



# wwPDB X-ray Structure Validation Summary Report

Sep 17, 2023 – 07:58 AM EDT

PDB ID : 4TKN  
Title : Structure of the SNX17 FERM domain bound to the second NPxF motif of KRIT1  
Authors : Stiegler, A.L.; Zhang, R.; Liu, W.; Boggon, T.J.  
Deposited on : 2014-05-27  
Resolution : 3.00 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references](#) ) 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

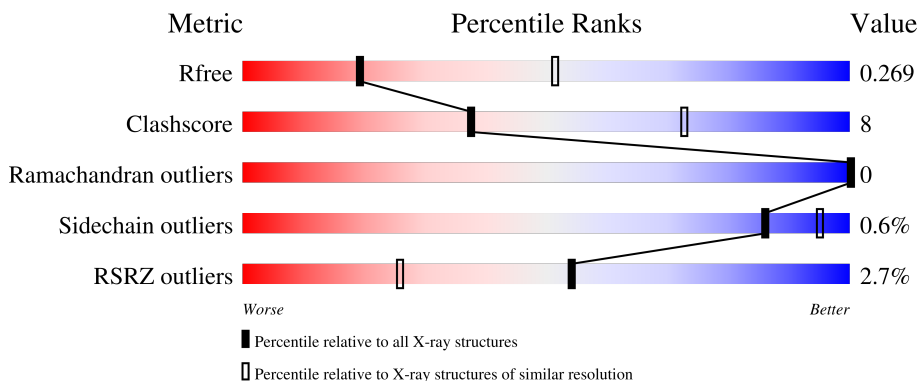
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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  $\geq 3$ , 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	A	286	
1	B	286	
1	C	286	
2	D	19	
2	E	19	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	19	 16% 32% 53%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6419 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 Sorting nexin-17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	256	Total 2075	C 1321	N 351	O 392	S 11	0	0	0
1	B	247	Total 1996	C 1272	N 335	O 379	S 10	0	0	0
1	C	257	Total 2081	C 1326	N 350	O 394	S 11	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	106	GLY	-	expression tag	UNP Q15036
A	107	SER	-	expression tag	UNP Q15036
B	106	GLY	-	expression tag	UNP Q15036
B	107	SER	-	expression tag	UNP Q15036
C	106	GLY	-	expression tag	UNP Q15036
C	107	SER	-	expression tag	UNP Q15036

- Molecule 2 is a protein called Krev interaction trapped protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	10	Total 77	C 51	N 11	O 14	S 1	0	0	0
2	E	9	Total 71	C 48	N 10	O 12	S 1	0	0	0
2	F	9	Total 71	C 48	N 10	O 12	S 1	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	219	GLY	-	expression tag	UNP O00522
D	220	PRO	-	expression tag	UNP O00522

*Continued on next page...*

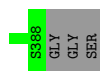
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	221	LEU	-	expression tag	UNP O00522
D	222	GLY	-	expression tag	UNP O00522
D	223	SER	-	expression tag	UNP O00522
D	224	PRO	-	expression tag	UNP O00522
E	219	GLY	-	expression tag	UNP O00522
E	220	PRO	-	expression tag	UNP O00522
E	221	LEU	-	expression tag	UNP O00522
E	222	GLY	-	expression tag	UNP O00522
E	223	SER	-	expression tag	UNP O00522
E	224	PRO	-	expression tag	UNP O00522
F	219	GLY	-	expression tag	UNP O00522
F	220	PRO	-	expression tag	UNP O00522
F	221	LEU	-	expression tag	UNP O00522
F	222	GLY	-	expression tag	UNP O00522
F	223	SER	-	expression tag	UNP O00522
F	224	PRO	-	expression tag	UNP O00522

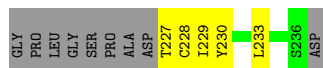
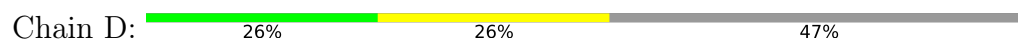
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	15	Total O 15 15	0	0
3	B	17	Total O 17 17	0	0
3	C	16	Total O 16 16	0	0

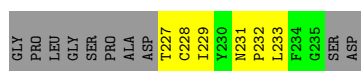
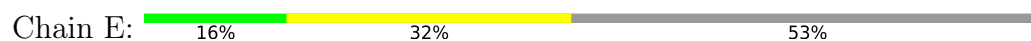




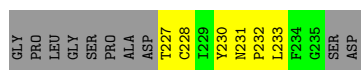
- Molecule 2: Krev interaction trapped protein 1



- Molecule 2: Krev interaction trapped protein 1



- Molecule 2: Krev interaction trapped protein 1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.92Å 191.20Å 46.59Å 90.00° 110.25° 90.00°	Depositor
Resolution (Å)	47.88 – 3.00 47.88 – 3.00	Depositor EDS
% Data completeness (in resolution range)	92.5 (47.88-3.00) 92.2 (47.88-3.00)	Depositor EDS
$R_{merge}$	0.24	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.65 (at 3.01Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.218 , 0.268 0.222 , 0.269	Depositor DCC
$R_{free}$ test set	1912 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	86.8	Xtrriage
Anisotropy	0.233	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 73.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.052 for -h-2*1,-k,l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6419	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	126.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<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.



## 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	0.32	0/2108	0.61	0/2844
1	B	0.32	0/2025	0.62	0/2731
1	C	0.30	0/2114	0.62	0/2853
2	D	0.28	0/79	0.51	0/107
2	E	0.31	0/73	0.63	0/99
2	F	0.27	0/73	0.58	0/99
All	All	0.31	0/6472	0.62	0/8733

There are no bond length outliers.

There are no bond angle outliers.

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	A	2075	0	2103	32	1
1	B	1996	0	2024	34	0
1	C	2081	0	2109	31	1
2	D	77	0	72	4	0
2	E	71	0	66	7	0
2	F	71	0	67	5	0
3	A	15	0	0	0	0
3	B	17	0	0	0	0
3	C	16	0	0	0	0
All	All	6419	0	6441	99	1

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 8.

The worst 5 of 99 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:177:ARG:NH1	1:B:183:GLU:OE1	2.16	0.78
1:C:322:ARG:HB2	2:F:228:CYS:HB3	1.67	0.75
1:B:310:GLU:OE2	1:C:267:ARG:NH2	2.21	0.74
1:B:166:ARG:HH12	1:B:168:LYS:HD2	1.52	0.73
1:C:177:ARG:HH12	1:C:183:GLU:CD	1.97	0.67

All (1) 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:A:193:ARG:NH1	1:C:116:GLU:O[4_556]	2.15	0.05

## 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	A	250/286 (87%)	247 (99%)	3 (1%)	0	100	100
1	B	239/286 (84%)	236 (99%)	3 (1%)	0	100	100
1	C	251/286 (88%)	248 (99%)	3 (1%)	0	100	100
2	D	8/19 (42%)	8 (100%)	0	0	100	100
2	E	7/19 (37%)	7 (100%)	0	0	100	100
2	F	7/19 (37%)	7 (100%)	0	0	100	100
All	All	762/915 (83%)	753 (99%)	9 (1%)	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	236/258 (92%)	235 (100%)	1 (0%)	91	97
1	B	227/258 (88%)	226 (100%)	1 (0%)	91	97
1	C	237/258 (92%)	235 (99%)	2 (1%)	81	93
2	D	9/15 (60%)	9 (100%)	0	100	100
2	E	8/15 (53%)	8 (100%)	0	100	100
2	F	8/15 (53%)	8 (100%)	0	100	100
All	All	725/819 (88%)	721 (99%)	4 (1%)	86	95

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	302	ARG
1	B	387	LYS
1	C	218	ARG
1	C	358	ARG

Sometimes 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 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](#)

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	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	256/286 (89%)	0.01	2 (0%) 86 65	80, 117, 167, 193	0
1	B	247/286 (86%)	0.16	8 (3%) 47 20	83, 117, 169, 192	0
1	C	257/286 (89%)	0.19	11 (4%) 35 13	87, 129, 188, 241	0
2	D	10/19 (52%)	-0.03	0 100 100	107, 122, 140, 152	0
2	E	9/19 (47%)	0.43	0 100 100	104, 122, 133, 135	0
2	F	9/19 (47%)	0.50	0 100 100	141, 159, 171, 173	0
All	All	788/915 (86%)	0.12	21 (2%) 54 26	80, 123, 174, 241	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	278	VAL	5.8
1	C	285	ASP	5.4
1	C	280	ASP	4.7
1	B	277	CYS	3.9
1	B	310	GLU	3.8

### 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 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.