wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 06:11 pm BST

PDB ID	:	5JTI
Title	:	Crystal structure of the human Tankyrase 1 (TNKS) SAM domain (D1055R), crystal form 2
Authors	:	Guetter, S.; Mariotti, L.; Cronin, N.
Deposited on	:	2016-05-09
Resolution	:	2.90 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:

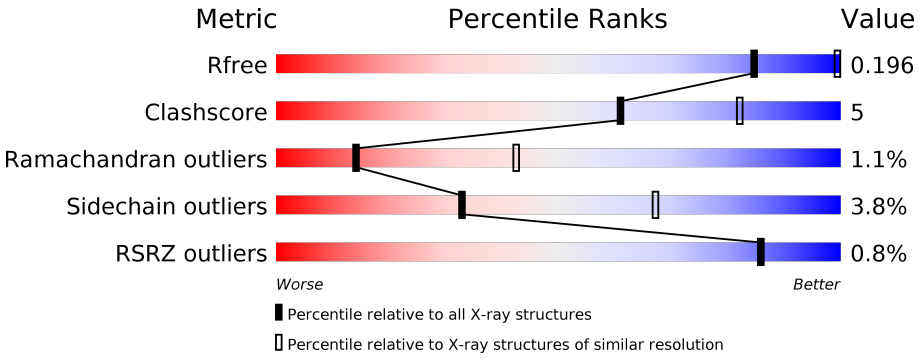
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	A	79	% 65% 14% • 20%
1	B	79	3% 65% 13% • 22%
1	C	79	70% 6% 24%
1	D	79	71% • 25%
1	E	79	63% 11% • 24%
1	F	79	63% 9% • 25%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2924 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 Tankyrase-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	63	Total	C	N	O	S	0	0	0
			502	318	91	92	1			
1	B	62	Total	C	N	O	S	0	1	0
			509	323	93	91	2			
1	C	60	Total	C	N	O	S	0	0	0
			481	307	86	86	2			
1	D	59	Total	C	N	O	S	0	0	0
			448	288	77	82	1			
1	E	60	Total	C	N	O	S	0	0	0
			468	297	85	84	2			
1	F	59	Total	C	N	O	S	0	0	0
			454	293	78	82	1			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1015	SER	-	expression tag	UNP O95271
A	1016	ASN	-	expression tag	UNP O95271
A	1017	ALA	-	expression tag	UNP O95271
A	1055	ARG	ASP	engineered mutation	UNP O95271
B	1015	SER	-	expression tag	UNP O95271
B	1016	ASN	-	expression tag	UNP O95271
B	1017	ALA	-	expression tag	UNP O95271
B	1055	ARG	ASP	engineered mutation	UNP O95271
C	1015	SER	-	expression tag	UNP O95271
C	1016	ASN	-	expression tag	UNP O95271
C	1017	ALA	-	expression tag	UNP O95271
C	1055	ARG	ASP	engineered mutation	UNP O95271
D	1015	SER	-	expression tag	UNP O95271
D	1016	ASN	-	expression tag	UNP O95271
D	1017	ALA	-	expression tag	UNP O95271
D	1055	ARG	ASP	engineered mutation	UNP O95271
E	1015	SER	-	expression tag	UNP O95271

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	1016	ASN	-	expression tag	UNP O95271
E	1017	ALA	-	expression tag	UNP O95271
E	1055	ARG	ASP	engineered mutation	UNP O95271
F	1015	SER	-	expression tag	UNP O95271
F	1016	ASN	-	expression tag	UNP O95271
F	1017	ALA	-	expression tag	UNP O95271
F	1055	ARG	ASP	engineered mutation	UNP O95271

- Molecule 2 is water.

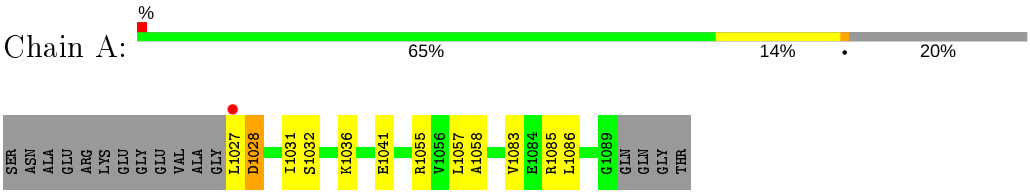
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	9	Total O 9 9	0	0
2	B	15	Total O 15 15	0	0
2	C	10	Total O 10 10	0	0
2	D	8	Total O 8 8	0	0
2	E	13	Total O 13 13	0	0
2	F	7	Total O 7 7	0	0

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.

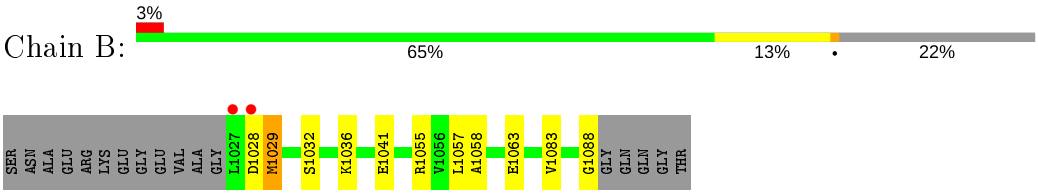
- Molecule 1: Tankyrase-1

Chain A:



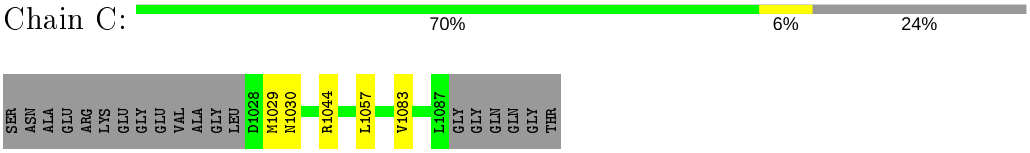
- Molecule 1: Tankyrase-1

Chain B:



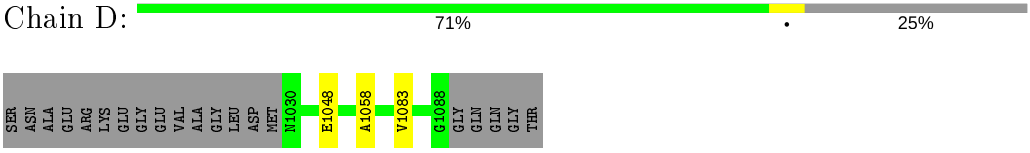
- Molecule 1: Tankyrase-1

Chain C:



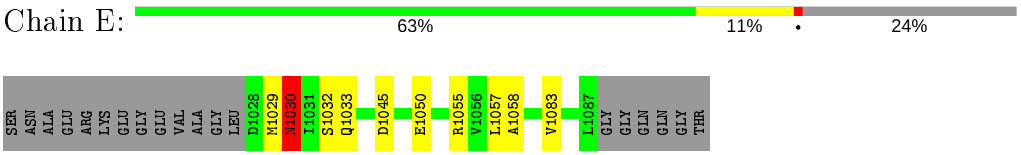
- Molecule 1: Tankyrase-1

Chain D:



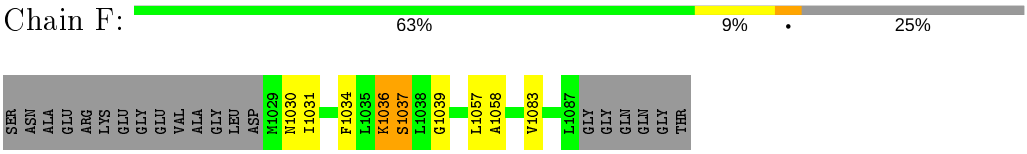
- Molecule 1: Tankyrase-1

Chain E:



- Molecule 1: Tankyrase-1

Chain F:


63%
9%
25%
SER
ASN
ALA
GLU
ARG
LYS
GLU
GLY
GLU
VAL
ALA
GLY
LEU
ASP
M1029
N1030
I1031
F1034
L1035
K1036
S1037
L1038
G1039
L1057
A1058
V1083
L1087
GLY
GLY
GLN
GLN
GLY
THR

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.93Å 55.48Å 79.41Å 90.00° 102.90° 90.00°	Depositor
Resolution (Å)	77.41 – 2.90 77.41 – 2.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (77.41-2.90) 100.0 (77.41-2.80)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 2.82Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.193 , 0.232 0.203 , 0.196	Depositor DCC
R_{free} test set	787 reflections (5.22%)	wwPDB-VP
Wilson B-factor (Å^2)	48.1	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 54.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	2924	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.60% 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	0.52	0/508	0.72	0/679
1	B	0.50	0/518	0.70	0/691
1	C	0.53	0/487	0.67	0/651
1	D	0.47	0/454	0.62	0/610
1	E	0.45	0/474	0.70	1/636 (0.2%)
1	F	0.50	0/460	0.77	2/618 (0.3%)
All	All	0.50	0/2901	0.70	3/3885 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	1036	LYS	C-N-CA	6.79	138.68	121.70
1	E	1030	ASN	C-N-CA	5.49	135.42	121.70
1	F	1030	ASN	C-N-CA	5.16	134.59	121.70

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	502	0	511	9	0
1	B	509	0	528	6	0
1	C	481	0	488	3	0
1	D	448	0	438	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	468	0	460	6	0
1	F	454	0	449	5	0
2	A	9	0	0	1	0
2	B	15	0	0	0	0
2	C	10	0	0	0	0
2	D	8	0	0	0	0
2	E	13	0	0	0	0
2	F	7	0	0	0	0
All	All	2924	0	2874	26	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 5.

The worst 5 of 26 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:F:1034:PHE:O	1:F:1037:SER:HB3	1.84	0.78
1:B:1036:LYS:HD2	1:B:1041:GLU:HG2	1.76	0.68
1:F:1036:LYS:HG3	1:F:1039:GLY:HA2	1.81	0.62
1:C:1029:MET:HG2	1:C:1030:ASN:N	2.16	0.61
1:F:1036:LYS:HA	1:F:1039:GLY:H	1.67	0.60

There are no symmetry-related clashes.

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	61/79 (77%)	59 (97%)	2 (3%)	0	100 100
1	B	61/79 (77%)	59 (97%)	1 (2%)	1 (2%)	9 32
1	C	58/79 (73%)	58 (100%)	0	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	57/79 (72%)	56 (98%)	1 (2%)	0	100 100
1	E	58/79 (73%)	55 (95%)	2 (3%)	1 (2%)	9 31
1	F	57/79 (72%)	53 (93%)	2 (4%)	2 (4%)	3 14
All	All	352/474 (74%)	340 (97%)	8 (2%)	4 (1%)	14 42

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1029	MET
1	F	1031	ILE
1	F	1037	SER
1	E	1029	MET

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	53/65 (82%)	51 (96%)	2 (4%)	33 67
1	B	55/65 (85%)	50 (91%)	5 (9%)	9 28
1	C	50/65 (77%)	49 (98%)	1 (2%)	55 82
1	D	44/65 (68%)	43 (98%)	1 (2%)	50 80
1	E	47/65 (72%)	44 (94%)	3 (6%)	17 45
1	F	45/65 (69%)	45 (100%)	0	100 100
All	All	294/390 (75%)	282 (96%)	12 (4%)	33 64

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1055[B]	ARG
1	B	1063	GLU
1	E	1030	ASN
1	B	1055[A]	ARG
1	D	1048	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1042	HIS
1	E	1042	HIS
1	F	1042	HIS

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

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^{th} 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	63/79 (79%)	0.16	1 (1%) 72 71	27, 42, 62, 80	0
1	B	62/79 (78%)	0.16	2 (3%) 47 43	25, 40, 68, 93	0
1	C	60/79 (75%)	0.01	0 100 100	24, 44, 83, 100	0
1	D	59/79 (74%)	0.17	0 100 100	29, 54, 88, 102	0
1	E	60/79 (75%)	0.03	0 100 100	35, 54, 100, 106	0
1	F	59/79 (74%)	0.22	0 100 100	39, 65, 101, 111	0
All	All	363/474 (76%)	0.13	3 (0%) 86 86	24, 50, 92, 111	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1028	ASP	5.5
1	A	1027	LEU	4.5
1	B	1027	LEU	3.4

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 carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.