



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 06:35 pm BST

PDB ID : 3CNH  
Title : Crystal structure of predicted hydrolase of haloacid dehalogenase-like superfamily (NP\_295428.1) from *Deinococcus radiodurans* at 1.66 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2008-03-25  
Resolution : 1.66 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 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	:	2.11
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

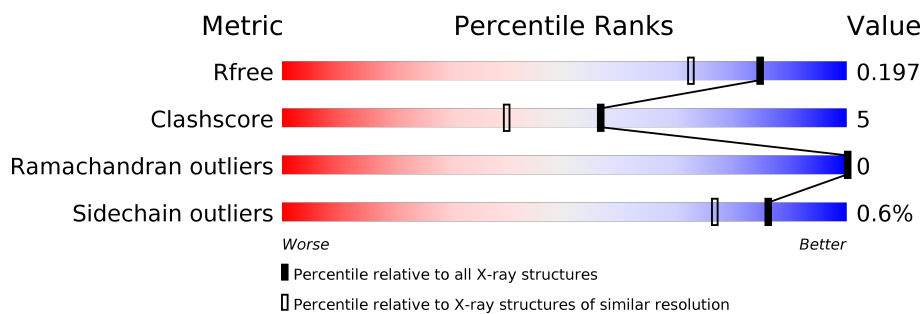
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.66 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	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)

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 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	200	
1	B	200	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 3798 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 protein called Hydrolase family protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	Se	0	20	0
			1720	1089	312	309	1	9			
1	B	199	Total	C	N	O	S	Se	0	19	0
			1706	1081	311	304	1	9			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	LEADER SEQUENCE	UNP Q9RTQ1
B	0	GLY	-	LEADER SEQUENCE	UNP Q9RTQ1

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

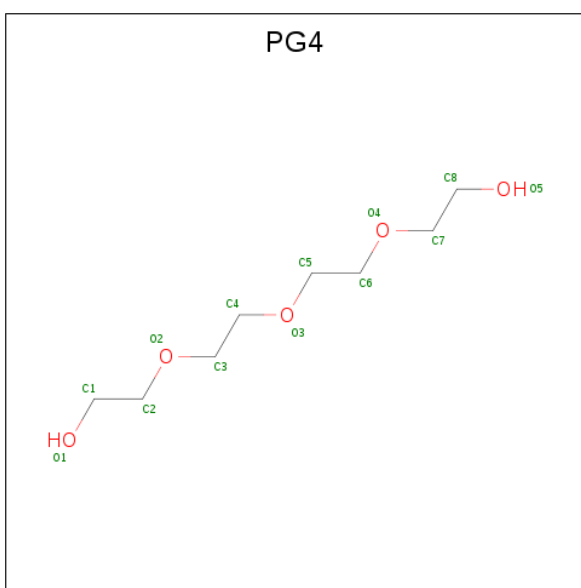
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Na	0	0
			1	1		
2	A	1	Total	Na	0	0
			1	1		

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			6	3	3		

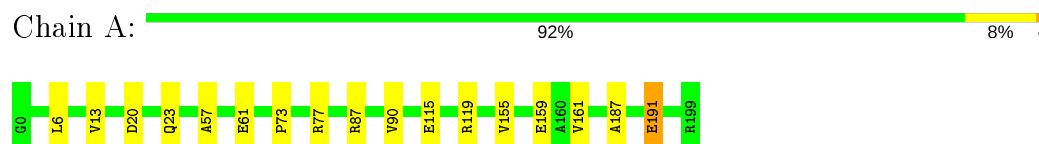
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	177	Total	O	0	12
			190	190		
6	B	146	Total	O	0	8
			154	154		

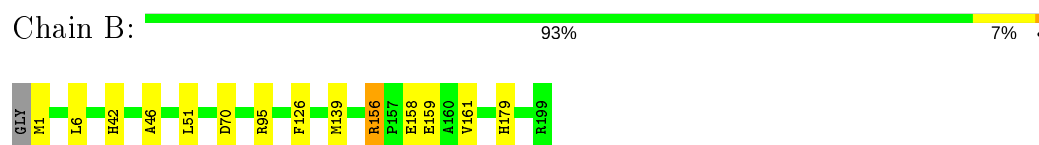
### 3 Residue-property plots [i](#)

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. 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. 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: Hydrolase family protein



- Molecule 1: Hydrolase family protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.11Å 74.11Å 160.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.35 – 1.66 35.36 – 1.66	Depositor EDS
% Data completeness (in resolution range)	99.9 (35.35-1.66) 99.9 (35.36-1.66)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.00 (at 1.66Å)	Xtriage
Refinement program	REFMAC 5.4.0067, PHENIX	Depositor
R, $R_{free}$	0.160 , 0.187 0.167 , 0.197	Depositor DCC
$R_{free}$ test set	2734 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.0	Xtriage
Anisotropy	0.159	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 51.4	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	0.95	EDS
Total number of atoms	3798	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% of the height of the origin peak. No significant pseudotranslation is detected.*

<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: GOL, PO4, PG4, NA

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	0.71	0/1799	0.86	0/2409
1	B	0.70	0/1781	0.87	1/2390 (0.0%)
All	All	0.70	0/3580	0.86	1/4799 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	70	ASP	CB-CG-OD2	-5.04	113.76	118.30

There are no chirality outliers.

There are no planarity outliers.

### 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	A	1720	0	1754	18	0
1	B	1706	0	1743	16	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	5	0	0	0	0
3	B	5	0	0	1	0
4	A	10	0	13	0	0
5	B	6	0	8	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	190	0	0	0	0
6	B	154	0	0	1	0
All	All	3798	0	3518	35	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 5.

All (35) 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:B:1[A]:MSE:HE3	1:B:158:GLU:CG	1.47	1.41
1:B:1[A]:MSE:CE	1:B:158:GLU:HG3	1.57	1.32
1:B:1[A]:MSE:HE1	1:B:158:GLU:OE2	1.23	1.25
1:B:1[A]:MSE:CE	1:B:158:GLU:CG	2.19	1.12
1:A:20:ASP:H	1:A:23[B]:GLN:HE21	1.13	0.96
1:B:1[A]:MSE:CE	1:B:158:GLU:OE2	2.13	0.95
1:B:1[A]:MSE:HE1	1:B:158:GLU:CD	1.91	0.91
1:A:20:ASP:CG	1:A:23[B]:GLN:NE2	2.25	0.89
1:A:20:ASP:H	1:A:23[B]:GLN:NE2	1.74	0.84
1:A:20:ASP:N	1:A:23[B]:GLN:HE21	1.83	0.77
3:B:201:PO4:O2	6:B:304:HOH:O	2.06	0.74
1:B:1[B]:MSE:HE2	1:B:179:HIS:CE1	2.25	0.72
1:B:1[A]:MSE:CE	1:B:158:GLU:CD	2.52	0.71
1:A:20:ASP:CG	1:A:23[B]:GLN:HE21	1.92	0.70
1:B:1[A]:MSE:HE3	1:B:158:GLU:HG3	0.73	0.70
1:B:6:LEU:HG	1:B:161[B]:VAL:CG1	2.27	0.65
1:B:1[A]:MSE:HE3	1:B:158:GLU:HG2	1.70	0.63
1:A:73:PRO:O	1:A:77[B]:ARG:HG3	1.99	0.63
1:B:95[A]:ARG:HG2	1:B:126:PHE:CD2	2.40	0.57
1:A:20:ASP:CB	1:A:23[B]:GLN:HE21	2.18	0.56
1:A:87[B]:ARG:HH12	1:A:90:VAL:HG21	1.70	0.56
1:A:6:LEU:HG	1:A:161[B]:VAL:CG1	2.36	0.56
1:A:187:ALA:O	1:A:191[B]:GLU:HG3	2.08	0.53
1:B:1[B]:MSE:HE2	1:B:179:HIS:HE1	1.70	0.52
1:A:20:ASP:OD1	1:A:23[B]:GLN:NE2	2.42	0.52
1:A:155[B]:VAL:HG22	1:A:159:GLU:HB2	1.92	0.51
1:A:20:ASP:CA	1:A:23[B]:GLN:HE21	2.28	0.47
1:A:155[B]:VAL:CG2	1:A:159:GLU:HB2	2.45	0.47
1:A:115:GLU:HB3	1:A:119[B]:ARG:HH12	1.79	0.47
1:A:57:ALA:O	1:A:61[B]:GLU:HG3	2.15	0.46
1:A:13:VAL:O	1:A:87[B]:ARG:NH1	2.46	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:51:LEU:HD11	1:B:139:MSE:HE3	2.00	0.43
1:B:156:ARG:HB2	1:B:159[A]:GLU:HG3	2.00	0.43
1:B:42:HIS:CE1	1:B:46:ALA:HB2	2.56	0.40

There are no symmetry-related clashes.

## 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	
1	A	217/200 (108%)	211 (97%)	6 (3%)	0	100	100
1	B	215/200 (108%)	209 (97%)	6 (3%)	0	100	100
All	All	432/400 (108%)	420 (97%)	12 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	
1	A	181/156 (116%)	179 (99%)	2 (1%)	73	57
1	B	179/156 (115%)	178 (99%)	1 (1%)	86	76
All	All	360/312 (115%)	357 (99%)	3 (1%)	86	70

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	191[A]	GLU
1	A	191[B]	GLU
1	B	156	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	179	HIS

### 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](#)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

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	$\# Z  > 2$
3	PO4	A	201	2	4,4,4	0.98	0	6,6,6	1.42	1 (16%)
4	PG4	A	202	-	9,9,12	0.38	0	8,8,11	0.46	0
5	GOL	B	202	-	5,5,5	0.34	0	5,5,5	0.61	0
3	PO4	B	201	2	4,4,4	1.10	0	6,6,6	1.78	2 (33%)

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
4	PG4	A	202	-	-	5/7/7/10	-
5	GOL	B	202	-	-	4/4/4/4	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	201	PO4	O3-P-O1	-2.76	100.81	110.89
3	A	201	PO4	O4-P-O2	-2.22	100.84	107.97
3	B	201	PO4	O4-P-O2	2.10	114.72	107.97

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	202	GOL	O1-C1-C2-C3
5	B	202	GOL	C1-C2-C3-O3
4	A	202	PG4	O2-C3-C4-O3
4	A	202	PG4	O3-C5-C6-O4
5	B	202	GOL	O2-C2-C3-O3
4	A	202	PG4	O1-C1-C2-O2
5	B	202	GOL	O1-C1-C2-O2
4	A	202	PG4	C6-C5-O3-C4
4	A	202	PG4	C1-C2-O2-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	201	PO4	1	0

## 5.7 Other polymers ⓘ

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

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.