wwPDB X-ray Structure Validation Summary Report (i)

Aug 15, 2023 – 11:13 AM EDT

PDB ID	:	1XO0
Title	:	High resolution structure of the holliday junction intermediate in cre-loxp site-specific recombination
Authors	:	Ghosh, K.; Lau, C.K.; Guo, F.; Segall, A.M.; Van Duyne, G.D.
Deposited on	:	2004-10-05
Resolution	:	2.00 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 (i) 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 (1)) were used in the production of this report:

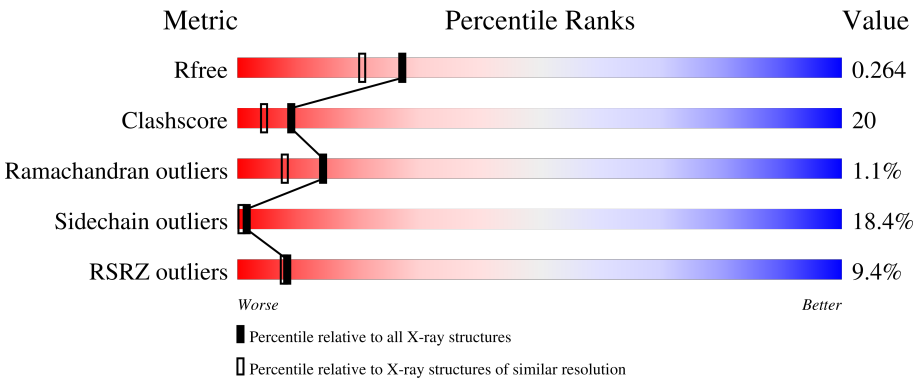
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 >=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 <=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	9%, 43%, 54%, .
2	D	35	3%, 29%, 60%, 11%
3	A	324	10%, 59%, 35%, 6%, .
3	B	324	9%, 63%, 28%, 6%, . .

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7017 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.

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

- Molecule 2 is a DNA chain called loxP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	35	713	345	126	208	34	0	0	0

- Molecule 3 is a protein called Recombinase CRE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	322	2548	1584	484	465	15	0	0	0
3	B	322	2548	1584	484	465	15	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	173	LYS	ARG	engineered mutation	UNP P06956
B	173	LYS	ARG	engineered mutation	UNP P06956

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	C	60	60	60	0	0
4	D	88	88	88	0	0
4	A	154	154	154	0	0

Continued on next page...

Continued from previous page...

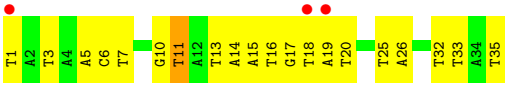
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	B	191	191	191	0	0

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: loxP

Chain C:

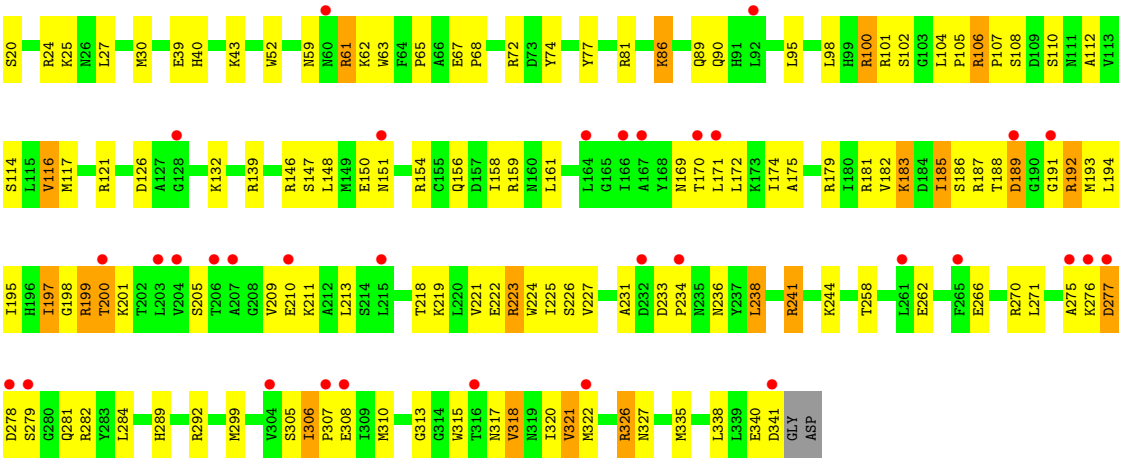


- Molecule 2: loxP

Chain D:

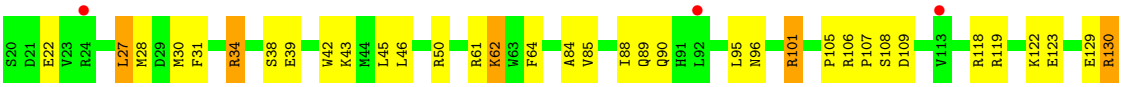
- Molecule 3: Recombinase CRE

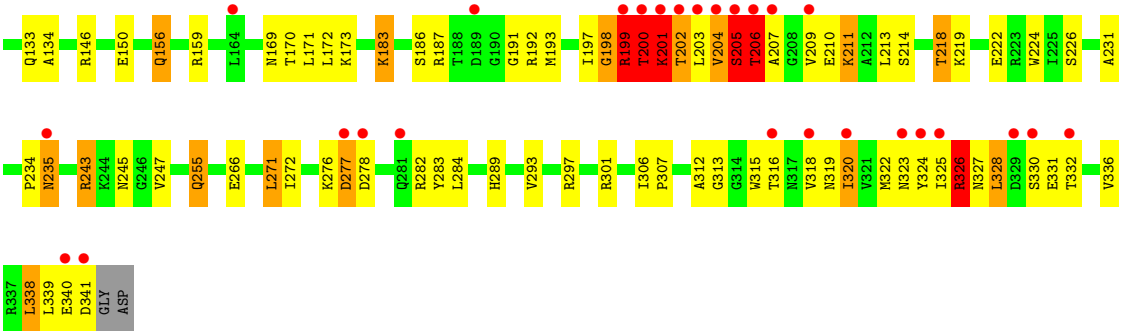
Chain A:



- Molecule 3: Recombinase CRE

Chain B:




Q133
A134
R146
E150
Q156
R159
L164
N169
T170
L171
L172
K173
K183
S186
R187
T188
D189
G190
G191
R192
M193
I197
G198
R199
T200
K201
T202
L203
V204
S205
T206
A207
G208
V209
E210
K211
A212
L213
S214
T218
K219
E222
R223
W224
I225
S226
A231
P234
N235
R243
K244
N245
G246
V247
Q255
E266
L271
I272
K276
D277
D278
Q281
R282
Y283
L284
H289
V293
R297
R301
I306
P307
A312
G313
G314
W315
T316
N317
V318
N319
I320
V321
M322
N323
Y324
I325
R326
N327
L328
D329
S330
E331
T332
V336
R337
L338
L339
E340
D341
GLY
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	107.40Å 123.00Å 180.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.00 – 2.00 26.68 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (26.00-2.00) 82.9 (26.68-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.13 (at 1.99Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.219 , 0.265 0.225 , 0.264	Depositor DCC
R_{free} test set	3452 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	36.8	Xtriage
Anisotropy	0.274	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 84.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7017	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.06% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	C	1.39	11/801 (1.4%)	1.68	15/1235 (1.2%)
2	D	1.93	11/799 (1.4%)	1.78	15/1231 (1.2%)
3	A	0.40	0/2589	0.57	0/3490
3	B	1.01	6/2589 (0.2%)	1.13	15/3490 (0.4%)
All	All	1.06	28/6778 (0.4%)	1.17	45/9446 (0.5%)

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
3	B	0	3

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	DT	O3'-P	42.24	2.11	1.61
3	B	199	ARG	CD-NE	-25.77	1.02	1.46
1	C	1	DT	O3'-P	24.51	1.90	1.61
3	B	202	THR	C-N	22.71	1.86	1.34
3	B	206	THR	C-N	-17.07	0.94	1.34

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	1	DT	O3'-P-O5'	35.94	172.28	104.00
2	D	1	DT	OP1-P-O3'	-30.19	38.78	105.20
2	D	1	DT	P-O3'-C3'	25.10	149.82	119.70
3	B	198	GLY	O-C-N	-24.75	83.10	122.70
3	B	201	LYS	C-N-CA	-21.53	67.88	121.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	198	GLY	Mainchain
3	B	199	ARG	Sidechain
3	B	201	LYS	Mainchain

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	C	715	0	401	29	0
2	D	713	0	400	32	0
3	A	2548	0	2571	88	1
3	B	2548	0	2565	142	1
4	A	154	0	0	9	0
4	B	191	0	0	5	0
4	C	60	0	0	1	0
4	D	88	0	0	5	0
All	All	7017	0	5937	253	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 20.

The worst 5 of 253 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)
3:B:202:THR:HG1	3:B:203:LEU:N	0.96	1.43
2:D:24:DA:C5'	3:B:201:LYS:HG2	1.56	1.33
1:C:17:DG:H5"	3:B:202:THR:CB	1.62	1.30
3:B:202:THR:C	3:B:203:LEU:N	1.86	1.28
1:C:17:DG:C5'	3:B:202:THR:HB	1.66	1.26

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)
3:A:126:ASP:OD2	3:B:199:ARG:NH2[4_566]	2.17	0.03

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
3	A	320/324 (99%)	301 (94%)	16 (5%)	3 (1%)	17 11
3	B	320/324 (99%)	302 (94%)	14 (4%)	4 (1%)	12 6
All	All	640/648 (99%)	603 (94%)	30 (5%)	7 (1%)	14 8

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	281	GLN
3	B	200	THR
3	B	201	LYS
3	B	277	ASP
3	A	275	ALA

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
3	A	269/270 (100%)	215 (80%)	54 (20%)	1 0
3	B	269/270 (100%)	224 (83%)	45 (17%)	2 1
All	All	538/540 (100%)	439 (82%)	99 (18%)	1 1

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	30	MET
3	B	156	GLN
3	B	38	SER
3	B	95	LEU
3	B	186	SER

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

Mol	Chain	Res	Type
3	B	133	GLN
3	B	235	ASN
3	B	311	GLN
3	A	144	GLN
3	A	89	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)

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	B	4
2	D	1
1	C	1

The worst 5 of 6 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	1:DT	O3'	2:DA	P	2.11
1	C	1:DT	O3'	2:DA	P	1.90
1	B	202:THR	C	203:LEU	N	1.86
1	B	205:SER	C	206:THR	N	1.05
1	B	198:GLY	C	199:ARG	N	0.95

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%)	0.29	3 (8%)	10	9	25, 39, 75, 90	0
2	D	35/35 (100%)	0.04	1 (2%)	51	50	24, 36, 60, 70	0
3	A	322/324 (99%)	0.64	33 (10%)	6	6	22, 47, 82, 99	0
3	B	322/324 (99%)	0.68	30 (9%)	8	8	19, 35, 83, 99	0
All	All	714/718 (99%)	0.61	67 (9%)	8	8	19, 41, 83, 99	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	203	LEU	13.4
3	B	205	SER	13.3
3	B	204	VAL	12.9
3	B	202	THR	12.6
3	B	200	THR	11.3

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 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.