wwPDB X-ray Structure Validation Summary Report (i)

May 25, 2020 – 12:04 pm BST

PDB ID	:	4XUU
Title	:	The hSac2 domain from human phosphoinositide phosphatase Sac2
Authors	:	Hsu, F.; Mao, Y.
Deposited on	:	2015-01-26
Resolution	:	2.62 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:

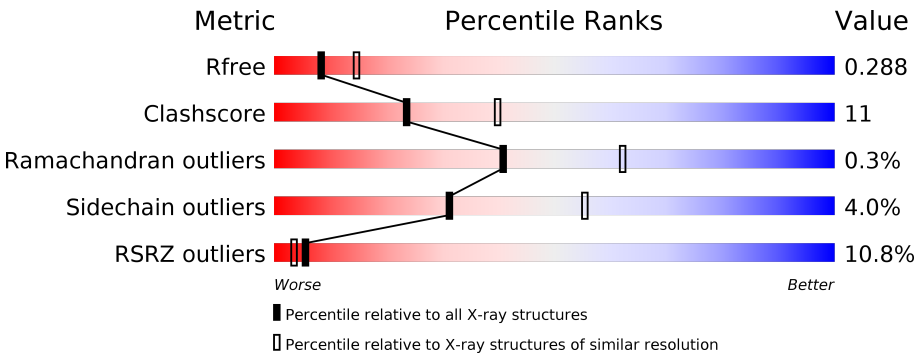
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.62 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	3797 (2.64-2.60)
Clashscore	141614	4168 (2.64-2.60)
Ramachandran outliers	138981	4093 (2.64-2.60)
Sidechain outliers	138945	4093 (2.64-2.60)
RSRZ outliers	127900	3731 (2.64-2.60)

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	169	7% 82% 17% •
1	B	169	14% 82% 17% •
1	C	169	11% 83% 13% • •
1	D	169	9% 80% 17% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5456 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 Phosphatidylinositide phosphatase SAC2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	169	1368	868	228	264	3	5	0	0	0
1	B	169	1368	868	228	264	3	5	0	0	0
1	C	164	1330	845	221	256	3	5	0	0	0
1	D	169	1368	868	228	264	3	5	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP Q9Y2H2
A	159	ALA	ARG	conflict	UNP Q9Y2H2
A	161	ALA	SER	conflict	UNP Q9Y2H2
B	1	SER	-	expression tag	UNP Q9Y2H2
B	159	ALA	ARG	conflict	UNP Q9Y2H2
B	161	ALA	SER	conflict	UNP Q9Y2H2
C	1	SER	-	expression tag	UNP Q9Y2H2
C	159	ALA	ARG	conflict	UNP Q9Y2H2
C	161	ALA	SER	conflict	UNP Q9Y2H2
D	1	SER	-	expression tag	UNP Q9Y2H2
D	159	ALA	ARG	conflict	UNP Q9Y2H2
D	161	ALA	SER	conflict	UNP Q9Y2H2

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	8	8	8	0	0
2	B	6	6	6	0	0

Continued on next page...

Continued from previous page...

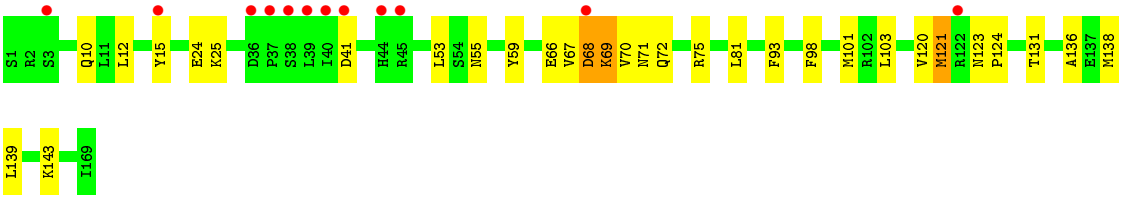
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	6	Total O 6 6	0	0
2	D	2	Total O 2 2	0	0

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.

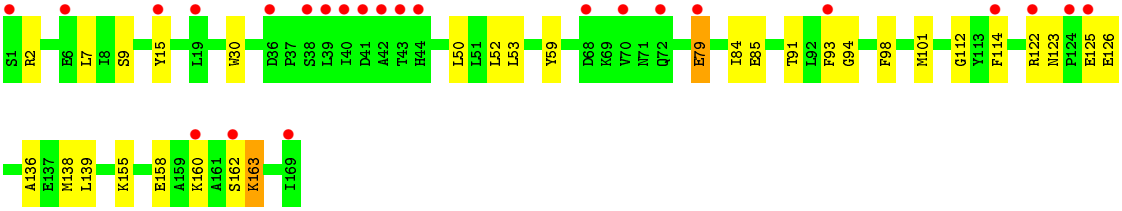
- Molecule 1: Phosphatidylinositide phosphatase SAC2

Chain A:



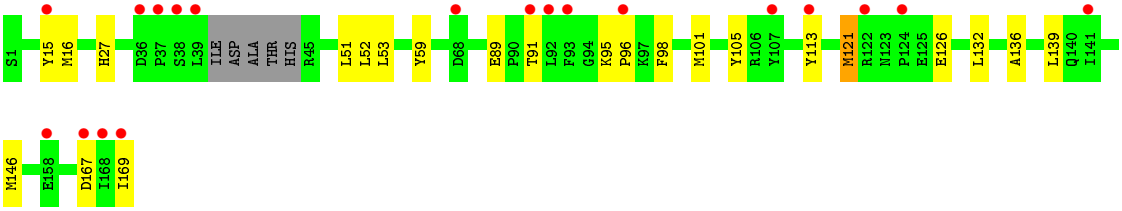
- Molecule 1: Phosphatidylinositide phosphatase SAC2

Chain B:



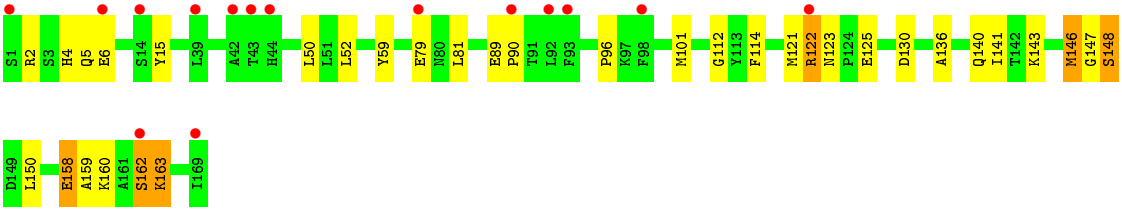
- Molecule 1: Phosphatidylinositide phosphatase SAC2

Chain C:



- Molecule 1: Phosphatidylinositide phosphatase SAC2

Chain D:


S1
R2
S3
H4
Q5
E6
S14
Y15
L39
A42
T43
H44
L50
L51
L52
Y59
E79
N80
L81
E89
P90
T91
L92
F93
P96
K97
F98
M101
G112
Y113
F114
M121
R122
N123
P124
E125
D130
A136
Q140
I141
T142
K143
M146
G147
S148

D149
L150
E158
A159
K160
A161
S162
K163
I169

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.77Å 82.85Å 157.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.62 49.52 – 2.62	Depositor EDS
% Data completeness (in resolution range)	98.9 (30.00-2.62) 98.9 (49.52-2.62)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	6.69 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.248 , 0.287 0.253 , 0.288	Depositor DCC
R_{free} test set	1262 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	63.5	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 70.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	5456	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 20.40 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.7604e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (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	0.42	0/1391	0.52	0/1868
1	B	0.31	0/1391	0.50	0/1868
1	C	0.30	0/1351	0.49	0/1811
1	D	0.41	0/1391	0.55	0/1868
All	All	0.36	0/5524	0.52	0/7415

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	1368	0	1355	42	0
1	B	1368	0	1355	31	0
1	C	1330	0	1320	17	0
1	D	1368	0	1355	42	0
2	A	8	0	0	0	0
2	B	6	0	0	0	0
2	C	6	0	0	0	0
2	D	2	0	0	1	0
All	All	5456	0	5385	121	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 121 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:A:67:VAL:HB	1:A:69:LYS:CE	1.63	1.29
1:A:67:VAL:HB	1:A:69:LYS:HE2	1.17	1.17
1:D:114:PHE:HZ	1:D:160:LYS:O	1.34	1.10
1:D:90:PRO:HB2	1:D:159:ALA:HB2	1.30	1.09
1:A:67:VAL:CG1	1:A:69:LYS:HZ3	1.65	1.08

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	167/169 (99%)	164 (98%)	3 (2%)	0	100	100
1	B	167/169 (99%)	162 (97%)	4 (2%)	1 (1%)	25	45
1	C	160/169 (95%)	158 (99%)	2 (1%)	0	100	100
1	D	167/169 (99%)	162 (97%)	4 (2%)	1 (1%)	25	45
All	All	661/676 (98%)	646 (98%)	13 (2%)	2 (0%)	41	62

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	163	LYS
1	B	163	LYS

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	152/147 (103%)	143 (94%)	9 (6%)	19	37
1	B	152/147 (103%)	147 (97%)	5 (3%)	38	62
1	C	148/147 (101%)	143 (97%)	5 (3%)	37	61
1	D	152/147 (103%)	147 (97%)	5 (3%)	38	62
All	All	604/588 (103%)	580 (96%)	24 (4%)	31	55

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

Mol	Chain	Res	Type
1	B	91	THR
1	B	126	GLU
1	D	158	GLU
1	B	98	PHE
1	B	122	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	44	HIS
1	B	71	ASN
1	D	4	HIS
1	B	27	HIS
1	C	144	GLN

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)

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, 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	A	164/169 (97%)	0.69	12 (7%)	15	11	31, 62, 114, 140	0
1	B	164/169 (97%)	0.78	24 (14%)	2	1	33, 73, 117, 168	0
1	C	159/169 (94%)	0.78	19 (11%)	4	3	38, 69, 124, 152	0
1	D	164/169 (97%)	0.75	15 (9%)	9	6	38, 75, 119, 151	0
All	All	651/676 (96%)	0.75	70 (10%)	5	3	31, 71, 122, 168	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	39	LEU	10.4
1	C	39	LEU	8.9
1	B	42	ALA	8.6
1	C	38	SER	8.5
1	D	39	LEU	7.5

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.