



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 12:53 am GMT

PDB ID : 2CRX  
Title : STRUCTURE OF THE HOLLIDAY JUNCTION INTERMEDIATE IN CRE-  
LOXP SITE-SPECIFIC RECOMBINATION  
Authors : Gopaul, D.N.; Guo, F.; Vanduyne, G.D.  
Deposited on : 1998-06-19  
Resolution : 2.50 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

<http://wwpdb.org/validation/2016/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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

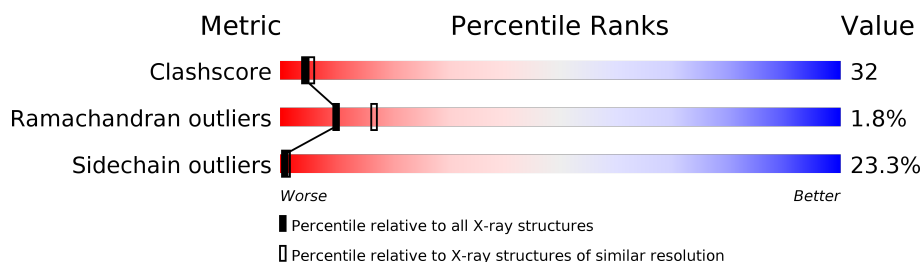
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 Å.



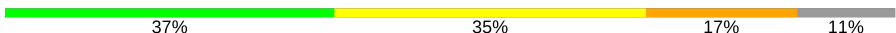

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	35	
1	D	35	
2	A	343	
2	B	343	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6639 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 DNA 35-MER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	35	Total	C	N	O	P	0	0	0
			711	345	126	207	33			
1	D	34	Total	C	N	O	P	0	0	0
			694	335	124	202	33			

- Molecule 2 is a protein called PROTEIN (CRE RECOMBINASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	305	Total	C	N	O	S	0	0	0
			2424	1510	460	439	15			
2	B	309	Total	C	N	O	S	0	0	0
			2449	1524	468	442	15			

- Molecule 3 is water.

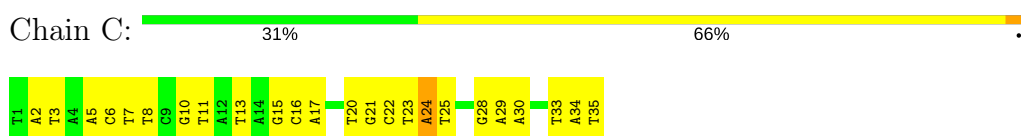
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	126	Total	O	0	0
			126	126		
3	B	150	Total	O	0	0
			150	150		
3	C	31	Total	O	0	0
			31	31		
3	D	54	Total	O	0	0
			54	54		

### 3 Residue-property plots

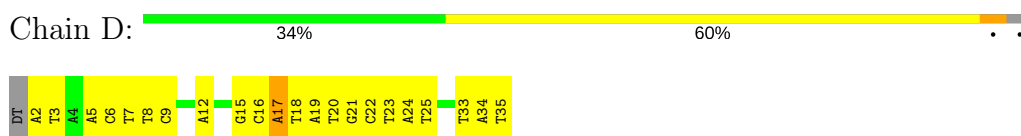
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

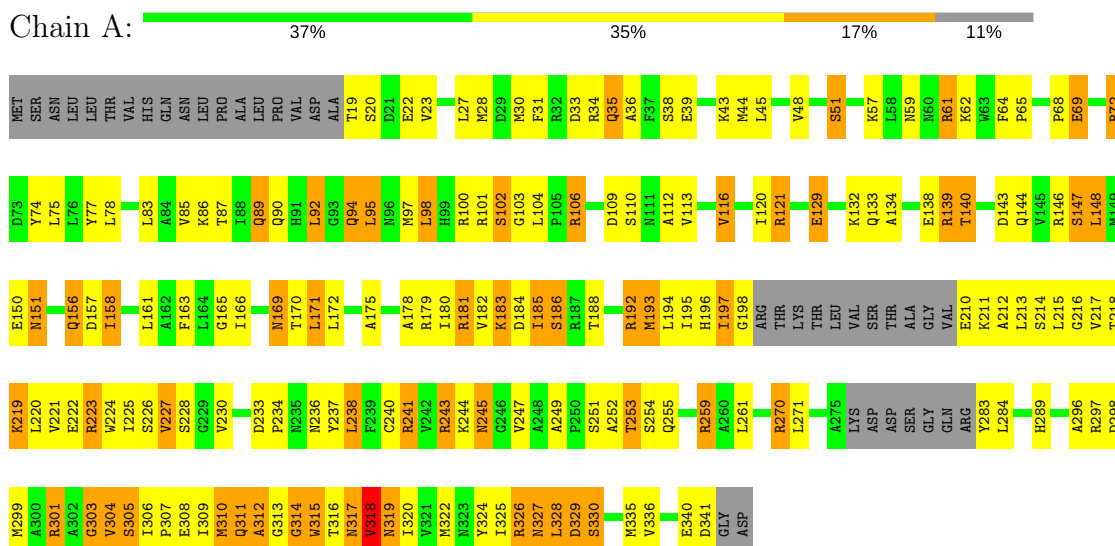
#### • Molecule 1: DNA 35-MER



#### • Molecule 1: DNA 35-MER



#### • Molecule 2: PROTEIN (CRE RECOMBINASE)



#### • Molecule 2: PROTEIN (CRE RECOMBINASE)



MET	K86	D157	D233	T316
SER	T87	I158	F294	N317
ASN	I88	R159	N235	V318
LEU	Q89	N160	N236	N319
LEU	Q90	L161	Y237	I320
THR	H91	L238	L238	V321
VAL	L92	I166	R241	M322
HIS	L95	A167	Y242	I325
GLN	N96	Y168	R243	R326
ASN	M97	N169	R244	ASN
LEU		T170	N245	LEU
PRO		L171	G246	ASP
ALA	R100	L172	V247	SER
LEU	R101	R173	A248	GLU
PRO	S102	I177	A249	THR
VAL	G103		A252	G333
ASP	L104		T253	A334
ALA	P105	I180	R181	M335
T19	R106	V182	S254	V336
	P107	K183	Q255	R337
V23	S108	D184	L256	L338
R24	D109	I185	S257	L339
K25	S110	S186	T258	E340
N26	N27	R187	R259	D341
L27	H28	T188	A260	GLY
	M30	R192	L261	ASP
	R34	M193	E266	
Q35	R118	L194	L271	
	R119			
	I120	I197	D277	
S38	R121	G198	D278	
E39	F122	ARG	S279	
	E123	THR	G280	
W42	N124	LYS	Q281	
		THR	E282	
L45	E129	LEU	Y283	
L46	R130	VAL	L284	
	Q133	SER		
S51	A134	THR	A291	
K57		A207	R292	
		G208	V293	
M60	R139	V209	D298	
R61	T140	E210	M299	
		K211	A300	
F64	D143	A212	V304	
P65	Q144	L213	S305	
	V145	S214	I306	
K66	R146		P307	
E67	S147	T218	E308	
P68	L148	K219	I309	
	M149		M310	
R72	E150	E222		
	N151	R223	G313	
L78	S152	W224	G314	
Q79	D153	I225	W315	
	R154	S226		
L83	C155	V227		
	Q156			

## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.90Å 122.50Å 180.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.30 – 2.50	Depositor
% Data completeness (in resolution range)	94.4 (27.30-2.50)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.198 , 0.268	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6639	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

## 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	0.48	0/796	0.80	0/1225
1	D	0.58	0/777	0.84	1/1195 (0.1%)
2	A	0.38	0/2463	0.64	1/3319 (0.0%)
2	B	0.43	0/2488	0.64	1/3350 (0.0%)
All	All	0.44	0/6524	0.69	3/9089 (0.0%)

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
1	D	0	1
All	All	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	17	DA	C5'-C4'-C3'	-7.04	101.42	114.10
2	B	207	ALA	N-CA-C	-5.78	95.38	111.00
2	A	102	SER	N-CA-C	-5.04	97.40	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	24	DA	Sidechain
1	D	24	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	711	0	402	23	0
1	D	694	0	389	28	0
2	A	2424	0	2437	219	0
2	B	2449	0	2465	126	0
3	A	126	0	0	18	0
3	B	150	0	0	9	0
3	C	31	0	0	1	0
3	D	54	0	0	5	0
All	All	6639	0	5693	380	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 32.

The worst 5 of 380 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:A:19:THR:HA	2:A:22:GLU:HG3	1.31	1.10
1:D:19:DA:H2''	1:D:20:DT:H5''	1.32	1.10
2:B:333:GLY:HA2	2:B:336:VAL:HB	1.27	1.05
2:A:132:LYS:HB3	2:A:283:TYR:CE2	1.96	1.00
2:B:333:GLY:O	2:B:334:ALA:C	2.02	0.94

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	
2	A	299/343 (87%)	262 (88%)	29 (10%)	8 (3%)	6	9
2	B	303/343 (88%)	286 (94%)	14 (5%)	3 (1%)	18	32
All	All	602/686 (88%)	548 (91%)	43 (7%)	11 (2%)	10	17

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	334	ALA
2	A	103	GLY
2	A	303	GLY
2	A	312	ALA
2	A	318	VAL

### 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	255/287 (89%)	189 (74%)	66 (26%)	0	1
2	B	256/287 (89%)	203 (79%)	53 (21%)	1	2
All	All	511/574 (89%)	392 (77%)	119 (23%)	1	1

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	305	SER
2	B	19	THR
2	B	305	SER
2	A	311	GLN
2	A	320	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	245	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	A	311	GLN
2	B	281	GLN
2	A	156	GLN
2	B	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 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

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.