecule 3: DNA (26-MER)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.37Å 67.19Å 169.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.02 – 1.50 42.02 – 1.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (42.02-1.50) 100.0 (42.02-1.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.12 (at 1.50Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.185 , 0.206 0.191 , 0.208	Depositor DCC
$R_{free}$ test set	4055 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.8	Xtrriage
Anisotropy	0.487	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3983	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.01% 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](#)

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	0.51	0/2388	0.60	0/3221
2	B	0.89	0/569	1.17	3/873 (0.3%)
3	C	0.87	0/619	1.06	2/957 (0.2%)
All	All	0.65	0/3576	0.82	5/5051 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	9	DG	O4'-C1'-N9	8.56	113.99	108.00
2	B	17	DC	O4'-C4'-C3'	-7.93	101.24	106.00
2	B	17	DC	C1'-O4'-C4'	-5.60	104.50	110.10
2	B	8	DC	C1'-O4'-C4'	-5.26	104.84	110.10
3	C	9	DG	C1'-O4'-C4'	-5.01	105.09	110.10

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	2340	0	2351	32	0
2	B	513	0	299	2	0
3	C	547	0	292	2	0
4	A	342	0	0	10	0
4	B	125	0	0	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	116	0	0	1	0
All	All	3983	0	2942	36	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 6.

All (36) 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:192:LEU:HD11	1:A:287:LEU:HD22	1.55	0.85
1:A:22:GLU:OE2	4:A:402:HOH:O	1.97	0.81
1:A:52:LYS:HD3	1:A:303:PHE:CZ	2.22	0.75
3:C:21:DG:OP1	4:C:101:HOH:O	2.08	0.71
1:A:186:ILE:HD12	1:A:195:GLN:CD	2.10	0.71
1:A:186:ILE:HB	1:A:195:GLN:HB3	1.75	0.69
1:A:154:ILE:C	1:A:154:ILE:HD12	2.15	0.67
1:A:7:GLU:N	4:A:405:HOH:O	2.31	0.62
1:A:186:ILE:O	1:A:194:VAL:HA	2.01	0.61
1:A:186:ILE:HD12	1:A:195:GLN:OE1	2.01	0.60
1:A:186:ILE:CD1	1:A:195:GLN:OE1	2.49	0.60
1:A:151:PRO:HD2	4:A:413:HOH:O	2.05	0.56
1:A:148:LYS:N	4:A:407:HOH:O	2.40	0.54
1:A:192:LEU:CD1	1:A:287:LEU:HD22	2.32	0.53
1:A:192:LEU:O	1:A:192:LEU:HD12	2.09	0.52
1:A:130:GLU:OE1	1:A:133:ARG:NH2	2.43	0.51
2:B:15:DG:OP1	4:B:101:HOH:O	2.20	0.50
1:A:64:TRP:HB3	1:A:66:VAL:HG22	1.95	0.47
1:A:128:ILE:O	1:A:132:VAL:HG23	2.15	0.47
1:A:151:PRO:CD	4:A:413:HOH:O	2.62	0.46
1:A:186:ILE:HD13	1:A:195:GLN:OE1	2.16	0.46
1:A:289:GLU:OE1	4:A:403:HOH:O	2.20	0.44
1:A:194:VAL:O	1:A:194:VAL:HG13	2.17	0.44
1:A:231:GLU:N	1:A:231:GLU:OE1	2.49	0.44
1:A:296:ASN:HB3	4:A:596:HOH:O	2.17	0.44
1:A:154:ILE:C	1:A:154:ILE:CD1	2.86	0.44
3:C:23:DA:H2"	3:C:24:DA:H8	1.83	0.44
1:A:271:LYS:NZ	4:A:411:HOH:O	2.50	0.43
1:A:225:LYS:HD3	1:A:238:VAL:HG21	2.00	0.43
1:A:105:GLY:HA2	1:A:108:MET:HE3	2.01	0.42
1:A:151:PRO:N	4:A:413:HOH:O	2.52	0.42
1:A:192:LEU:C	1:A:192:LEU:HD12	2.40	0.42

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:150:PHE:HA	4:A:413:HOH:O	2.19	0.42
2:B:23:DA:H4'	4:B:205:HOH:O	2.20	0.41
1:A:52:LYS:HD3	1:A:303:PHE:CE1	2.54	0.41
1:A:154:ILE:HD12	1:A:155:SER:N	2.35	0.40

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	293/307 (95%)	280 (96%)	12 (4%)	1 (0%)	41 18

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	148	LYS

### 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	255/276 (92%)	254 (100%)	1 (0%)	91 82

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	192	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	119	ASN
1	A	296	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](#)

There are no ligands in this entry.

### 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	295/307 (96%)	0.57	20 (6%) 17 18	12, 22, 49, 65	0
2	B	26/26 (100%)	-0.06	0 100 100	14, 28, 37, 45	0
3	C	26/26 (100%)	0.26	2 (7%) 13 14	15, 23, 48, 53	0
All	All	347/359 (96%)	0.50	22 (6%) 20 21	12, 23, 48, 65	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	154	ILE	11.8
1	A	155	SER	7.5
1	A	192	LEU	6.4
1	A	33	ASN	6.2
1	A	303	PHE	5.3
1	A	190	SER	5.1
1	A	188	SER	5.1
1	A	153	ILE	4.3
1	A	39	TYR	4.1
1	A	186	ILE	4.1
1	A	189	LYS	3.5
1	A	37	VAL	3.2
3	C	26	DG	3.1
1	A	193	GLY	3.1
1	A	151	PRO	3.0
1	A	123	LEU	3.0
1	A	230	SER	2.7
1	A	31	ASN	2.7
1	A	125	ILE	2.7
3	C	25	DG	2.7
1	A	147	LYS	2.6
1	A	140	TRP	2.3

## 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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.