Full wwPDB X-ray Structure Validation Report (i)

Jun 17, 2024 – 04:43 AM EDT

PDB ID	:	5Q8E
Title	:	PanDDA analysis group deposition – Crystal Structure of DCLRE1A after initial refinement with no ligand modelled (structure 223)
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	:	1.63 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

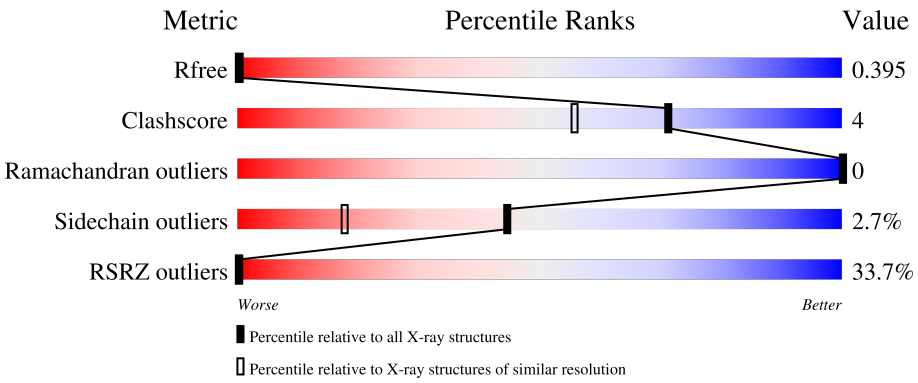
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

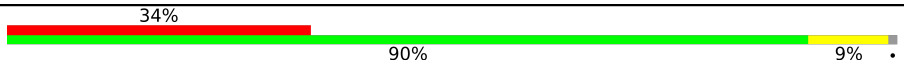
The reported resolution of this entry is 1.63 Å.

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	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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 >=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 <=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)

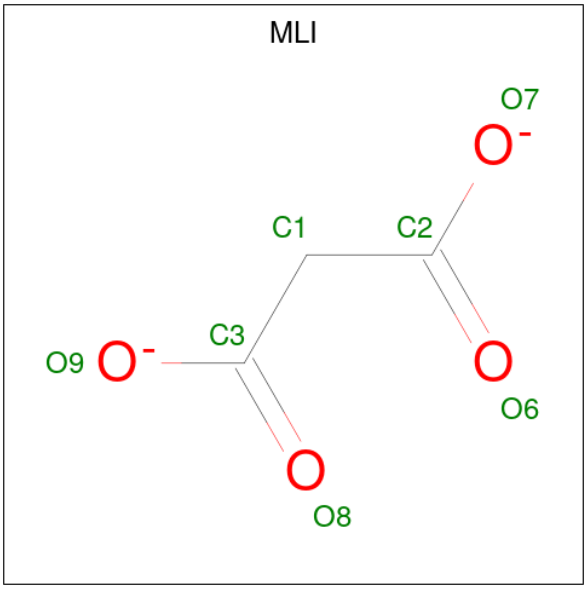
There are 4 unique types of molecules in this entry. The entry contains 3044 atoms, of which 0 are hydrogens and 0 are deuteriums.

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
			Total	Ni		
3	A	1	1	1	0	0

- Molecule 4 is water.

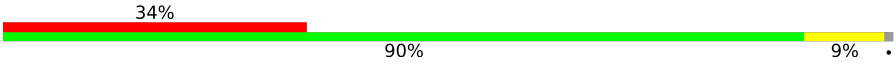
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	310	310	310	0	0

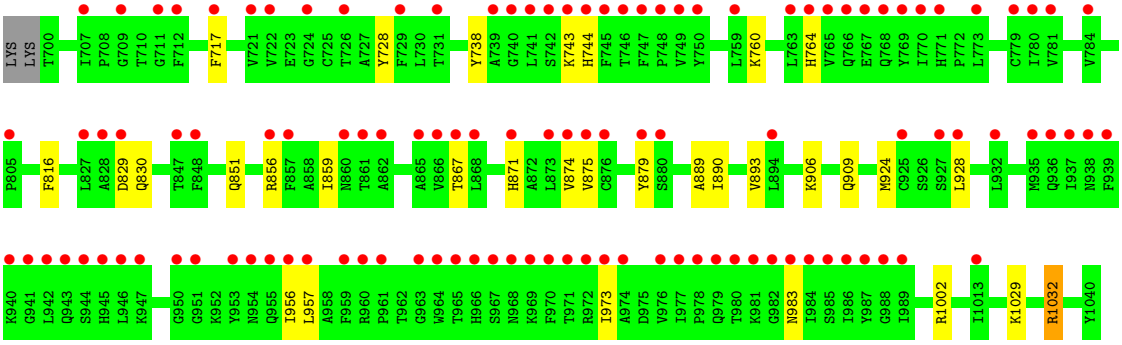
3 Residue-property plots (i)

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: DCLRE1A

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	51.80Å 57.89Å 114.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.63 – 1.63 28.63 – 1.63	Depositor EDS
% Data completeness (in resolution range)	99.9 (28.63-1.63) 99.9 (28.63-1.63)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.46 (at 1.63Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.312 , 0.374 0.331 , 0.395	Depositor DCC
R_{free} test set	2114 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å^2)	23.4	Xtriage
Anisotropy	0.482	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 50.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	3044	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.17% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.63	0/2802	0.76	2/3809 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1032	ARG	NE-CZ-NH1	-5.23	117.69	120.30
1	A	1002	ARG	NE-CZ-NH2	-5.17	117.72	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	19	0
2	A	7	0	2	0	0
3	A	1	0	0	0	0
4	A	310	0	0	12	1
All	All	3044	0	2678	19	1

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 (19) 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:1202:HOH:O	2.11	0.84
1:A:867:THR:O	4:A:1201:HOH:O	2.08	0.70
1:A:906:LYS:NZ	4:A:1211:HOH:O	2.42	0.53
1:A:1032:ARG:NE	4:A:1216:HOH:O	2.44	0.50
1:A:743:LYS:HG3	1:A:744:HIS:CE1	2.46	0.50
1:A:875:VAL:HG23	1:A:956:ILE:HG23	1.95	0.48
1:A:1032:ARG:CZ	4:A:1216:HOH:O	2.62	0.48
1:A:871:HIS:CE1	4:A:1338:HOH:O	2.66	0.47
1:A:851:GLN:NE2	4:A:1223:HOH:O	2.47	0.47
1:A:760:LYS:O	1:A:764[A]:HIS:HA	2.15	0.46
1:A:983:ASN:ND2	4:A:1219:HOH:O	2.45	0.45
1:A:1029:LYS:NZ	4:A:1207:HOH:O	2.37	0.44
1:A:973:ILE:N	4:A:1227:HOH:O	2.50	0.44
1:A:924:MET:N	4:A:1218:HOH:O	2.45	0.44
1:A:973:ILE:HG22	4:A:1227:HOH:O	2.17	0.44
1:A:717:PHE:CE1	1:A:728:TYR:HB3	2.54	0.43
1:A:859:ILE:HD11	1:A:889:ALA:HB1	2.00	0.42
1:A:874:VAL:HG11	1:A:890:ILE:HG21	2.03	0.41
1:A:856:ARG:HA	1:A:856:ARG:HE	1.86	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:1376:HOH:O	4:A:1420:HOH:O[3_545]	2.17	0.03

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%)	327 (95%)	16 (5%)	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	299/305 (98%)	291 (97%)	8 (3%)	44	18

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

Mol	Chain	Res	Type
1	A	738	TYR
1	A	816	PHE
1	A	829	ASP
1	A	830	GLN
1	A	879	TYR
1	A	893	VAL
1	A	928	LEU
1	A	957	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	744	HIS
1	A	851	GLN
1	A	909	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 monosaccharides 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	6,6,6	1.26	0	7,7,7	1.00	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	MLI	A	1101	3	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1101	MLI	C3-C1-C2-O6
2	A	1101	MLI	C3-C1-C2-O7
2	A	1101	MLI	C2-C1-C3-O8
2	A	1101	MLI	C2-C1-C3-O9

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, 95th 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^2)	Q<0.9
1	A	341/343 (99%)	2.03	115 (33%) 0 0	20, 40, 105, 153	0

All (115) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	970	PHE	20.4
1	A	966	HIS	19.0
1	A	984	ILE	15.5
1	A	744	HIS	14.7
1	A	768	GLN	10.7
1	A	939	PHE	10.6
1	A	967	SER	10.3
1	A	980	THR	10.0
1	A	769	TYR	10.0
1	A	977	ILE	9.8
1	A	987	TYR	9.6
1	A	745	PHE	8.9
1	A	973	ILE	8.9
1	A	979	GLN	8.7
1	A	972	ARG	8.7
1	A	986	ILE	8.6
1	A	937	ILE	8.5
1	A	942	LEU	8.0
1	A	941	GLY	7.5
1	A	765	VAL	7.1
1	A	985	SER	7.0
1	A	721	VAL	6.9
1	A	956	ILE	6.9
1	A	982	GLY	6.9
1	A	981	LYS	6.6
1	A	943	GLN	6.6
1	A	954	ASN	6.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	968	ASN	6.4
1	A	711	GLY	6.2
1	A	971	THR	6.1
1	A	866	VAL	5.7
1	A	976	VAL	5.5
1	A	747	PHE	5.5
1	A	743	LYS	5.3
1	A	978	PRO	5.2
1	A	781	VAL	5.1
1	A	750	TYR	5.1
1	A	983	ASN	5.1
1	A	974	ALA	4.9
1	A	865	ALA	4.7
1	A	938	ASN	4.6
1	A	969	LYS	4.5
1	A	741	LEU	4.4
1	A	946	LEU	4.4
1	A	957	LEU	4.3
1	A	928	LEU	4.2
1	A	935	MET	4.2
1	A	767	GLU	4.2
1	A	724	GLY	4.2
1	A	960	ARG	4.1
1	A	770	ILE	4.1
1	A	955	GLN	4.1
1	A	959	PHE	4.1
1	A	944	SER	4.0
1	A	950	GLY	3.9
1	A	951	GLY	3.9
1	A	868	LEU	3.8
1	A	940	LYS	3.8
1	A	988	GLY	3.8
1	A	856	ARG	3.8
1	A	963	GLY	3.7
1	A	764[A]	HIS	3.6
1	A	947	LYS	3.5
1	A	876[A]	CYS	3.5
1	A	780	ILE	3.5
1	A	746	THR	3.5
1	A	932	LEU	3.4
1	A	874	VAL	3.4
1	A	936	GLN	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	766	GLN	3.4
1	A	722	VAL	3.3
1	A	759	LEU	3.3
1	A	763	LEU	3.3
1	A	873	LEU	3.3
1	A	717	PHE	3.2
1	A	739	ALA	3.2
1	A	742	SER	3.2
1	A	871	HIS	3.2
1	A	828	ALA	3.2
1	A	779	CYS	3.1
1	A	879	TYR	3.1
1	A	945	HIS	3.1
1	A	729	PHE	3.1
1	A	953	TYR	3.0
1	A	961	PRO	3.0
1	A	862	ALA	2.9
1	A	726	THR	2.9
1	A	857	PHE	2.7
1	A	964	TRP	2.7
1	A	749	VAL	2.7
1	A	829	ASP	2.6
1	A	847	THR	2.6
1	A	894	LEU	2.5
1	A	989	ILE	2.5
1	A	927	SER	2.5
1	A	861	THR	2.4
1	A	848	PHE	2.4
1	A	860	ASN	2.3
1	A	1013	ILE	2.3
1	A	880	SER	2.2
1	A	707	ILE	2.2
1	A	731	THR	2.2
1	A	784	VAL	2.2
1	A	748	PRO	2.2
1	A	771	HIS	2.2
1	A	965	THR	2.2
1	A	925	CYS	2.1
1	A	712	PHE	2.1
1	A	867	THR	2.1
1	A	805	PRO	2.1
1	A	709	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	773	LEU	2.0
1	A	827	LEU	2.0
1	A	875	VAL	2.0
1	A	740	GLY	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, 95th 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^2)	Q<0.9
2	MLI	A	1101	7/7	0.93	0.12	28,36,40,43	0
3	NI	A	1102	1/1	1.00	0.04	24,24,24,24	0

6.5 Other polymers (i)

There are no such residues in this entry.