

# Full wwPDB X-ray Structure Validation Report (i)

May 25, 2020 - 04:13 am BST

PDB ID	:	5NKP
$\operatorname{Title}$	:	$\label{eq:crystal} Crystal \ structure \ of \ the \ human \ KLHL3 \ Kelch \ domain \ in \ complex \ with \ a \ WNK3$
		peptide
Authors	:	Chen, Z.; Sorrell, F.J.; Pinkas, D.M.; Williams, E.; Mathea, S.; Goubin, S.;
		von Delft, F.; Arrowsmith, C.H.; Edwards, A.M.; Burgess-Brown, N.; Bountra,
		C.; Bullock, A.
Deposited on	:	2017-03-31
Resolution	:	2.80  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 following versions of software and data (see references (1)) were used in the production of this report:

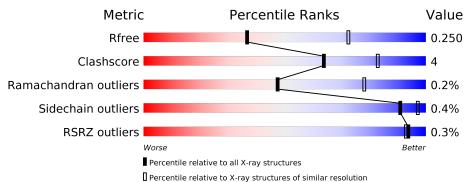
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
$\operatorname{Refmac}$	:	5.8.0158
$\operatorname{CCP4}$	:	$7.0.044 (\mathrm{Gargrove})$
Ideal geometry (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 2.80 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R <sub>free</sub>	130704	3140 (2.80-2.80)
Clashscore	141614	3569(2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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% 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	А	292	% • 86%	12%	•
1	В	292	90%	9%	•
2	С	11	82%	18%	_
2	D	11	82%	9% 9%	-



## 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4556 atoms, of which 5 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 protein 3.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	А	286	Total 2151	C 1348	N 377	0 412	S 14	0	0	0
1	В	289	Total 2170	$\mathrm{C}$ 1357	N 382	O 417	S 14	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

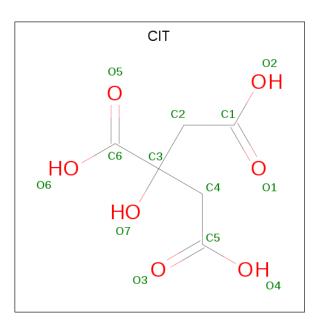
Chain	Residue	Modelled	Actual	Comment	Reference
А	296	SER	-	expression tag	UNP Q9UH77
А	297	MET	-	expression tag	UNP Q9UH77
В	296	SER	-	expression tag	UNP Q9UH77
В	297	MET	-	expression tag	UNP Q9UH77

• Molecule 2 is a protein called Serine/threonine-protein kinase WNK3.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	п	10	Total	С	Ν	Ο	S	0	0	0
	D	10	83	47	13	22	1	0	0	0
0	С	11	Total	С	Ν	Ο	S	0	0	0
	U	11	88	50	14	23	1	0	U	U

• Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula:  $C_6H_8O_7$ ).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	А	1	Total 18	С 6	H5	0 7	0	0

• Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	В	1	Total Cl 1 1	0	0
4	А	2	Total Cl 2 2	0	0

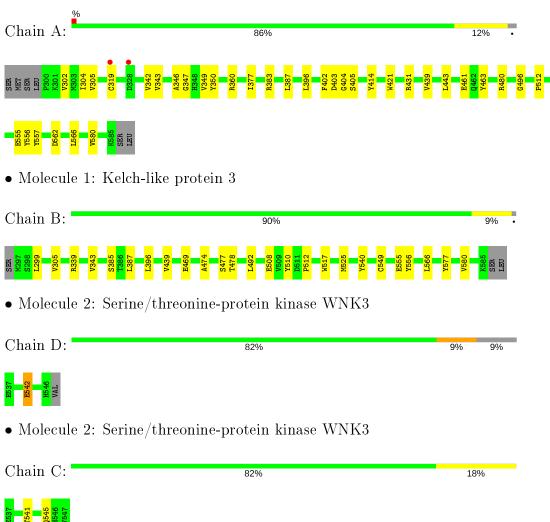
• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	24	TotalO2424	0	0
5	В	16	Total         O           16         16	0	0
5	С	3	Total O 3 3	0	0



## 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 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 protein 3



## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants	84.66Å $169.72$ Å $123.86$ Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	50.02 - 2.80	Depositor
Resolution (A)	50.03 - 2.80	EDS
% Data completeness	99.9 (50.02-2.80)	Depositor
(in resolution range)	91.7(50.03-2.80)	EDS
R <sub>merge</sub>	0.39	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.41 (at 2.81 \text{\AA})$	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
D D.	0.229 , $0.247$	Depositor
$R, R_{free}$	0.232 , $0.250$	DCC
$R_{free}$ test set	1102 reflections $(4.93%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	37.4	Xtriage
Anisotropy	0.238	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.30 , $13.5$	EDS
L-test for twinning <sup>2</sup>	$ \langle L  \rangle = 0.47, \langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	4556	wwPDB-VP
Average B, all atoms $(Å^2)$	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.03% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical 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.



 $<sup>^1 {\</sup>rm Intensities}$  estimated from amplitudes.

## 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: CL, CIT

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	Bo	nd lengths	Bond angles		
	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.44	0/2196	0.46	0/2985	
1	В	0.39	1/2215~(0.0%)	0.43	1/3013~(0.0%)	
2	С	1.58	0/88	0.66	0/118	
2	D	1.32	0/83	0.65	0/111	
All	All	0.50	1/4582~(0.0%)	0.45	1/6227~(0.0%)	

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
1	В	469	GLU	C-O	-5.12	1.13	1.23

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	В	299	LEU	C-N-CD	5.04	138.98	128.40

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	А	2151	0	2065	22	0
1	В	2170	0	2081	13	2

Continued on next page...



Mol		Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	С	88	0	65	1	0
2	D	83	0	63	2	0
3	А	13	5	5	1	1
4	А	2	0	0	0	0
4	В	1	0	0	1	0
5	А	24	0	0	0	0
5	В	16	0	0	0	0
5	С	3	0	0	0	0
All	All	4551	5	4279	37	2

Continued from previous page...

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 (37) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

A 4 and 1	A.4 a.m. D	Interatomic	Clash
Atom-1	Atom-2	${ m distance}~({ m \AA})$	overlap (Å)
1:A:402:PHE:CE1	2:D:542:GLU:HG3	2.03	0.93
1:A:403:ASP:HB3	1:A:405:SER:H	1.37	0.88
1:A:302:VAL:HG21	1:A:319:CYS:SG	2.21	0.80
1:A:302:VAL:CG2	1:A:319:CYS:SG	2.73	0.77
1:A:402:PHE:CZ	2:D:542:GLU:HG3	2.23	0.73
1:A:403:ASP:N	1:A:404:GLY:HA2	2.14	0.62
1:A:480:ARG:NH1	1:A:496:GLY:O	2.32	0.61
1:B:492:LEU:HB3	1:B:510:TYR:HB3	1.83	0.61
1:B:305:VAL:HG22	1:B:580:VAL:HG12	1.85	0.58
1:A:305:VAL:HG22	1:A:580:VAL:HG12	1.90	0.52
1:A:360:ARG:O	1:A:383:ARG:NH1	2.36	0.51
1:A:557:TYR:OH	1:A:562:ASP:OD1	2.20	0.49
1:A:343:VAL:HG21	1:A:396:LEU:HD13	1.97	0.47
1:B:540:TYR:CD2	1:B:556:TYR:HB3	2.51	0.46
1:A:346:ALA:HA	1:A:347:GLY:HA2	1.45	0.46
1:A:387:LEU:H	1:A:387:LEU:HD23	1.81	0.46
1:B:387:LEU:H	1:B:387:LEU:HD23	1.82	0.45
1:B:343:VAL:HG21	1:B:396:LEU:HD13	1.97	0.45
1:B:385:SER:OG	4:B:601:CL:CL	2.59	0.45
1:A:439:VAL:HG11	1:A:512:PRO:HG3	1.98	0.45
1:A:540:TYR:CD2	1:A:556:TYR:HB3	2.52	0.44
1:A:304:ILE:HD11	1:A:349:VAL:HG21	2.00	0.44
1:A:431:ARG:HH21	1:A:461:GLU:HG3	1.81	0.44
1:B:510:TYR:HB2	1:B:517:TRP:CH2	2.52	0.44
1:A:555:GLU:HB3	1:A:566:LEU:HD23	2.01	0.43

Continued on next page...



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:477:SER:OG	1:B:508:GLU:OE2	2.34	0.43
2:C:541:THR:O	2:C:545:GLN:HG3	2.19	0.42
1:A:443:LEU:HB3	1:A:463:TYR:HB3	2.01	0.42
3:A:601:CIT:C6	3:A:601:CIT:O3	2.68	0.42
1:A:403:ASP:HB3	1:A:405:SER:N	2.19	0.41
1:B:555:GLU:HB3	1:B:566:LEU:HD23	2.03	0.41
1:B:339:ARG:HG2	1:B:577:TYR:CG	2.55	0.41
1:A:342:VAL:HA	1:A:350:TYR:O	2.21	0.41
1:B:439:VAL:HG11	1:B:512:PRO:HG2	2.03	0.41
1:A:414:TYR:HB2	1:A:421:TRP:CE2	2.56	0.40
1:B:474:ALA:O	1:B:517:TRP:NE1	2.49	0.40
1:B:510:TYR:HB2	1:B:517:TRP:CZ3	2.57	0.40

Continued from previous page...

All (2) 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)
1:B:549:CYS:CB	$1:B:549:CYS:SG[3_554]$	1.81	0.39
1:B:478:THR:OG1	3:A:601:CIT:O5[6_554]	1.85	0.35

### 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	Perce	ntiles
1	А	284/292~(97%)	271~(95%)	12~(4%)	1 (0%)	34	66
1	В	287/292~(98%)	278~(97%)	9~(3%)	0	100	100
2	С	9/11~(82%)	9 (100%)	0	0	100	100
2	D	8/11~(73%)	8 (100%)	0	0	100	100
All	All	588/606~(97%)	566~(96%)	21 (4%)	1 (0%)	47	78

All (1) Ramachandran outliers are listed below:



Mol	Chain	Res	Type	
1	А	377	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	А	220/233~(94%)	220~(100%)	0	100 100
1	В	224/233~(96%)	223~(100%)	1 (0%)	91 97
2	С	10/11~(91%)	10~(100%)	0	100 100
2	D	10/11~(91%)	9~(90%)	1 (10%)	7 22
All	All	464/488~(95%)	462~(100%)	2(0%)	91 97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	525	MET
2	D	542	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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 carbohydrates in this entry.



### 5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 3 are 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).

Mal	Tuno	Chain	Dog	Tink	B	ond leng	$\mathbf{gths}$	B	ond ang	gles
MOI	туре	e Chain	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	CIT	А	601	-	$3,\!12,\!12$	2.13	2 (66%)	$3,\!17,\!17$	2.05	1 (33%)

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	CIT	А	601	-	-	4/6/16/16	-

All (2) bond length outliers are listed below:

Mol	Chain	$\mathbf{Res}$	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
3	А	601	CIT	C4-C3	-2.84	1.50	1.54
3	А	601	CIT	C2-C3	-2.11	1.51	1.54

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
3	A	601	CIT	C3-C4-C5	3.45	120.50	114.98

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	А	601	CIT	C1-C2-C3-C4
3	А	601	CIT	C1-C2-C3-C6
3	А	601	CIT	C1-C2-C3-O7
3	А	601	CIT	C6-C3-C4-C5



There are no ring outliers.

1 monomer is involved in 2 short contacts:

	Mol	Chain	Res	Type	Clashes	Symm-Clashes
ſ	3	A	601	CIT	1	1

### 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, 95<sup>th</sup> 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 $>$	$\# RSRZ {>}2$	$\mathbf{OWAB}(\mathbf{A}^2)$	Q<0.9
1	А	286/292~(97%)	0.01	2 (0%) 87 84	30, 37, 48, 55	0
1	В	289/292~(98%)	-0.13	0 100 100	30, 36, 48, 59	0
2	С	11/11~(100%)	-0.34	0 100 100	30, 32, 33, 34	0
2	D	10/11~(90%)	-0.09	0 100 100	35, 36, 39, 40	0
All	All	596/606~(98%)	-0.07	2 (0%) 94 93	30, 37, 48, 59	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	319	CYS	2.2
1	А	328	ASP	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 carbohydrates 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,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathbf{\AA}^2)$	$Q{<}0.9$
3	CIT	А	601	13/13	0.85	0.23	$41,\!43,\!52,\!52$	0
4	CL	А	602	1/1	0.94	0.17	$33,\!33,\!33,\!33$	0
4	CL	В	601	1/1	0.94	0.10	$31,\!31,\!31,\!31$	0
4	CL	А	603	1/1	0.97	0.08	$36,\!36,\!36,\!36$	0

## 6.5 Other polymers (i)

There are no such residues in this entry.

