



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 08:27 AM EDT

PDB ID : 5GNW  
Title : Crystal structure of Uracil DNA glycosylase-Uracil complex from *Bradyrhizobium diazoefficiens*.  
Authors : Patil, V.V.; Chembazhi, U.V.; Varshney, U.; Woo, E.  
Deposited on : 2016-07-25  
Resolution : 2.87 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.37.1  
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) : 2.37.1

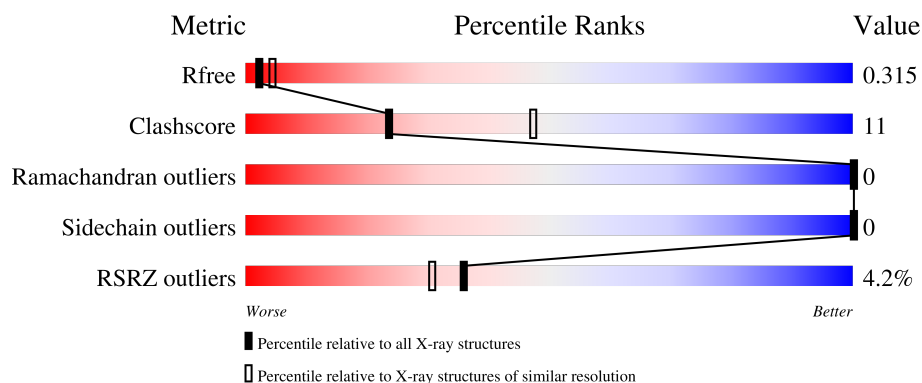
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.87 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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	272	<div> <div>6%</div> <div>87%</div> <div>13%</div> </div>
1	B	272	<div> <div>3%</div> <div>85%</div> <div>14%</div> </div>
1	C	272	<div> <div>6%</div> <div>74%</div> <div>26%</div> </div>
1	D	272	<div> <div>2%</div> <div>79%</div> <div>20%</div> </div>
1	E	272	<div> <div>0%</div> <div>81%</div> <div>18%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	272	
1	G	272	
1	H	272	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	URA	A	301	-	X	-	-
2	URA	B	301	-	X	-	-
2	URA	C	301	-	X	-	-
2	URA	D	301	-	X	-	-
2	URA	E	301	-	X	-	-
2	URA	F	301	-	X	-	-
2	URA	G	301	-	X	-	-
2	URA	H	301	-	X	-	-

## 2 Entry composition

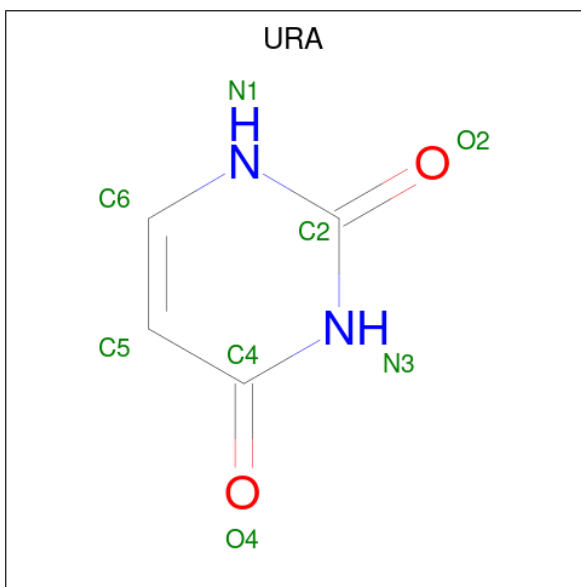
There are 2 unique types of molecules in this entry. The entry contains 32863 atoms, of which 16007 are hydrogens and 0 are deuteriums.

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 Blr0248 protein.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	C	272	Total	C	H	N	O	S	Se	0	0	0
			4082	1344	1983	368	381	1	5			
1	D	272	Total	C	H	N	O	S	Se	0	0	0
			4103	1344	2004	368	381	1	5			
1	E	272	Total	C	H	N	O	S	Se	0	0	0
			4117	1344	2018	368	381	1	5			
1	F	272	Total	C	H	N	O	S	Se	0	0	0
			4110	1344	2011	368	381	1	5			
1	A	272	Total	C	H	N	O	S	Se	0	0	0
			4124	1344	2025	368	381	1	5			
1	B	272	Total	C	H	N	O	S	Se	0	0	0
			4073	1344	1974	368	381	1	5			
1	G	272	Total	C	H	N	O	S	Se	0	0	0
			4084	1344	1985	368	381	1	5			
1	H	272	Total	C	H	N	O	S	Se	0	0	0
			4074	1344	1975	368	381	1	5			

- Molecule 2 is URACIL (three-letter code: URA) (formula: C<sub>4</sub>H<sub>4</sub>N<sub>2</sub>O<sub>2</sub>).

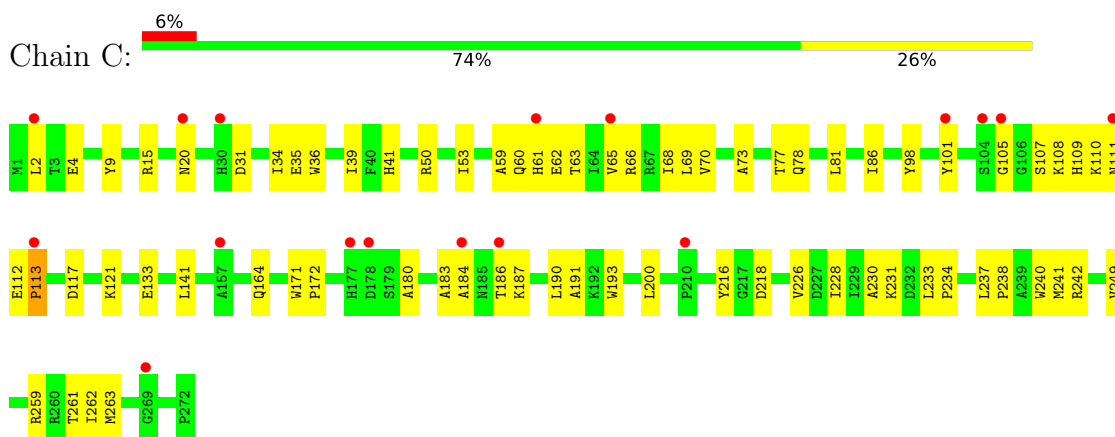


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	H	N	O	0	0
			12	4	4	2	2		
2	D	1	Total	C	H	N	O	0	0
			12	4	4	2	2		
2	E	1	Total	C	H	N	O	0	0
			12	4	4	2	2		
2	F	1	Total	C	H	N	O	0	0
			12	4	4	2	2		
2	A	1	Total	C	H	N	O	0	0
			12	4	4	2	2		
2	B	1	Total	C	H	N	O	0	0
			12	4	4	2	2		
2	G	1	Total	C	H	N	O	0	0
			12	4	4	2	2		
2	H	1	Total	C	H	N	O	0	0
			12	4	4	2	2		

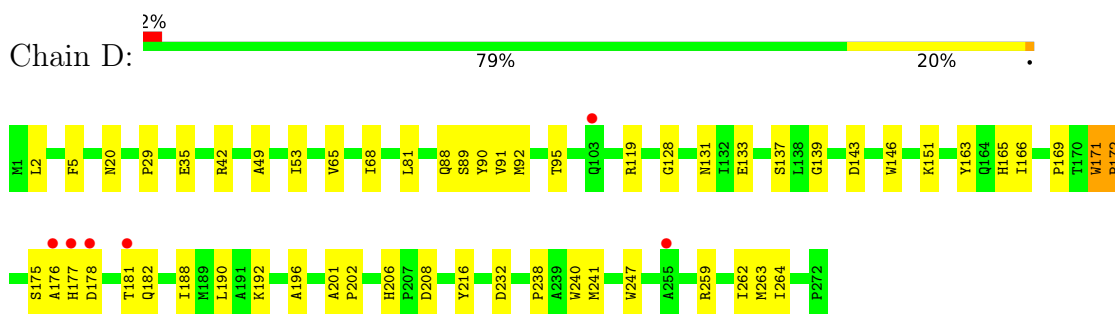
### 3 Residue-property plots [i](#)

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.

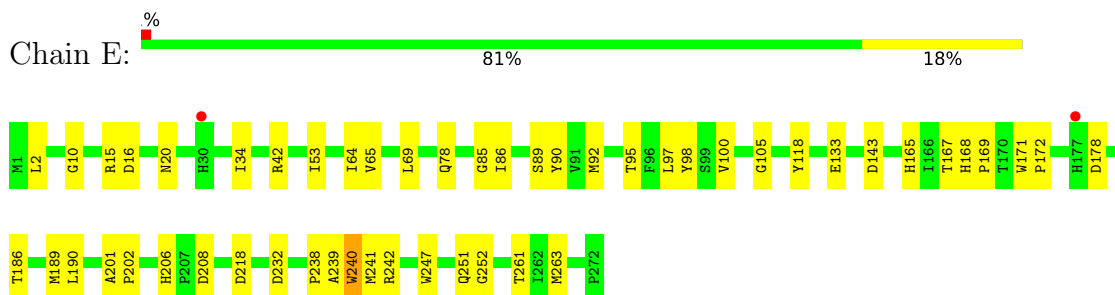
- Molecule 1: Blr0248 protein



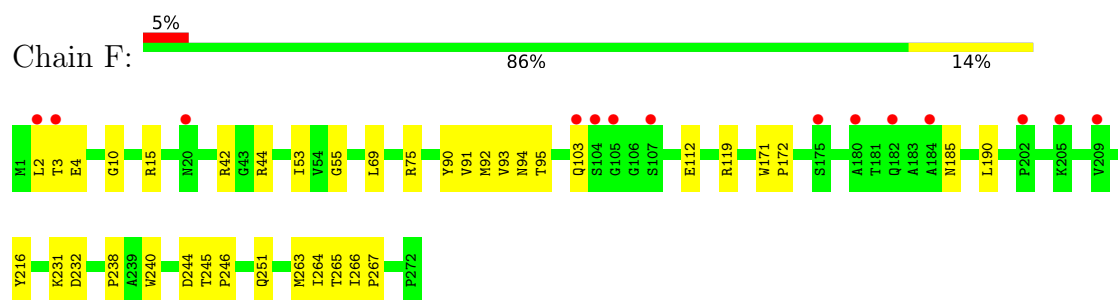
- Molecule 1: Blr0248 protein



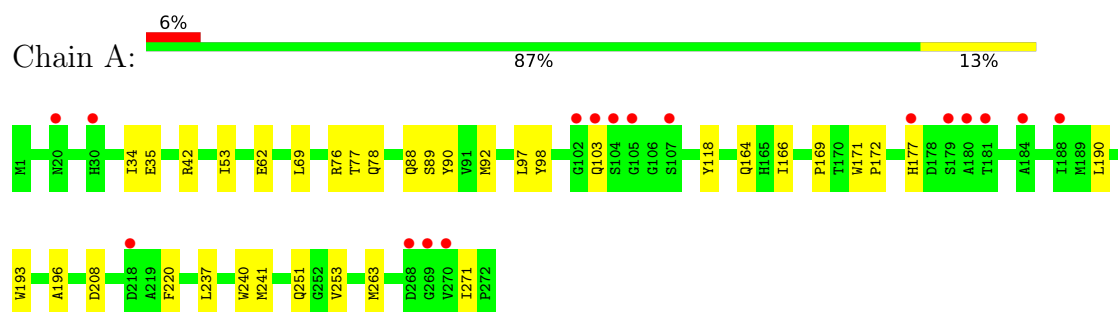
- Molecule 1: Blr0248 protein



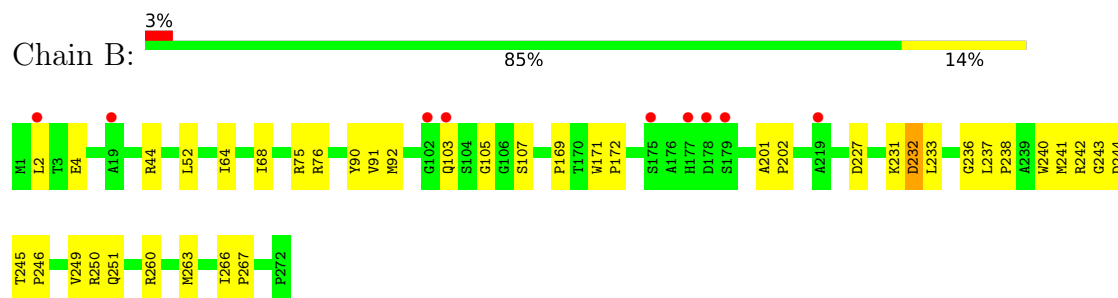
- Molecule 1: Blr0248 protein



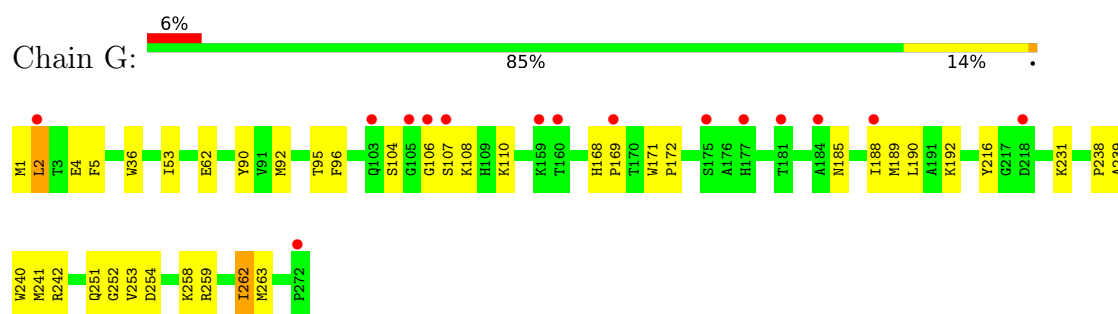
- Molecule 1: Blr0248 protein



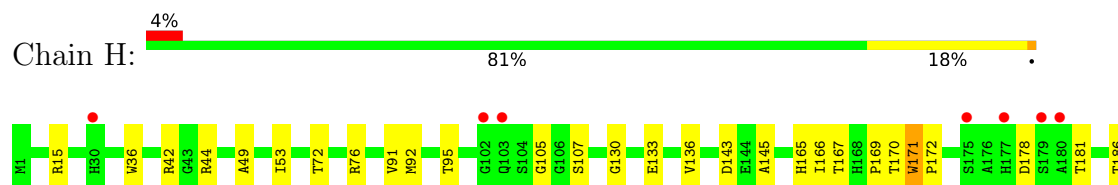
- Molecule 1: Blr0248 protein

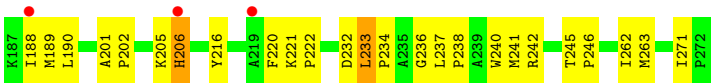


- Molecule 1: Blr0248 protein



- Molecule 1: Blr0248 protein







## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	209.54Å 89.63Å 143.78Å 90.00° 96.20° 90.00°	Depositor
Resolution (Å)	29.62 – 2.87 29.62 – 2.87	Depositor EDS
% Data completeness (in resolution range)	83.2 (29.62-2.87) 83.2 (29.62-2.87)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.62 (at 2.85Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.251 , 0.302 0.272 , 0.315	Depositor DCC
$R_{free}$ test set	1998 reflections (3.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.5	Xtriage
Anisotropy	0.899	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 33.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	32863	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	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 20.52 % 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 8.5452e-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, 0.333 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: URA

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	0.57	1/2155 (0.0%)	0.51	0/2932
1	B	0.58	2/2155 (0.1%)	0.55	3/2932 (0.1%)
1	C	0.59	3/2155 (0.1%)	0.52	1/2932 (0.0%)
1	D	0.85	3/2155 (0.1%)	0.62	2/2932 (0.1%)
1	E	0.64	2/2155 (0.1%)	0.51	0/2932
1	F	0.54	3/2155 (0.1%)	0.55	2/2932 (0.1%)
1	G	0.75	2/2155 (0.1%)	0.58	2/2932 (0.1%)
1	H	0.72	3/2155 (0.1%)	0.64	6/2932 (0.2%)
All	All	0.66	19/17240 (0.1%)	0.56	16/23456 (0.1%)

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	169	PRO	N-CD	-22.18	1.16	1.47
1	D	172	PRO	N-CD	-11.35	1.31	1.47
1	C	35	GLU	CD-OE1	-6.70	1.18	1.25
1	E	240	TRP	CZ3-CH2	-5.96	1.30	1.40
1	C	35	GLU	CD-OE2	-5.69	1.19	1.25

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	2	LEU	CA-CB-CG	7.60	132.79	115.30
1	G	262	ILE	CG1-CB-CG2	-6.69	96.69	111.40
1	H	206	HIS	N-CA-C	-6.12	94.47	111.00
1	H	245	THR	C-N-CD	6.12	141.25	128.40
1	C	112	GLU	C-N-CD	5.95	140.90	128.40

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	2099	2025	2072	39	0
1	B	2099	1974	2072	43	0
1	C	2099	1983	2072	74	0
1	D	2099	2004	2072	70	0
1	E	2099	2018	2072	42	0
1	F	2099	2011	2072	45	0
1	G	2099	1985	2072	49	0
1	H	2099	1975	2072	59	0
2	A	8	4	3	2	0
2	B	8	4	3	0	0
2	C	8	4	3	0	0
2	D	8	4	3	0	0
2	E	8	4	3	0	0
2	F	8	4	3	1	0
2	G	8	4	3	1	0
2	H	8	4	3	0	0
All	All	16856	16007	16600	376	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 11.

The worst 5 of 376 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:H:53:ILE:HB	1:H:92:MSE:HG2	1.23	1.10
1:C:61:HIS:O	1:C:70:VAL:CG2	2.06	1.02
1:D:53:ILE:HB	1:D:92:MSE:HG2	1.41	1.01
1:A:53:ILE:HB	1:A:92:MSE:HG2	1.44	0.97
1:G:53:ILE:HB	1:G:92:MSE:HG2	1.42	0.97

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	270/272 (99%)	256 (95%)	14 (5%)	0	100	100
1	B	270/272 (99%)	254 (94%)	16 (6%)	0	100	100
1	C	270/272 (99%)	254 (94%)	16 (6%)	0	100	100
1	D	270/272 (99%)	257 (95%)	13 (5%)	0	100	100
1	E	270/272 (99%)	257 (95%)	13 (5%)	0	100	100
1	F	270/272 (99%)	254 (94%)	16 (6%)	0	100	100
1	G	270/272 (99%)	254 (94%)	16 (6%)	0	100	100
1	H	270/272 (99%)	250 (93%)	20 (7%)	0	100	100
All	All	2160/2176 (99%)	2036 (94%)	124 (6%)	0	100	100

There are no Ramachandran outliers to report.

### 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	213/208 (102%)	213 (100%)	0	100	100
1	B	213/208 (102%)	213 (100%)	0	100	100
1	C	213/208 (102%)	213 (100%)	0	100	100
1	D	213/208 (102%)	213 (100%)	0	100	100
1	E	213/208 (102%)	213 (100%)	0	100	100
1	F	213/208 (102%)	213 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	213/208 (102%)	213 (100%)	0	100	100
1	H	213/208 (102%)	213 (100%)	0	100	100
All	All	1704/1664 (102%)	1704 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	165	HIS
1	A	251	GLN
1	B	165	HIS
1	G	185	ASN
1	G	251	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](#)

8 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	URA	H	301	-	8,8,8	1.72	3 (37%)	9,10,10	3.22	6 (66%)
2	URA	E	301	-	8,8,8	1.72	3 (37%)	9,10,10	3.22	6 (66%)
2	URA	A	301	-	8,8,8	1.71	3 (37%)	9,10,10	3.23	6 (66%)
2	URA	C	301	-	8,8,8	1.72	3 (37%)	9,10,10	3.23	6 (66%)
2	URA	G	301	-	8,8,8	1.71	3 (37%)	9,10,10	3.24	6 (66%)
2	URA	B	301	-	8,8,8	1.71	3 (37%)	9,10,10	3.25	6 (66%)
2	URA	F	301	-	8,8,8	1.73	3 (37%)	9,10,10	3.23	6 (66%)
2	URA	D	301	-	8,8,8	1.70	2 (25%)	9,10,10	3.24	6 (66%)

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	URA	H	301	-	-	-	0/1/1/1
2	URA	E	301	-	-	-	0/1/1/1
2	URA	A	301	-	-	-	0/1/1/1
2	URA	C	301	-	-	-	0/1/1/1
2	URA	G	301	-	-	-	0/1/1/1
2	URA	B	301	-	-	-	0/1/1/1
2	URA	F	301	-	-	-	0/1/1/1
2	URA	D	301	-	-	-	0/1/1/1

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	301	URA	C4-N3	-2.50	1.34	1.38
2	B	301	URA	C4-N3	-2.47	1.34	1.38
2	H	301	URA	C4-N3	-2.46	1.34	1.38
2	C	301	URA	C4-N3	-2.46	1.34	1.38
2	A	301	URA	C4-N3	-2.46	1.34	1.38

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	301	URA	N1-C2-N3	5.33	121.17	115.13
2	B	301	URA	N1-C2-N3	5.32	121.16	115.13
2	C	301	URA	N1-C2-N3	5.31	121.15	115.13
2	D	301	URA	N1-C2-N3	5.31	121.14	115.13

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	301	URA	N1-C2-N3	5.31	121.14	115.13

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	URA	2	0
2	G	301	URA	1	0
2	F	301	URA	1	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 [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	267/272 (98%)	0.41	17 (6%) 19 15	12, 25, 49, 74	0
1	B	267/272 (98%)	0.36	9 (3%) 45 39	11, 22, 38, 65	0
1	C	267/272 (98%)	0.56	17 (6%) 19 15	11, 26, 44, 66	0
1	D	267/272 (98%)	0.23	6 (2%) 62 59	8, 18, 35, 72	0
1	E	267/272 (98%)	0.15	2 (0%) 87 87	5, 14, 33, 60	0
1	F	267/272 (98%)	0.55	14 (5%) 27 22	9, 28, 45, 73	0
1	G	267/272 (98%)	0.43	15 (5%) 24 20	10, 24, 50, 74	0
1	H	267/272 (98%)	0.42	10 (3%) 41 36	12, 27, 42, 85	0
All	All	2136/2176 (98%)	0.39	90 (4%) 36 31	5, 24, 43, 85	0

The worst 5 of 90 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	103	GLN	5.7
1	B	2	LEU	5.5
1	C	2	LEU	5.1
1	B	102	GLY	5.0
1	A	104	SER	5.0

### 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
2	URA	D	301	8/8	0.89	0.27	8,11,18,22	0
2	URA	F	301	8/8	0.90	0.19	20,26,32,36	0
2	URA	A	301	8/8	0.91	0.21	18,22,29,30	0
2	URA	C	301	8/8	0.93	0.24	11,16,24,25	0
2	URA	B	301	8/8	0.93	0.24	12,17,21,25	0
2	URA	H	301	8/8	0.93	0.19	12,21,26,27	0
2	URA	E	301	8/8	0.94	0.22	7,13,20,22	0
2	URA	G	301	8/8	0.95	0.23	9,14,18,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.