W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Sep 11, 2023 – 06:24 AM EDT

PDB ID	:	4LDX
Title	:	Crystal structure of the DNA binding domain of arabidopsis thaliana auxin response factor 1 (ARF1) in complex with protomor-like sequence ER7
Authors	:	Boer, D.R.; Freire-Rios, A.; van den Berg, W.M.A.; Weijers, D.; Coll, M.
Deposited on	:	2013-06-25
Resolution	:	2.90 Å(reported)

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

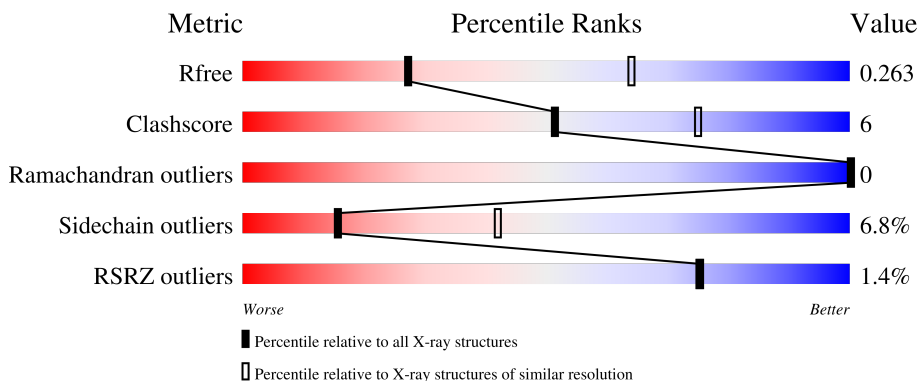|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.90 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	Percentile Ranks	Value
Rfree		0.263
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		6.8%
RSRZ outliers		1.4%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	363	% — 74% / 15% · 9%
1	B	363	2% — 75% / 18% · 7%
2	C	21	67% / 33%
3	D	21	5% — 71% / 29%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 6274 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 Auxin response factor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	332	Total	C	N	O	S	0	0	0
			2671	1684	481	492	14			
1	B	337	Total	C	N	O	S	0	0	0
			2715	1713	487	501	14			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	356	ASN	-	expression tag	UNP Q8L7G0
A	357	SER	-	expression tag	UNP Q8L7G0
A	358	TYR	-	expression tag	UNP Q8L7G0
A	359	SER	-	expression tag	UNP Q8L7G0
A	360	GLN	-	expression tag	UNP Q8L7G0
A	361	SER	-	expression tag	UNP Q8L7G0
A	362	MET	-	expression tag	UNP Q8L7G0
A	363	CYS	-	expression tag	UNP Q8L7G0
B	356	ASN	-	expression tag	UNP Q8L7G0
B	357	SER	-	expression tag	UNP Q8L7G0
B	358	TYR	-	expression tag	UNP Q8L7G0
B	359	SER	-	expression tag	UNP Q8L7G0
B	360	GLN	-	expression tag	UNP Q8L7G0
B	361	SER	-	expression tag	UNP Q8L7G0
B	362	MET	-	expression tag	UNP Q8L7G0
B	363	CYS	-	expression tag	UNP Q8L7G0

- Molecule 2 is a DNA chain called ER7, forward sequence.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	21	Total	C	N	O	P	0	0	0
			426	205	74	127	20			

- Molecule 3 is a DNA chain called ER7, reverse sequence.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	21	Total 429	C 205	N 83	O 121	P 20	0	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	16	Total 16	O 16	0	0
4	B	15	Total 15	O 15	0	0
4	D	2	Total 2	O 2	0	0

3 Residue-property plots ⓘ

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.
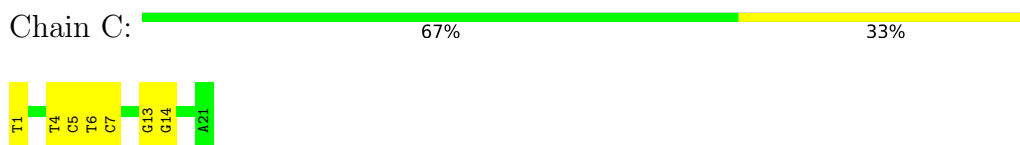
• Molecule 1: Auxin response factor 1

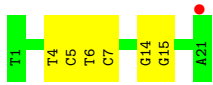Chain A:

• Molecule 1: Auxin response factor 1

Chain B:
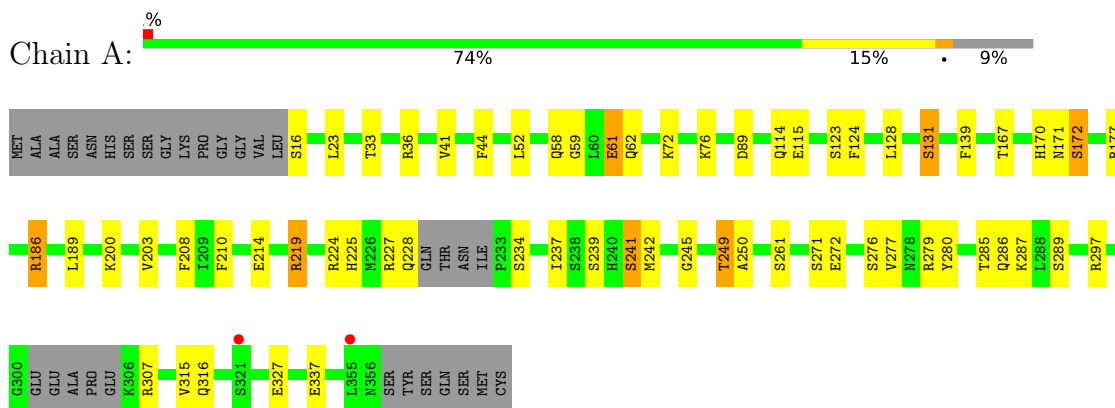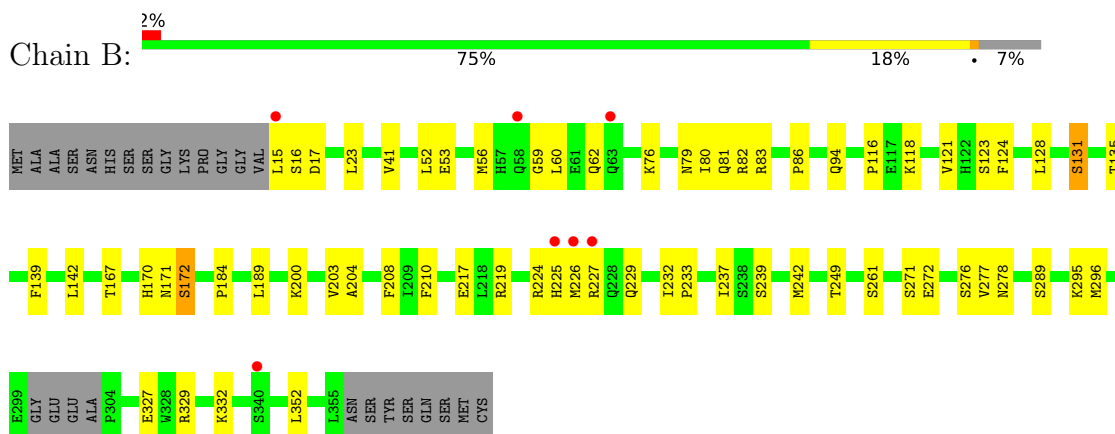
• Molecule 2: ER7, forward sequence

Chain C:

• Molecule 3: ER7, reverse sequence

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	43.52Å 105.19Å 127.91Å 90.00° 98.14° 90.00°	Depositor
Resolution (Å)	33.81 – 2.90 33.79 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.4 (33.81-2.90) 98.6 (33.79-2.90)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.80 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.214 , 0.259 0.218 , 0.263	Depositor DCC
R_{free} test set	1203 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å^2)	42.2	Xtriage
Anisotropy	0.104	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 41.7	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.047 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	6274	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.62% 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)

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	0.64	0/2742	0.79	3/3715 (0.1%)
1	B	0.58	0/2788	0.78	1/3781 (0.0%)
2	C	0.45	0/476	0.85	0/733
3	D	0.46	0/482	0.83	0/742
All	All	0.59	0/6488	0.79	4/8971 (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
1	B	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	219	ARG	NE-CZ-NH1	-6.68	116.96	120.30
1	A	219	ARG	NE-CZ-NH2	5.62	123.11	120.30
1	A	89	ASP	CB-CG-OD1	5.53	123.27	118.30
1	B	329	ARG	NE-CZ-NH1	5.04	122.82	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	116	PRO	Peptide

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	A	2671	0	2607	31	0
1	B	2715	0	2655	31	0
2	C	426	0	240	5	0
3	D	429	0	237	4	0
4	A	16	0	0	4	0
4	B	15	0	0	3	0
4	D	2	0	0	0	0
All	All	6274	0	5739	67	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 6.

The worst 5 of 67 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:245:GLY:O	1:A:249:THR:HB	1.92	0.69
1:A:186:ARG:HH11	1:A:186:ARG:HB2	1.58	0.68
1:B:200:LYS:O	1:B:224:ARG:NH2	2.29	0.66
1:B:296:MET:HE2	1:B:352:LEU:HD23	1.77	0.66
1:A:297:ARG:HD3	1:A:307:ARG:NH2	2.11	0.65

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	
1	A	326/363 (90%)	313 (96%)	13 (4%)	0	100	100
1	B	333/363 (92%)	318 (96%)	15 (4%)	0	100	100
All	All	659/726 (91%)	631 (96%)	28 (4%)	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	298/323 (92%)	277 (93%)	21 (7%)	15	41
1	B	304/323 (94%)	284 (93%)	20 (7%)	16	44
All	All	602/646 (93%)	561 (93%)	41 (7%)	16	42

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	121	VAL
1	B	239	SER
1	B	131	SER
1	B	172	SER
1	B	271	SER

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	114	GLN
1	B	79	ASN
1	B	229	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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 ⓘ

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(Å²)	Q<0.9
1	A	332/363 (91%)	-0.43	2 (0%)	89	89	19, 30, 76, 120	0
1	B	337/363 (92%)	-0.26	7 (2%)	63	61	21, 46, 91, 123	0
2	C	21/21 (100%)	-0.36	0	100	100	35, 77, 88, 102	0
3	D	21/21 (100%)	-0.14	1 (4%)	30	27	28, 68, 88, 122	0
All	All	711/768 (92%)	-0.34	10 (1%)	75	75	19, 38, 88, 123	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	225	HIS	3.0
1	B	340	SER	2.7
1	B	58	GLN	2.5
1	B	15	LEU	2.5
3	D	21	DA	2.3

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.5 Other polymers ⓘ

There are no such residues in this entry.