



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 31, 2021 – 12:04 PM EDT

PDB ID : 7K2L
Title : Kelch domain of human KEAP1 bound to Nrf2 cyclic peptide, c[BAL-NPETGE]
Authors : Muellers, S.N.; Allen, K.N.
Deposited on : 2020-09-08
Resolution : 1.98 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 ⓘ 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	:	2.18
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)	:	2.18

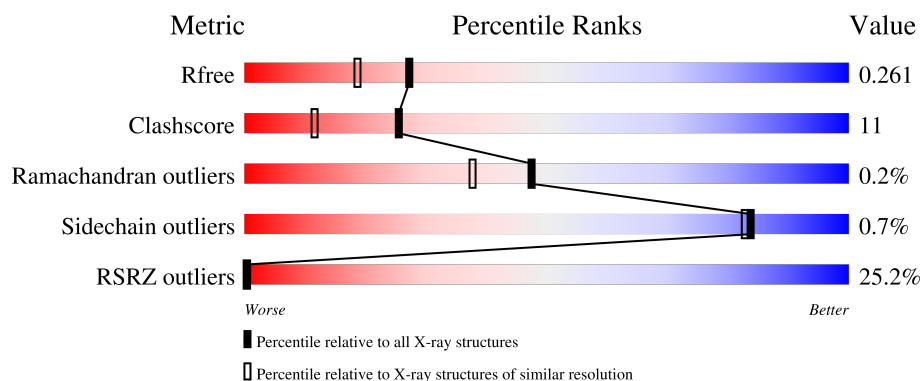
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.98 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 $\leq 5\%$. 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	B	289	<div> <div>12%</div> <div>84%</div> <div>14%</div> </div>
2	A	284	<div> <div>39%</div> <div>82%</div> <div>18%</div> </div>
3	P	7	<div> <div>29%</div> <div>57%</div> <div>14%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	BAL	P	76	-	X	-	-

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4547 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 Kelch-like ECH-associated protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	289	Total	C	N	O	S	0	1	0
			2219	1380	401	421	17			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	540	ALA	GLU	conflict	UNP Q14145
B	542	ALA	GLU	conflict	UNP Q14145
B	613	SER	CYS	conflict	UNP Q14145

- Molecule 2 is a protein called Kelch-like ECH-associated protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	284	Total	C	N	O	S	0	1	0
			2126	1325	381	405	15			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	540	ALA	GLU	conflict	UNP Q14145
A	542	ALA	GLU	conflict	UNP Q14145

- Molecule 3 is a protein called Nrf2 cyclic peptide,c[BAL-NPETGE].

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	P	7	Total	C	N	O	0	0	0
			49	28	8	13			

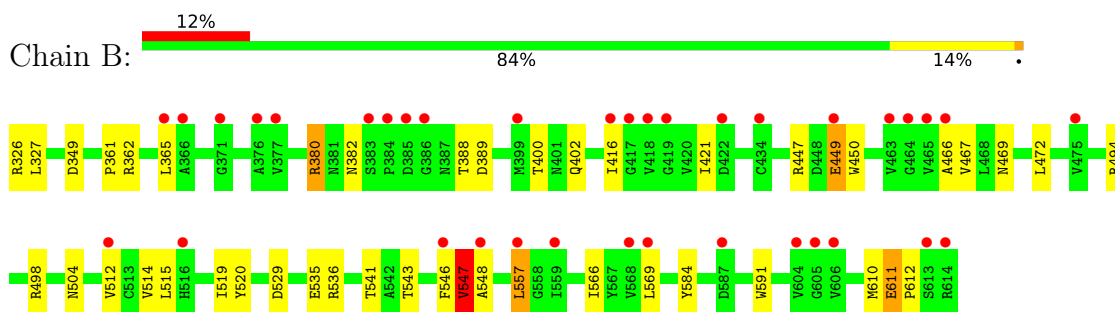
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	122	Total 122	O 122	0	0
4	A	28	Total 28	O 28	0	0
4	P	3	Total 3	O 3	0	0

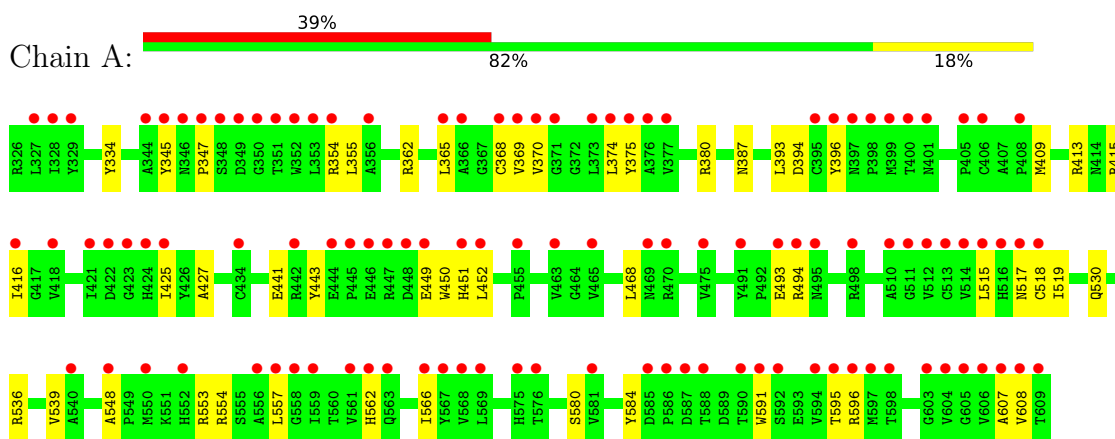
3 Residue-property plots

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: Kelch-like ECH-associated protein 1



- Molecule 2: Kelch-like ECH-associated protein 1



- Molecule 3: Nrf2 cyclic peptide,c[BAL-NPETGE]



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	162.49Å 68.68Å 77.35Å 90.00° 117.93° 90.00°	Depositor
Resolution (Å)	29.50 – 1.98 29.50 – 1.98	Depositor EDS
% Data completeness (in resolution range)	98.1 (29.50-1.98) 98.1 (29.50-1.98)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.05 (at 1.98Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.221 , 0.261 0.221 , 0.261	Depositor DCC
R_{free} test set	1998 reflections (3.90%)	wwPDB-VP
Wilson B-factor (Å ²)	34.2	Xtriage
Anisotropy	0.397	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 41.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.007 for -h-2*k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4547	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.40% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: BAL

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	B	0.77	2/2276 (0.1%)	0.83	4/3100 (0.1%)
2	A	0.51	0/2180	0.66	0/2977
3	P	1.41	0/44	0.84	0/59
All	All	0.67	2/4500 (0.0%)	0.75	4/6136 (0.1%)

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
3	P	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	449	GLU	CG-CD	6.58	1.61	1.51
1	B	449	GLU	CD-OE2	5.26	1.31	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	380	ARG	NE-CZ-NH2	-7.39	116.61	120.30
1	B	349	ASP	CB-CG-OD1	6.09	123.78	118.30
1	B	380	ARG	NE-CZ-NH1	5.57	123.08	120.30
1	B	611	GLU	OE1-CD-OE2	5.07	129.38	123.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	P	76	BAL	Mainchain,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	B	2219	0	2113	44	0
2	A	2126	0	1990	47	0
3	P	49	0	36	5	0
4	A	28	0	0	0	0
4	B	122	0	0	2	0
4	P	3	0	0	0	0
All	All	4547	0	4139	90	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 11.

The worst 5 of 90 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:B:504:ASN:HB2	1:B:546:PHE:CZ	1.82	1.13
1:B:504:ASN:CB	1:B:546:PHE:CE2	2.45	0.99
1:B:504:ASN:HB2	1:B:546:PHE:CE2	2.08	0.88
1:B:504:ASN:CB	1:B:546:PHE:CZ	2.57	0.87
1:B:466:ALA:HB1	1:B:514:VAL:HG23	1.60	0.83

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	B	288/289 (100%)	280 (97%)	7 (2%)	1 (0%)	41	29
2	A	283/284 (100%)	272 (96%)	11 (4%)	0	100	100
3	P	5/7 (71%)	5 (100%)	0	0	100	100
All	All	576/580 (99%)	557 (97%)	18 (3%)	1 (0%)	47	38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	547	VAL

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	B	234/234 (100%)	232 (99%)	2 (1%)	78	77
2	A	217/229 (95%)	216 (100%)	1 (0%)	88	87
3	P	5/5 (100%)	5 (100%)	0	100	100
All	All	456/468 (97%)	453 (99%)	3 (1%)	84	83

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	547	VAL
1	B	557	LEU
2	A	494	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	517	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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	$\# Z > 2$
3	BAL	P	76	3	4,4,5	2.37	1 (25%)	3,3,5	20.87	2 (66%)

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
3	BAL	P	76	3	-	1/1/2/3	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	P	76	BAL	O-C	4.40	1.44	1.19

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	P	76	BAL	CB-CA-C	35.82	164.61	111.42
3	P	76	BAL	O-C-CA	-4.77	95.55	126.89

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	P	76	BAL	C-CA-CB-N

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	P	76	BAL	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	P	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	P	76:BAL	C	77:ASN	N	1.07

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	B	289/289 (100%)	0.76	36 (12%) 3 4	24, 32, 52, 76	0
2	A	284/284 (100%)	1.68	110 (38%) 0 0	25, 54, 96, 119	0
3	P	6/7 (85%)	0.20	0 100 100	36, 38, 43, 44	0
All	All	579/580 (99%)	1.21	146 (25%) 0 0	24, 39, 91, 119	0

The worst 5 of 146 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	399	MET	7.8
2	A	351	THR	6.4
2	A	353	LEU	6.2
2	A	423	GLY	5.7
2	A	575	HIS	5.4

6.2 Non-standard residues in protein, DNA, RNA chains [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
3	BAL	P	76	5/6	0.72	0.25	38,42,55,84	0

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.