wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 05:04 AM GMT

PDB ID	:	2P8Q
Title	:	Crystal Structure of human Importin beta bound to the Snurportin1 IBB-domain
Authors	:	Mitrousis, G.; Cingolani, G.
Deposited on	:	2007-03-22
Resolution	:	2.35 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 (1)) were used in the production of this report:

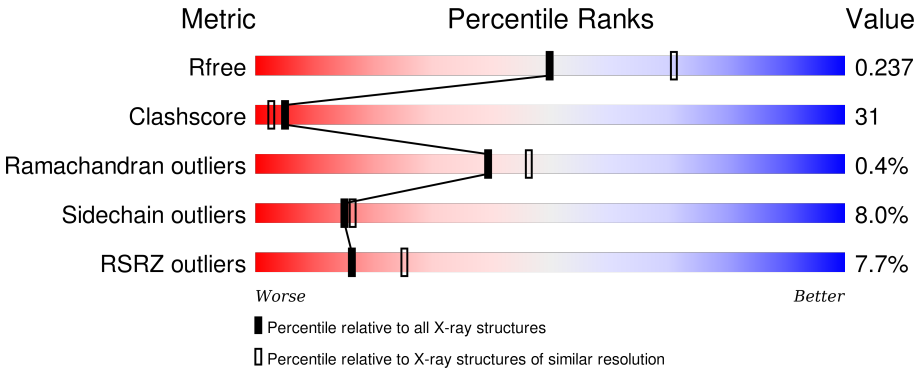
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.35 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}	91344	1352 (2.38-2.34)
Clashscore	102246	1456 (2.38-2.34)
Ramachandran outliers	100387	1435 (2.38-2.34)
Sidechain outliers	100360	1436 (2.38-2.34)
RSRZ outliers	91569	1358 (2.38-2.34)

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	876	7% 60% 33% 5%
2	B	40	18% 38% 50% 5% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7471 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 beta-1 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	873	Total	C	N	O	S	0	0	0
			6782	4273	1136	1326	47			

- Molecule 2 is a protein called Snurportin-1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	39	Total	C	N	O	0	0	0
			344	210	73	61			

- Molecule 3 is water.

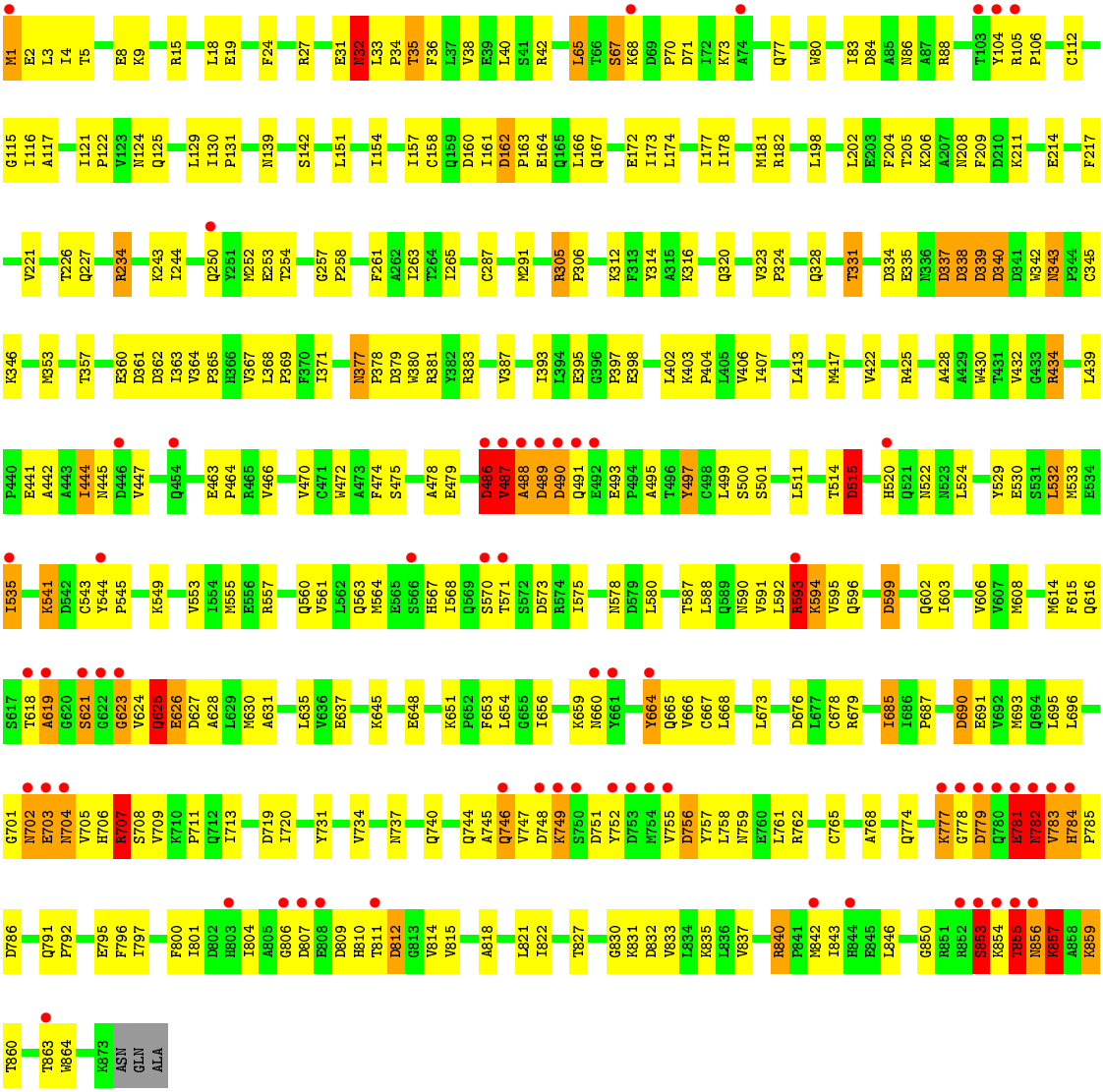
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	332	Total	O	0	0
			332	332		
3	B	13	Total	O	0	0
			13	13		

3 Residue-property plots (i)

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.

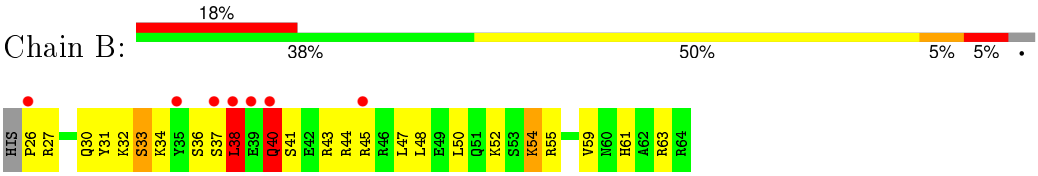
- Molecule 1: Importin beta-1 subunit

Chain A:



- Molecule 2: Snurportin-1

Chain B:


18%
38%
50%
5%
5%
HIS
P26
R27
Q30
Y31
K32
S33
K34
Y35
S36
S37
L38
E39
Q40
S41
E42
R43
R44
R45
R46
L47
L48
E49
L50
Q51
K52
S53
K54
R55
V59
W60
H61
A62
R63
R64

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	64.75Å 97.89Å 84.29Å 90.00° 90.86° 90.00°	Depositor
Resolution (Å)	40.00 – 2.35 39.04 – 2.31	Depositor EDS
% Data completeness (in resolution range)	98.7 (40.00-2.35) 94.2 (39.04-2.31)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 2.31Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.227 , 0.250 0.236 , 0.237	Depositor DCC
R_{free} test set	4282 reflections (10.03%)	DCC
Wilson B-factor (Å^2)	33.4	Xtriage
Anisotropy	0.390	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 45.8	EDS
Estimated twinning fraction	0.027 for h,-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	1 of 43297 reflections (0.002%)	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7471	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 5.22% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

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	0.32	0/6893	0.56	18/9360 (0.2%)
2	B	0.25	0/348	0.42	0/458
All	All	0.32	0/7241	0.56	18/9818 (0.2%)

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
1	A	0	24
2	B	0	2
All	All	0	26

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	515	ASP	CB-CG-OD2	5.28	123.05	118.30
1	A	338	ASP	CB-CG-OD2	5.28	123.05	118.30
1	A	361	ASP	CB-CG-OD2	5.27	123.04	118.30
1	A	334	ASP	CB-CG-OD2	5.27	123.04	118.30
1	A	690	ASP	CB-CG-OD2	5.25	123.03	118.30

There are no chirality outliers.

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	487	VAL	Peptide
1	A	488	ALA	Peptide
1	A	489	ASP	Peptide
1	A	490	ASP	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	593	ARG	Peptide

5.2 Too-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 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	6782	0	6781	417	0
2	B	344	0	359	41	0
3	A	332	0	0	56	0
3	B	13	0	0	4	0
All	All	7471	0	7140	441	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 31.

The worst 5 of 441 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:702:ASN:CB	1:A:703:GLU:HG3	1.47	1.41
1:A:702:ASN:HB3	1:A:703:GLU:CG	1.48	1.40
2:B:37:SER:C	2:B:38:LEU:HD13	1.48	1.31
1:A:316:LYS:HD2	3:A:1193:HOH:O	1.40	1.21
1:A:702:ASN:O	1:A:705:VAL:HG23	1.37	1.19

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	871/876 (99%)	809 (93%)	58 (7%)	4 (0%)	34 39
2	B	37/40 (92%)	31 (84%)	6 (16%)	0	100 100
All	All	908/916 (99%)	840 (92%)	64 (7%)	4 (0%)	39 46

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	664	TYR
1	A	32	ASN
1	A	444	ILE
1	A	687	PRO

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	749/751 (100%)	691 (92%)	58 (8%)	16 17
2	B	38/39 (97%)	33 (87%)	5 (13%)	5 4
All	All	787/790 (100%)	724 (92%)	63 (8%)	15 16

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	541	LYS
1	A	599	ASP
2	B	33	SER
1	A	561	VAL
1	A	580	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	343	ASN
1	A	408	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	780	GLN
1	A	377	ASN
1	A	401	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, 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	873/876 (99%)	0.46	63 (7%) 18 28	16, 42, 88, 160	0
2	B	39/40 (97%)	0.97	7 (17%) 2 3	34, 62, 104, 128	0
All	All	912/916 (99%)	0.48	70 (7%) 16 25	16, 42, 91, 160	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	780	GLN	21.3
1	A	783	VAL	12.3
1	A	488	ALA	10.2
1	A	487	VAL	9.7
1	A	490	ASP	9.3

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

There are no such residues in this entry.