WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 10:45 am BST

PDB ID	:	1QIU
Title	:	A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for biological fibres
Authors	:	van Raaij, M.J.; Lavigne, G.; Mitraki, A.; Cusack, S.
Deposited on	:	1999-06-16
Resolution	:	2.40 Å(reported)

The following versions of software and data (see references (1)) were used in the production of this report:

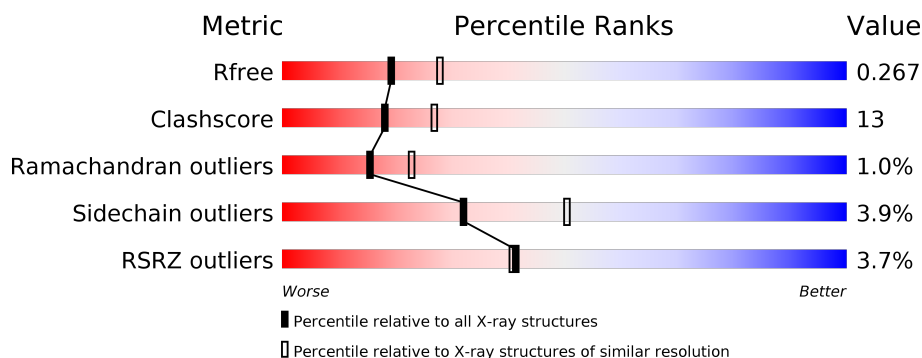|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.40 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		0.267
Clashscore		13
Ramachandran outliers		1.0%
Sidechain outliers		3.9%
RSRZ outliers		3.7%

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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 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	264	4% / 74% / 24% / •
1	B	264	4% / 69% / 30% / •
1	C	264	4% / 76% / 22% / •
1	D	264	4% / 75% / 21% / •
1	E	264	4% / 71% / 26% / •
1	F	264	2% / 74% / 23% / •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12621 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 ADENOVIRUS FIBRE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	264	Total 2007	C 1254	N 331	O 413	S 9	0	0	0
1	B	264	Total 2007	C 1254	N 331	O 413	S 9	0	0	0
1	C	264	Total 2007	C 1254	N 331	O 413	S 9	0	0	0
1	D	264	Total 2007	C 1254	N 331	O 413	S 9	0	0	0
1	E	264	Total 2007	C 1254	N 331	O 413	S 9	0	0	0
1	F	264	Total 2007	C 1254	N 331	O 413	S 9	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	99	Total 99	O 99	0	0
2	B	96	Total 96	O 96	0	0
2	C	85	Total 85	O 85	0	0
2	D	95	Total 95	O 95	0	0
2	E	92	Total 92	O 92	0	0
2	F	112	Total 112	O 112	0	0

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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotat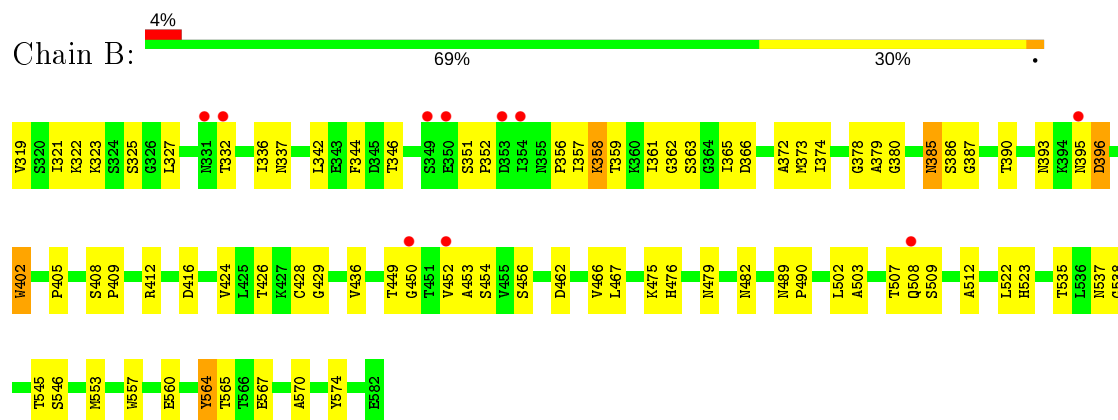ed by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 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.
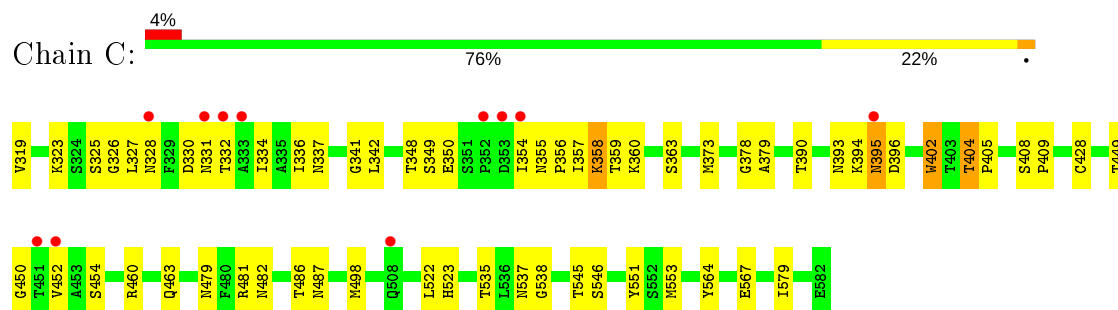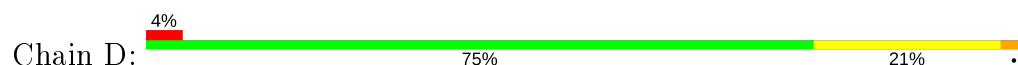
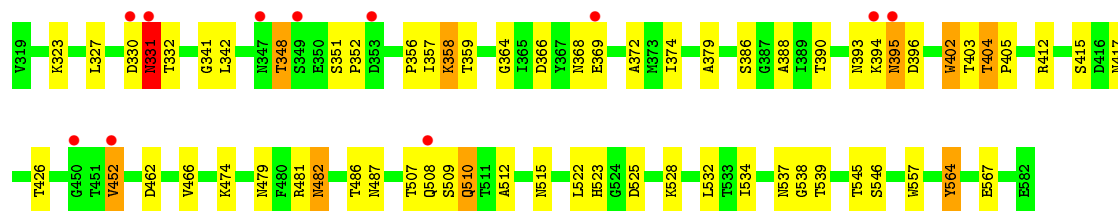- Molecule 1: ADENOVIRUS FIBRE

- Molecule 1: ADENOVIRUS FIBRE

- Molecule 1: ADENOVIRUS FIBRE

- Molecule 1: ADENOVIRUS FIBRE

V319 K323 L327 D330 N331 T332 G341 L342 N347 T348 S349 E350 S351 P352 D353 P356 I357 K358 T359 G364 I365 D366 Y367 N368 E369 A372 M373 I374 A379 S386 G387 A388 I389 T390 N393 K394 N395 D396 W402 T403 T404 P405 R412 S415 D416 N417

T426 G450 I451 V452 D462 V466 K474 N479 F480 R481 N482 T486 N487 T507 Q508 S509 Q510 T511 A512 N515 L522 H523 G524 D525 K528 L532 T533 I534 N537 G538 T539 T545 S546 W557 Y564 E567 E582

• Molecule 1: ADENOVIRUS FIBRE

Chain E:

4%
71%   26%

V319 S320 I321 L327 N328 F329 D330 N331 T332 A333 I334 A335 I336 L342 N347 T348 S349 E350 S351 P352 D353 I354 K358 T359 K360 I361 G362 S363 Y367 A372 M373 K376 L377 G378 A379 F383 D384 N385 S386 G387 A388 I389 T390 K394 N395 D396 L399 W402 T403

T404 P405 D406 P407 C419 K420 T426 K427 C428 T435 L445 T449 V452 A453 S454 D462 Q463 V466 E469 F480 R481 N482 Y491 M498 P499 Q508 S509 Q510 T511 N514 T535 L536 N537 G538 E544 T545 S546 Y551 S552 M553 E560

K563 Y564 E567 Y574 E582

• Molecule 1: ADENOVIRUS FIBRE

Chain F:

2%
74%   23%

V319 S320 I321 K322 K323 S324 L327 N328 F329 D330 N331 T332 A333 I334 A335 I336 L342 E343 F344 D345 T348 S349 E350 P356 I357 K358 T359 K360 D366 Y367 N368 E369 N370 G371 A372 T375 A379 G380 L381 I389 T390 N393 K394 N395 W402 T403 T404 P405 S408

P409 C428 L445 T449 G450 T451 V452 A453 S454 D462 R481 T486 N487 A488 N489 M498 A503 T507 Q508 S509 Q510 K526 T535 G538 T539 T545 S546 Y551 S552 M553 E560 K563 Y564 T565 F569 N572 E582

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	165.51Å 95.87Å 211.77Å 90.00° 106.83° 90.00°	Depositor
Resolution (Å)	25.00 – 2.40 25.06 – 2.41	Depositor EDS
% Data completeness (in resolution range)	84.0 (25.00-2.40) 78.0 (25.06-2.41)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	CNS 0.4	Depositor
R, R_{free}	0.232 , 0.265 0.232 , 0.267	Depositor DCC
R_{free} test set	827 reflections (0.80%)	wwPDB-VP
Wilson B-factor (Å^2)	(Not available)	Xtriage
Anisotropy	(Not available)	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 22.1	EDS
L-test for twinning[1]	$< |L| > =$ (Not available), $< L^2 > =$ (Not available)	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	12621	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *(Not available)*

---

[1]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	0.37	0/2046	0.65	0/2780
1	B	0.37	0/2046	0.64	0/2780
1	C	0.36	0/2046	0.65	0/2780
1	D	0.37	0/2046	0.66	0/2780
1	E	0.37	0/2046	0.65	0/2780
1	F	0.37	0/2046	0.64	0/2780
All	All	0.37	0/12276	0.65	0/16680

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	2007	0	1949	63	0
1	B	2007	0	1949	73	0
1	C	2007	0	1949	60	0
1	D	2007	0	1949	59	0
1	E	2007	0	1949	71	0
1	F	2007	0	1949	68	0
2	A	99	0	0	4	0
2	B	96	0	0	3	0
2	C	85	0	0	0	0
2	D	95	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	92	0	0	1	0
2	F	112	0	0	7	0
All	All	12621	0	11694	317	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 13.

The worst 5 of 317 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:321:ILE:HD11	1:B:327:LEU:HB3	1.45	0.98
1:A:321:ILE:HD11	1:A:334:ILE:HG12	1.51	0.92
1:B:336:ILE:HD11	1:C:326:GLY:HA3	1.47	0.92
1:A:336:ILE:HD11	1:B:336:ILE:HG22	1.51	0.91
1:A:560:GLU:HB2	1:A:563:LYS:HD2	1.56	0.88

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	262/264 (99%)	241 (92%)	21 (8%)	0	100	100
1	B	262/264 (99%)	230 (88%)	30 (12%)	2 (1%)	19	29
1	C	262/264 (99%)	234 (89%)	26 (10%)	2 (1%)	19	29
1	D	262/264 (99%)	235 (90%)	23 (9%)	4 (2%)	10	14
1	E	262/264 (99%)	238 (91%)	21 (8%)	3 (1%)	14	20
1	F	262/264 (99%)	242 (92%)	16 (6%)	4 (2%)	10	14
All	All	1572/1584 (99%)	1420 (90%)	137 (9%)	15 (1%)	15	23

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	395	ASN
1	E	452	VAL
1	F	395	ASN
1	D	331	ASN
1	D	395	ASN

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	228/228 (100%)	218 (96%)	10 (4%)	28	45
1	B	228/228 (100%)	216 (95%)	12 (5%)	22	37
1	C	228/228 (100%)	224 (98%)	4 (2%)	59	76
1	D	228/228 (100%)	219 (96%)	9 (4%)	32	50
1	E	228/228 (100%)	219 (96%)	9 (4%)	32	50
1	F	228/228 (100%)	219 (96%)	9 (4%)	32	50
All	All	1368/1368 (100%)	1315 (96%)	53 (4%)	32	50

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	358	LYS
1	D	402	TRP
1	F	393	ASN
1	C	402	TRP
1	D	331	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	476	HIS
1	D	393	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	385	ASN
1	C	487	ASN
1	C	515	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 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	264/264 (100%)	-0.08	10 (3%)	40	39	7, 29, 68, 89	0
1	B	264/264 (100%)	-0.06	10 (3%)	40	39	10, 31, 68, 87	0
1	C	264/264 (100%)	-0.00	11 (4%)	36	35	11, 30, 75, 89	0
1	D	264/264 (100%)	-0.06	11 (4%)	36	35	9, 31, 68, 90	0
1	E	264/264 (100%)	-0.02	11 (4%)	36	35	10, 31, 75, 89	0
1	F	264/264 (100%)	-0.12	6 (2%)	60	58	10, 30, 66, 76	0
All	All	1584/1584 (100%)	-0.06	59 (3%)	41	41	7, 30, 72, 90	0

The worst 5 of 59 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	395	ASN	4.4
1	D	349	SER	4.2
1	A	321	ILE	4.0
1	C	352	PRO	3.8
1	D	394	LYS	3.8

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.