



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2021 – 12:51 PM EDT

PDB ID : 2H8C
Title : Structure of RusA D70N in complex with DNA
Authors : Macmaster, R.A.
Deposited on : 2006-06-07
Resolution : 3.10 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

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	:	2.23.2
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)	:	2.23.2

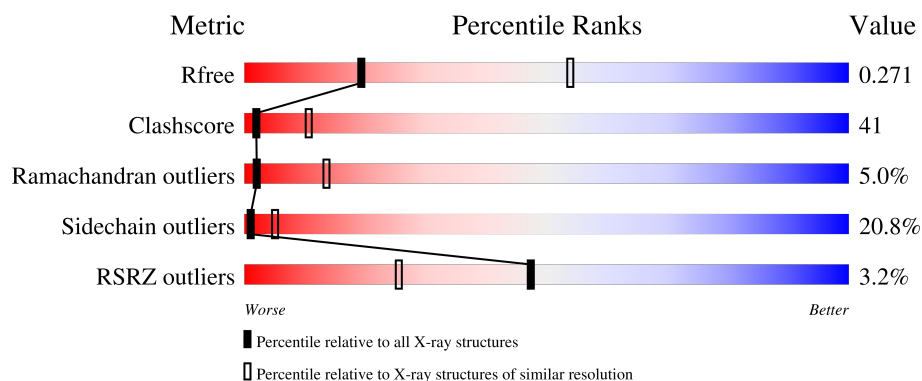
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 ≥ 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	W	11	
1	X	11	
1	Y	11	
1	Z	11	
2	A	120	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	120	
2	C	120	
2	D	120	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4357 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 DNA chain called 5'-D(*CP*CP*GP*GP*TP*AP*CP*CP*GP*GP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Y	11	Total	C	N	O	P	0	0	0
			222	106	41	65	10			
1	Z	10	Total	C	N	O	P	0	0	0
			202	96	39	58	9			
1	W	11	Total	C	N	O	P	0	0	0
			222	106	41	65	10			
1	X	11	Total	C	N	O	P	0	0	0
			222	106	41	65	10			

- Molecule 2 is a protein called Crossover junction endodeoxyribonuclease rusA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	117	Total	C	N	O	S	0	3	0
			927	583	171	166	7			
2	B	116	Total	C	N	O	S	0	1	0
			885	559	158	162	6			
2	C	111	Total	C	N	O	S	0	1	0
			830	525	144	155	6			
2	D	112	Total	C	N	O	S	0	2	0
			845	537	143	159	6			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	70	ASN	ASP	engineered mutation	UNP P0AG74
B	70	ASN	ASP	engineered mutation	UNP P0AG74
C	70	ASN	ASP	engineered mutation	UNP P0AG74
D	70	ASN	ASP	engineered mutation	UNP P0AG74

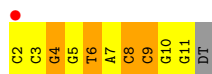
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total 1	O 1	0	0
3	D	1	Total 1	O 1	0	0

- Molecule 1: 5'-D(*CP*CP*GP*GP*TP*AP*CP*CP*GP*GP*T)-3'



- Chain Z:  9% 55% 36% 9%



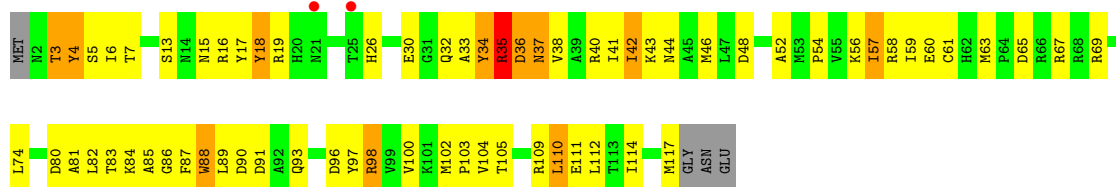
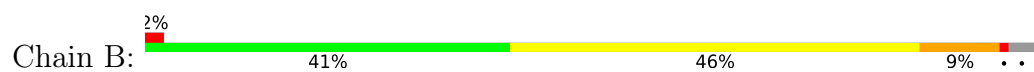
- | |
|-----|
| C2 |
| C3 |
| G4 |
| G5 |
| T6 |
| A7 |
| C8 |
| C9 |
| G10 |
| G11 |
| T12 |

- | |
|-----|
| C2 |
| C3 |
| G4 |
| G5 |
| T6 |
| A7 |
| C8 |
| C9 |
| G10 |
| G11 |
| T12 |

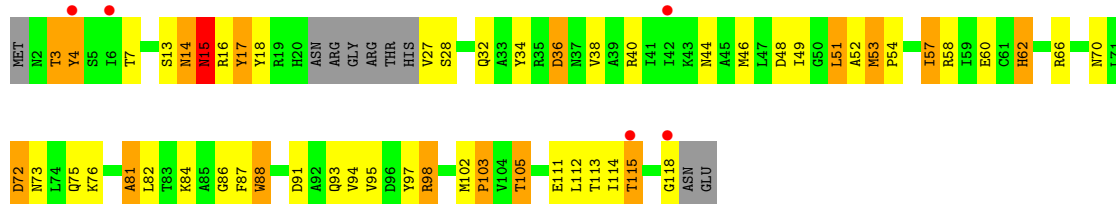
- | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | N2 | T3 | I6 | W10 | P11 | P12 | S13 | N14 | N15 | R16 | Y17 | Y18 | N21 | R22 | G23 | R24 | T25 | G31 | Y34 | R35 | V38 | A39 | R40 | N44 | A45 | M46 | L47 | D48 | I49 | G50 | L51 | A52 | M53 | K56 | I57 | R58 | I59 | E60 | G61 | G62 | M63 | P64 | D65 | R69 | L74 | K75 | Q76 | F70 |
|-----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|



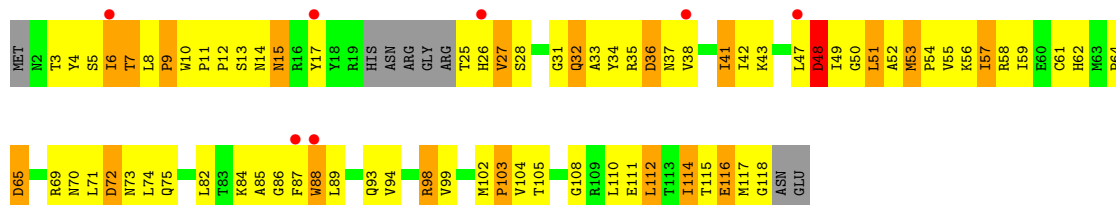
- 
- WORLD WIDE
PDB
PROTEIN DATA BANK



• Molecule 2: Crossover junction endodeoxyribonuclease *rusA*



• Molecule 2: Crossover junction endodeoxyribonuclease *rusA*



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	64.62Å 59.48Å 90.72Å 90.00° 101.59° 90.00°	Depositor
Resolution (Å)	44.46 – 3.10 44.44 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.46-3.10) 100.0 (44.44-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.59 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.247 , 0.289 0.236 , 0.271	Depositor DCC
R_{free} test set	601 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å ²)	90.2	Xtriage
Anisotropy	0.146	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 122.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	4357	wwPDB-VP
Average B, all atoms (Å ²)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.41% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	W	0.56	0/248	1.44	2/381 (0.5%)
1	X	0.63	0/248	1.55	6/381 (1.6%)
1	Y	0.59	0/248	1.44	5/381 (1.3%)
1	Z	0.64	0/226	1.57	5/347 (1.4%)
2	A	0.28	0/951	0.41	0/1289
2	B	0.25	0/903	0.42	0/1231
2	C	0.25	0/847	0.41	0/1156
2	D	0.25	0/861	0.41	0/1173
All	All	0.36	0/4532	0.81	18/6339 (0.3%)

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
2	B	0	1
2	C	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	Y	3	DC	O4'-C4'-C3'	-8.15	101.11	106.00
1	Z	6	DT	N3-C2-O2	-8.11	117.44	122.30
1	X	8	DC	O4'-C1'-N1	8.00	113.60	108.00
1	Z	9	DC	C1'-O4'-C4'	-6.71	103.39	110.10
1	X	7	DA	P-O3'-C3'	6.33	127.30	119.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	19	ARG	Peptide
2	C	48	ASP	Peptide

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	W	222	0	125	22	0
1	X	222	0	125	18	1
1	Y	222	0	125	19	0
1	Z	202	0	113	22	0
2	A	927	0	914	58	1
2	B	885	0	846	62	0
2	C	830	0	751	66	0
2	D	845	0	790	75	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
All	All	4357	0	3789	336	1

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 41.

The worst 5 of 336 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:B:90:ASP:O	2:B:93:GLN:HG3	1.39	1.21
1:Z:9:DC:H2''	1:Z:10:DG:C8	1.83	1.12
2:B:37:ASN:HA	2:B:40:ARG:HE	1.07	1.12
2:A:6:ILE:HD11	2:A:47:LEU:HD12	1.17	1.09
2:D:49:ILE:HG13	2:D:50:GLY:H	1.13	1.07

All (1) 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:X:3:DC:OP1	2:A:109[B]:ARG:NE[2_555]	2.12	0.08

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	
2	A	117/120 (98%)	93 (80%)	19 (16%)	5 (4%)	2	16
2	B	114/120 (95%)	84 (74%)	27 (24%)	3 (3%)	5	26
2	C	107/120 (89%)	78 (73%)	23 (22%)	6 (6%)	2	11
2	D	108/120 (90%)	86 (80%)	14 (13%)	8 (7%)	1	6
All	All	446/480 (93%)	341 (76%)	83 (19%)	22 (5%)	2	14

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	117	MET
2	C	14	ASN
2	C	15	ASN
2	C	53	MET
2	D	9	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	
2	A	95/103 (92%)	81 (85%)	14 (15%)	3	13
2	B	88/103 (85%)	72 (82%)	16 (18%)	1	7
2	C	79/103 (77%)	61 (77%)	18 (23%)	1	3
2	D	82/103 (80%)	59 (72%)	23 (28%)	0	1
All	All	344/412 (84%)	273 (79%)	71 (21%)	1	5

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	48	ASP
2	D	53	MET
2	D	104	VAL
2	B	58	ARG
2	B	57	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	62	HIS
2	B	93	GLN
2	D	73	ASN
2	D	32	GLN
2	D	70	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](#)

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 ⓘ

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, 95th 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 ²)	Q<0.9
1	W	11/11 (100%)	0.26	0 100 100	93, 93, 93, 93	0
1	X	11/11 (100%)	0.43	0 100 100	93, 93, 93, 93	0
1	Y	11/11 (100%)	0.70	0 100 100	93, 93, 93, 93	0
1	Z	10/11 (90%)	1.11	1 (10%) 7 2	93, 93, 93, 93	0
2	A	117/120 (97%)	-0.07	1 (0%) 84 69	93, 93, 93, 93	1 (0%)
2	B	116/120 (96%)	0.07	2 (1%) 70 49	93, 93, 93, 93	1 (0%)
2	C	111/120 (92%)	0.02	5 (4%) 33 16	93, 93, 93, 93	1 (0%)
2	D	112/120 (93%)	0.24	7 (6%) 20 8	93, 93, 93, 93	0
All	All	499/524 (95%)	0.11	16 (3%) 47 25	93, 93, 93, 93	3 (0%)

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	88	TRP	6.4
2	D	87	PHE	4.6
2	B	21	ASN	3.9
2	B	25	THR	3.3
2	D	38	VAL	3.2

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.