



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:51 AM GMT

PDB ID : 3NBY  
Title : Crystal structure of the PKI NES-CRM1-RanGTP nuclear export complex  
Authors : Guttler, T.; Madl, T.; Neumann, P.; Deichsel, D.; Corsini, L.; Monecke, T.;  
Ficner, R.; Sattler, M.; Gorlich, D.  
Deposited on : 2010-06-04  
Resolution : 3.42 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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

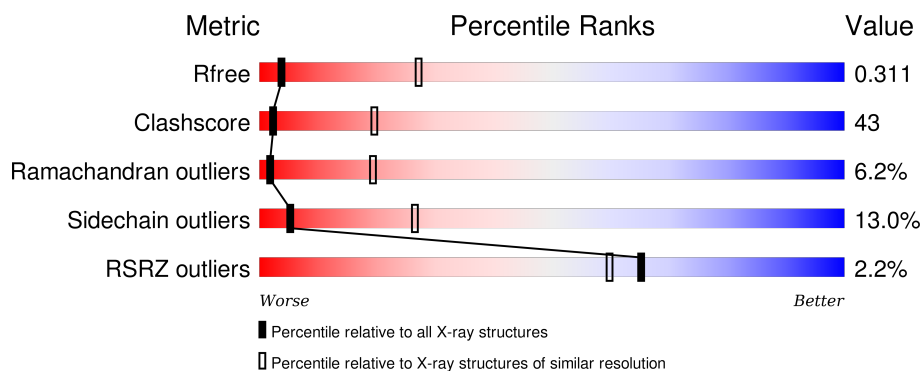
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.42 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	1049 (3.52-3.32)
Clashscore	102246	1032 (3.50-3.34)
Ramachandran outliers	100387	1002 (3.50-3.34)
Sidechain outliers	100360	1003 (3.50-3.34)
RSRZ outliers	91569	1054 (3.52-3.32)

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. 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  $\leq 5\%$ . 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	B	361	<div> <div>2%</div> <div>22% 40% 12% 25%</div> </div>
1	E	361	<div> <div>2%</div> <div>30% 36% 10% 23%</div> </div>
2	C	176	<div> <div>35% 52% 9%</div> </div>
2	F	176	<div> <div>39% 46% 11%</div> </div>
3	A	1073	<div> <div>2%</div> <div>33% 52% 11%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	1073	<div><div></div><div>3%</div><div>32%</div><div>53%</div><div>11%</div><div></div></div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 24084 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 Snurportin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	272	Total	C	N	O	S	0	0	0
			2193	1398	376	405	14			
1	E	277	Total	C	N	O	S	0	0	0
			2221	1416	380	410	15			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP O95149
B	0	SER	-	expression tag	UNP O95149
B	1	LEU	-	expression tag	UNP O95149
B	2	ASN	-	expression tag	UNP O95149
B	3	GLU	-	expression tag	UNP O95149
B	4	LEU	-	expression tag	UNP O95149
B	5	ALA	-	expression tag	UNP O95149
B	6	LEU	-	expression tag	UNP O95149
B	7	LYS	-	expression tag	UNP O95149
B	8	LEU	-	expression tag	UNP O95149
B	9	ALA	-	expression tag	UNP O95149
B	10	GLY	-	expression tag	UNP O95149
B	11	LEU	-	expression tag	UNP O95149
B	12	ASP	-	expression tag	UNP O95149
B	13	ILE	-	expression tag	UNP O95149
E	-1	GLY	-	expression tag	UNP O95149
E	0	SER	-	expression tag	UNP O95149
E	1	LEU	-	expression tag	UNP O95149
E	2	ASN	-	expression tag	UNP O95149
E	3	GLU	-	expression tag	UNP O95149
E	4	LEU	-	expression tag	UNP O95149
E	5	ALA	-	expression tag	UNP O95149
E	6	LEU	-	expression tag	UNP O95149
E	7	LYS	-	expression tag	UNP O95149
E	8	LEU	-	expression tag	UNP O95149

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	9	ALA	-	expression tag	UNP O95149
E	10	GLY	-	expression tag	UNP O95149
E	11	LEU	-	expression tag	UNP O95149
E	12	ASP	-	expression tag	UNP O95149
E	13	ILE	-	expression tag	UNP O95149

- Molecule 2 is a protein called GTP-binding nuclear protein Ran.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	171	Total	C	N	O	S	0	0	0
			1389	904	243	237	5			
2	F	171	Total	C	N	O	S	0	0	0
			1389	904	243	237	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	69	LEU	GLN	engineered	UNP P62826
F	69	LEU	GLN	engineered	UNP P62826

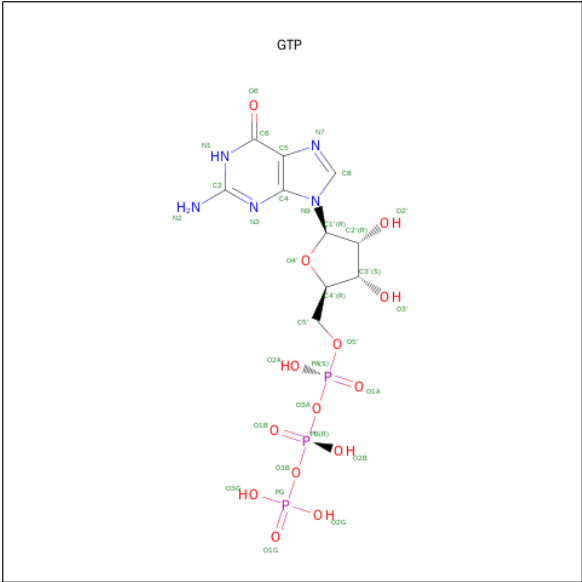
- Molecule 3 is a protein called Exportin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	1041	Total	C	N	O	S	0	0	0
			8413	5397	1413	1549	54			
3	D	1041	Total	C	N	O	S	0	0	0
			8413	5397	1413	1549	54			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q6P5F9
A	0	SER	-	expression tag	UNP Q6P5F9
D	-1	GLY	-	expression tag	UNP Q6P5F9
D	0	SER	-	expression tag	UNP Q6P5F9

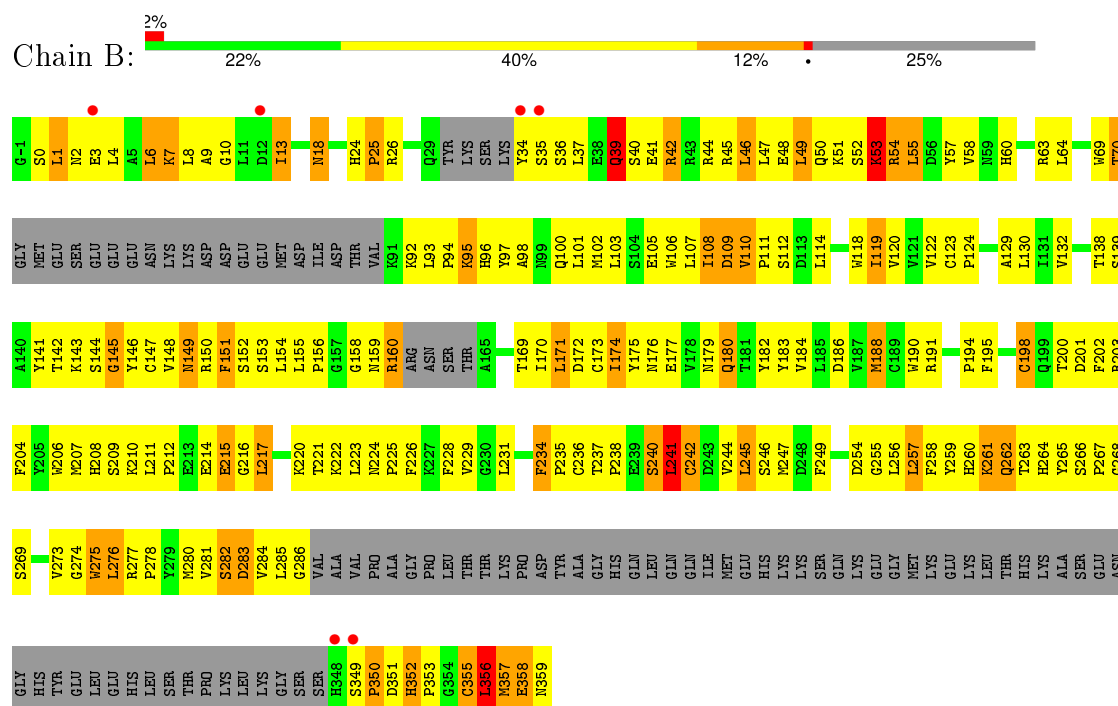
- Molecule 4 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>14</sub>P<sub>3</sub>).



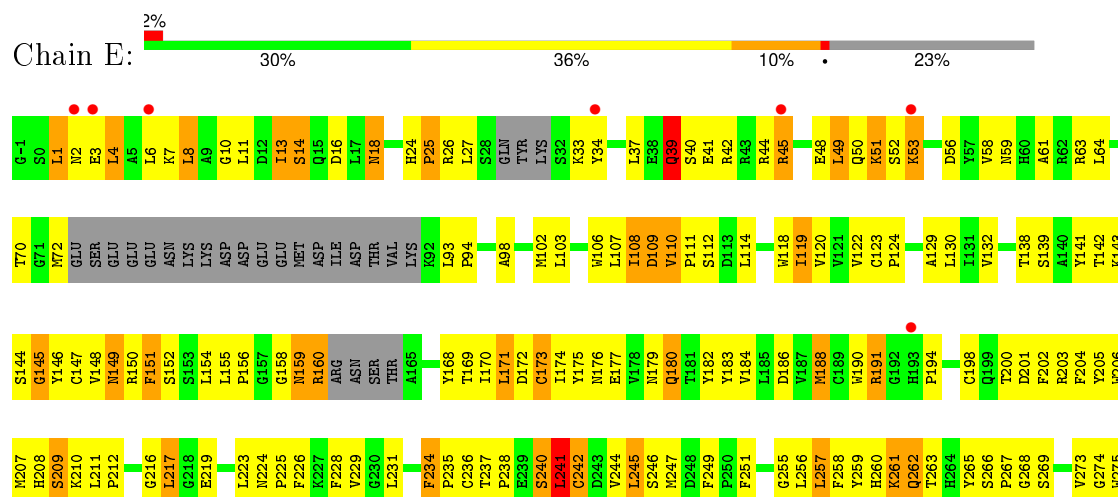
### 3 Residue-property plots

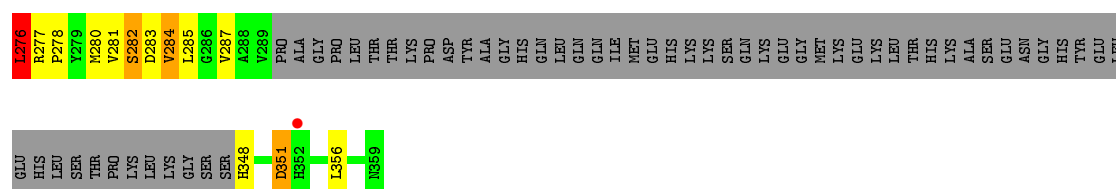
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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 ( $\text{RSRZ} > 2$ ). 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: Snurportin-1



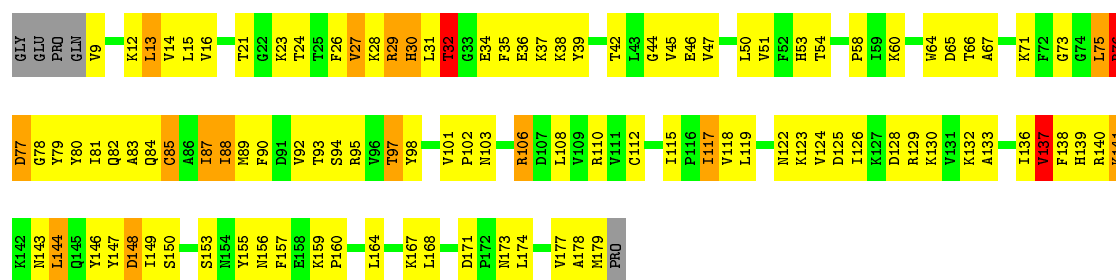
#### • Molecule 1: Snurportin-1





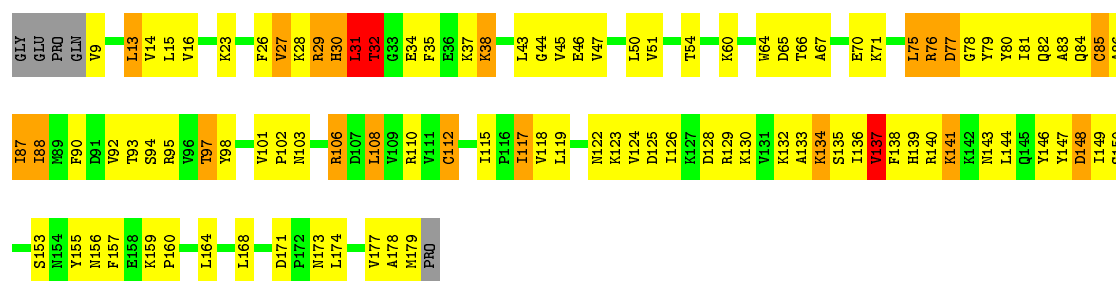
• Molecule 2: GTP-binding nuclear protein Ran

Chain C: 35% 52% 9% . .



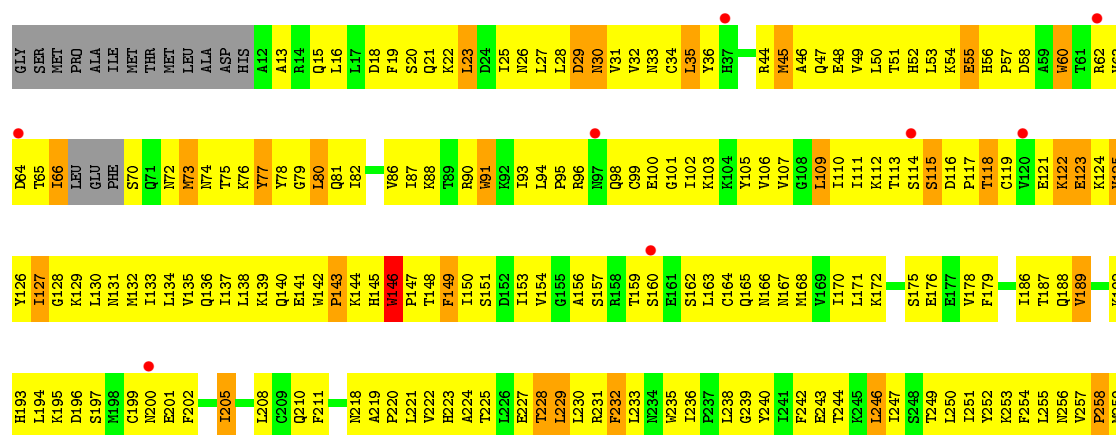
• Molecule 2: GTP-binding nuclear protein Ran

Chain F: 39% 46% 11% . .

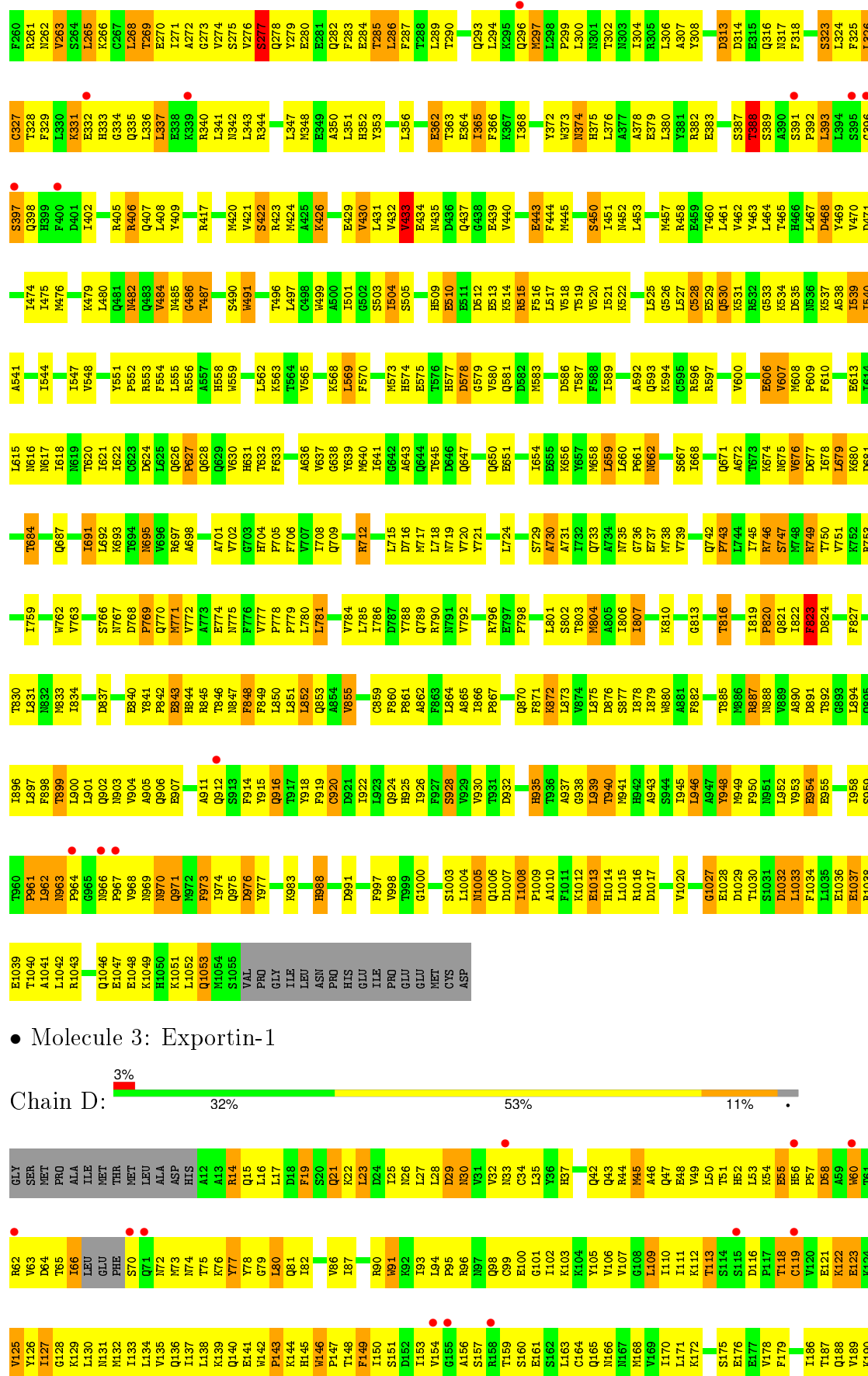


• Molecule 3: Exportin-1

Chain A: 2% 33% 52% 11% .









## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.09 Å   223.73 Å   163.06 Å 90.00°   100.63°   90.00°	Depositor
Resolution (Å)	38.63 – 3.42 38.63 – 2.95	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.63-3.42) 88.9 (38.63-2.95)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.80 (at 2.95 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, $R_{free}$	0.258 , 0.315 0.252 , 0.311	Depositor DCC
$R_{free}$ test set	3408 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.9	Xtriage
Anisotropy	0.412	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 6.6	EDS
Estimated twinning fraction	0.115 for h,-k,-h-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 104746 reflections	Xtriage
$F_o, F_c$ correlation	0.82	EDS
Total number of atoms	24084	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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

Bond lengths and bond angles in the following residue types are not validated in this section: GTP, MG

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	B	0.62	0/2250	0.85	3/3046 (0.1%)
1	E	0.59	1/2278 (0.0%)	0.81	4/3084 (0.1%)
2	C	0.58	0/1423	0.74	0/1921
2	F	0.57	0/1423	0.74	1/1921 (0.1%)
3	A	0.56	2/8584 (0.0%)	0.73	1/11627 (0.0%)
3	D	0.57	0/8584	0.74	0/11627
All	All	0.58	3/24542 (0.0%)	0.76	9/33226 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	173	CYS	CB-SG	-5.71	1.72	1.81
3	A	491	TRP	NE1-CE2	5.53	1.44	1.37
3	A	528	CYS	CB-SG	-5.08	1.73	1.81

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	42	ARG	NE-CZ-NH1	-11.62	114.49	120.30
1	B	42	ARG	NE-CZ-NH2	-11.22	114.69	120.30
1	B	42	ARG	NE-CZ-NH1	10.86	125.73	120.30
1	E	42	ARG	NE-CZ-NH2	10.83	125.72	120.30
2	F	31	LEU	CA-CB-CG	5.93	128.93	115.30

There are no chirality outliers.

There are no planarity outliers.

## 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	B	2193	0	2138	241	0
1	E	2221	0	2170	208	0
2	C	1389	0	1419	128	0
2	F	1389	0	1419	121	0
3	A	8413	0	8480	753	0
3	D	8413	0	8480	731	0
4	C	32	0	12	8	0
4	F	32	0	12	0	0
5	C	1	0	0	0	0
5	F	1	0	0	0	0
All	All	24084	0	24130	2091	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 43.

The worst 5 of 2091 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)
3:A:328:THR:HA	3:A:331:LYS:HD3	1.27	1.14
3:D:328:THR:HA	3:D:331:LYS:HD3	1.33	1.09
3:D:961:PRO:HG2	3:D:973:PHE:HD2	1.16	1.09
1:B:3:GLU:O	1:B:7:LYS:HB2	1.53	1.07
3:D:996:LEU:HD13	3:D:1035:LEU:HD12	1.37	1.06

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	B	262/361 (73%)	206 (79%)	38 (14%)	18 (7%)	1	17
1	E	267/361 (74%)	211 (79%)	39 (15%)	17 (6%)	2	18
2	C	169/176 (96%)	139 (82%)	23 (14%)	7 (4%)	3	32
2	F	169/176 (96%)	138 (82%)	24 (14%)	7 (4%)	3	32
3	A	1037/1073 (97%)	760 (73%)	206 (20%)	71 (7%)	1	17
3	D	1037/1073 (97%)	755 (73%)	220 (21%)	62 (6%)	2	20
All	All	2941/3220 (91%)	2209 (75%)	550 (19%)	182 (6%)	2	19

5 of 182 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	39	GLN
1	B	145	GLY
1	B	180	GLN
1	B	217	LEU
2	C	76	ARG

### 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	B	244/323 (76%)	199 (82%)	45 (18%)	2	10
1	E	247/323 (76%)	210 (85%)	37 (15%)	3	20
2	C	150/154 (97%)	129 (86%)	21 (14%)	4	23
2	F	150/154 (97%)	125 (83%)	25 (17%)	3	15
3	A	945/973 (97%)	838 (89%)	107 (11%)	7	33
3	D	945/973 (97%)	831 (88%)	114 (12%)	6	29
All	All	2681/2900 (92%)	2332 (87%)	349 (13%)	5	26

5 of 349 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	935	HIS
1	E	198	CYS
3	D	781	LEU
3	A	969	ASN
1	E	14	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 101 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	924	GLN
1	E	199	GLN
3	D	906	GLN
3	A	970	ASN
3	A	1046	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules 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](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GTP	C	217	5	25,34,34	0.98	1 (4%)	34,54,54	2.28	10 (29%)
4	GTP	F	217	5	25,34,34	0.96	1 (4%)	34,54,54	2.09	9 (26%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GTP	C	217	5	-	0/18/38/38	0/3/3/3
4	GTP	F	217	5	-	0/18/38/38	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	217	GTP	C6-N1	2.77	1.38	1.33
4	F	217	GTP	C6-N1	2.82	1.38	1.33

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	217	GTP	PA-O3A-PB	-5.74	116.60	132.73
4	C	217	GTP	PA-O3A-PB	-5.27	117.93	132.73
4	C	217	GTP	C4-C5-N7	-5.05	104.83	109.48
4	F	217	GTP	N3-C2-N1	-4.75	120.22	127.44
4	C	217	GTP	N3-C2-N1	-4.43	120.69	127.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	217	GTP	8	0

## 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	B	272/361 (75%)	0.01	6 (2%) 65 60	7, 30, 72, 91	4 (1%)
1	E	277/361 (76%)	0.05	8 (2%) 55 49	7, 30, 80, 111	7 (2%)
2	C	171/176 (97%)	-0.12	0 100 100	10, 28, 58, 84	1 (0%)
2	F	171/176 (97%)	-0.23	0 100 100	10, 27, 56, 80	0
3	A	1041/1073 (97%)	0.01	20 (1%) 70 64	4, 35, 90, 141	6 (0%)
3	D	1041/1073 (97%)	0.03	30 (2%) 55 49	5, 35, 89, 147	13 (1%)
All	All	2973/3220 (92%)	-0.00	64 (2%) 65 60	4, 32, 85, 147	31 (1%)

The worst 5 of 64 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	70	SER	8.8
3	A	967	PRO	8.7
3	A	966	ASN	7.7
3	D	966	ASN	5.3
3	D	397	SER	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 [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	GTP	C	217	32/32	0.94	0.19	-0.50	10,21,47,63	0
5	MG	C	218	1/1	0.96	0.19	-0.57	0,0,0,0	0
4	GTP	F	217	32/32	0.96	0.17	-0.71	7,18,30,35	0
5	MG	F	218	1/1	0.96	0.13	-2.05	6,6,6,6	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.