wwPDB X-ray Structure Validation Summary Report (i)

Oct 22, 2014 – 03:22 PM EDT

PDB ID	:	4PVZ
Title	:	Structure of yeast importin a bound to the membrane protein Nuclear Localization Signal sequence of INM protein Heh2
Authors	:	Lokareddy, R.K.; Hapsari, A.R.; van Rheenen, M.; Bhardwaj, A.; Veenhoff, L.M.; Cingolani, C.
Deposited on	:	2014-03-18
Resolution	:	2.50 Å(reported)

DISCLAIMER
This is a preliminary version of the new style of wwPDB validation report.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at http://wwpdb.org/ValidationPDFNotes.html

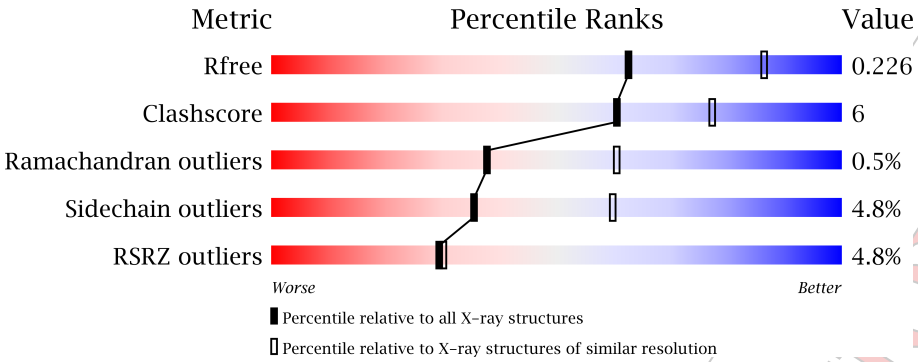
The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24055
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)	:	stable24055

1 Overall quality at a glance

The reported resolution of this entry is 2.50 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	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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 >=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	422	
1	B	422	
2	C	49	
2	D	49	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7282 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 Importin subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	422	Total	C	N	O	S	0	0	0
			3278	2074	554	634	16			
1	B	422	Total	C	N	O	S	0	0	0
			3278	2074	554	634	16			

- Molecule 2 is a protein called Inner nuclear membrane protein HEH2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	35	Total	C	N	O	0	0	0
			282	167	60	55			
2	D	19	Total	C	N	O	0	0	0
			167	101	41	25			

- Molecule 3 is water.

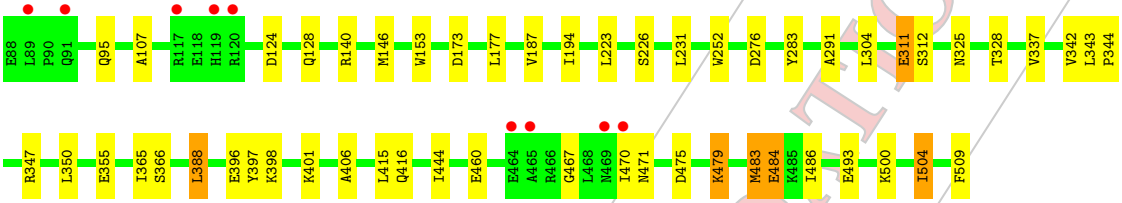
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	127	Total	O	0	0
			127	127		
3	B	132	Total	O	0	0
			132	132		
3	C	11	Total	O	0	0
			11	11		
3	D	7	Total	O	0	0
			7	7		

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.

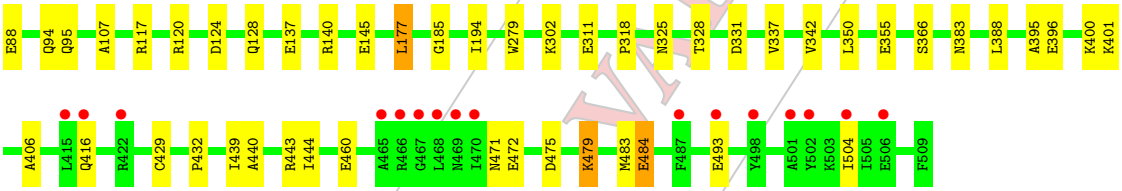
- Molecule 1: Importin subunit alpha

Chain A:



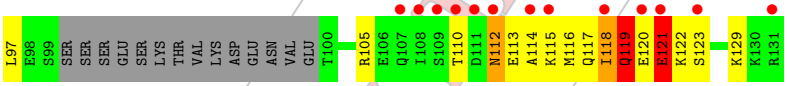
- Molecule 1: Importin subunit alpha

Chain B:



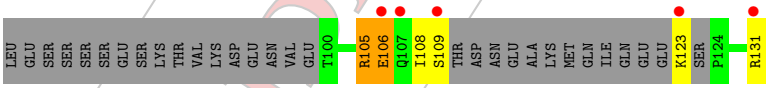
- Molecule 2: Inner nuclear membrane protein HEH2

Chain C:



- Molecule 2: Inner nuclear membrane protein HEH2

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	49.47Å 105.32Å 224.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.55 – 2.50 49.62 – 2.50	Depositor EDS
% Data completeness (in resolution range)	96.9 (30.55-2.50) 96.6 (49.62-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	8.80	Depositor
$< I/\sigma(I) >$ [1]	1.04 (at 2.48Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.189 , 0.227 0.191 , 0.226	Depositor DCC
R_{free} test set	2020 reflections (5.00%)	DCC
Wilson B-factor (Å^2)	32.8	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 30.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 40538 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7282	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.45% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] 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	0.36	0/3330	0.49	0/4530
1	B	0.33	0/3330	0.48	0/4530
2	C	0.42	0/282	0.60	0/369
2	D	0.34	0/166	0.46	0/210
All	All	0.35	0/7108	0.49	0/9639

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	C	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	119	GLN	Peptide

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	3278	0	3330	39	0
1	B	3278	0	3330	33	0
2	C	282	0	289	25	0
2	D	167	0	195	1	0
3	A	127	0	0	3	0
3	B	132	0	0	7	0
3	C	11	0	0	2	0
3	D	7	0	0	0	0
All	All	7282	0	7144	84	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.

The worst 5 of 84 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:C:117:GLN:O	2:C:120:GLU:HB2	1.53	1.08
2:C:121:GLU:HA	2:C:121:GLU:OE1	1.65	0.95
1:A:483:MET:CE	1:B:120:ARG:HD2	2.00	0.91
1:A:483:MET:HE3	1:B:120:ARG:HD2	1.57	0.83
2:C:121:GLU:CA	2:C:121:GLU:OE1	2.30	0.79

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	420/422 (100%)	410 (98%)	9 (2%)	1 (0%)	56 79
1	B	420/422 (100%)	412 (98%)	8 (2%)	0	100 100
2	C	33/49 (67%)	28 (85%)	4 (12%)	1 (3%)	7 9
2	D	15/49 (31%)	13 (87%)	0	2 (13%)	0 0
All	All	888/942 (94%)	863 (97%)	21 (2%)	4 (0%)	38 60

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	467	GLY
2	D	106	GLU
2	D	108	ILE
2	C	121	GLU

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	363/364 (100%)	349 (96%)	14 (4%)	43	70
1	B	363/364 (100%)	352 (97%)	11 (3%)	53	80
2	C	30/48 (62%)	23 (77%)	7 (23%)	1	1
2	D	18/48 (38%)	13 (72%)	5 (28%)	0	0
All	All	774/824 (94%)	737 (95%)	37 (5%)	35	60

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	311	GLU
1	B	479	LYS
2	D	109	SER
1	B	328	THR
1	B	350	LEU

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

5.3.3 RNA (i)

There are no RNA chains 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)

There are no ligands in this entry.

5.7 Other polymers (i)

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	422/422 (100%)	-0.09	9 (2%) 60 63	19, 41, 78, 109	0
1	B	422/422 (100%)	0.04	16 (3%) 38 40	19, 37, 95, 138	0
2	C	35/49 (71%)	1.40	13 (37%) 1 1	26, 78, 129, 134	0
2	D	19/49 (38%)	0.93	5 (26%) 1 1	34, 48, 87, 103	0
All	All	898/942 (95%)	0.05	43 (4%) 29 30	19, 40, 95, 138	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	107	GLN	6.1
1	B	468	LEU	5.6
1	B	465	ALA	5.4
1	B	466	ARG	5.3
1	B	467	GLY	5.3

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 (i)

There are no such residues in this entry.