



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 10:28 PM GMT

PDB ID : 2R05  
Title : Crystal Structure of ALIX/AIP1 in complex with the HIV-1 YPLASL Late Domain  
Authors : Hill, C.P.; Zhai, Q.; Fisher, R.D.  
Deposited on : 2007-08-17  
Resolution : 2.55 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

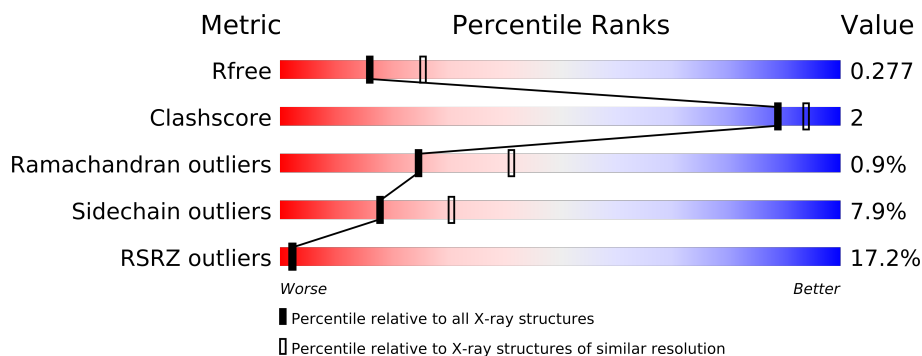
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.55 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))
$R_{free}$	66092	3413 (2.58-2.50)
Clashscore	79885	4284 (2.58-2.50)
Ramachandran outliers	78287	4193 (2.58-2.50)
Sidechain outliers	78261	4195 (2.58-2.50)
RSRZ outliers	66119	3414 (2.58-2.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	697	
2	B	11	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5591 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Programmed cell death 6-interacting protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	697	Total	C	N	O	S	0	0	0
			5486	3461	938	1069	18			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	268	TYR	LYS	ENGINEERED	UNP Q8WUM4
A	269	TYR	LYS	ENGINEERED	UNP Q8WUM4

- Molecule 2 is a protein called p6-Gag.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	11	Total	C	N	O	0	0	0
			88	58	14	16			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	17	Total	O	0	0
			17	17		

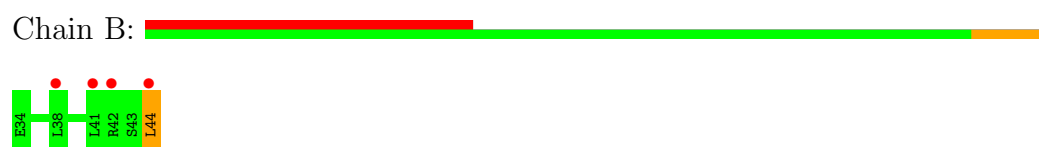
### 3 Residue-property plots

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: Programmed cell death 6-interacting protein



- Molecule 2: p6-Gag



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	146.09Å 99.44Å 73.54Å 90.00° 107.29° 90.00°	Depositor
Resolution (Å)	27.90 – 2.55 24.33 – 2.54	Depositor EDS
% Data completeness (in resolution range)	98.7 (27.90-2.55) 98.2 (24.33-2.54)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.37 (at 2.53Å)	Xtriage
Refinement program	REFMAC 5.3.0008	Depositor
R, $R_{free}$	0.225 , 0.280 0.223 , 0.277	Depositor DCC
$R_{free}$ test set	1622 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	69.1	Xtriage
Anisotropy	0.355	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 69.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 32600 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5591	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	95.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.43% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0.57	1/5570 (0.0%)	0.64	0/7523
2	B	2.35	1/89 (1.1%)	0.60	0/120
All	All	0.63	2/5659 (0.0%)	0.64	0/7643

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	44	LEU	C-O	21.85	1.64	1.23
1	A	246	ALA	CA-CB	5.29	1.63	1.52

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5486	0	0	12	0
2	B	88	0	0	1	0
3	A	17	0	0	0	0
All	All	5591	0	0	13	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

All (13) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:44:LEU:C	2:B:44:LEU:O	1.64	1.35
1:A:386:ARG:NH2	1:A:570:GLU:OE2	2.35	0.58
1:A:517:ARG:CG	1:A:517:ARG:NH1	2.69	0.55
1:A:462:GLU:OE1	1:A:482:ASN:ND2	2.40	0.54
1:A:434:ASP:OD1	1:A:513:TYR:OH	2.27	0.52
1:A:71:TYR:OH	1:A:119:GLU:OE2	2.30	0.49
1:A:322:ARG:NH1	1:A:323:VAL:O	2.51	0.43
1:A:233:TYR:O	1:A:234:LYS:CG	2.67	0.42
1:A:456:ARG:NE	1:A:460:GLU:OE1	2.53	0.42
1:A:98:ALA:N	1:A:178:ASP:OD2	2.53	0.41
1:A:409:SER:OG	1:A:411:ASP:CB	2.68	0.41
1:A:462:GLU:OE1	1:A:482:ASN:CG	2.59	0.41
1:A:197:GLU:OE2	1:A:248:LYS:NZ	2.53	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	695/697 (100%)	651 (94%)	38 (6%)	6 (1%)	25	41
2	B	9/11 (82%)	8 (89%)	1 (11%)	0	100	100
All	All	704/708 (99%)	659 (94%)	39 (6%)	6 (1%)	25	41

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	81	LYS
1	A	544	GLN
1	A	87	ASN
1	A	88	GLN
1	A	82	PHE
1	A	54	VAL

### 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	600/600 (100%)	552 (92%)	48 (8%)	17	30
2	B	10/10 (100%)	10 (100%)	0	100	100
All	All	610/610 (100%)	562 (92%)	48 (8%)	18	30

All (48) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	THR
1	A	38	GLN
1	A	54	VAL
1	A	75	ILE
1	A	78	ILE
1	A	86	GLU
1	A	88	GLN
1	A	104	LEU
1	A	109	VAL
1	A	140	LEU
1	A	167	VAL
1	A	172	SER
1	A	184	VAL
1	A	201	LEU
1	A	239	LYS
1	A	289	SER
1	A	297	VAL
1	A	312	LYS
1	A	322	ARG
1	A	337	LEU
1	A	338	VAL
1	A	350	LYS
1	A	352	THR
1	A	357	LYS
1	A	408	VAL
1	A	417	ILE
1	A	471	LYS
1	A	473	LYS
1	A	478	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	484	LEU
1	A	494	ASN
1	A	517	ARG
1	A	544	GLN
1	A	547	GLU
1	A	569	LEU
1	A	573	LEU
1	A	585	LEU
1	A	596	GLU
1	A	597	GLU
1	A	618	SER
1	A	628	ASN
1	A	631	VAL
1	A	634	GLN
1	A	648	LEU
1	A	650	GLU
1	A	651	GLU
1	A	653	LEU
1	A	682	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains 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<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	697/697 (100%)	1.11	118 (16%) 2 2	50, 93, 142, 190	2 (0%)
2	B	11/11 (100%)	1.67	4 (36%) 1 0	102, 121, 173, 182	0
All	All	708/708 (100%)	1.12	122 (17%) 2 2	50, 94, 144, 190	2 (0%)

All (122) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	29	TYR	9.1
1	A	103	SER	7.8
1	A	696	PHE	7.2
1	A	88	GLN	7.0
1	A	540	ALA	6.6
1	A	541	LYS	6.4
1	A	631	VAL	6.0
1	A	30	PRO	5.9
1	A	593	VAL	5.6
1	A	33	GLY	5.5
1	A	60	LYS	5.3
1	A	545	GLY	5.3
1	A	2	ALA	5.3
1	A	104	LEU	5.2
1	A	57	PRO	5.1
1	A	58	LEU	4.9
1	A	35	GLU	4.8
1	A	90	CYS	4.7
1	A	543	MET	4.6
1	A	144	GLU	4.6
1	A	542	THR	4.4
1	A	34	GLU	4.4
1	A	107	GLY	4.4
1	A	426	GLU	4.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	590	GLN	4.2
1	A	122	CYS	4.2
1	A	127	CYS	4.2
1	A	128	ALA	4.1
1	A	87	ASN	4.1
1	A	126	ASN	3.9
1	A	408	VAL	3.9
1	A	539	PRO	3.8
1	A	39	TYR	3.8
1	A	106	GLY	3.8
1	A	105	PHE	3.8
1	A	641	GLN	3.7
1	A	646	ALA	3.7
1	A	595	ASN	3.7
1	A	36	GLN	3.7
1	A	173	ARG	3.7
1	A	642	SER	3.7
1	A	589	ALA	3.5
1	A	592	GLY	3.5
1	A	31	SER	3.5
1	A	172	SER	3.5
1	A	644	ASN	3.5
1	A	327	LYS	3.4
1	A	125	PHE	3.4
1	A	317	PHE	3.4
1	A	124	LEU	3.4
1	A	37	ALA	3.4
1	A	123	VAL	3.3
1	A	86	GLU	3.3
1	A	23	LYS	3.3
1	A	96	LYS	3.3
1	A	548	VAL	3.2
2	B	42	ARG	3.2
1	A	544	GLN	3.2
1	A	653	LEU	3.2
1	A	129	ALA	3.1
1	A	409	SER	3.1
1	A	546	SER	3.1
1	A	652	VAL	3.1
1	A	56	ARG	3.1
1	A	594	ILE	3.0
1	A	16	ASP	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	89	ILE	3.0
1	A	453	GLU	3.0
1	A	61	HIS	3.0
2	B	41	LEU	2.9
1	A	643	ASN	2.9
1	A	306	ARG	2.7
1	A	108	SER	2.7
1	A	597	GLU	2.7
1	A	102	GLY	2.7
1	A	118	TYR	2.7
1	A	480	PRO	2.7
1	A	601	VAL	2.6
1	A	41	ARG	2.6
1	A	68	LEU	2.6
1	A	293	GLU	2.6
1	A	538	ASN	2.6
1	A	640	LYS	2.6
1	A	51	ARG	2.6
1	A	547	GLU	2.6
1	A	328	ASP	2.6
2	B	38	LEU	2.6
1	A	55	GLY	2.6
1	A	19	LYS	2.6
1	A	634	GLN	2.6
1	A	121	SER	2.5
1	A	190	ILE	2.5
1	A	22	VAL	2.5
1	A	169	SER	2.4
1	A	44	GLU	2.4
1	A	694	ILE	2.4
1	A	613	THR	2.4
1	A	598	ALA	2.4
1	A	638	LYS	2.4
1	A	156	ALA	2.4
2	B	44	LEU	2.4
1	A	71	TYR	2.3
1	A	130	LEU	2.3
1	A	431	GLN	2.3
1	A	486	LYS	2.3
1	A	193	ALA	2.3
1	A	109	VAL	2.3
1	A	119	GLU	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	655	ASN	2.2
1	A	368	ALA	2.2
1	A	191	MET	2.2
1	A	639	MET	2.2
1	A	380	ARG	2.2
1	A	485	TYR	2.1
1	A	294	TYR	2.1
1	A	657	ALA	2.1
1	A	26	GLN	2.0
1	A	233	TYR	2.0
1	A	698	ARG	2.0
1	A	483	GLU	2.0
1	A	114	ALA	2.0
1	A	372	GLN	2.0

## 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 ⓘ

There are no such residues in this entry.