



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 15, 2024 – 03:51 pm GMT

PDB ID : 6TT5  
Title : Crystal structure of DCLRE1C/Artemis  
Authors : Yosaatmadja, Y.; Goubin, S.; Newman, J.A.; Mukhopadhyay, S.M.M.; Dannerfjord, A.A.; Burgess-Brown, N.A.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.; Bountra, C.; Gileadi, O.  
Deposited on : 2019-12-23  
Resolution : 1.50 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

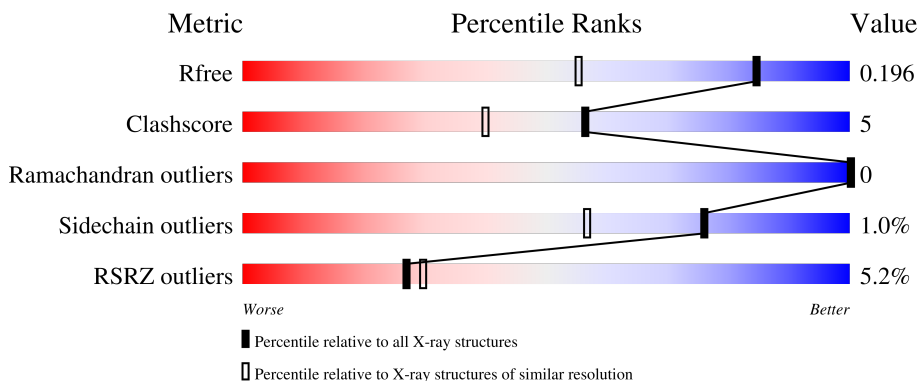
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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))
$R_{free}$	130704	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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 of 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  $\leq 5\%$ . 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	AAA	362	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3178 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 Protein artemis.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	348	2927	1884	507	514	22	0	30	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	SER	-	expression tag	UNP Q96SD1

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	AAA	1	4	2	2	0	0
2	AAA	1	4	2	2	0	0
2	AAA	1	4	2	2	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total C O 4 2 2	0	0
2	AAA	1	Total C O 4 2 2	0	0
2	AAA	1	Total C O 4 2 2	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	2	Total Zn 2 2	0	0

- Molecule 4 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total Ni 1 1	0	0

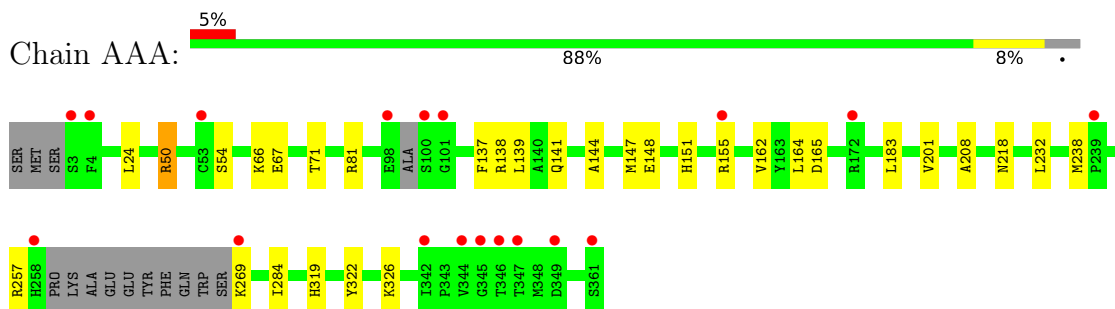
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	224	Total O 224 224	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: Protein artemis



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	35.87Å 47.99Å 48.25Å 82.61° 76.37° 85.98°	Depositor
Resolution (Å)	47.55 – 1.50 47.55 – 1.50	Depositor EDS
% Data completeness (in resolution range)	94.6 (47.55-1.50) 94.6 (47.55-1.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.97 (at 1.50Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.166 , 0.192 0.175 , 0.196	Depositor DCC
$R_{free}$ test set	2199 reflections (4.68%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.3	Xtrriage
Anisotropy	0.469	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 43.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3178	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.59% 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: NI, EDO, ZN

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	AAA	0.62	0/3079	0.70	0/4173

There are no bond length outliers.

There are no bond angle outliers.

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	AAA	2927	0	2957	30	0
2	AAA	24	0	36	0	0
3	AAA	2	0	0	0	0
4	AAA	1	0	0	0	0
5	AAA	224	0	0	7	0
All	All	3178	0	2993	30	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 (30) 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:AAA:165:ASP:OD2	1:AAA:319:HIS:HD2	1.73	0.72
1:AAA:50:ARG:HD3	1:AAA:50:ARG:C	2.12	0.70
1:AAA:201[A]:VAL:CG2	1:AAA:284:ILE:HG12	2.31	0.60
1:AAA:139[A]:LEU:HD21	1:AAA:147[A]:MET:HE1	1.84	0.59
1:AAA:54:SER:O	1:AAA:81:ARG:NH1	2.35	0.59
1:AAA:257[A]:ARG:HH11	1:AAA:257[A]:ARG:HG2	1.67	0.58
1:AAA:269:LYS:HB3	5:AAA:507:HOH:O	2.04	0.58
1:AAA:24:LEU:O	1:AAA:50:ARG:HG3	2.04	0.56
1:AAA:322:TYR:CE2	1:AAA:326:LYS:HD3	2.42	0.55
1:AAA:66[A]:LYS:NZ	5:AAA:501:HOH:O	0.69	0.54
1:AAA:67:GLU:O	1:AAA:71:THR:HG23	2.07	0.54
1:AAA:257[A]:ARG:HH11	1:AAA:257[A]:ARG:CG	2.22	0.53
1:AAA:183[A]:LEU:HD11	1:AAA:218:ASN:HB3	1.92	0.50
1:AAA:257[A]:ARG:NH2	5:AAA:503:HOH:O	2.43	0.50
1:AAA:50:ARG:HD3	1:AAA:50:ARG:O	2.13	0.49
1:AAA:139[A]:LEU:HD21	1:AAA:147[A]:MET:CE	2.42	0.48
1:AAA:162:VAL:HG12	1:AAA:164[A]:LEU:HB3	1.96	0.47
1:AAA:319:HIS:HE1	5:AAA:703:HOH:O	1.97	0.47
1:AAA:269:LYS:HD3	5:AAA:507:HOH:O	2.13	0.47
1:AAA:144:ALA:HA	1:AAA:147[A]:MET:HE2	1.97	0.46
1:AAA:257[B]:ARG:CZ	1:AAA:257[B]:ARG:HB3	2.44	0.46
1:AAA:66[B]:LYS:NZ	5:AAA:501:HOH:O	2.45	0.46
1:AAA:148:GLU:HA	1:AAA:151[B]:HIS:CD2	2.51	0.45
1:AAA:208:ALA:HB2	1:AAA:257[B]:ARG:CZ	2.46	0.45
1:AAA:155:ARG:HD2	5:AAA:639:HOH:O	2.17	0.44
1:AAA:67:GLU:HG2	1:AAA:238:MET:HE2	2.00	0.44
1:AAA:269:LYS:HE2	1:AAA:284:ILE:HD12	2.01	0.42
1:AAA:137:PHE:CD1	1:AAA:139[B]:LEU:HD23	2.55	0.42
1:AAA:257[A]:ARG:CG	1:AAA:257[A]:ARG:NH1	2.82	0.41

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	AAA	372/362 (103%)	371 (100%)	1 (0%)	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	AAA	330/329 (100%)	327 (99%)	3 (1%)	78 61

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

Mol	Chain	Res	Type
1	AAA	50	ARG
1	AAA	138	ARG
1	AAA	232	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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

Of 9 ligands modelled in this entry, 3 are monoatomic - leaving 6 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
2	EDO	AAA	406	-	3,3,3	0.08	0	2,2,2	0.30	0
2	EDO	AAA	405	-	3,3,3	0.06	0	2,2,2	0.18	0
2	EDO	AAA	401	-	3,3,3	0.07	0	2,2,2	0.08	0
2	EDO	AAA	402	-	3,3,3	0.05	0	2,2,2	0.21	0
2	EDO	AAA	404	-	3,3,3	0.10	0	2,2,2	0.29	0
2	EDO	AAA	403	-	3,3,3	0.11	0	2,2,2	0.15	0

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	EDO	AAA	406	-	-	1/1/1/1	-
2	EDO	AAA	405	-	-	0/1/1/1	-
2	EDO	AAA	401	-	-	0/1/1/1	-
2	EDO	AAA	402	-	-	0/1/1/1	-
2	EDO	AAA	404	-	-	0/1/1/1	-
2	EDO	AAA	403	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	406	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	AAA	348/362 (96%)	0.44	18 (5%) <span style="border: 1px solid red; padding: 2px;">27</span> <span style="border: 1px solid red; padding: 2px;">30</span>	8, 16, 32, 49	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	344	VAL	6.8
1	AAA	3	SER	6.4
1	AAA	101	GLY	5.0
1	AAA	361	SER	4.7
1	AAA	100	SER	4.0
1	AAA	342	ILE	3.8
1	AAA	347	THR	3.7
1	AAA	239	PRO	3.3
1	AAA	53	CYS	3.2
1	AAA	345	GLY	2.9
1	AAA	346	THR	2.5
1	AAA	349	ASP	2.5
1	AAA	269	LYS	2.5
1	AAA	98	GLU	2.4
1	AAA	4	PHE	2.4
1	AAA	155	ARG	2.2
1	AAA	258	HIS	2.1
1	AAA	172	ARG	2.0

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	EDO	AAA	406	4/4	0.73	0.17	39,40,42,42	0
2	EDO	AAA	404	4/4	0.83	0.23	25,27,29,30	0
2	EDO	AAA	401	4/4	0.83	0.11	27,28,28,29	0
2	EDO	AAA	405	4/4	0.86	0.15	35,36,37,40	0
2	EDO	AAA	403	4/4	0.89	0.21	27,27,28,30	0
2	EDO	AAA	402	4/4	0.90	0.15	24,24,26,26	0
3	ZN	AAA	408	1/1	0.98	0.14	13,13,13,13	1
3	ZN	AAA	407	1/1	1.00	0.07	12,12,12,12	0
4	NI	AAA	409	1/1	1.00	0.07	11,11,11,11	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.