



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 5, 2022 – 12:15 PM EDT

PDB ID : 7RCH  
Title : Crystal structure of NS1-ED of Vietnam influenza A virus in complex with the p85-beta-iSH2 domain of human PI3K  
Authors : Kim, I.; Zhao, B.; Li, P.; Cho, J.H.  
Deposited on : 2021-07-07  
Resolution : 3.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	2.29
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)	:	2.29

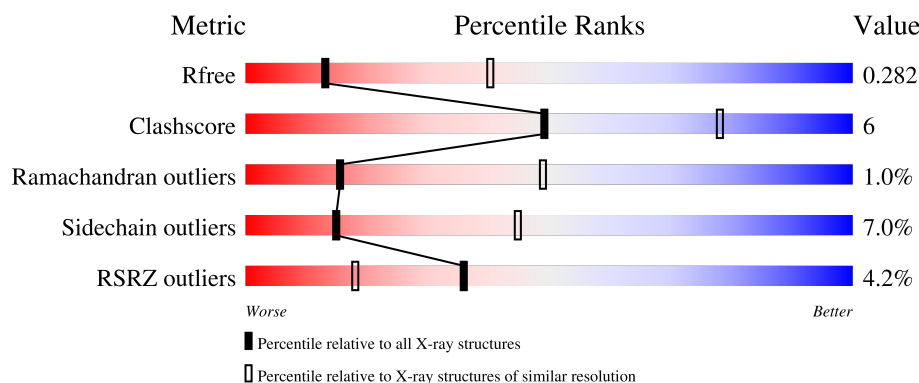
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	126	<div> <div>6%</div> <div> <div></div> <div>80%</div> <div>20%</div> </div> </div>
1	B	126	<div> <div></div> <div> <div>79%</div> <div>14%</div> <div>6%</div> </div> </div>
2	C	165	<div> <div>4%</div> <div> <div></div> <div>62%</div> <div>14%</div> <div>22%</div> </div> </div>
2	D	165	<div> <div>5%</div> <div> <div></div> <div>62%</div> <div>24%</div> <div>13%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4204 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 Non-structural protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	126	Total	C	N	O	S	0	0	0
			985	621	168	189	7			
1	B	119	Total	C	N	O	S	0	0	0
			931	593	157	174	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	182	ALA	TRP	engineered mutation	UNP A5A5U1
B	182	ALA	TRP	engineered mutation	UNP A5A5U1

- Molecule 2 is a protein called Phosphatidylinositol 3-kinase regulatory subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	129	Total	C	N	O	S	0	0	0
			1081	671	196	211	3			
2	D	144	Total	C	N	O	S	0	0	0
			1207	747	221	236	3			

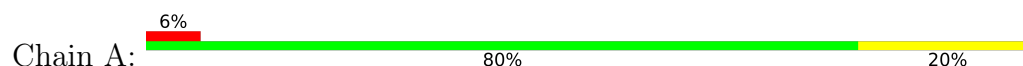
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	433	GLY	-	expression tag	UNP O00459
C	434	SER	-	expression tag	UNP O00459
D	433	GLY	-	expression tag	UNP O00459
D	434	SER	-	expression tag	UNP O00459

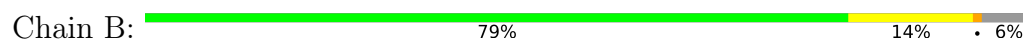
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

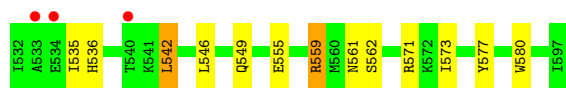
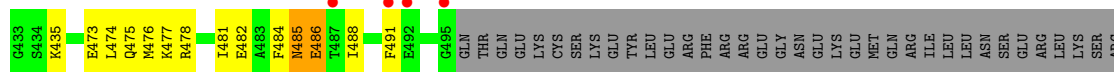
- Molecule 1: Non-structural protein 1



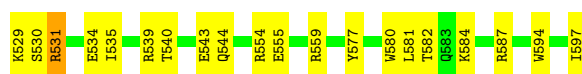
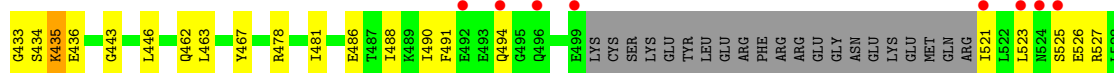
- Molecule 1: Non-structural protein 1



- Molecule 2: Phosphatidylinositol 3-kinase regulatory subunit beta



- Molecule 2: Phosphatidylinositol 3-kinase regulatory subunit beta



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.44Å 92.84Å 67.12Å 90.00° 105.03° 90.00°	Depositor
Resolution (Å)	50.00 – 3.10 50.00 – 3.10	Depositor EDS
% Data completeness (in resolution range)	97.9 (50.00-3.10) 97.9 (50.00-3.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122, PHENIX 1.19.1_4122	Depositor
R, $R_{free}$	0.243 , 0.276 0.248 , 0.282	Depositor DCC
$R_{free}$ test set	625 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	74.3	Xtriage
Anisotropy	0.285	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 32.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	4204	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.28% 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	0.23	0/999	0.49	0/1348
1	B	0.24	0/945	0.49	0/1276
2	C	0.24	0/1094	0.45	0/1462
2	D	0.26	0/1220	0.48	0/1630
All	All	0.24	0/4258	0.48	0/5716

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	985	0	996	14	0
1	B	931	0	955	12	0
2	C	1081	0	1074	13	0
2	D	1207	0	1208	19	0
All	All	4204	0	4233	50	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 6.

All (50) 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)
2:C:555:GLU:OE2	2:C:559:ARG:NH1	2.28	0.67
1:B:127:ALA:HB2	1:B:142:LEU:HD12	1.80	0.62
1:A:177:ILE:HD13	1:A:200:ASN:HB2	1.82	0.62
2:D:435:LYS:HD3	2:D:436:GLU:HG3	1.83	0.60
1:A:132:ILE:HG22	1:A:137:GLU:HB2	1.88	0.55
2:D:527:ARG:HG3	2:D:531:ARG:HH21	1.72	0.55
1:A:177:ILE:HG12	1:A:187:VAL:HG11	1.89	0.54
2:C:580:TRP:CG	1:B:162:PRO:HG3	2.43	0.53
1:A:160:SER:HA	1:B:160:SER:HA	1.92	0.52
2:C:482:GLU:O	2:C:486:GLU:HB2	2.09	0.52
1:B:143:ARG:HD3	1:B:151:ILE:HD13	1.92	0.51
2:D:523:LEU:HD21	2:D:527:ARG:HH21	1.75	0.51
1:A:177:ILE:HD12	1:A:199:ARG:HB2	1.91	0.51
1:B:114:MET:HE3	1:B:180:LEU:HD23	1.92	0.51
2:D:463:LEU:HD21	2:D:559:ARG:HH21	1.76	0.49
2:C:577:TYR:CZ	1:B:162:PRO:HB2	2.47	0.49
1:B:83:ARG:NH2	1:B:134:ASP:OD1	2.46	0.48
2:C:475:GLN:OE1	2:C:478:ARG:NH1	2.38	0.48
2:D:523:LEU:HD12	2:D:526:GLU:HB3	1.96	0.48
2:D:467:TYR:OH	2:D:554:ARG:NH2	2.47	0.48
2:C:473:GLU:HA	2:C:476:MET:HE2	1.95	0.47
2:C:562:SER:O	1:B:135:ARG:NH2	2.47	0.47
1:A:86:THR:OG1	2:C:571:ARG:NH2	2.49	0.46
1:A:191:GLU:O	1:A:195:ARG:HG3	2.15	0.46
1:B:121:LYS:N	1:B:185:ASN:OD1	2.39	0.46
2:C:485:ASN:HA	2:C:488:ILE:HB	1.97	0.45
2:D:488:ILE:HA	2:D:491:PHE:HB2	1.99	0.45
2:D:433:GLY:O	2:D:435:LYS:N	2.50	0.44
1:B:193:ILE:HG23	1:B:197:ALA:HB3	1.99	0.44
1:A:122:THR:HG22	1:A:186:THR:HB	2.00	0.44
1:A:125:LEU:HD12	1:A:189:VAL:HG12	1.98	0.44
2:C:481:ILE:HG13	2:C:542:LEU:HD12	1.99	0.44
2:C:473:GLU:O	2:C:477:LYS:HG3	2.18	0.44
1:B:170:LYS:HE3	1:B:170:LYS:HB2	1.86	0.43
2:D:481:ILE:HD12	2:D:543:GLU:HB2	1.99	0.43
1:B:103:LYS:HD2	1:B:116:GLN:OE1	2.19	0.43
1:A:162:PRO:HB2	2:D:577:TYR:CE1	2.54	0.43
2:D:443:GLY:HA2	2:D:581:LEU:HD11	2.00	0.43
2:D:435:LYS:HB2	2:D:435:LYS:HE2	1.81	0.42
2:D:531:ARG:H	2:D:531:ARG:HG2	1.52	0.42
2:D:446:LEU:HD22	2:D:581:LEU:HD12	2.02	0.42
1:A:84:TYR:HE2	2:C:571:ARG:HG2	1.85	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:443:GLY:HA3	2:D:594:TRP:CE2	2.54	0.42
2:D:486:GLU:O	2:D:490:ILE:HG12	2.20	0.42
1:A:162:PRO:HG3	2:D:580:TRP:CG	2.54	0.41
2:C:474:LEU:HD21	2:C:549:GLN:HB3	2.01	0.41
1:A:194:GLN:HA	1:A:198:TRP:HB3	2.03	0.41
2:D:580:TRP:O	2:D:584:LYS:HG2	2.20	0.41
1:A:86:THR:HG22	1:A:128:ASN:HB3	2.03	0.41
2:D:525:SER:O	2:D:529:LYS:HG2	2.21	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	124/126 (98%)	113 (91%)	8 (6%)	3 (2%)	6	27
1	B	117/126 (93%)	112 (96%)	5 (4%)	0	100	100
2	C	125/165 (76%)	122 (98%)	3 (2%)	0	100	100
2	D	140/165 (85%)	135 (96%)	3 (2%)	2 (1%)	11	40
All	All	506/582 (87%)	482 (95%)	19 (4%)	5 (1%)	15	49

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	204	ASP
2	D	434	SER
2	D	530	SER
1	A	82	SER
1	A	205	GLY



### 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	108/108 (100%)	105 (97%)	3 (3%)	43	73
1	B	102/108 (94%)	100 (98%)	2 (2%)	55	80
2	C	116/151 (77%)	104 (90%)	12 (10%)	7	27
2	D	131/151 (87%)	116 (88%)	15 (12%)	5	22
All	All	457/518 (88%)	425 (93%)	32 (7%)	15	45

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	83	ARG
1	A	120	ASP
1	A	180	LEU
2	C	435	LYS
2	C	484	PHE
2	C	485	ASN
2	C	486	GLU
2	C	491	PHE
2	C	535	ILE
2	C	536	HIS
2	C	542	LEU
2	C	546	LEU
2	C	559	ARG
2	C	561	ASN
2	C	573	ILE
1	B	89	THR
1	B	135	ARG
2	D	435	LYS
2	D	462	GLN
2	D	478	ARG
2	D	494	GLN
2	D	521	ILE
2	D	531	ARG
2	D	534	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	D	535	ILE
2	D	539	ARG
2	D	540	THR
2	D	544	GLN
2	D	555	GLU
2	D	582	THR
2	D	587	ARG
2	D	597	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	116	GLN
2	C	494	GLN
2	C	549	GLN
2	C	592	ASN
2	D	544	GLN
2	D	588	GLN

### 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 monosaccharides 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<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	126/126 (100%)	0.10	7 (5%) 24 11	41, 67, 113, 128	0
1	B	119/126 (94%)	-0.15	0 100 100	38, 61, 87, 109	0
2	C	129/165 (78%)	0.27	7 (5%) 25 12	49, 80, 177, 212	0
2	D	144/165 (87%)	0.10	8 (5%) 24 11	39, 74, 190, 212	0
All	All	518/582 (89%)	0.08	22 (4%) 36 18	38, 70, 167, 212	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	533	ALA	5.4
2	D	525	SER	4.1
1	A	205	GLY	4.0
2	D	524	ASN	3.8
2	C	492	GLU	3.8
2	C	534	GLU	3.8
2	C	491	PHE	3.8
1	A	200	ASN	3.7
2	D	496	GLN	3.6
2	C	487	THR	3.2
2	D	499	GLU	3.2
2	D	521	ILE	2.9
1	A	206	ARG	2.9
2	D	492	GLU	2.9
2	D	523	LEU	2.7
2	C	540	THR	2.5
1	A	203	GLU	2.5
2	D	494	GLN	2.5
1	A	202	ASP	2.4
1	A	204	ASP	2.2
1	A	201	SER	2.1
2	C	495	GLY	2.0

## 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.