W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Sep 3, 2023 – 01:03 AM EDT

PDB ID	:	3R50
Title	:	Structure analysis of a wound-inducible lectin ipomoelin from sweet potato
Authors	:	Liu, K.L.; Chang, W.C.; Jeng, S.T.; Cheng, Y.S.
Deposited on	:	2011-03-18
Resolution	:	2.27 Å(reported)

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

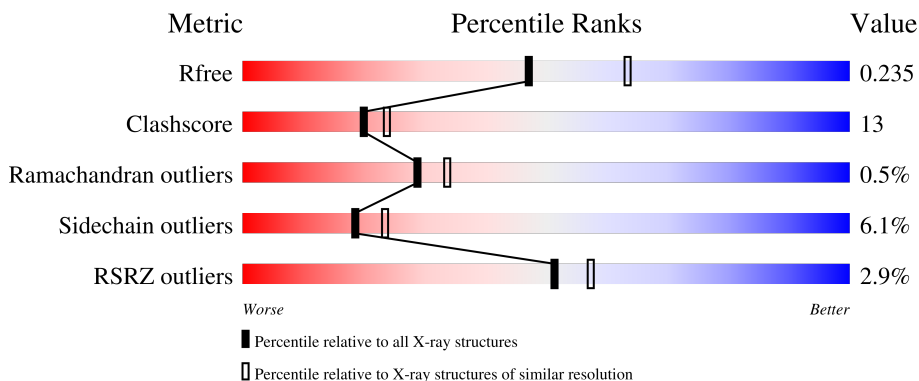|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1   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 2.27 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		0.235
Clashscore		13
Ramachandran outliers		0.5%
Sidechain outliers		6.1%
RSRZ outliers		2.9%

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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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	160	3% / 70% / 22% / · ·
1	B	160	4% / 69% / 26% / · ·
1	C	160	4% / 68% / 26% / · ·
1	D	160	.% / 74% / 20% / · ·
1	E	160	2% / 76% / 14% / · / 6%

2 Entity composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6164 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 Ipomoelin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	153	Total	C	N	O	S	0	0	0
			1160	744	192	223	1			
1	B	154	Total	C	N	O	S	0	0	0
			1167	749	193	223	2			
1	C	153	Total	C	N	O	S	0	0	0
			1160	744	192	223	1			
1	D	153	Total	C	N	O	S	0	0	0
			1160	744	192	223	1			
1	E	151	Total	C	N	O	S	0	0	0
			1147	735	190	221	1			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	HIS	-	expression tag	UNP P93193
A	-4	HIS	-	expression tag	UNP P93193
A	-3	HIS	-	expression tag	UNP P93193
A	-2	HIS	-	expression tag	UNP P93193
A	-1	HIS	-	expression tag	UNP P93193
A	0	HIS	-	expression tag	UNP P93193
B	-5	HIS	-	expression tag	UNP P93193
B	-4	HIS	-	expression tag	UNP P93193
B	-3	HIS	-	expression tag	UNP P93193
B	-2	HIS	-	expression tag	UNP P93193
B	-1	HIS	-	expression tag	UNP P93193
B	0	HIS	-	expression tag	UNP P93193
C	-5	HIS	-	expression tag	UNP P93193
C	-4	HIS	-	expression tag	UNP P93193
C	-3	HIS	-	expression tag	UNP P93193
C	-2	HIS	-	expression tag	UNP P93193
C	-1	HIS	-	expression tag	UNP P93193
C	0	HIS	-	expression tag	UNP P93193
D	-5	HIS	-	expression tag	UNP P93193

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-4	HIS	-	expression tag	UNP P93193
D	-3	HIS	-	expression tag	UNP P93193
D	-2	HIS	-	expression tag	UNP P93193
D	-1	HIS	-	expression tag	UNP P93193
D	0	HIS	-	expression tag	UNP P93193
E	-5	HIS	-	expression tag	UNP P93193
E	-4	HIS	-	expression tag	UNP P93193
E	-3	HIS	-	expression tag	UNP P93193
E	-2	HIS	-	expression tag	UNP P93193
E	-1	HIS	-	expression tag	UNP P93193
E	0	HIS	-	expression tag	UNP P93193

- Molecule 2 is water.

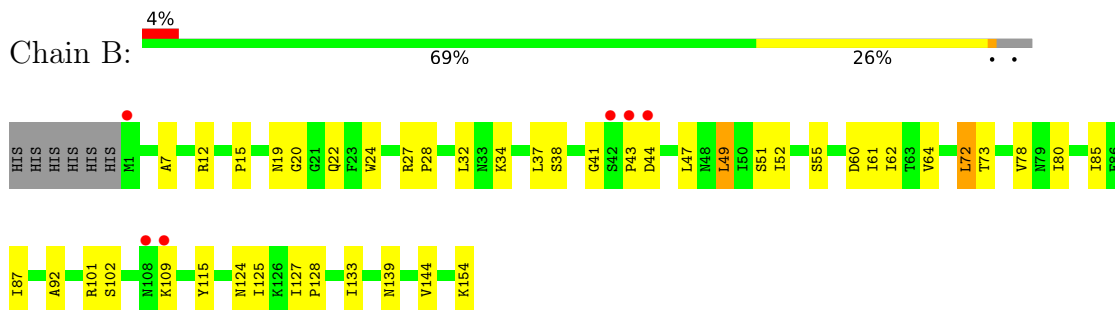| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	49	Total 49	O 49	0	0
2	B	54	Total 54	O 54	0	0
2	C	62	Total 62	O 62	0	0
2	D	103	Total 103	O 103	0	0
2	E	102	Total 102	O 102	0	0

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.
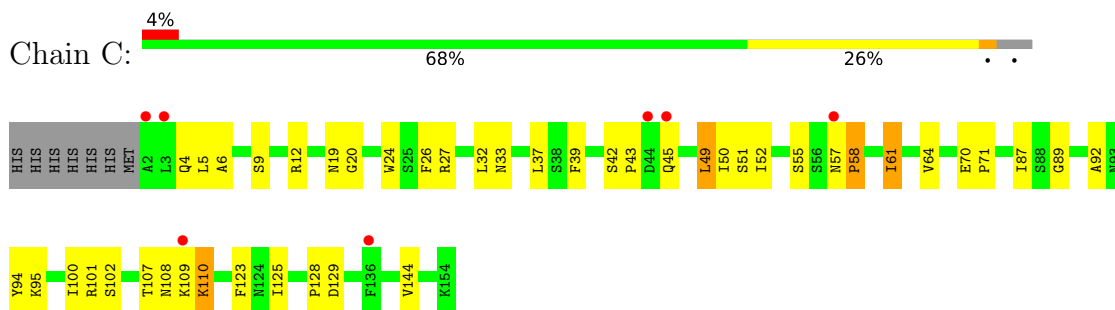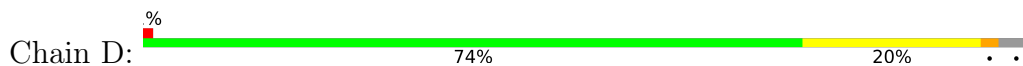
- Molecule 1: Ipomoelin
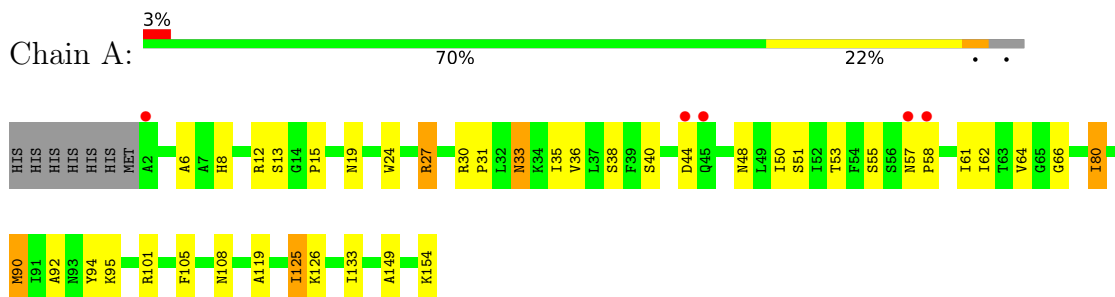
Chain A:

- Molecule 1: Ipomoelin

Chain B:

- Molecule 1: Ipomoelin

Chain C:

- Molecule 1: Ipomoelin

Chain D:

HIS HIS HIS HIS HIS MET A2 L3 Q4 L5 A6 A7 H8 R12 P15 V16 G17 S18 N19 F23 W24 R27 R30 S38 F39 S40 G41 S42 Q45 T46 L49 S56 N57 V64 L72 I87 I91 Y94 K95 I100 K104 N118

A119 I125 K126 V144 K154

• Molecule 1: Ipomoelin

Chain E:

2%

76%   14%   •   6%

HIS HIS HIS HIS HIS MET ALA LEU Q4 P15 W24 R27 S38 F39 Q45 T46 L49 I50 N57 P58 T59 D60 I61 V64 L72 E76 T77 V78 N79 I80 I85 E86 I87 A92 I100 R101 K109 K110 Y115 N124 I125 G141

K154
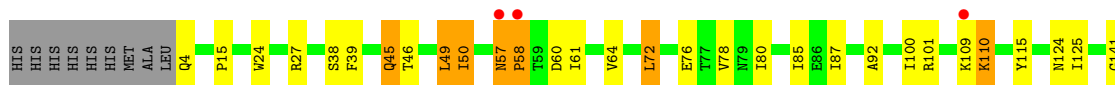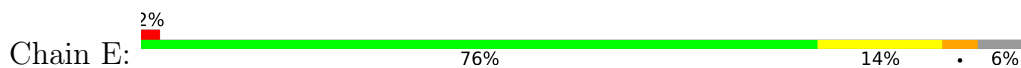
4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	87.49Å 139.47Å 189.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.21 – 2.27 26.21 – 2.27	Depositor EDS
% Data completeness (in resolution range)	91.7 (26.21-2.27) 91.8 (26.21-2.27)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.36 (at 2.29Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.202 , 0.234 0.202 , 0.235	Depositor DCC
R_{free} test set	2472 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å²)	33.3	Xtriage
Anisotropy	0.357	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.29 , 35.3	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6164	wwPDB-VP
Average B, all atoms (Å²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% 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 ⓘ

5.1 Standard geometry ⓘ

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	0.66	0/1190	0.75	0/1619
1	B	0.67	0/1197	0.79	0/1628
1	C	0.67	0/1190	0.76	0/1619
1	D	0.74	0/1190	0.83	1/1619 (0.1%)
1	E	0.73	0/1177	0.83	1/1601 (0.1%)
All	All	0.69	0/5944	0.79	2/8086 (0.0%)

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	E	141	GLY	N-CA-C	-5.32	99.80	113.10
1	D	40	SER	N-CA-C	-5.24	96.85	111.00

There are no chirality outliers.

There are no planarity outliers.

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	1160	0	1136	34	0
1	B	1167	0	1145	38	0
1	C	1160	0	1136	40	0
1	D	1160	0	1136	28	0
1	E	1147	0	1120	26	0
2	A	49	0	0	2	0

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	54	0	0	1	0
2	C	62	0	0	2	0
2	D	103	0	0	2	0
2	E	102	0	0	3	0
All	All	6164	0	5673	145	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 145 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:C:129:ASP:HB3	2:C:219:HOH:O	1.61	1.01
1:C:33:ASN:HD21	1:C:57:ASN:H	1.17	0.93
1:C:107:THR:HG23	1:C:109:LYS:H	1.37	0.88
1:B:20:GLY:H	1:B:139:ASN:HD22	1.31	0.77
1:C:50:ILE:HD12	1:C:51:SER:N	2.00	0.77

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	151/160 (94%)	140 (93%)	11 (7%)	0	100	100
1	B	152/160 (95%)	148 (97%)	4 (3%)	0	100	100
1	C	151/160 (94%)	144 (95%)	6 (4%)	1 (1%)	22	25
1	D	151/160 (94%)	145 (96%)	4 (3%)	2 (1%)	12	11
1	E	149/160 (93%)	145 (97%)	3 (2%)	1 (1%)	22	25
All	All	754/800 (94%)	722 (96%)	28 (4%)	4 (0%)	29	34

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	58	PRO
1	D	119	ALA
1	E	58	PRO
1	D	57	ASN

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	
1	A	124/131 (95%)	115 (93%)	9 (7%)	14	16
1	B	124/131 (95%)	117 (94%)	7 (6%)	21	27
1	C	124/131 (95%)	118 (95%)	6 (5%)	25	34
1	D	124/131 (95%)	117 (94%)	7 (6%)	21	27
1	E	123/131 (94%)	114 (93%)	9 (7%)	14	16
All	All	619/655 (94%)	581 (94%)	38 (6%)	18	23

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	126	LYS
1	E	72	LEU
1	E	15	PRO
1	E	49	LEU
1	E	124	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	33	ASN
1	D	45	GLN
1	E	124	ASN
1	D	124	ASN
1	E	57	ASN

5.3.3 RNA ⓘ

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(Å^2)	Q<0.9
1	A	153/160 (95%)	-0.27	5 (3%)	46	52	20, 37, 61, 75	1 (0%)
1	B	154/160 (96%)	-0.27	6 (3%)	39	44	22, 33, 57, 72	1 (0%)
1	C	153/160 (95%)	0.04	7 (4%)	32	38	23, 42, 70, 81	1 (0%)
1	D	153/160 (95%)	-0.44	1 (0%)	87	90	21, 28, 49, 65	1 (0%)
1	E	151/160 (94%)	-0.38	3 (1%)	65	70	19, 28, 45, 63	1 (0%)
All	All	764/800 (95%)	-0.26	22 (2%)	51	57	19, 32, 60, 81	5 (0%)

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	3	LEU	7.3
1	B	42	SER	4.9
1	D	2	ALA	4.6
1	B	43	PRO	4.4
1	E	57	ASN	4.2

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.

6.5 Other polymers ⓘ

There are no such residues in this entry.