



# Full wwPDB X-ray Structure Validation Report ⓘ

May 24, 2020 – 04:31 am BST

PDB ID : 5Q4V  
Title : PanDDA analysis group deposition – Crystal Structure of DCLRE1A after initial refinement with no ligand modelled (structure 96)  
Authors : Newman, J.A.; Aitkenhead, H.; Lee, S.Y.; Kupinska, K.; Burgess-Brown, N.; Tallon, R.; Krojer, T.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.; Bountra, C.; Gileadi, O.  
Deposited on : 2017-05-25  
Resolution : 1.59 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

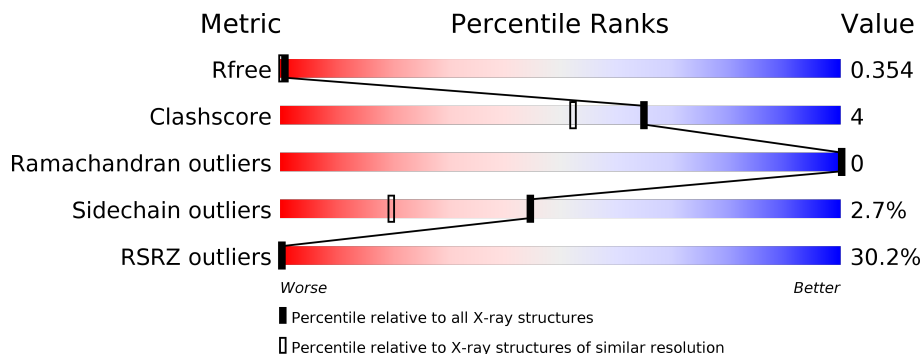
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.59 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (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  $\leq 5\%$ . 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	A	343	

## 2 Entry composition [i](#)

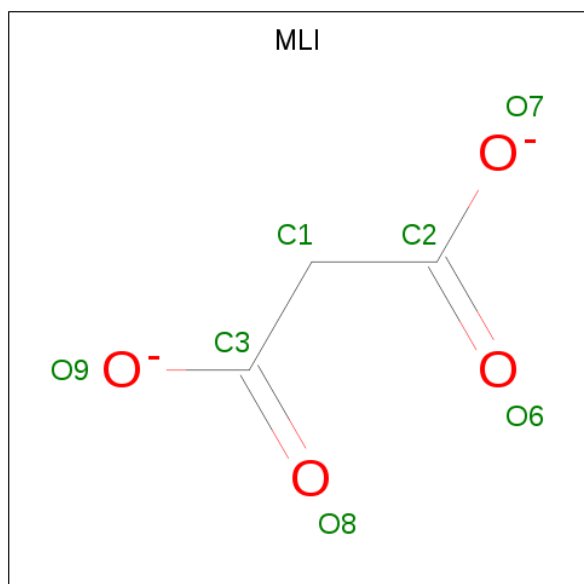
There are 4 unique types of molecules in this entry. The entry contains 3044 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 DCLRE1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	341	2726	1765	450	491	20	0	4	0

- Molecule 2 is MALONATE ION (three-letter code: MLI) (formula:  $C_3H_2O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	7	3	4	0	0

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ni	0	0
			1	1		

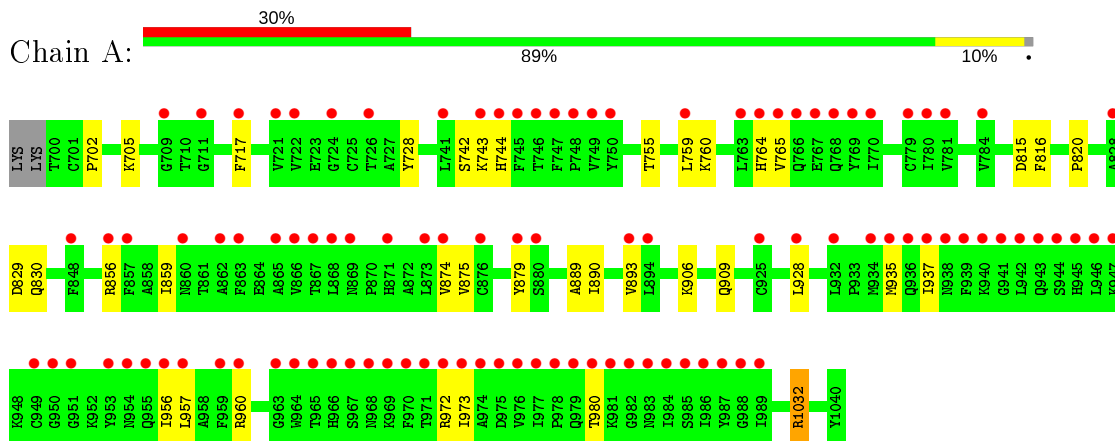
- Molecule 4 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	A	310	Total 310	O 310	0	0

### 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 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: DCLRE1A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.87Å 58.04Å 114.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	57.26 – 1.59 57.27 – 1.59	Depositor EDS
% Data completeness (in resolution range)	99.9 (57.26-1.59) 100.0 (57.27-1.59)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 1.59Å)	Xtrriage
Refinement program	REFMAC 5.8.0155	Depositor
R, $R_{free}$	0.297 , 0.344 0.309 , 0.354	Depositor DCC
$R_{free}$ test set	2278 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.2	Xtrriage
Anisotropy	0.252	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 41.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3044	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.72% 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, MLI

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	0/2802	0.80	1/3809 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	1032	ARG	NE-CZ-NH1	-5.91	117.35	120.30

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	2726	0	2676	22	0
2	A	7	0	2	0	0
3	A	1	0	0	0	0
4	A	310	0	0	11	0
All	All	3044	0	2678	22	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 4.

All (22) 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:A:909:GLN:NE2	4:A:1201:HOH:O	2.04	0.91
1:A:1032:ARG:NH2	4:A:1208:HOH:O	2.34	0.58
1:A:1032:ARG:CZ	4:A:1225:HOH:O	2.55	0.55
1:A:906:LYS:NZ	4:A:1214:HOH:O	2.40	0.55
1:A:973:ILE:HG23	4:A:1447:HOH:O	2.07	0.54
1:A:755:THR:O	1:A:759:LEU:HG	2.09	0.53
1:A:742:SER:O	1:A:765:VAL:HG22	2.09	0.53
1:A:743:LYS:HG3	1:A:744:HIS:CE1	2.50	0.47
1:A:815:ASP:OD1	4:A:1202:HOH:O	2.19	0.47
1:A:1032:ARG:NE	4:A:1225:HOH:O	2.48	0.46
1:A:717:PHE:CE1	1:A:728:TYR:HB3	2.53	0.44
1:A:859:ILE:HD11	1:A:889:ALA:HB1	2.00	0.44
1:A:874:VAL:HG11	1:A:890:ILE:HG21	2.01	0.43
1:A:937:ILE:HG21	1:A:960:ARG:HG2	2.00	0.43
1:A:856:ARG:HA	1:A:856:ARG:HE	1.83	0.43
1:A:760:LYS:O	1:A:764[A]:HIS:HA	2.19	0.42
1:A:875:VAL:HG23	1:A:956:ILE:HG23	2.01	0.42
1:A:702:PRO:HG2	1:A:705:LYS:HG3	2.01	0.42
1:A:973:ILE:HG22	4:A:1236:HOH:O	2.19	0.41
1:A:935:MET:HG2	4:A:1480:HOH:O	2.20	0.41
1:A:972:ARG:HB3	4:A:1236:HOH:O	2.21	0.41
1:A:980:THR:HG23	4:A:1303:HOH:O	2.20	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	A	343/343 (100%)	320 (93%)	23 (7%)	0	<b>100</b> <b>100</b>

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	299/305 (98%)	291 (97%)	8 (3%)	44 20

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	816	PHE
1	A	820	PRO
1	A	829	ASP
1	A	830	GLN
1	A	879	TYR
1	A	893	VAL
1	A	928	LEU
1	A	957	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	744	HIS
1	A	851	GLN
1	A	869	ASN
1	A	909	GLN
1	A	979	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](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	MLI	A	1101	3	0,6,6	0.00	-	0,7,7	0.00	-

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	MLI	A	1101	3	-	0/0/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

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	A	341/343 (99%)	1.80	103 (30%) <b>0</b> <b>0</b>	14, 34, 92, 145	0

All (103) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	966	HIS	20.5
1	A	970	PHE	20.1
1	A	984	ILE	14.8
1	A	939	PHE	12.2
1	A	744	HIS	11.9
1	A	980	THR	11.0
1	A	768	GLN	10.3
1	A	937	ILE	9.8
1	A	769	TYR	9.4
1	A	977	ILE	9.1
1	A	972	ARG	9.0
1	A	973	ILE	8.9
1	A	974	ALA	8.7
1	A	941	GLY	8.7
1	A	986	ILE	8.7
1	A	956	ILE	8.0
1	A	942	LEU	7.7
1	A	765	VAL	7.6
1	A	943	GLN	6.9
1	A	938	ASN	6.9
1	A	987	TYR	6.9
1	A	985	SER	6.8
1	A	781	VAL	6.7
1	A	967	SER	6.6
1	A	745	PHE	6.5
1	A	940	LYS	6.5
1	A	747	PHE	6.3

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	971	THR	6.0
1	A	979	GLN	5.8
1	A	950	GLY	5.7
1	A	954	ASN	5.6
1	A	866	VAL	5.5
1	A	981	LYS	5.3
1	A	928	LEU	5.2
1	A	988	GLY	4.7
1	A	780	ILE	4.7
1	A	946	LEU	4.6
1	A	770	ILE	4.6
1	A	968	ASN	4.6
1	A	957	LEU	4.5
1	A	764[A]	HIS	4.5
1	A	721	VAL	4.4
1	A	767	GLU	4.4
1	A	868	LEU	4.3
1	A	873	LEU	4.3
1	A	746	THR	4.3
1	A	982	GLY	4.2
1	A	955	GLN	4.2
1	A	879	TYR	4.2
1	A	862	ALA	4.2
1	A	750	TYR	4.1
1	A	935	MET	4.1
1	A	976	VAL	4.0
1	A	951	GLY	3.8
1	A	936	GLN	3.8
1	A	856	ARG	3.7
1	A	969	LYS	3.7
1	A	983	ASN	3.7
1	A	963	GLY	3.7
1	A	759	LEU	3.6
1	A	947	LYS	3.6
1	A	978	PRO	3.6
1	A	959	PHE	3.4
1	A	779	CYS	3.4
1	A	711	GLY	3.4
1	A	722	VAL	3.3
1	A	894	LEU	3.3
1	A	741	LEU	3.3
1	A	953	TYR	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	766	GLN	3.2
1	A	893	VAL	3.2
1	A	960	ARG	3.2
1	A	945	HIS	3.2
1	A	965	THR	3.1
1	A	964	TRP	3.0
1	A	863	PHE	3.0
1	A	876[A]	CYS	2.9
1	A	848	PHE	2.9
1	A	871	HIS	2.9
1	A	743	LYS	2.8
1	A	749	VAL	2.7
1	A	763	LEU	2.7
1	A	865	ALA	2.6
1	A	867	THR	2.6
1	A	874	VAL	2.6
1	A	944	SER	2.6
1	A	724	GLY	2.6
1	A	989	ILE	2.5
1	A	748	PRO	2.4
1	A	857	PHE	2.3
1	A	949	CYS	2.3
1	A	709	GLY	2.3
1	A	717	PHE	2.3
1	A	726	THR	2.3
1	A	932	LEU	2.1
1	A	828	ALA	2.1
1	A	934	MET	2.1
1	A	975	ASP	2.1
1	A	880	SER	2.1
1	A	925	CYS	2.1
1	A	860	ASN	2.0
1	A	869	ASN	2.0
1	A	784	VAL	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 carbohydrates 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( $\text{\AA}^2$ )	Q<0.9
2	MLI	A	1101	7/7	0.94	0.09	20,26,29,30	0
3	NI	A	1102	1/1	0.99	0.02	18,18,18,18	0

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

There are no such residues in this entry.