W O R L D W I D E wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2023 – 04:22 AM EDT

PDB ID	:	2Q5D
Title	:	Crystal Structure of Human Importin Beta bound to the Snurportin1 IBB-domain second crystal form
Authors	:	Mitrousis, G.; Cingolani, G.
Deposited on	:	2007-05-31
Resolution	:	3.20 Å(reported)

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

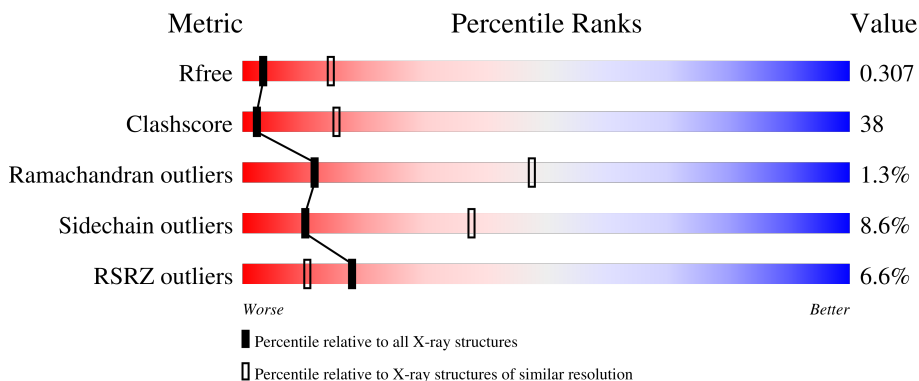|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 3.20 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	Percentile Ranks	Value
Rfree		0.307
Clashscore		38
Ramachandran outliers		1.3%
Sidechain outliers		8.6%
RSRZ outliers		6.6%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	3% 60% 33% 5% ··
1	B	876	10% 51% 33% 11% · ·
2	C	40	50% 30% 10% 10%
2	D	40	10% 48% 10% 42%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 13750 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	871	Total	C	N	O	S	1	0	0
			6762	4259	1135	1321	47			
1	B	845	Total	C	N	O	S	0	0	0
			6550	4129	1100	1275	46			

- Molecule 2 is a protein called Snurportin-1.

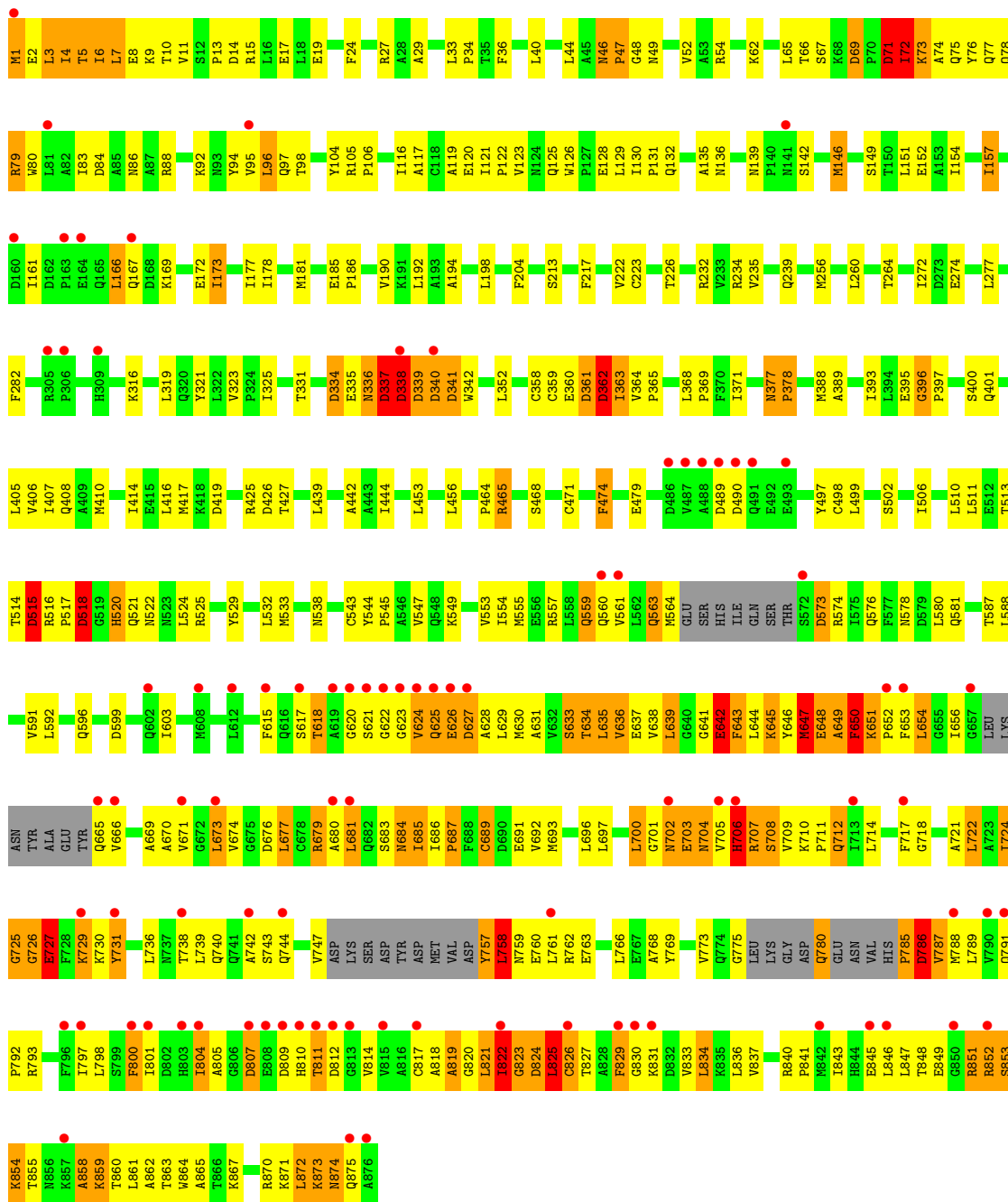| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	36	Total	C	N	O	0	0	0
			323	198	70	55			
2	D	23	Total	C	N	O	0	0	0
			115	69	23	23			

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.
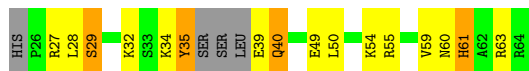
• Molecule 1: Importin beta-1 subunit
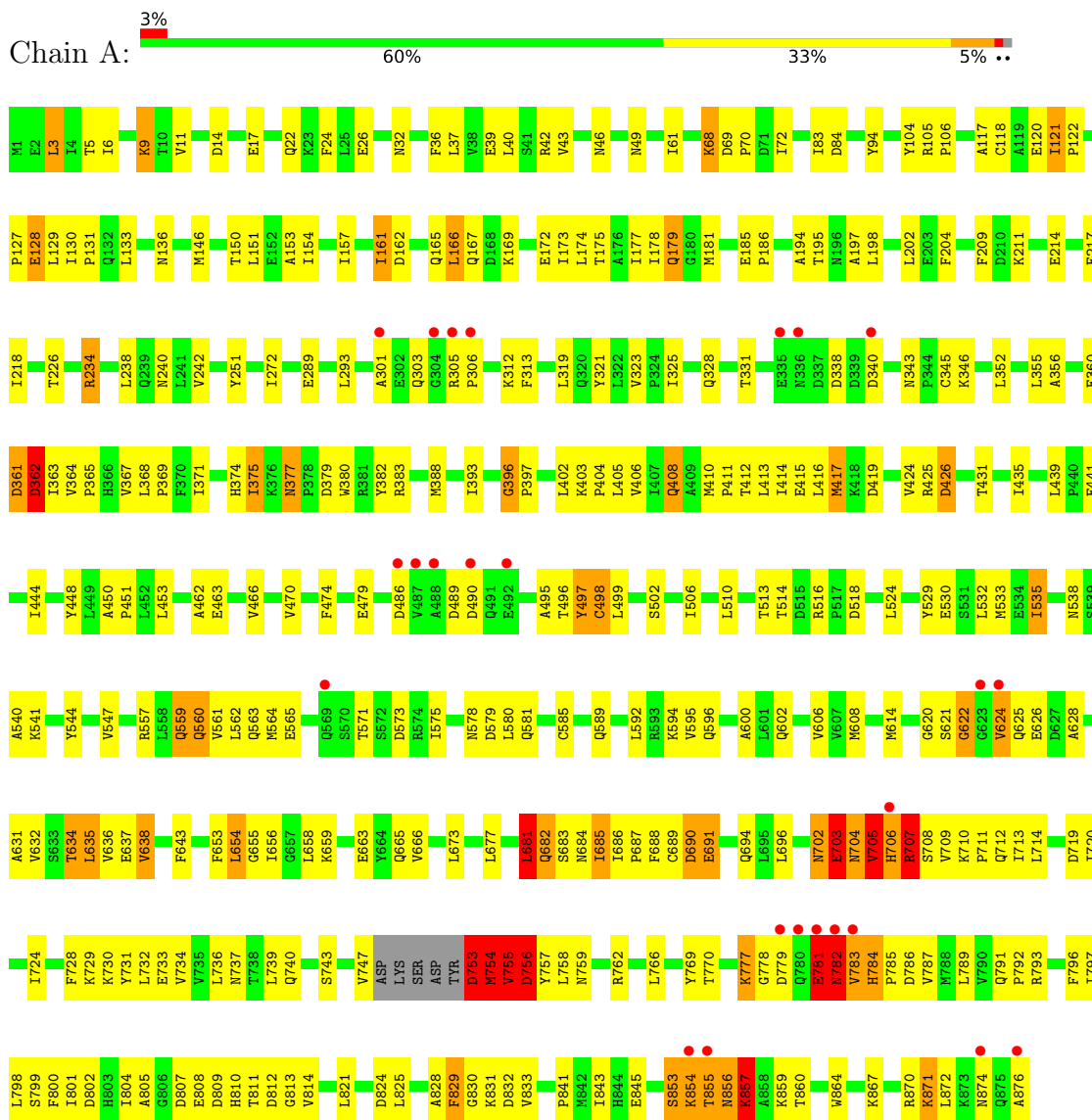
• Molecule 1: Importin beta-1 subunit

• Molecule 2: Snurportin-1

Chain C:

• Molecule 2: Snurportin-1

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	102.98Å 100.70Å 108.68Å 90.00° 109.23° 90.00°	Depositor
Resolution (Å)	40.00 – 3.20 39.61 – 3.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.20) 91.3 (39.61-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.98 (at 3.18Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.303 , 0.328 0.296 , 0.307	Depositor DCC
R_{free} test set	1598 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	70.4	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 73.8	EDS
L-test for twinning [2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	13750	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.47	0/6871	0.68	11/9328 (0.1%)
1	B	0.42	0/6651	0.67	19/9023 (0.2%)
2	C	0.50	0/326	0.67	0/428
2	D	0.47	0/114	0.37	0/158
All	All	0.44	0/13962	0.67	30/18937 (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	25
1	B	0	48
All	All	0	73

There are no bond length outliers.

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	396	GLY	C-N-CD	-5.68	108.11	120.60
1	A	779	ASP	CB-CG-OD2	5.28	123.05	118.30
1	A	690	ASP	CB-CG-OD2	5.28	123.05	118.30
1	A	832	ASP	CB-CG-OD2	5.28	123.05	118.30
1	B	337	ASP	CB-CG-OD2	5.27	123.05	118.30

There are no chirality outliers.

5 of 73 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	251	TYR	Peptide
1	A	396	GLY	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	498	CYS	Peptide
1	A	622	GLY	Peptide
1	A	681	LEU	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	6762	0	6764	426	0
1	B	6550	0	6568	655	0
2	C	323	0	337	15	0
2	D	115	0	48	6	0
All	All	13750	0	13717	1047	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 38.

The worst 5 of 1047 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:B:686:ILE:CG1	1:B:687:PRO:HD3	1.08	1.56
1:B:686:ILE:CG1	1:B:687:PRO:CD	2.02	1.37
1:B:647:MET:CA	1:B:649:ALA:HB2	1.52	1.35
1:B:647:MET:SD	1:B:649:ALA:HB3	1.65	1.35
1:A:130:ILE:CG2	1:A:131:PRO:HD3	1.61	1.30

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	867/876 (99%)	761 (88%)	97 (11%)	9 (1%)	15	54
1	B	834/876 (95%)	712 (85%)	110 (13%)	12 (1%)	11	46
2	C	32/40 (80%)	27 (84%)	4 (12%)	1 (3%)	4	26
2	D	21/40 (52%)	16 (76%)	5 (24%)	0	100	100
All	All	1754/1832 (96%)	1516 (86%)	216 (12%)	22 (1%)	12	47

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	441	GLU
1	A	560	GLN
1	A	732	LEU
1	B	47	PRO
1	B	72	ILE

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	746/751 (99%)	695 (93%)	51 (7%)	16	49
1	B	722/751 (96%)	648 (90%)	74 (10%)	7	29
2	C	35/39 (90%)	31 (89%)	4 (11%)	5	24
All	All	1503/1541 (98%)	1374 (91%)	129 (9%)	10	38

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	800	PHE
1	B	825	LEU
1	A	855	THR
1	A	782	ASN
1	B	834	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	141	ASN
1	B	408	GLN
1	B	744	GLN
1	B	179	GLN
1	B	320	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 monosaccharides 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	871/876 (99%)	-0.21	25 (2%)	51	36	11, 51, 141, 200	1 (0%)
1	B	845/876 (96%)	0.58	88 (10%)	6	4	14, 102, 189, 199	0
2	C	36/40 (90%)	-0.16	0	100	100	31, 63, 103, 133	0
2	D	23/40 (57%)	0.68	4 (17%)	1	1	58, 70, 82, 90	0
All	All	1775/1832 (96%)	0.18	117 (6%)	18	11	11, 67, 178, 200	1 (0%)

The worst 5 of 117 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	812	ASP	13.8
1	B	830	GLY	13.5
1	B	623	GLY	10.9
1	B	811	THR	10.4
1	B	800	PHE	10.1

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 monosaccharides 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.