wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 06:43 am BST

PDB ID	:	5XKN
Title	:	Crystal structure of plant receptor ERL2 in complexe with EPFL4
Authors	:	Chai, J.; Lin, G.
Deposited on	:	2017-05-08
Resolution	:	3.65 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see references (1)) were used in the production of this report:

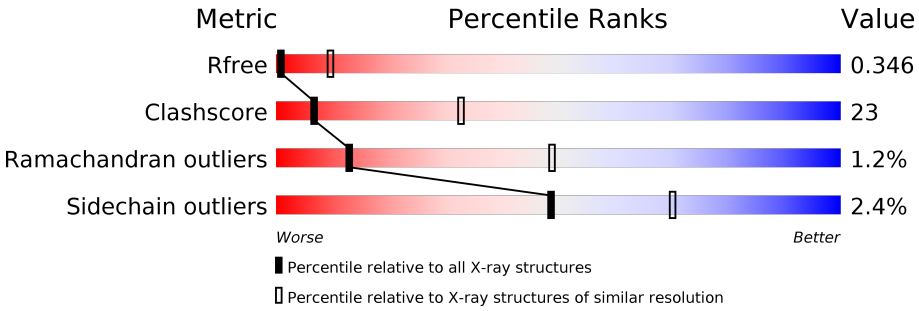
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.65 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	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)

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 >=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%

Mol	Chain	Length	Quality of chain
1	E	51	45% 22% 33%
1	F	51	37% 29% 33%
2	A	552	65% 31% ..
2	B	552	65% 32% ..

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8824 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 EPIDERMAL PATTERNING FACTOR-like protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	34	Total	C	N	O	S	0	0	0
			269	175	45	45	4			
1	F	34	Total	C	N	O	S	0	0	0
			269	175	45	45	4			

- Molecule 2 is a protein called LRR receptor-like serine/threonine-protein kinase ERL2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	540	Total	C	N	O	S	0	0	0
			4143	2646	696	785	16			
2	B	540	Total	C	N	O	S	0	0	0
			4143	2646	696	785	16			

There are 12 discrepancies between the modelled and reference sequences:

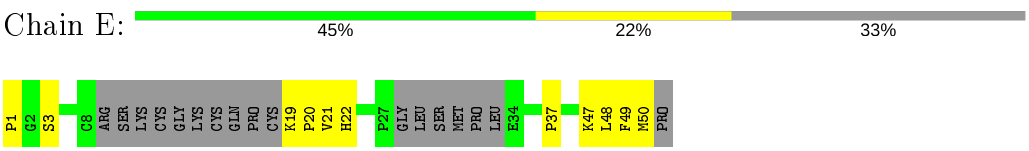
Chain	Residue	Modelled	Actual	Comment	Reference
A	576	HIS	-	expression tag	UNP Q6XAT2
A	577	HIS	-	expression tag	UNP Q6XAT2
A	578	HIS	-	expression tag	UNP Q6XAT2
A	579	HIS	-	expression tag	UNP Q6XAT2
A	580	HIS	-	expression tag	UNP Q6XAT2
A	581	HIS	-	expression tag	UNP Q6XAT2
B	576	HIS	-	expression tag	UNP Q6XAT2
B	577	HIS	-	expression tag	UNP Q6XAT2
B	578	HIS	-	expression tag	UNP Q6XAT2
B	579	HIS	-	expression tag	UNP Q6XAT2
B	580	HIS	-	expression tag	UNP Q6XAT2
B	581	HIS	-	expression tag	UNP Q6XAT2

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. 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. 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.

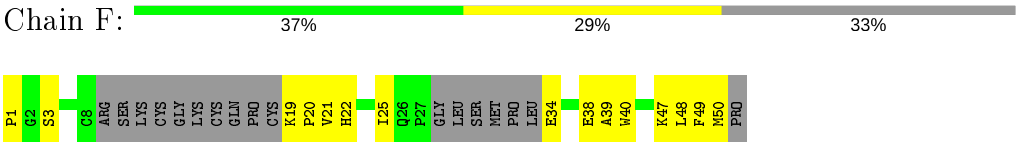
- Molecule 1: EPIDERMAL PATTERNING FACTOR-like protein 4

Chain E:



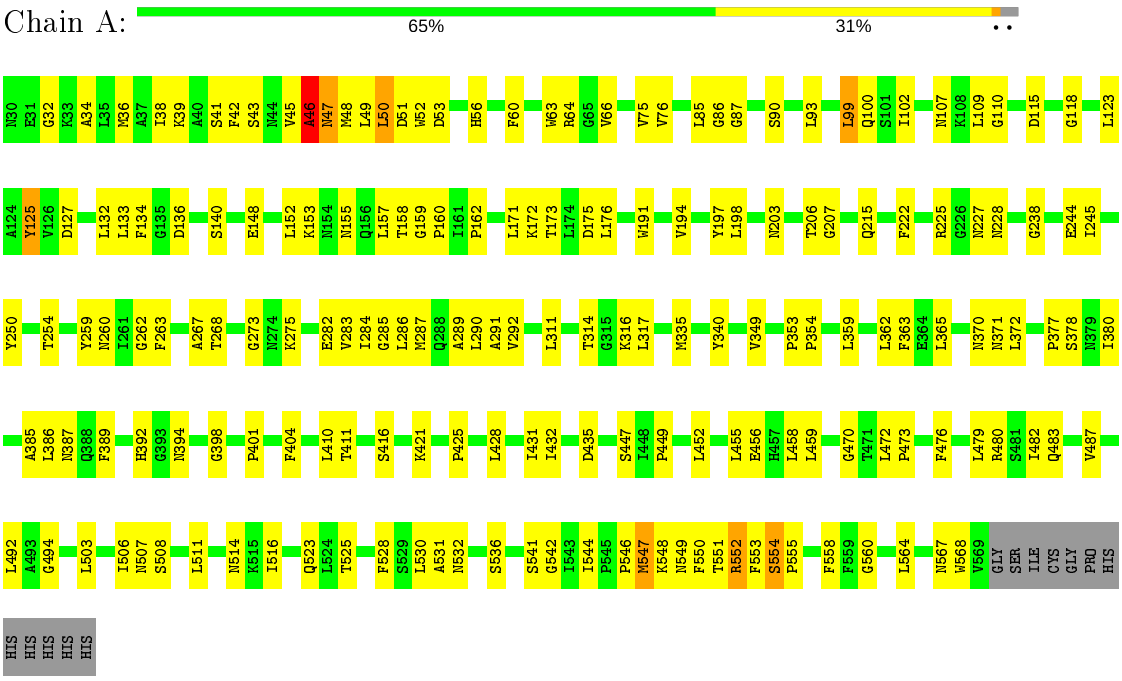
- Molecule 1: EPIDERMAL PATTERNING FACTOR-like protein 4

Chain F:



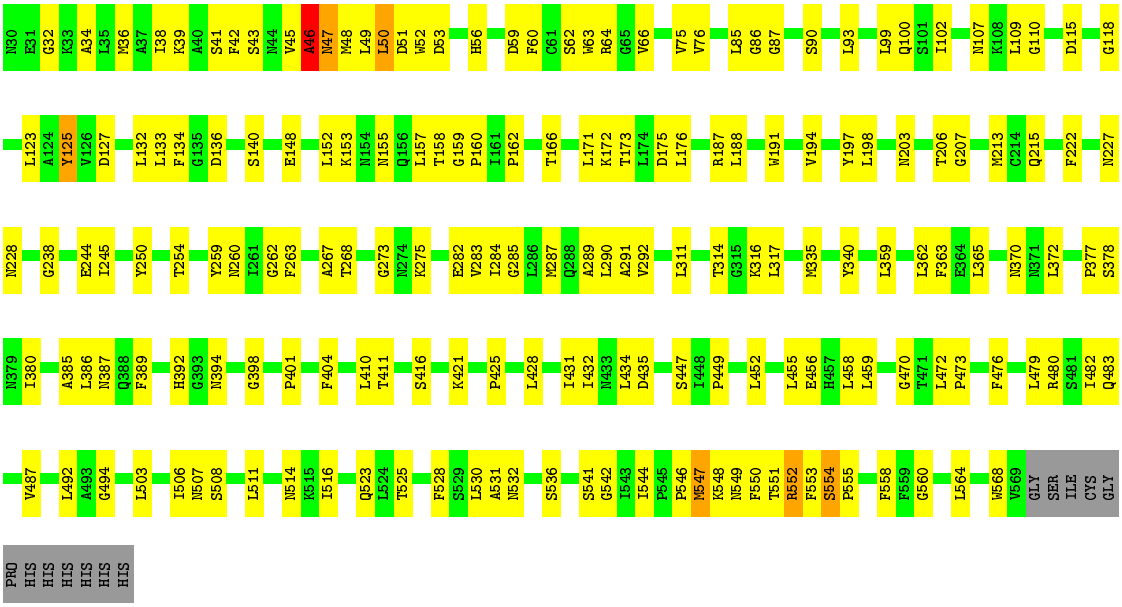
- Molecule 2: LRR receptor-like serine/threonine-protein kinase ERL2

Chain A:



- Molecule 2: LRR receptor-like serine/threonine-protein kinase ERL2

Chain B:


N30 E31 G32 K33 A34 L35 M36 A37 I38 K39 A40 S41 F42 S43 N44 V45 A46 N47 M48 L49 L50 D51 W52 D53 H56 D59 F60 G61 S62 W63 R64 G65 V66 V75 V76 L85 G86 G87 S90 L93 L99 Q100 S101 I102 N107 K108 L109 G110 D115 G118
L123 A124 Y125 V126 D127 L132 L133 F134 G135 D136 S140 E148 L152 K153 N154 N155 Q156 L157 T158 G159 P160 I161 P162 T166 L171 K172 T173 L174 D175 L176 R187 L188 W191 V194 Y197 L198 N203 T206 G207 M213 C214 Q215 F222 N227
N228 G238 E244 I245 Y250 T254 Y259 N260 I261 G262 F263 A267 T268 G273 N274 K275 E282 V283 I284 G285 L286 M287 Q288 A289 L290 A291 V292 L311 T314 G315 K316 L317 M335 Y340 L359 L362 F363 E364 L365 N370 N371 L372 P377 S378
N379 I380 A385 L386 N387 Q388 F389 H392 G393 N394 G398 P401 F404 L410 T411 S416 K421 P425 L428 I431 I432 N433 L434 D435 S447 I448 P449 L452 L455 E456 H457 L458 L459 G470 T471 L472 P473 F476 L479 R480 S481 I482 Q483
V487 L492 A493 G494 L503 I506 N507 S508 L511 N514 K515 I516 Q523 L524 T525 F528 S529 L530 A531 N532 S536 S541 G542 I543 I544 P545 P546 M547 K548 N549 F550 T551 R552 F553 S554 P555 F558 F559 G560 L564 W568 V569 GLY SER ILE CYS GLY
PRO HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	109.00Å 112.33Å 175.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.30 – 3.65 47.29 – 3.65	Depositor EDS
% Data completeness (in resolution range)	88.7 (47.30-3.65) 88.4 (47.29-3.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.85 (at 3.67Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.282 , 0.312 0.336 , 0.346	Depositor DCC
R_{free} test set	1096 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	93.4	Xtriage
Anisotropy	0.662	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 55.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	8824	wwPDB-VP
Average B, all atoms (Å^2)	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% 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	E	0.34	0/279	0.59	0/377
1	F	0.34	0/279	0.59	0/377
2	A	0.33	0/4224	0.60	2/5750 (0.0%)
2	B	0.33	0/4224	0.60	2/5750 (0.0%)
All	All	0.33	0/9006	0.60	4/12254 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	46	ALA	CB-CA-C	8.65	123.08	110.10
2	A	46	ALA	CB-CA-C	8.64	123.06	110.10
2	A	51	ASP	N-CA-C	5.60	126.11	111.00
2	B	51	ASP	N-CA-C	5.59	126.10	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	E	269	0	261	16	0
1	F	269	0	261	21	0
2	A	4143	0	4155	197	4
2	B	4143	0	4158	195	4
All	All	8824	0	8835	413	4

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 23.

The worst 5 of 413 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)
2:B:525:THR:CB	2:B:550:PHE:HZ	1.33	1.40
2:A:525:THR:CB	2:A:550:PHE:CZ	2.04	1.40
2:A:525:THR:CB	2:A:550:PHE:HZ	1.33	1.38
2:B:525:THR:CB	2:B:550:PHE:CZ	2.04	1.36
2:A:554:SER:HB2	2:A:555:PRO:CD	1.59	1.32

All (4) 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)
2:A:349:VAL:CG1	2:B:187:ARG:NE[2_555]	1.63	0.57
2:A:349:VAL:CG1	2:B:187:ARG:CZ[2_555]	1.90	0.30
2:A:567:ASN:O	2:B:48:MET:SD[3_645]	1.95	0.25
2:A:371:ASN:CB	2:B:187:ARG:NH1[2_555]	2.01	0.19

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	E	28/51 (55%)	23 (82%)	4 (14%)	1 (4%)	3	29
1	F	28/51 (55%)	23 (82%)	4 (14%)	1 (4%)	3	29
2	A	538/552 (98%)	476 (88%)	56 (10%)	6 (1%)	14	51
2	B	538/552 (98%)	476 (88%)	56 (10%)	6 (1%)	14	51
All	All	1132/1206 (94%)	998 (88%)	120 (11%)	14 (1%)	13	49

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	46	ALA
2	A	546	PRO
2	B	46	ALA
2	B	546	PRO
2	A	432	ILE

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	E	31/46 (67%)	31 (100%)	0	100	100
1	F	31/46 (67%)	31 (100%)	0	100	100
2	A	471/481 (98%)	459 (98%)	12 (2%)	47	69
2	B	471/481 (98%)	459 (98%)	12 (2%)	47	69
All	All	1004/1054 (95%)	980 (98%)	24 (2%)	49	70

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	549	ASN
2	B	53	ASP
2	B	549	ASN
2	A	552	ARG
2	B	47	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	549	ASN
2	B	549	ASN
2	B	394	ASN
2	A	394	ASN
2	B	387	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 carbohydrates 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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.