



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 11:40 am BST

PDB ID : 1EON  
Title : ECORV BOUND TO 3'-S-PHOSPHOROTHIOATE DNA AND CA<sup>2+</sup>  
Authors : Horton, N.C.; Connolly, B.A.; Perona, J.J.  
Deposited on : 2000-03-23  
Resolution : 1.60 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

<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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

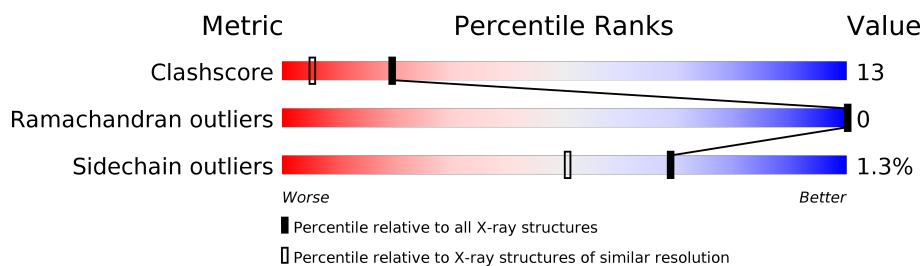
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	11	64% 36%
2	D	11	27% 36% 18% 18%
3	A	245	82% 14% . .
3	B	245	82% 13% .

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	CL	A	501	-	-	X	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	CL	A	504	-	-	X	-
4	CL	B	502	-	-	X	-
4	CL	B	503	-	-	X	-
5	ACY	A	601	-	-	X	-
5	ACY	B	602	-	-	X	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 4587 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 (5'-D(\*AP\*AP\*AP\*GP\*AP\*(TSP)P\*AP\*TP\*CP\*TP\*T)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	C	11	Total	C	N	O	P	S	0	0	0
			223	109	41	62	10	1			

- Molecule 2 is a DNA chain called DNA (5'-D(\*CP\*AP\*AP\*GP\*AP\*(TSP)P\*AP\*TP\*CP\*TP\*T)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	11	Total	C	N	O	P	S	0	0	0
			221	108	39	63	10	1			

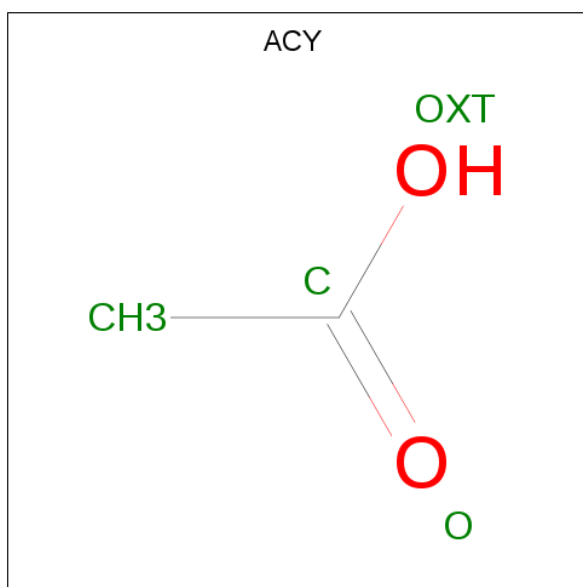
- Molecule 3 is a protein called TYPE II RESTRICTION ENZYME ECORV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	236	Total	C	N	O	S	0	2	0
			1887	1224	312	350	1			
3	B	234	Total	C	N	O	S	0	1	0
			1887	1221	310	355	1			

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	2	Total	Cl	0	0
			2	2		
4	A	2	Total	Cl	0	0
			2	2		

- Molecule 5 is ACETIC ACID (three-letter code: ACY) (formula: C<sub>2</sub>H<sub>4</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	26	Total	O	0	0
			26	26		
6	D	43	Total	O	0	0
			43	43		
6	A	169	Total	O	0	0
			169	169		
6	B	119	Total	O	0	0
			119	119		

### 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.

Note EDS was not executed.

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

Chain C: 




- Molecule 2: DNA (5'-D(\*CP\*AP\*AP\*GP\*AP\*(TSP)P\*AP\*TP\*CP\*TP\*T)-3')

Chain D: 




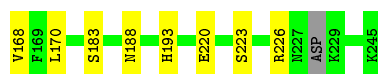
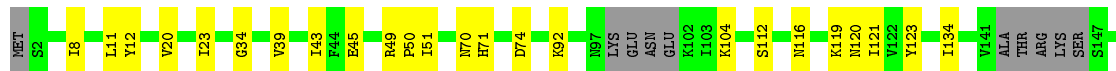
- Molecule 3: TYPE II RESTRICTION ENZYME ECORV

Chain A: 



- Molecule 3: TYPE II RESTRICTION ENZYME ECORV

Chain B: 



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.90 Å 48.60 Å 63.90 Å 96.90° 108.90° 106.80°	Depositor
Resolution (Å)	4.80 – 1.60	Depositor
% Data completeness (in resolution range)	95.1 (4.80-1.60)	Depositor
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.229 , 0.281	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4587	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

## 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: ACY, TSP, CL

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	0.49	0/227	0.84	0/346
2	D	0.71	0/224	1.40	2/341 (0.6%)
3	A	0.42	0/1944	0.65	0/2643
3	B	0.41	0/1939	0.63	0/2634
All	All	0.44	0/4334	0.72	2/5964 (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
2	D	0	5

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	7	DA	O4'-C1'-C2'	5.58	110.36	105.90
2	D	2	DA	O4'-C1'-C2'	5.25	110.10	105.90

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	1	DC	Sidechain
2	D	10	DT	Sidechain
2	D	2	DA	Sidechain
2	D	7	DA	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
2	D	8	DT	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	223	0	127	9	0
2	D	221	0	127	12	0
3	A	1887	0	1777	42	0
3	B	1887	0	1774	28	0
4	A	2	0	0	14	0
4	B	2	0	0	12	0
5	A	4	0	4	14	0
5	B	4	0	4	3	0
6	A	169	0	0	10	0
6	B	119	0	0	4	0
6	C	26	0	0	3	0
6	D	43	0	0	2	0
All	All	4587	0	3813	102	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 13.

The worst 5 of 102 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)
5:B:602:ACY:C	5:B:602:ACY:CH3	1.93	1.45
5:B:602:ACY:C	5:B:602:ACY:OXT	1.81	1.26
3:A:15:ASN:OD1	5:A:601:ACY:H2	1.71	0.87
4:B:503:CL:CL	6:B:706:HOH:O	2.32	0.84
2:D:1:DC:H5"	3:A:180:LEU:HD13	1.60	0.81

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	
3	A	232/245 (95%)	224 (97%)	8 (3%)	0	100	100
3	B	227/245 (93%)	218 (96%)	9 (4%)	0	100	100
All	All	459/490 (94%)	442 (96%)	17 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	193/221 (87%)	189 (98%)	4 (2%)	53	29
3	B	195/221 (88%)	194 (100%)	1 (0%)	88	80
All	All	388/442 (88%)	383 (99%)	5 (1%)	69	50

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	42	THR
3	A	58	LYS
3	A	96	THR
3	A	114	ILE
3	B	71	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	71	HIS
3	A	97	ASN
3	B	120	ASN
3	B	193	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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$
2	TSP	D	6	1,2	14,21,22	1.07	3 (21%)	12,30,33	4.45	2 (16%)
1	TSP	C	6	1,2	14,21,22	1.19	1 (7%)	12,30,33	4.42	3 (25%)

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
2	TSP	D	6	1,2	-	0/4/21/22	0/2/2/2
1	TSP	C	6	1,2	-	0/4/21/22	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	6	TSP	C4-N3	3.29	1.38	1.33
2	D	6	TSP	C4-N3	2.61	1.37	1.33
2	D	6	TSP	C4-C5	2.04	1.45	1.41
2	D	6	TSP	C6-C5	-2.03	1.34	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	6	TSP	C4-N3-C2	15.00	127.81	115.14
1	C	6	TSP	C4-N3-C2	14.28	127.19	115.14
1	C	6	TSP	C2'-C1'-N1	-4.33	104.27	114.27
1	C	6	TSP	C5M-C5-C6	2.23	123.39	118.68
2	D	6	TSP	C5M-C5-C6	2.16	123.24	118.68

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	$\# Z  > 2$
5	ACY	B	602	-	1,3,3	35.03	1 (100%)	0,3,3	0.00	-
5	ACY	A	601	-	1,3,3	126.03	1 (100%)	0,3,3	0.00	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	ACY	CH3-C	126.03	3.09	1.48
5	B	602	ACY	CH3-C	35.03	1.93	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	602	ACY	3	0
5	A	601	ACY	14	0

## 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.