



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 21, 2016 – 10:03 PM EDT

PDB ID : 5C79  
Title : PH domain of ASAP1 in complex with diC4-PtdIns(4,5)P2  
Authors : Xia, D.; Tang, W.K.  
Deposited on : 2015-06-24  
Resolution : 1.60 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  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0122  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

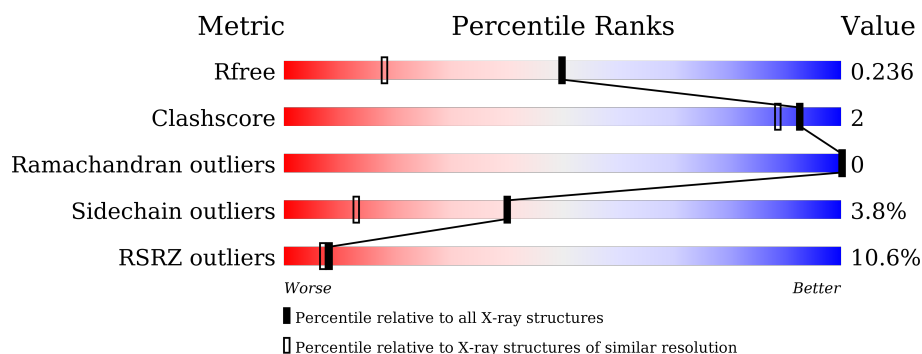
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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	2475 (1.60-1.60)
Clashscore	102246	2732 (1.60-1.60)
Ramachandran outliers	100387	2654 (1.60-1.60)
Sidechain outliers	100360	2653 (1.60-1.60)
RSRZ outliers	91569	2479 (1.60-1.60)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	150	<div> <div>7%</div> <div>61%</div> <div>7%</div> <div>31%</div> </div>
1	B	150	<div> <div>7%</div> <div>63%</div> <div>6%</div> <div>31%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2005 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 Arf-GAP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	104	Total	C	N	O	S	0	0	0
			835	523	152	157	3			
1	B	103	Total	C	N	O	S	0	0	0
			830	520	151	156	3			

There are 46 discrepancies between the modelled and reference sequences:

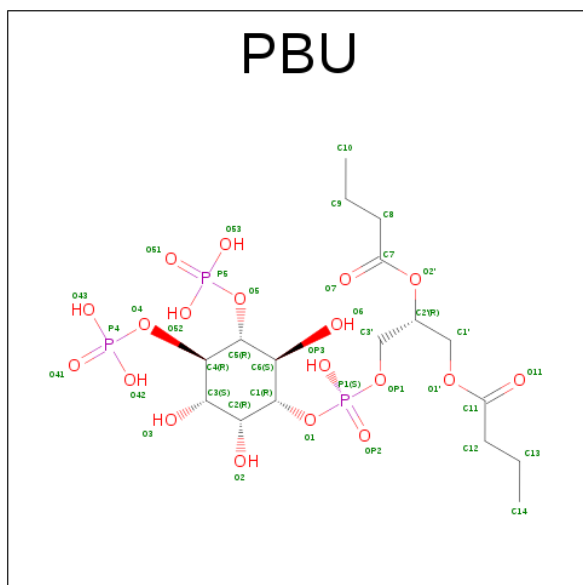
Chain	Residue	Modelled	Actual	Comment	Reference
A	302	MET	-	expression tag	UNP Q9QWY8
A	303	GLY	-	expression tag	UNP Q9QWY8
A	304	HIS	-	expression tag	UNP Q9QWY8
A	305	HIS	-	expression tag	UNP Q9QWY8
A	306	HIS	-	expression tag	UNP Q9QWY8
A	307	HIS	-	expression tag	UNP Q9QWY8
A	308	HIS	-	expression tag	UNP Q9QWY8
A	309	HIS	-	expression tag	UNP Q9QWY8
A	310	HIS	-	expression tag	UNP Q9QWY8
A	311	HIS	-	expression tag	UNP Q9QWY8
A	312	HIS	-	expression tag	UNP Q9QWY8
A	313	HIS	-	expression tag	UNP Q9QWY8
A	314	SER	-	expression tag	UNP Q9QWY8
A	315	SER	-	expression tag	UNP Q9QWY8
A	316	GLY	-	expression tag	UNP Q9QWY8
A	317	HIS	-	expression tag	UNP Q9QWY8
A	318	ILE	-	expression tag	UNP Q9QWY8
A	319	ASP	-	expression tag	UNP Q9QWY8
A	320	ASP	-	expression tag	UNP Q9QWY8
A	321	ASP	-	expression tag	UNP Q9QWY8
A	322	LYS	-	expression tag	UNP Q9QWY8
A	323	HIS	-	expression tag	UNP Q9QWY8
A	324	MET	-	expression tag	UNP Q9QWY8
B	302	MET	-	expression tag	UNP Q9QWY8
B	303	GLY	-	expression tag	UNP Q9QWY8

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	304	HIS	-	expression tag	UNP Q9QWY8
B	305	HIS	-	expression tag	UNP Q9QWY8
B	306	HIS	-	expression tag	UNP Q9QWY8
B	307	HIS	-	expression tag	UNP Q9QWY8
B	308	HIS	-	expression tag	UNP Q9QWY8
B	309	HIS	-	expression tag	UNP Q9QWY8
B	310	HIS	-	expression tag	UNP Q9QWY8
B	311	HIS	-	expression tag	UNP Q9QWY8
B	312	HIS	-	expression tag	UNP Q9QWY8
B	313	HIS	-	expression tag	UNP Q9QWY8
B	314	SER	-	expression tag	UNP Q9QWY8
B	315	SER	-	expression tag	UNP Q9QWY8
B	316	GLY	-	expression tag	UNP Q9QWY8
B	317	HIS	-	expression tag	UNP Q9QWY8
B	318	ILE	-	expression tag	UNP Q9QWY8
B	319	ASP	-	expression tag	UNP Q9QWY8
B	320	ASP	-	expression tag	UNP Q9QWY8
B	321	ASP	-	expression tag	UNP Q9QWY8
B	322	LYS	-	expression tag	UNP Q9QWY8
B	323	HIS	-	expression tag	UNP Q9QWY8
B	324	MET	-	expression tag	UNP Q9QWY8

- Molecule 2 is (2R)-3-[[[(R)-HYDROXY{[(1R,2R,3S,4R,5R,6S)-2,3,6-TRIHYDROXY-4,5-BIS(PHOSPHONOOXY)CYCLOHEXYL]OXY}PHOSPHORYL]OXY}PROPANE-1,2-DIYL DIBUTANOATE (three-letter code: PBU) (formula: C<sub>17</sub>H<sub>33</sub>O<sub>19</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	P	0	0
			39	17	19	3		
2	B	1	Total	C	O	P	0	0
			39	17	19	3		
2	B	1	Total	C	O	P	0	0
			39	17	19	3		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

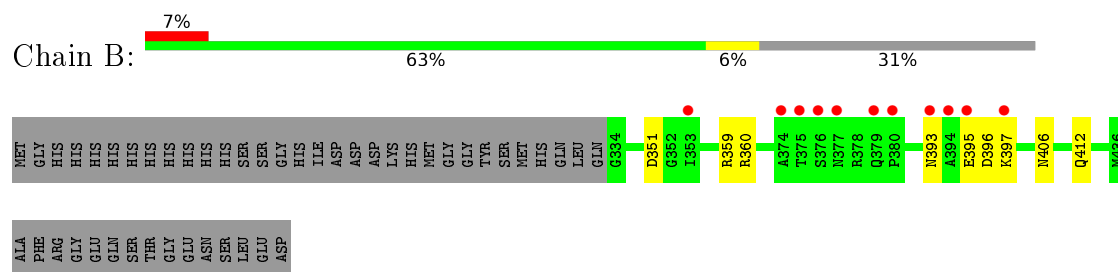
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	106	Total	O	0	0
			106	106		
4	B	116	Total	O	0	0
			116	116		



- Molecule 1: Arf-GAP



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	37.54Å 64.70Å 44.43Å 90.00° 95.63° 90.00°	Depositor
Resolution (Å)	44.21 – 1.60 27.24 – 1.60	Depositor EDS
% Data completeness (in resolution range)	97.1 (44.21-1.60) 97.1 (27.24-1.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.34 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.8.0107	Depositor
R, $R_{free}$	0.198 , 0.233 0.206 , 0.236	Depositor DCC
$R_{free}$ test set	1373 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.7	Xtriage
Anisotropy	0.083	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 36.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 27387 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2005	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.60% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 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: PBU, CL

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	0.65	0/849	0.83	1/1139 (0.1%)
1	B	0.67	0/844	0.85	3/1132 (0.3%)
All	All	0.66	0/1693	0.84	4/2271 (0.2%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	360	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	B	360	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	B	359	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	B	351	ASP	CB-CG-OD2	-5.19	113.62	118.30

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	835	0	843	6	0
1	B	830	0	838	1	0
2	A	39	0	28	2	0
2	B	78	0	56	0	0
3	B	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	106	0	0	1	0
4	B	116	0	0	0	0
All	All	2005	0	1765	7	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 2.

The worst 5 of 7 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:395:GLU:HG2	1:A:396:ASP:N	2.01	0.75
1:A:353:ILE:HB	2:A:800:PBU:H11	1.78	0.66
1:A:393:ASN:HB2	1:A:399:SER:O	2.15	0.47
1:A:398:LYS:HD3	1:A:415:ASP:HA	1.97	0.46
1:B:396:ASP:C	1:B:397:LYS:HG3	2.38	0.43

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	102/150 (68%)	100 (98%)	2 (2%)	0	100	100
1	B	101/150 (67%)	98 (97%)	3 (3%)	0	100	100
All	All	203/300 (68%)	198 (98%)	5 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	91/131 (70%)	88 (97%)	3 (3%)	45	17
1	B	91/131 (70%)	87 (96%)	4 (4%)	35	10
All	All	182/262 (70%)	175 (96%)	7 (4%)	40	13

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	393	ASN
1	B	412	GLN
1	B	395	GLU
1	A	395	GLU
1	B	406	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	366	ASN
1	B	393	ASN
1	B	405	HIS
1	B	406	ASN

### 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

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	PBU	A	800	-	37,39,39	1.28	4 (10%)	53,57,57	2.01	19 (35%)
2	PBU	B	502	-	37,39,39	1.24	2 (5%)	53,57,57	1.62	11 (20%)
2	PBU	B	503	-	37,39,39	1.23	2 (5%)	53,57,57	1.23	7 (13%)

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
2	PBU	A	800	-	-	0/36/60/60	0/1/1/1
2	PBU	B	502	-	-	0/36/60/60	0/1/1/1
2	PBU	B	503	-	-	0/36/60/60	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	800	PBU	P5-O53	-2.12	1.47	1.54
2	A	800	PBU	O4-C4	-2.06	1.37	1.44
2	A	800	PBU	O1'-C11	4.12	1.45	1.33
2	B	503	PBU	O2'-C7	4.19	1.46	1.34
2	A	800	PBU	O2'-C7	4.39	1.47	1.34

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	800	PBU	O5-P5-O51	-3.87	98.24	107.48
2	A	800	PBU	C6-C5-C4	-3.59	103.80	111.39
2	B	502	PBU	O4-P4-O41	-3.58	98.93	107.48
2	B	502	PBU	O5-P5-O51	-3.21	99.82	107.48
2	B	503	PBU	O5-P5-O51	-3.05	100.19	107.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	800	PBU	2	0

## 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	104/150 (69%)	0.67	11 (10%) 8 7	17, 23, 52, 66	0
1	B	103/150 (68%)	0.84	11 (10%) 8 7	16, 22, 61, 84	0
All	All	207/300 (69%)	0.75	22 (10%) 8 7	16, 23, 53, 84	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	376	SER	14.8
1	B	377	ASN	9.0
1	B	397	LYS	7.0
1	B	375	THR	6.4
1	A	394	ALA	6.3

### 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. 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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	PBU	A	800	39/39	0.92	0.14	0.49	27,37,80,80	0
2	PBU	B	502	39/39	0.93	0.16	0.19	26,37,91,92	0
2	PBU	B	503	39/39	0.94	0.12	0.01	18,22,49,52	0
3	CL	B	501	1/1	0.85	0.10	-0.23	53,53,53,53	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.