



# Full wwPDB X-ray Structure Validation Report

Sep 16, 2014 – 03:56 PM EDT

PDB ID : 4MZ5  
Title : Structure of importin-alpha: dUTPase NLS complex  
Authors : Marfori, M.; Rona, G.; Vertessy, B.G.; Kobe, B.  
Deposited on : 2013-09-29  
Resolution : 2.10 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>

---

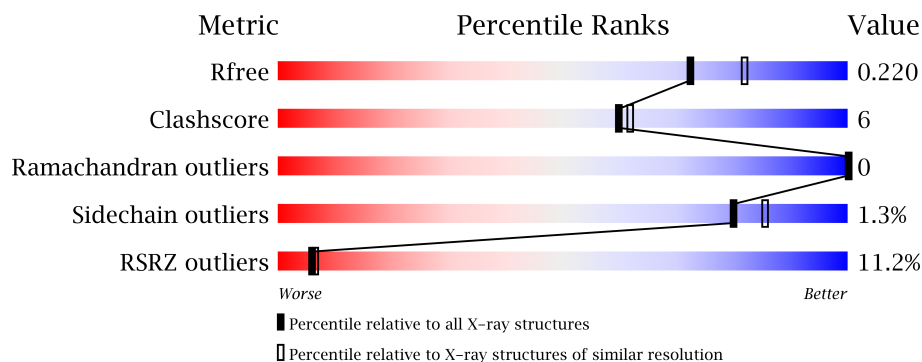
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23489  
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) : stable23489

# 1 Overall quality at a glance




The reported resolution of this entry is 2.10 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))
$R_{free}$	66092	3012 (2.10-2.10)
Clashscore	79885	3649 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (2.10-2.10)
RSRZ outliers	66119	3013 (2.10-2.10)

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 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	13	
1	C	13	
2	E	509	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3596 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 Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	8	Total	C	N	O	0	0	0
			62	37	15	10			
1	C	12	Total	C	N	O	0	0	0
			89	54	19	16			

- Molecule 2 is a protein called Importin subunit alpha-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	427	Total	C	N	O	S	0	0	0
			3253	2071	552	620	10			

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	20	MET	-	EXPRESSION TAG	UNP P52293
E	21	HIS	-	EXPRESSION TAG	UNP P52293
E	22	HIS	-	EXPRESSION TAG	UNP P52293
E	23	HIS	-	EXPRESSION TAG	UNP P52293
E	24	HIS	-	EXPRESSION TAG	UNP P52293
E	25	HIS	-	EXPRESSION TAG	UNP P52293
E	26	HIS	-	EXPRESSION TAG	UNP P52293
E	27	SER	-	EXPRESSION TAG	UNP P52293
E	28	SER	-	EXPRESSION TAG	UNP P52293
E	29	GLY	-	EXPRESSION TAG	UNP P52293
E	30	LEU	-	EXPRESSION TAG	UNP P52293
E	31	VAL	-	EXPRESSION TAG	UNP P52293
E	32	PRO	-	EXPRESSION TAG	UNP P52293
E	33	ARG	-	EXPRESSION TAG	UNP P52293
E	34	GLY	-	EXPRESSION TAG	UNP P52293
E	35	SER	-	EXPRESSION TAG	UNP P52293
E	36	GLY	-	EXPRESSION TAG	UNP P52293
E	37	MET	-	EXPRESSION TAG	UNP P52293

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	38	LEU	-	EXPRESSION TAG	UNP P52293
E	39	GLU	-	EXPRESSION TAG	UNP P52293
E	40	THR	-	EXPRESSION TAG	UNP P52293
E	41	ALA	-	EXPRESSION TAG	UNP P52293
E	42	ALA	-	EXPRESSION TAG	UNP P52293
E	43	ALA	-	EXPRESSION TAG	UNP P52293
E	44	LEU	-	EXPRESSION TAG	UNP P52293
E	45	PHE	-	EXPRESSION TAG	UNP P52293
E	46	GLU	-	EXPRESSION TAG	UNP P52293
E	47	ARG	-	EXPRESSION TAG	UNP P52293
E	48	ASN	-	EXPRESSION TAG	UNP P52293
E	49	HIS	-	EXPRESSION TAG	UNP P52293
E	50	MET	-	EXPRESSION TAG	UNP P52293
E	51	ASP	-	EXPRESSION TAG	UNP P52293
E	52	SER	-	EXPRESSION TAG	UNP P52293
E	53	PRO	-	EXPRESSION TAG	UNP P52293
E	54	ASP	-	EXPRESSION TAG	UNP P52293
E	55	LEU	-	EXPRESSION TAG	UNP P52293
E	56	GLY	-	EXPRESSION TAG	UNP P52293
E	57	THR	-	EXPRESSION TAG	UNP P52293
E	58	ASP	-	EXPRESSION TAG	UNP P52293
E	59	ASP	-	EXPRESSION TAG	UNP P52293
E	60	ASP	-	EXPRESSION TAG	UNP P52293
E	61	ASP	-	EXPRESSION TAG	UNP P52293
E	62	LEU	-	EXPRESSION TAG	UNP P52293
E	63	ALA	-	EXPRESSION TAG	UNP P52293
E	64	MET	-	EXPRESSION TAG	UNP P52293
E	65	ALA	-	EXPRESSION TAG	UNP P52293
E	66	ASP	-	EXPRESSION TAG	UNP P52293
E	67	ILE	-	EXPRESSION TAG	UNP P52293
E	68	GLY	-	EXPRESSION TAG	UNP P52293
E	69	SER	-	EXPRESSION TAG	UNP P52293

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total O 2 2	0	0
3	C	11	Total O 11 11	0	0
3	E	179	Total O 179 179	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.88Å 90.00Å 99.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.00 – 2.10 41.99 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.4 (42.00-2.10) 99.5 (41.99-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 2.10Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.199 , 0.232 0.185 , 0.220	Depositor DCC
$R_{free}$ test set	2084 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	34.5	Xtriage
Anisotropy	0.676	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 54.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 41372 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3596	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.07% 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	0.33	0/63	0.65	0/83
1	C	0.29	0/90	0.48	0/120
2	E	0.33	0/3311	0.48	0/4512
All	All	0.33	0/3464	0.48	0/4715

There are no bond length outliers.

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	62	0	67	1	0
1	C	89	0	94	1	0
2	E	3253	0	3328	40	0
3	A	2	0	0	0	0
3	C	11	0	0	1	0
3	E	179	0	0	6	0
All	All	3596	0	3489	41	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 6.

All (41) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:E:482:GLU:HG3	2:E:485:TYR:HE2	1.28	0.94
2:E:482:GLU:HG3	2:E:485:TYR:CE2	2.02	0.94
2:E:88:ASN:OD1	3:E:779:HOH:O	2.02	0.77
2:E:71:GLN:O	2:E:71:GLN:HG2	1.84	0.76
2:E:207:ASP:OD1	2:E:251:GLN:NE2	2.25	0.69
2:E:434:THR:HG21	2:E:481:ASN:HD22	1.57	0.68
2:E:386:LEU:HD21	2:E:425:LEU:HD13	1.76	0.66
2:E:306:GLU:OE2	3:E:776:HOH:O	2.15	0.65
2:E:423:GLU:OE1	3:E:697:HOH:O	2.14	0.65
2:E:80:ASP:OD2	3:E:701:HOH:O	2.13	0.64
1:A:15:ARG:NH2	2:E:396:GLU:OE1	2.28	0.64
2:E:440:ILE:O	2:E:444:ILE:HD12	2.04	0.57
2:E:481:ASN:OD1	2:E:482:GLU:N	2.35	0.56
2:E:71:GLN:OE1	2:E:71:GLN:N	2.39	0.55
2:E:71:GLN:O	2:E:74:VAL:HG23	2.07	0.54
2:E:484:VAL:O	2:E:488:SER:OG	2.25	0.53
2:E:350:ASN:OD1	3:E:731:HOH:O	2.19	0.51
1:C:11:SER:OG	3:C:105:HOH:O	2.20	0.50
2:E:284:GLU:O	2:E:288:MET:HG3	2.12	0.49
2:E:426:MET:HE1	2:E:444:ILE:HG12	1.97	0.47
2:E:321:VAL:HG21	2:E:358:THR:HG23	1.97	0.47
2:E:73:THR:HA	2:E:76:TRP:NE1	2.30	0.46
2:E:482:GLU:HA	2:E:485:TYR:CD2	2.51	0.46
2:E:482:GLU:CG	2:E:485:TYR:HE2	2.15	0.45
2:E:469:GLY:O	2:E:473:ILE:HG13	2.17	0.45
2:E:348:LYS:HA	2:E:348:LYS:HD3	1.78	0.44
2:E:366:ARG:HB2	2:E:368:ASP:OD1	2.18	0.43
2:E:299:VAL:O	2:E:302:LEU:HB2	2.19	0.42
2:E:426:MET:CE	2:E:444:ILE:HG12	2.49	0.42
2:E:175:HIS:HB3	2:E:177:HIS:CE1	2.55	0.42
2:E:215:VAL:O	2:E:258:ARG:NH2	2.52	0.41
2:E:423:GLU:HG3	2:E:467:CYS:HB2	2.02	0.41
2:E:425:LEU:HD12	2:E:425:LEU:HA	1.92	0.41
2:E:291:LYS:HA	2:E:291:LYS:HD2	1.76	0.41
2:E:421:ILE:O	2:E:425:LEU:HB2	2.21	0.41
2:E:273:TRP:CD2	2:E:312:PRO:HB3	2.56	0.41
2:E:401:ILE:HG23	2:E:416:LEU:HD21	2.03	0.41
2:E:456:GLU:O	2:E:459:LYS:HG2	2.20	0.41
2:E:452:GLU:C	2:E:454:LEU:H	2.24	0.41
2:E:360:SER:HA	2:E:400:ALA:HA	2.03	0.40
2:E:347:PRO:O	3:E:762:HOH:O	2.22	0.40

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	6/13 (46%)	6 (100%)	0	0	100	100
1	C	10/13 (77%)	10 (100%)	0	0	100	100
2	E	425/509 (84%)	419 (99%)	6 (1%)	0	100	100
All	All	441/535 (82%)	435 (99%)	6 (1%)	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	7/10 (70%)	7 (100%)	0	100	100
1	C	9/10 (90%)	9 (100%)	0	100	100
2	E	358/425 (84%)	353 (99%)	5 (1%)	78	83
All	All	374/445 (84%)	369 (99%)	5 (1%)	80	85

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	75	ASN
2	E	225	TYR
2	E	483	SER
2	E	485	TYR
2	E	488	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such

sidechains are listed below:

Mol	Chain	Res	Type
2	E	71	GLN

### 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	8/13 (61%)	0.68	1 (12%) 5 5	52, 59, 85, 89	0
1	C	12/13 (92%)	0.77	3 (25%) 1 1	35, 45, 81, 107	0
2	E	427/509 (83%)	0.60	46 (10%) 6 7	27, 42, 93, 131	0
All	All	447/535 (83%)	0.61	50 (11%) 6 6	27, 43, 93, 131	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	485	TYR	8.1
2	E	476	LEU	7.4
2	E	484	VAL	6.0
2	E	489	LEU	5.3
2	E	475	ALA	5.1
2	E	477	GLN	4.9
2	E	480	GLU	4.6
2	E	479	HIS	4.1
1	C	9	ALA	4.1
2	E	491	LEU	4.0
2	E	492	ILE	3.9
1	C	20	GLU	3.9
2	E	486	LYS	3.8
2	E	496	PHE	3.8
2	E	482	GLU	3.7
2	E	107	GLU	3.6
2	E	71	GLN	3.6
2	E	478	ARG	3.5
2	E	431	ALA	3.5
2	E	495	TYR	3.3
2	E	432	LYS	3.3
2	E	483	SER	3.2
2	E	487	ALA	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	E	488	SER	3.0
2	E	199	LEU	3.0
2	E	430	SER	3.0
2	E	493	GLU	2.8
2	E	378	LEU	2.7
2	E	468	GLY	2.7
2	E	434	THR	2.7
2	E	108	LYS	2.7
2	E	449	GLN	2.6
2	E	379	VAL	2.6
2	E	473	ILE	2.5
1	A	12	PRO	2.5
2	E	444	ILE	2.5
2	E	497	SER	2.4
2	E	122	PRO	2.4
2	E	200	VAL	2.4
1	C	10	ILE	2.3
2	E	471	ASP	2.2
2	E	419	CYS	2.2
2	E	164	ILE	2.1
2	E	72	GLY	2.1
2	E	474	GLU	2.1
2	E	490	ASN	2.1
2	E	494	LYS	2.1
2	E	481	ASN	2.1
2	E	157	ALA	2.1
2	E	383	VAL	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.