wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 09:07 AM GMT

PDB ID	:	3HAJ
Title	:	Crystal structure of human PACSIN2 F-BAR domain (p212121 lattice)
Authors	:	Wang, Q.; Navarro, M.V.A.S.; Peng, G.; Rajashankar, K.R.; Sondermann, H.
Deposited on	:	2009-05-01
Resolution	:	2.78 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
http://wwpdb.org/validation/2016/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:

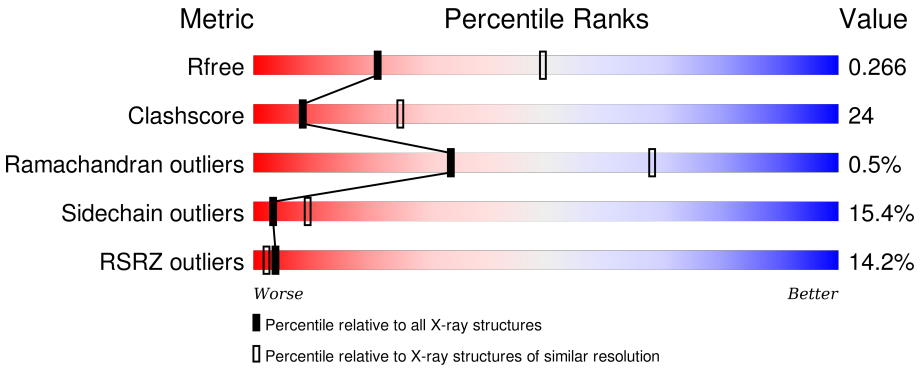
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.78 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}	91344	3004 (2.80-2.76)
Clashscore	102246	3480 (2.80-2.76)
Ramachandran outliers	100387	3423 (2.80-2.76)
Sidechain outliers	100360	3425 (2.80-2.76)
RSRZ outliers	91569	3016 (2.80-2.76)

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. 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	486	10% 35% 20% • 41%
1	B	486	7% 32% 19% 7% 42%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4753 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 human PACSIN2 F-BAR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	286	Total	C	N	O	S	0	0	0
			2368	1489	427	437	15			
1	B	284	Total	C	N	O	S	0	0	0
			2347	1475	424	433	15			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ca	0	0
			1	1		
2	A	1	Total	Ca	0	0
			1	1		

- Molecule 3 is water.

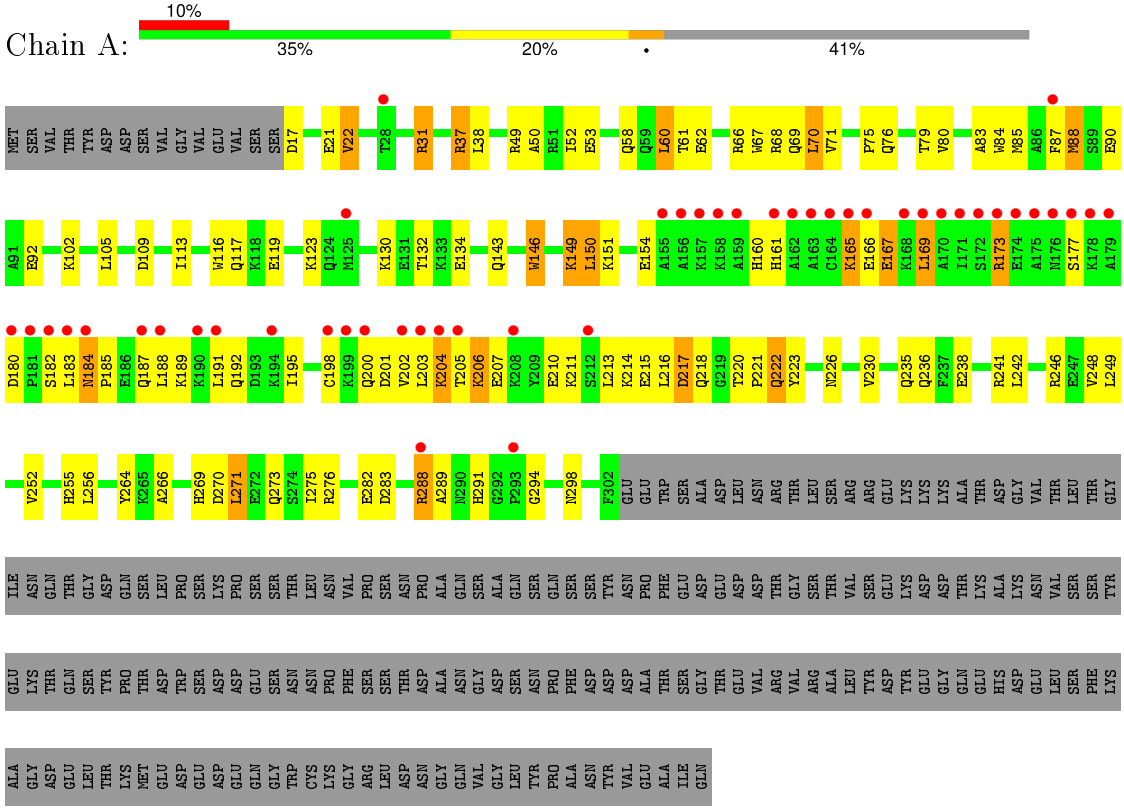
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	14	Total	O	0	0
			14	14		
3	B	22	Total	O	0	0
			22	22		

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 errors 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.

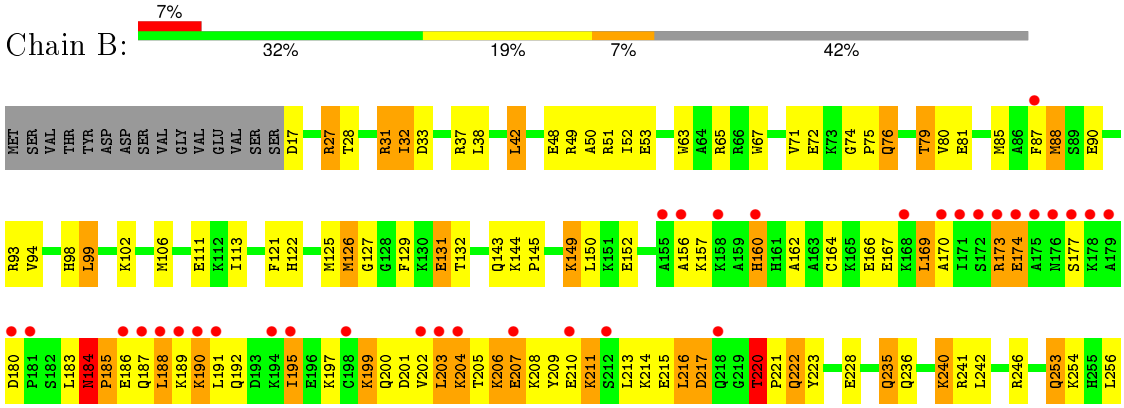
- Molecule 1: human PACSIN2 F-BAR

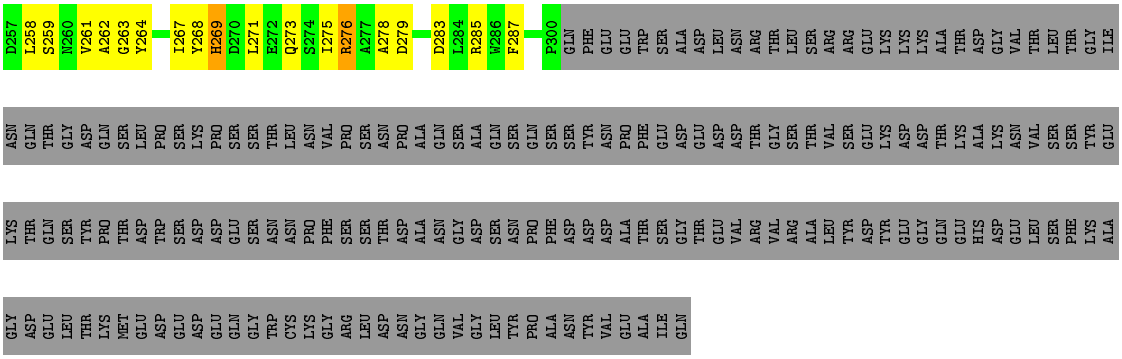
Chain A:



- Molecule 1: human PACSIN2 F-BAR

Chain B:




D257 L258 S259 N260 V261 A262 G263 Y264 I267 Y268 H269 D270 L271 E272 Q273 S274 I275 R276 A277 A278 D279 D283 L284 R285 W286 F287 P300 GLN PHE GLU GLU TRP SER ALA ASP LEU ASN ARG THR LEU SER ARG ARG GLU LYS LYS LYS ALA THR ASP GLY VAL THR LEU THR GLY ILE
ASN GLN THR GLY ASP GLN SER LEU PRO SER LYS PRO SER SER THR LEU ASN VAL PRO SER ASN PRO ALA GLN SER ALA GLN SER GLN SER SER TYR ASN PRO PHE GLU ASP GLU ASP ASP THR GLY SER THR VAL SER GLU LYS ASP ASP THR LYS ALA LYS ASN VAL SER SER TYR GLU
LYS THR GLN SER TYR PRO THR ASP TRP SER ASP ASP GLU SER ASN ASN PRO PHE SER SER THR ASP ALA ASN GLY ASP SER ASN PRO PHE ASP ASP ASP ALA THR SER GLY THR GLU VAL ARG VAL ARG ALA LEU TYR ASP TYR GLU GLY GLN GLU HIS ASP GLU LEU SER PHE LYS ALA
GLY ASP GLU LEU THR LYS MET GLU ASP GLU ASP GLU GLN GLY TRP CYS LYS GLY ARG LEU ASP ASN GLY GLN VAL GLY LEU TYR PRO ALA ASN TYR VAL GLU ALA ILE GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	31.30Å 88.39Å 357.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.72 – 2.78 44.72 – 2.78	Depositor EDS
% Data completeness (in resolution range)	95.4 (44.72-2.78) 95.4 (44.72-2.78)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.81 (at 2.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.224 , 0.274 0.218 , 0.266	Depositor DCC
R_{free} test set	1258 reflections (5.00%)	DCC
Wilson B-factor (Å^2)	46.6	Xtriage
Anisotropy	0.271	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 77.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Outliers	0 of 25179 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	4753	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.85% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: CA

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	0.23	0/2417	0.37	0/3235
1	B	0.24	0/2395	0.41	0/3205
All	All	0.24	0/4812	0.39	0/6440

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

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	2368	0	2334	114	0
1	B	2347	0	2317	139	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	14	0	0	0	0
3	B	22	0	0	4	0
All	All	4753	0	4651	226	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 24.

The worst 5 of 226 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:192:GLN:HA	1:B:195:ILE:HD11	1.37	1.06
1:B:183:LEU:HB3	1:B:187:GLN:HG3	1.41	1.01
1:A:183:LEU:HB3	1:A:184:ASN:HB3	1.39	1.01
1:B:285:ARG:HG2	1:B:285:ARG:HH21	1.23	0.99
1:B:79:THR:HG21	1:B:279:ASP:H	1.27	0.97

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	284/486 (58%)	277 (98%)	7 (2%)	0	100	100
1	B	282/486 (58%)	268 (95%)	11 (4%)	3 (1%)	17	47
All	All	566/972 (58%)	545 (96%)	18 (3%)	3 (0%)	34	68

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	185	PRO
1	B	220	THR
1	B	184	ASN

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	247/424 (58%)	212 (86%)	35 (14%)	4 11
1	B	245/424 (58%)	204 (83%)	41 (17%)	3 7
All	All	492/848 (58%)	416 (85%)	76 (15%)	3 9

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	288	ARG
1	B	79	THR
1	B	235	GLN
1	B	17	ASP
1	B	31	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	236	GLN
1	A	273	GLN
1	B	253	GLN
1	A	255	HIS
1	A	260	ASN

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 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	286/486 (58%)	0.99	47 (16%)	2	1	27, 69, 159, 174	0
1	B	284/486 (58%)	0.82	34 (11%)	6	3	23, 63, 161, 177	0
All	All	570/972 (58%)	0.90	81 (14%)	4	2	23, 66, 160, 177	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	163	ALA	7.2
1	B	188	LEU	7.1
1	B	178	LYS	6.9
1	A	171	ILE	6.5
1	A	182	SER	6.5

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
2	CA	A	488	1/1	0.98	0.17	-1.10	69,69,69,69	0
2	CA	B	488	1/1	0.94	0.10	-3.88	68,68,68,68	0

6.5 Other polymers (i)

There are no such residues in this entry.