wwPDB X-ray Structure Validation Summary Report (i)

Dec 25, 2021 – 06:04 am GMT

PDB ID	:	6YO6
Title	:	Structure of iC3b1
Authors	:	Jensen, R.K.; Andersen, G.R.
Deposited on	:	2020-04-14
Resolution	:	6.00 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
https://www.wwpdb.org/validation/2017/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:

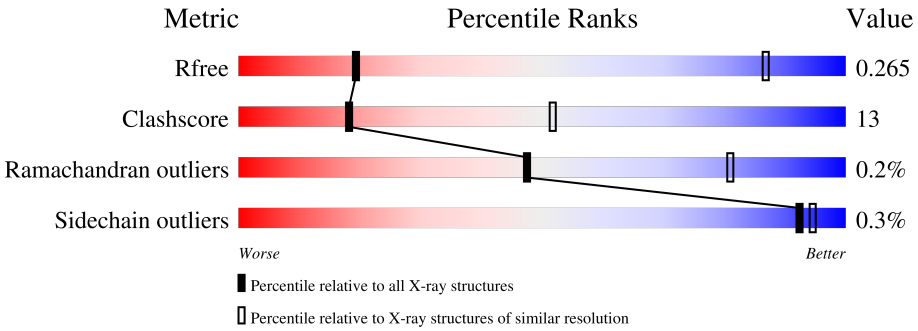
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.24
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.24

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 6.00 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}	130704	1000 (8.00-3.88)
Clashscore	141614	1049 (8.00-3.90)
Ramachandran outliers	138981	1016 (8.00-3.86)
Sidechain outliers	138945	1017 (8.00-3.82)

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%

Mol	Chain	Length	Quality of chain
1	A	645	78% 20% ··
2	B	915	83% 16% ·
3	C	130	85% 10% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 13131 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 iC3b1 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	638	4970	3165	843	947	15	0	0	0

- Molecule 2 is a protein called iC3b1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	903	7211	4570	1213	1390	38	4	0	0

- Molecule 3 is a protein called hC3Nb1.

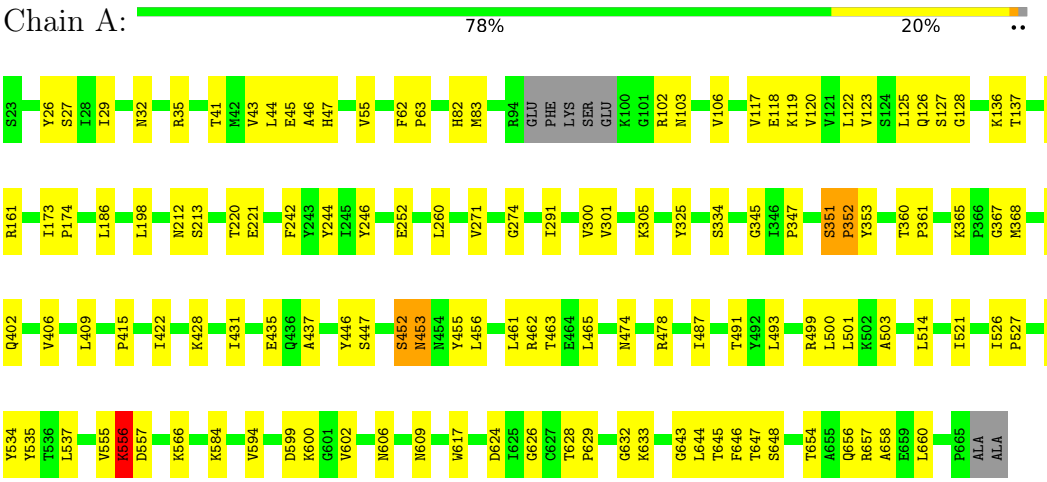
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	124	950	591	170	184	5	1	0	0

3 Residue-property plots (i)

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. 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. 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.

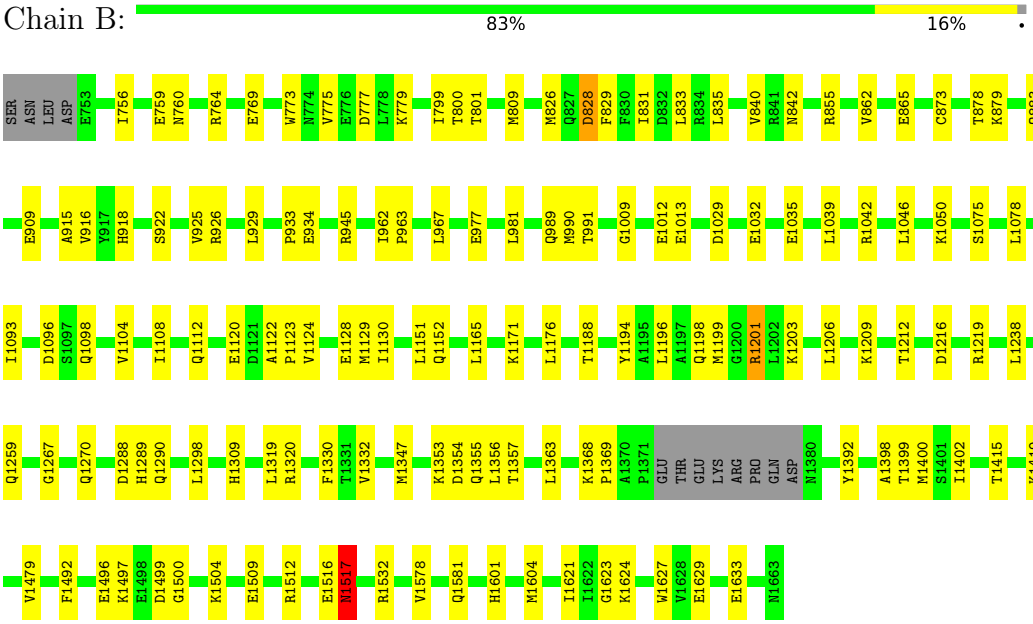
- Molecule 1: iC3b1 alpha chain

Chain A:



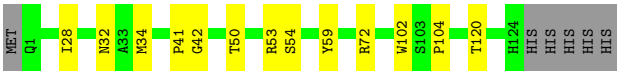
- Molecule 2: iC3b1 beta chain

Chain B:



- Molecule 3: hC3Nb1

Chain C:


MET
Q1
I28
N32
A33
M34
P41
G42
T50
R53
S54
Y59
R72
W102
S103
P104
T120
H124
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	87.45Å 100.60Å 136.28Å 94.23° 108.12° 113.81°	Depositor
Resolution (Å)	47.59 – 6.00 47.59 – 4.87	Depositor EDS
% Data completeness (in resolution range)	98.3 (47.59-6.00) 68.9 (47.59-4.87)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.02 (at 4.85Å)	Xtriage
Refinement program	PHENIX dev_2614	Depositor
R, R_{free}	0.241 , 0.266 0.240 , 0.265	Depositor DCC
R_{free} test set	1707 reflections (9.96%)	wwPDB-VP
Wilson B-factor (Å^2)	139.9	Xtriage
Anisotropy	0.733	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.033 for h,-h-k,-h-l	Xtriage
F_o,F_c correlation	0.84	EDS
Total number of atoms	13131	wwPDB-VP
Average B, all atoms (Å^2)	315.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% 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	0.67	3/5068 (0.1%)	0.69	6/6884 (0.1%)
2	B	0.48	2/7348 (0.0%)	0.54	6/9936 (0.1%)
3	C	0.25	0/971	0.46	0/1313
All	All	0.55	5/13387 (0.0%)	0.60	12/18133 (0.1%)

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
2	B	0	2
All	All	0	3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	351	SER	C-N	29.52	1.90	1.34
2	B	1517	ASN	C-N	-28.68	0.68	1.34
1	A	556	LYS	C-N	-27.98	0.69	1.34
2	B	828	ASP	C-N	-20.10	0.87	1.34
1	A	452	SER	C-N	16.07	1.71	1.34

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	452	SER	O-C-N	-25.48	81.94	122.70
2	B	1517	ASN	CA-C-N	-18.30	76.94	117.20
2	B	1517	ASN	O-C-N	-17.29	95.04	122.70
1	A	452	SER	CA-C-N	16.68	153.90	117.20
1	A	452	SER	C-N-CA	16.03	161.78	121.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	556	LYS	Mainchain
2	B	1517	ASN	Mainchain
2	B	828	ASP	Mainchain

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	4970	0	5030	192	7
2	B	7211	0	7129	204	17
3	C	950	0	904	12	10
All	All	13131	0	13063	350	17

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 13.

The worst 5 of 350 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:174:PRO:CD	2:B:1319:LEU:HD21	1.22	1.59
1:A:174:PRO:CG	2:B:1319:LEU:HD21	1.21	1.57
2:B:989:GLN:HA	2:B:1290:GLN:CG	1.41	1.47
2:B:989:GLN:CA	2:B:1290:GLN:HG2	1.44	1.45
1:A:452:SER:C	1:A:453:ASN:N	1.71	1.43

The worst 5 of 17 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:305:LYS:CE	2:B:1203:LYS:NZ[1_565]	1.30	0.90
2:B:1581:GLN:NE2	3:C:41:PRO:C[1_655]	1.34	0.86
1:A:305:LYS:CE	2:B:1203:LYS:CE[1_565]	1.35	0.85
1:A:246:TYR:CE1	2:B:1171:LYS:CE[1_565]	1.59	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:305:LYS:NZ	2:B:1203:LYS:CE[1_565]	1.62	0.58

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	630/645 (98%)	606 (96%)	21 (3%)	3 (0%)	29 69
2	B	886/915 (97%)	855 (96%)	30 (3%)	1 (0%)	51 85
3	C	122/130 (94%)	120 (98%)	2 (2%)	0	100 100
All	All	1638/1690 (97%)	1581 (96%)	53 (3%)	4 (0%)	47 81

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	453	ASN
2	B	1517	ASN
1	A	352	PRO
1	A	556	LYS

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	562/567 (99%)	561 (100%)	1 (0%)	93 96
2	B	799/811 (98%)	795 (100%)	4 (0%)	88 93

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	C	98/104 (94%)	98 (100%)	0	100 100
All	All	1459/1482 (98%)	1454 (100%)	5 (0%)	92 94

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	628	THR
2	B	967	LEU
2	B	1188	THR
2	B	1201	ARG
2	B	1499	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	32	ASN
1	A	154	HIS
2	B	760	ASN
2	B	842	ASN

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 monosaccharides 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	9
1	A	5

The worst 5 of 14 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	988:ALA	C	989:GLN	N	6.14
1	B	1287:PRO	C	1288:ASP	N	4.98
1	B	767:PHE	C	768:PRO	N	3.72
1	B	766:GLU	C	767:PHE	N	3.69
1	B	1354:ASP	C	1355:GLN	N	2.95

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.