W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2017 – 04:26 pm GMT

PDB ID	:	3TJ3
Title	:	Structure of importin a5 bound to the N-terminus of Nup50
Authors	:	Pumroy, R.; Nardozzi, J.D.; Hart, D.J.; Root, M.J.; Cingolani, G.
Deposited on	:	2011-08-23
Resolution	:	2.70 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*
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 2.70 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.
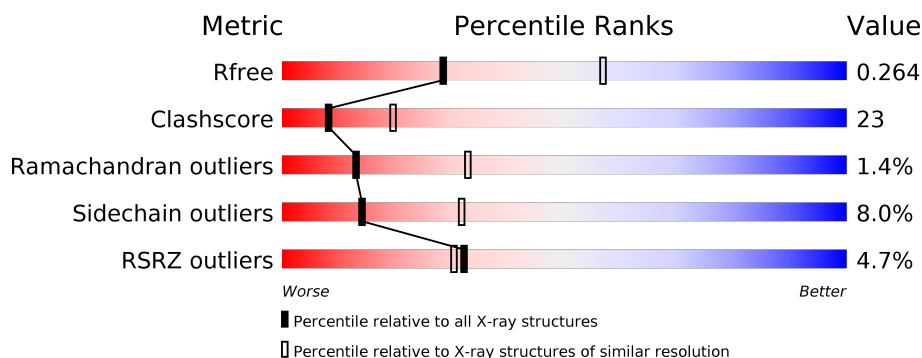
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	2259 (2.70-2.70)
Clashscore	112137	2590 (2.70-2.70)
Ramachandran outliers	110173	2550 (2.70-2.70)
Sidechain outliers	110143	2550 (2.70-2.70)
RSRZ outliers	101464	2275 (2.70-2.70)

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. 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 <=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	447	3% / 62% / 30% / • / 5%
1	B	447	4% / 63% / 28% / 5% / 5%
2	C	114	7% / 22% / 17% / 5% • / 55%
2	D	114	7% / 25% / 13% / 6% / 56%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 7545 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 Importin subunit alpha-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	426	Total 3309	C 2104	N 556	O 631	S 18	0	0	0
1	B	426	Total 3311	C 2105	N 556	O 632	S 18	0	0	0

- Molecule 2 is a protein called Nuclear pore complex protein Nup50.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	51	Total 393	C 238	N 77	O 76	S 2	0	0	0
2	D	50	Total 402	C 243	N 76	O 81	S 2	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-4	GLY	-	EXPRESSION TAG	UNP Q9UKX7
C	-3	PRO	-	EXPRESSION TAG	UNP Q9UKX7
C	-2	LEU	-	EXPRESSION TAG	UNP Q9UKX7
C	-1	GLY	-	EXPRESSION TAG	UNP Q9UKX7
C	0	SER	-	EXPRESSION TAG	UNP Q9UKX7
D	-4	GLY	-	EXPRESSION TAG	UNP Q9UKX7
D	-3	PRO	-	EXPRESSION TAG	UNP Q9UKX7
D	-2	LEU	-	EXPRESSION TAG	UNP Q9UKX7
D	-1	GLY	-	EXPRESSION TAG	UNP Q9UKX7
D	0	SER	-	EXPRESSION TAG	UNP Q9UKX7

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	65	Total 65	O 65	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	8	Total 8	O 8	0	0
3	B	49	Total 49	O 49	0	0
3	D	8	Total 8	O 8	0	0

3 Residue-property plots ⓘ
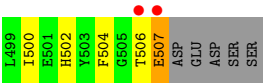
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.
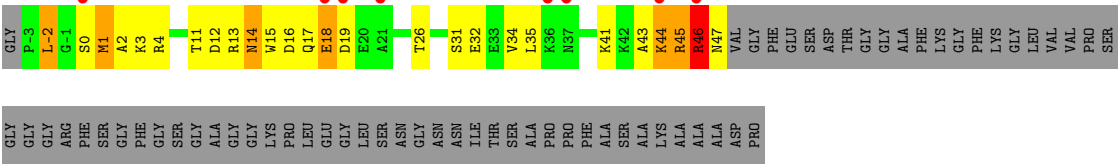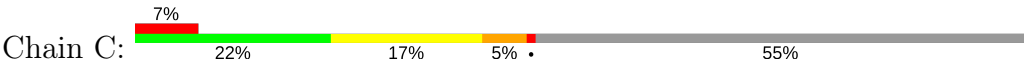
- Molecule 1: Importin subunit alpha-1

- Molecule 1: Importin subunit alpha-1

L499 I500 E501 H502 Y503 F504 G505 T506 E507 ASP GLU ASP SER SER

• Molecule 2: Nuclear pore complex protein Nup50

Chain C:

7%

22%   17%   5% •   55%

GLY P-3 L-2 G-1 S0 M1 A2 K3 R4 T11 D12 R13 N14 W15 D16 Q17 E18 D19 E20 A21 T26 S31 E32 E33 V34 L35 K36 N37 K41 K42 A43 K44 R45 R46 N47 VAL GLY PHE GLU SER ASP THR GLY GLY ALA PHE LYS GLY PHE LYS GLY LEU VAL VAL PRO SER

GLY GLY GLY ARG PHE SER GLY PHE GLY SER GLY ALA GLY GLY LYS PRO LEU GLU GLY LEU SER ASN GLY ASN ASN ILE THR SER ALA PRO PRO PHE ALA SER ALA LYS ALA ALA ASP PRO

• Molecule 2: Nuclear pore complex protein Nup50

Chain D:

7%

25%   13%   6%   56%

GLY P-3 M1 A2 K3 R4 D12 R13 D16 Q17 E18 D19 E20 A21 E22 E23 V24 G25 T26 S31 E32 E33 V34 L35 K36 N37 K41 K44 R45 R46 ASN VAL GLY PHE GLU SER ASP THR GLY GLY ALA PHE LYS GLY PHE LYS GLY LEU VAL VAL PRO SER GLY GLY

GLY ARG PHE SER GLY PHE GLY SER GLY ALA GLY GLY LYS PRO LEU GLU GLY LEU SER ASN GLY ASN ASN ILE THR SER ALA PRO PRO PHE ALA SER ALA LYS ALA ALA ALA ASP PRO

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	81.68Å 98.29Å 135.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.03 – 2.70 15.03 – 2.70	Depositor EDS
% Data completeness (in resolution range)	89.0 (15.03-2.70) 88.1 (15.03-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.53 (at 2.69Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.215 , 0.255 0.227 , 0.264	Depositor DCC
R_{free} test set	1464 reflections (5.77%)	DCC
Wilson B-factor (Å^2)	56.1	Xtriage
Anisotropy	0.486	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 47.9	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7545	wwPDB-VP
Average B, all atoms (Å^2)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 32.72 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 8.9849e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.29	0/3367	0.43	0/4576
1	B	0.29	0/3369	0.43	0/4579
2	C	0.28	0/396	0.42	0/524
2	D	0.27	0/405	0.51	1/537 (0.2%)
All	All	0.29	0/7537	0.44	1/10216 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	21	ALA	CB-CA-C	-5.60	101.71	110.10

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	3309	0	3346	150	0
1	B	3311	0	3351	140	0
2	C	393	0	391	49	0
2	D	402	0	402	35	0
3	A	65	0	0	13	0
3	B	49	0	0	12	0
3	C	8	0	0	1	0
3	D	8	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7545	0	7490	349	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 23.

The worst 5 of 349 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:152:THR:HG21	1:A:191:TRP:CH2	1.81	1.15
2:C:45:ARG:CG	2:C:45:ARG:HH11	1.64	1.09
2:D:36:LYS:HD2	2:D:36:LYS:C	1.71	1.09
1:B:438:LYS:HA	1:B:438:LYS:HE3	1.32	1.06
1:B:299:ARG:HD3	1:B:303:GLU:OE2	1.57	1.02

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	424/447 (95%)	401 (95%)	20 (5%)	3 (1%)	25	53
1	B	424/447 (95%)	399 (94%)	20 (5%)	5 (1%)	15	37
2	C	49/114 (43%)	37 (76%)	8 (16%)	4 (8%)	1	1
2	D	48/114 (42%)	39 (81%)	8 (17%)	1 (2%)	8	21
All	All	945/1122 (84%)	876 (93%)	56 (6%)	13 (1%)	13	33

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	19	ASP
1	B	114	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	462	GLY
1	A	158	ASN
2	C	14	ASN

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	369/386 (96%)	350 (95%)	19 (5%)	28	56
1	B	370/386 (96%)	343 (93%)	27 (7%)	16	38
2	C	39/82 (48%)	29 (74%)	10 (26%)	0	2
2	D	42/82 (51%)	32 (76%)	10 (24%)	1	2
All	All	820/936 (88%)	754 (92%)	66 (8%)	14	32

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	112	LYS
1	B	187	GLU
2	D	23	GLU
1	B	115	ASN
1	B	133	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	461	ASN
1	B	456	GLN
1	B	231	ASN
1	A	290	GLN
1	B	290	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 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, 95th 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	426/447 (95%)	-0.16	13 (3%)	49	49	36, 61, 110, 159	0
1	B	426/447 (95%)	-0.10	16 (3%)	41	39	35, 61, 109, 155	0
2	C	51/114 (44%)	0.75	8 (15%)	2	2	52, 100, 148, 160	0
2	D	50/114 (43%)	0.76	8 (16%)	2	1	61, 99, 152, 159	0
All	All	953/1122 (84%)	-0.04	45 (4%)	32	30	35, 64, 121, 160	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	37	ASN	6.4
1	B	506	THR	6.3
1	A	462	GLY	5.5
1	A	113	GLU	5.3
1	A	463	THR	5.1

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.