



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 12:20 am GMT

PDB ID : 3FEY
Title : Crystal structure of the CBC-importin alpha complex.
Authors : Dias, S.M.G.; Ambrosio, A.L.B.; Cerione, R.A.
Deposited on : 2008-12-01
Resolution : 2.20 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 ⓘ symbol.

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

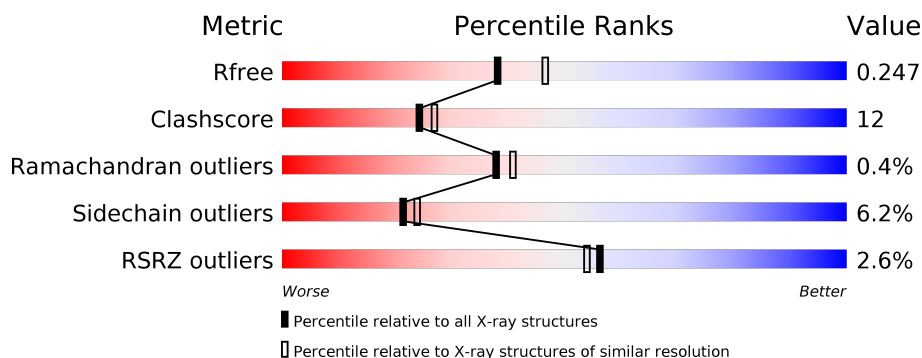
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	4002 (2.20-2.20)
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)
RSRZ outliers	101464	4033 (2.20-2.20)

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 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 $\leq 5\%$. 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	A	790	<div> <div>2%</div> <div> <div></div> <div>74%</div> <div>19%</div> <div>• •</div> </div> </div>
2	B	156	<div> <div>5%</div> <div> <div></div> <div>47%</div> <div>12%</div> <div>•</div> <div>38%</div> </div> </div>
3	C	467	<div> <div>2%</div> <div> <div></div> <div>68%</div> <div>22%</div> <div>•</div> <div>8%</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10746 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 Nuclear cap-binding protein subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	761	Total	C	N	O	S	0	3	0
			6242	4007	1067	1130	38			

- Molecule 2 is a protein called Nuclear cap-binding protein subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	97	Total	C	N	O	S	0	0	0
			792	501	129	157	5			

- Molecule 3 is a protein called Importin subunit alpha-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	428	Total	C	N	O	S	0	1	0
			3258	2074	547	626	11			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	69	MET	-	EXPRESSION TAG	UNP P52292
C	530	HIS	-	EXPRESSION TAG	UNP P52292
C	531	HIS	-	EXPRESSION TAG	UNP P52292
C	532	HIS	-	EXPRESSION TAG	UNP P52292
C	533	HIS	-	EXPRESSION TAG	UNP P52292
C	534	HIS	-	EXPRESSION TAG	UNP P52292
C	535	HIS	-	EXPRESSION TAG	UNP P52292

- Molecule 4 is water.

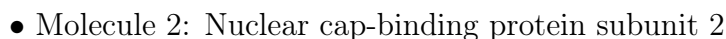
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	282	Total	O	0	0
			282	282		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	43	Total 43	O 43	0	0
4	C	129	Total 129	O 129	0	0

- Molecule 1: Nuclear cap-binding protein subunit 1





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	83.84Å 104.50Å 108.56Å 90.00° 109.04° 90.00°	Depositor
Resolution (Å)	15.00 – 2.20 14.98 – 2.20	Depositor EDS
% Data completeness (in resolution range)	97.2 (15.00-2.20) 97.2 (14.98-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.192 , 0.238 0.203 , 0.247	Depositor DCC
R_{free} test set	4376 reflections (5.30%)	DCC
Wilson B-factor (Å ²)	40.5	Xtriage
Anisotropy	0.083	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 51.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	10746	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.52	2/6407 (0.0%)	0.61	0/8685
2	B	0.60	1/804 (0.1%)	0.66	0/1075
3	C	0.53	0/3317	0.61	0/4525
All	All	0.53	3/10528 (0.0%)	0.62	0/14285

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	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	483	CYS	CB-SG	-6.32	1.71	1.82
1	A	616	CYS	CB-SG	5.18	1.91	1.82
2	B	81	CYS	CB-SG	-5.07	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	8	ASP	Peptide

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	6242	0	6221	136	0
2	B	792	0	756	20	0
3	C	3258	0	3323	108	0
4	A	282	0	0	12	0
4	B	43	0	0	4	0
4	C	129	0	0	13	0
All	All	10746	0	10300	255	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 12.

The worst 5 of 255 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:C:246:ILE:H	3:C:246:ILE:CD1	1.55	1.15
1:A:6[A]:HIS:NE2	3:C:354:GLU:OE2	1.83	1.11
3:C:219:SER:HB2	3:C:220:SER:CA	1.80	1.10
3:C:246:ILE:HD13	3:C:246:ILE:N	1.67	1.07
3:C:219:SER:CB	3:C:220:SER:HA	1.73	1.05

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	758/790 (96%)	740 (98%)	14 (2%)	4 (0%)	32	34
2	B	91/156 (58%)	91 (100%)	0	0	100	100
3	C	425/467 (91%)	416 (98%)	8 (2%)	1 (0%)	51	58
All	All	1274/1413 (90%)	1247 (98%)	22 (2%)	5 (0%)	38	41

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	491	SER
3	C	130	ARG
1	A	488	GLY
1	A	9	GLU
1	A	511	LYS

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	699/724 (96%)	652 (93%)	47 (7%)	19	21
2	B	84/130 (65%)	77 (92%)	7 (8%)	13	13
3	C	362/396 (91%)	344 (95%)	18 (5%)	28	34
All	All	1145/1250 (92%)	1073 (94%)	72 (6%)	21	23

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	584	LEU
1	A	696	GLN
3	C	339	VAL
1	A	597	HIS
1	A	629	SER

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	207	GLN
3	C	261	HIS
1	A	367	HIS
1	A	138	HIS
3	C	137	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](#)

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 [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	761/790 (96%)	-0.21	15 (1%) 65 63	9, 23, 30, 44	41 (5%)
2	B	97/156 (62%)	0.28	8 (8%) 12 11	11, 17, 93, 111	5 (5%)
3	C	428/467 (91%)	-0.21	10 (2%) 61 58	10, 20, 35, 42	25 (5%)
All	All	1286/1413 (91%)	-0.17	33 (2%) 56 54	9, 22, 33, 111	71 (5%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	15	VAL	7.7
1	A	491	SER	6.5
2	B	14	TYR	6.3
2	B	39	SER	5.8
1	A	8	ASP	5.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.