wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2023 – 12:21 PM EDT

PDB ID	:	2I3S
Title	:	Bub3 complex with Bub1 GLEBS motif
Authors	:	Larsen, N.A.; Harrison, S.C.
Deposited on	:	2006-08-20
Resolution	:	1.90 Å(reported)

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

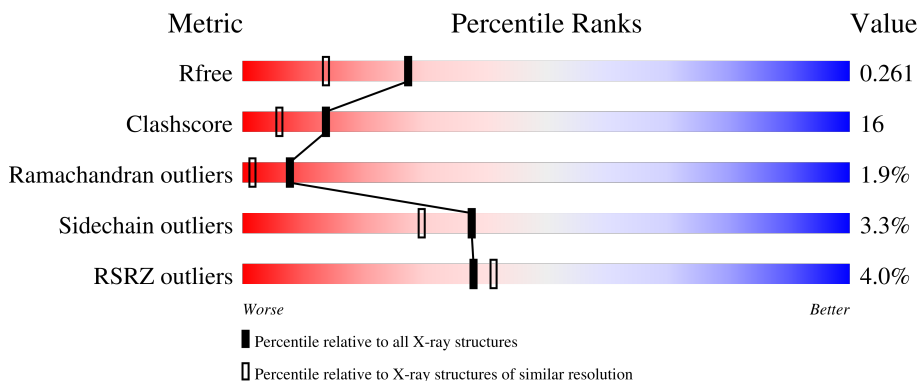|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 1.90 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.
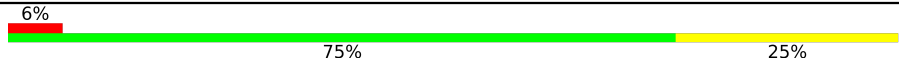
Metric	Percentile Ranks	Value
Rfree		0.261
Clashscore		16
Ramachandran outliers		1.9%
Sidechain outliers		3.3%
RSRZ outliers		4.0%

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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	349	2% — 76% / 19% · ·
1	C	349	3% — 66% / 26% · 6%
1	E	349	5% — 63% / 29% · 5%
2	B	36	11% — 72% / 25% ·
2	D	36	8% — 56% / 36% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	36	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 9442 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 Cell cycle arrest protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	334	Total	C	N	O	S	0	0	0
			2646	1676	441	516	13			
1	C	328	Total	C	N	O	S	0	0	0
			2600	1651	434	502	13			
1	E	332	Total	C	N	O	S	0	0	0
			2631	1670	441	507	13			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	342	GLU	-	cloning artifact	UNP P26449
A	343	LEU	-	cloning artifact	UNP P26449
A	344	HIS	-	expression tag	UNP P26449
A	345	HIS	-	expression tag	UNP P26449
A	346	HIS	-	expression tag	UNP P26449
A	347	HIS	-	expression tag	UNP P26449
A	348	HIS	-	expression tag	UNP P26449
A	349	HIS	-	expression tag	UNP P26449
C	342	GLU	-	cloning artifact	UNP P26449
C	343	LEU	-	cloning artifact	UNP P26449
C	344	HIS	-	expression tag	UNP P26449
C	345	HIS	-	expression tag	UNP P26449
C	346	HIS	-	expression tag	UNP P26449
C	347	HIS	-	expression tag	UNP P26449
C	348	HIS	-	expression tag	UNP P26449
C	349	HIS	-	expression tag	UNP P26449
E	342	GLU	-	cloning artifact	UNP P26449
E	343	LEU	-	cloning artifact	UNP P26449
E	344	HIS	-	expression tag	UNP P26449
E	345	HIS	-	expression tag	UNP P26449
E	346	HIS	-	expression tag	UNP P26449
E	347	HIS	-	expression tag	UNP P26449
E	348	HIS	-	expression tag	UNP P26449

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	349	HIS	-	expression tag	UNP P26449

• Molecule 2 is a protein called Checkpoint serine/threonine-protein kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	36	Total 307	C 200	N 47	O 59	S 1	0	0	0
2	D	36	Total 307	C 200	N 47	O 59	S 1	0	0	0
2	F	36	Total 307	C 200	N 47	O 59	S 1	0	0	0

• Molecule 3 is water.

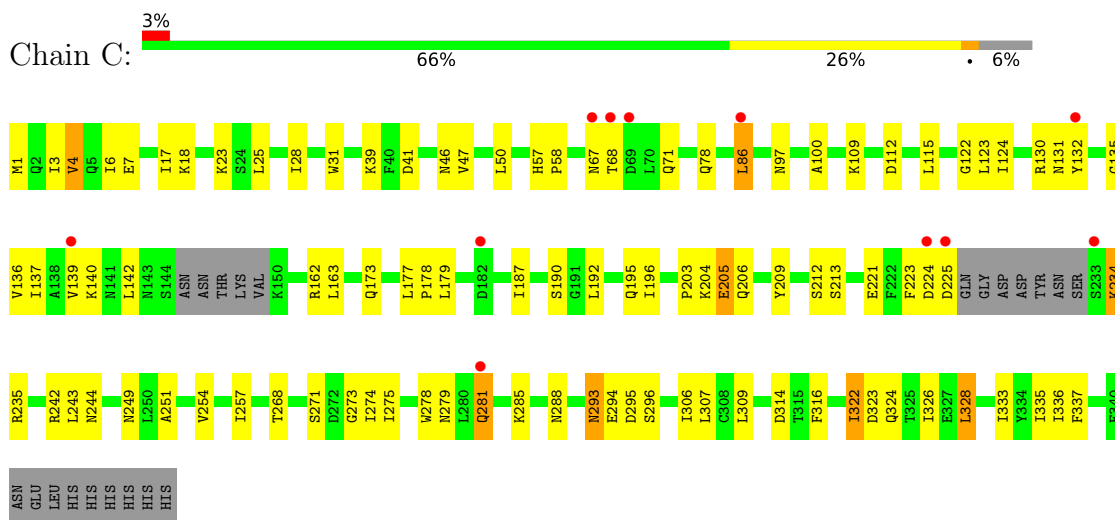| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	264	Total 264	O 264	0	0
3	B	21	Total 21	O 21	0	0
3	C	188	Total 188	O 188	0	0
3	D	27	Total 27	O 27	0	0
3	E	136	Total 136	O 136	0	0
3	F	8	Total 8	O 8	0	0

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.
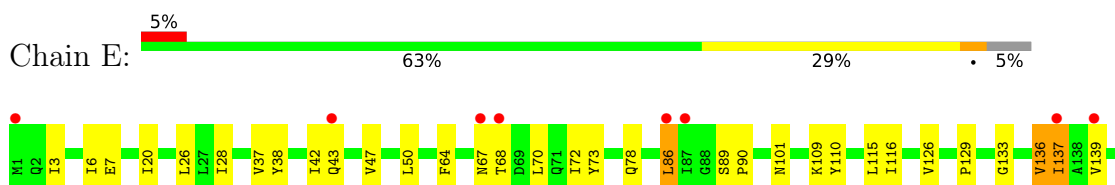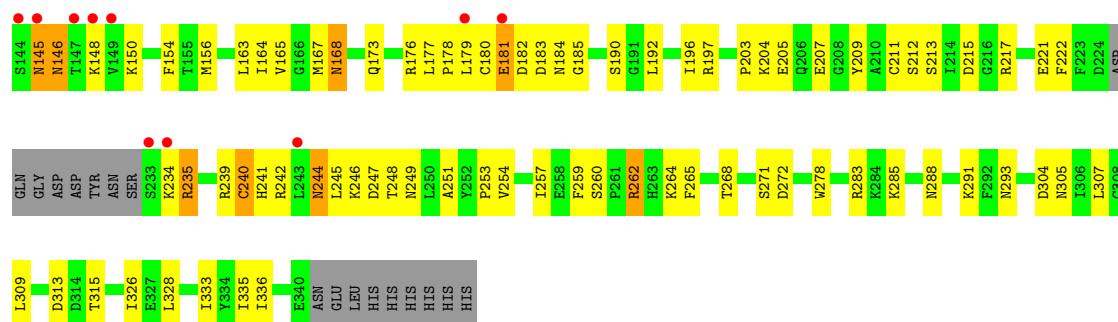
- Molecule 1: Cell cycle arrest protein

- Molecule 1: Cell cycle arrest protein

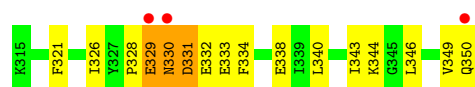- Molecule 1: Cell cycle arrest protein

S144 N145 N146 T147 K148 V149 K150 F154 T155 M156 L163 I164 V165 G166 M167 N168 Q173 R176 L177 P178 L179 C180 E181 D182 D183 N184 G185 S190 G191 L192 I196 R197 P203 K204 E205 Q206 E207 G208 Y209 A210 C211 S212 S213 I214 D215 G216 R217 E221 F222 F223 D224 ASP

GLN GLY ASP ASP TYR ASN SER S233 K234 R235 R239 C240 H241 R242 L243 N244 L245 K246 D247 T248 N249 L250 A251 Y252 P253 V254 I257 E258 F259 S260 P261 R262 H263 K264 F265 T268 S271 D272 W278 R283 K284 K285 N288 K291 F292 N293 D304 N305 I306 L307 C308

L309 D313 D314 T315 I326 E327 L328 I333 Y334 I335 I336 E340 ASN GLU LEU HIS HIS HIS HIS HIS HIS

• Molecule 2: Checkpoint serine/threonine-protein kinase

Chain B:

11%

72% 25% •

K315 R318 F321 I326 Y327 P328 E329 N330 D331 E332 V349 Q350

• Molecule 2: Checkpoint serine/threonine-protein kinase

Chain D:

8%

56% 36% 8%

K315 F321 I326 Y327 P328 E329 N330 D331 E332 E333 F334 E338 I339 L340 I343 K344 G345 L346 V349 Q350

• Molecule 2: Checkpoint serine/threonine-protein kinase

Chain F:

6%

75% 25%

K315 R318 I319 L325 N330 I339 M342 K348 V349 Q350

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	50.66Å 79.18Å 142.58Å 90.00° 98.34° 90.00°	Depositor
Resolution (Å)	50.00 – 1.90 49.55 – 1.90	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-1.90) 85.2 (49.55-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.29 (at 1.90Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.219 , 0.262 0.218 , 0.261	Depositor DCC
R_{free} test set	3811 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	25.1	Xtriage
Anisotropy	0.764	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 49.3	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9442	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% 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	0.35	0/2692	0.68	0/3644
1	C	0.34	0/2644	0.64	0/3576
1	E	0.32	0/2676	0.63	0/3621
2	B	0.38	0/313	0.53	0/420
2	D	0.42	0/313	0.60	0/420
2	F	0.34	0/313	0.53	0/420
All	All	0.34	0/8951	0.64	0/12101

There are no bond length outliers.

There are no bond angle outliers.

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	A	2646	0	2618	63	0
1	C	2600	0	2591	90	0
1	E	2631	0	2629	105	0
2	B	307	0	301	10	0
2	D	307	0	301	19	0
2	F	307	0	301	8	0
3	A	264	0	0	4	0
3	B	21	0	0	1	0
3	C	188	0	0	5	0
3	D	27	0	0	0	0

Continued on next page...

W O R L D W I D E

wwPDB

PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	136	0	0	4	0
3	F	8	0	0	1	0
All	All	9442	0	8741	278	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 16.

The worst 5 of 278 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)
2:D:329:GLU:HG2	2:D:330:ASN:H	1.24	1.01
1:E:239:ARG:HG2	1:E:242:ARG:HH21	1.27	0.99
1:C:86:LEU:H	1:C:86:LEU:HD23	1.32	0.94
1:E:47:VAL:HG11	1:E:335:ILE:HD12	1.51	0.93
1:A:322:ILE:HD13	1:A:322:ILE:H	1.36	0.91

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	330/349 (95%)	311 (94%)	15 (4%)	4 (1%)	13	4
1	C	322/349 (92%)	304 (94%)	17 (5%)	1 (0%)	41	31
1	E	328/349 (94%)	292 (89%)	24 (7%)	12 (4%)	3	0
2	B	34/36 (94%)	28 (82%)	5 (15%)	1 (3%)	4	1
2	D	34/36 (94%)	28 (82%)	4 (12%)	2 (6%)	1	0
2	F	34/36 (94%)	27 (79%)	6 (18%)	1 (3%)	4	1
All	All	1082/1155 (94%)	990 (92%)	71 (7%)	21 (2%)	8	1

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	204	LYS
2	B	328	PRO
1	C	204	LYS
2	D	329	GLU
1	E	145	ASN

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	298/313 (95%)	289 (97%)	9 (3%)	41	33
1	C	293/313 (94%)	280 (96%)	13 (4%)	28	19
1	E	297/313 (95%)	289 (97%)	8 (3%)	44	38
2	B	34/34 (100%)	33 (97%)	1 (3%)	42	35
2	D	34/34 (100%)	33 (97%)	1 (3%)	42	35
2	F	34/34 (100%)	33 (97%)	1 (3%)	42	35
All	All	990/1041 (95%)	957 (97%)	33 (3%)	38	29

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	205	GLU
1	E	240	CYS
2	F	325	LEU
1	C	71	GLN
1	C	7	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	151	ASN
1	E	244	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	249	ASN
1	E	329	ASN
1	E	146	ASN

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	334/349 (95%)	-0.09	6 (1%)	68	71	17, 30, 50, 81	0
1	C	328/349 (93%)	0.15	11 (3%)	45	48	20, 37, 60, 73	0
1	E	332/349 (95%)	0.31	18 (5%)	25	29	26, 43, 67, 83	0
2	B	36/36 (100%)	0.46	4 (11%)	5	6	21, 36, 66, 74	0
2	D	36/36 (100%)	0.16	3 (8%)	11	13	21, 30, 59, 69	0
2	F	36/36 (100%)	0.53	2 (5%)	24	27	30, 43, 73, 79	0
All	All	1102/1155 (95%)	0.15	44 (3%)	38	41	17, 36, 63, 83	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	139	VAL	9.5
1	A	325	THR	8.6
2	F	350	GLN	7.5
1	E	145	ASN	5.7
1	A	326	ILE	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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.