



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 04:47 am BST

PDB ID : 1XNS
Title : Peptide trapped Holliday junction intermediate in Cre-loxP recombination
Authors : Ghosh, K.; Lau, C.K.; Guo, F.; Segall, A.M.; Van Duyne, G.D.
Deposited on : 2004-10-05
Resolution : 2.80 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	:	2.11
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

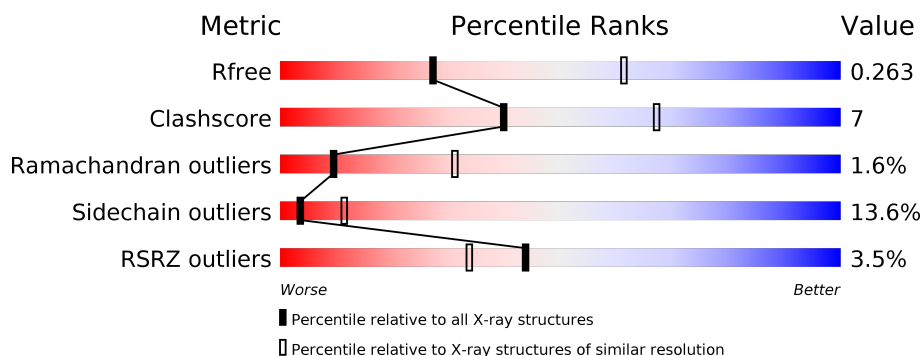
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 ≥ 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	C	35	<div> <div>11%</div> <div> <div>23%</div> <div>51%</div> <div>23%</div> </div> <div>•</div> </div>
2	D	34	<div> <div>21%</div> <div>47%</div> <div>32%</div> </div>
3	A	324	<div> <div>4%</div> <div> <div>71%</div> <div>24%</div> <div>5%</div> </div> <div>•</div> </div>
3	B	324	<div> <div>2%</div> <div> <div>76%</div> <div>19%</div> </div> <div>• •</div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6807 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 loxP DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	35	Total	C	N	O	P	1	0	0
			715	346	125	210	34			

- Molecule 2 is a DNA chain called loxP DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	34	Total	C	N	O	P	1	0	0
			696	335	124	203	34			

- Molecule 3 is a protein called Recombinase CRE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	322	Total	C	N	O	S	0	0	0
			2550	1584	486	465	15			
3	B	322	Total	C	N	O	S	0	0	0
			2550	1584	486	465	15			

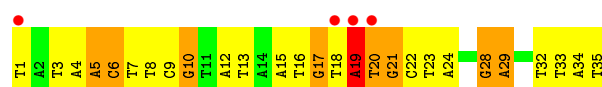
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	35	Total	O	0	0
			35	35		
4	D	36	Total	O	0	0
			36	36		
4	A	74	Total	O	0	0
			74	74		
4	B	151	Total	O	0	0
			151	151		

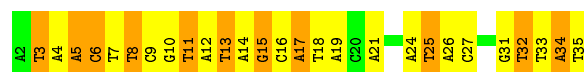
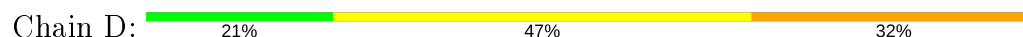
3 Residue-property plots

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.

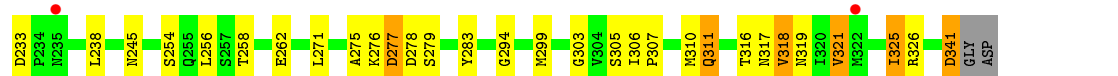
• Molecule 1: loxP DNA



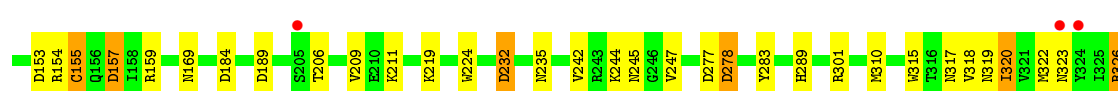
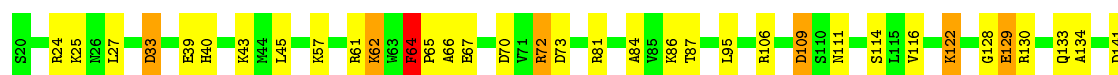
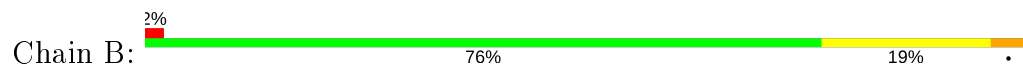
• Molecule 2: loxP DNA



• Molecule 3: Recombinase CRE



• Molecule 3: Recombinase CRE



R327				
L328				
D329				
S330				
E331				
T332				
R337				
L338				
L339				
F340				
D341				
G1Y				
ASP				

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	106.08 Å 121.53 Å 177.80 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.67 – 2.80 47.60 – 2.80	Depositor EDS
% Data completeness (in resolution range)	96.2 (47.67-2.80) 96.2 (47.60-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.18 (at 2.81 Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.196 , 0.265 0.200 , 0.263	Depositor DCC
R_{free} test set	1386 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	53.8	Xtriage
Anisotropy	0.174	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 59.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	6807	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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 ⓘ

5.1 Standard geometry ⓘ

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	C	1.82	6/801 (0.7%)	2.55	41/1235 (3.3%)
2	D	1.70	2/780 (0.3%)	2.20	40/1201 (3.3%)
3	A	0.60	0/2591	0.85	13/3493 (0.4%)
3	B	0.66	0/2591	0.89	14/3493 (0.4%)
All	All	1.01	8/6763 (0.1%)	1.42	108/9422 (1.1%)

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	C	0	1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	18	DT	C5-C7	32.01	1.69	1.50
2	D	18	DT	C5-C7	-28.25	1.33	1.50
1	C	19	DA	C1'-N9	-7.84	1.36	1.47
1	C	23	DT	C3'-O3'	-7.72	1.33	1.44
2	D	10	DG	C3'-O3'	-7.50	1.34	1.44

The worst 5 of 108 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	18	DT	C6-C5-C7	-30.18	104.80	122.90
1	C	19	DA	O4'-C1'-N9	28.67	128.07	108.00
1	C	18	DT	C4-C5-C7	28.57	136.15	119.00
2	D	17	DA	O4'-C1'-N9	-16.14	96.70	108.00
2	D	13	DT	O4'-C1'-N1	12.76	116.93	108.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	19	DA	Sidechain

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	C	715	0	401	10	0
2	D	696	0	387	13	0
3	A	2550	0	2571	37	0
3	B	2550	0	2571	37	0
4	A	74	0	0	3	0
4	B	151	0	0	14	1
4	C	35	0	0	2	1
4	D	36	0	0	5	0
All	All	6807	0	5930	90	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 7.

The worst 5 of 90 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:C:19:DA:H2''	1:C:20:DT:O5'	1.35	1.09
3:B:67:GLU:HG3	4:B:436:HOH:O	1.62	0.97
3:A:311:GLN:HG3	4:B:487:HOH:O	1.67	0.93
3:B:64:PHE:O	3:B:65:PRO:C	2.12	0.85
3:A:31:PHE:HB2	4:A:380:HOH:O	1.79	0.83

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)
4:C:39:HOH:O	4:B:365:HOH:O[4_555]	1.80	0.40

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	
3	A	320/324 (99%)	277 (87%)	38 (12%)	5 (2%)	9	31
3	B	320/324 (99%)	305 (95%)	10 (3%)	5 (2%)	9	31
All	All	640/648 (99%)	582 (91%)	48 (8%)	10 (2%)	9	31

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	317	ASN
3	B	64	PHE
3	B	329	ASP
3	A	232	ASP
3	A	275	ALA

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	
3	A	269/270 (100%)	223 (83%)	46 (17%)	2	6
3	B	269/270 (100%)	242 (90%)	27 (10%)	7	22
All	All	538/540 (100%)	465 (86%)	73 (14%)	3	11

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	276	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	318	VAL
3	B	289	HIS
3	A	305	SER
3	A	321	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	89	GLN
3	B	281	GLN
3	B	111	ASN
3	A	245	ASN
3	B	133	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 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 [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	C	35/35 (100%)	-0.00	4 (11%) 5 3	24, 44, 139, 182	0
2	D	34/34 (100%)	-0.52	0 100 100	26, 41, 91, 111	0
3	A	322/324 (99%)	0.07	13 (4%) 38 28	33, 60, 107, 122	0
3	B	322/324 (99%)	-0.21	8 (2%) 57 47	25, 43, 91, 121	0
All	All	713/717 (99%)	-0.09	25 (3%) 44 34	24, 51, 103, 182	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	19	DA	5.2
3	A	198	GLY	4.9
3	A	199	ARG	4.3
1	C	18	DT	3.8
3	A	207	ALA	3.5

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.