Full wwPDB X-ray Structure Validation Report (i)

May 30, 2020 – 12:17 am BST

PDB ID	:	4CW8
Title	:	Structure of the carboxy-terminal domain of the turkey type 3 siadenovirus fibre, virulent form
Authors	:	Singh, A.K.; van Raaij, M.J.
Deposited on	:	2014-04-01
Resolution	:	2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

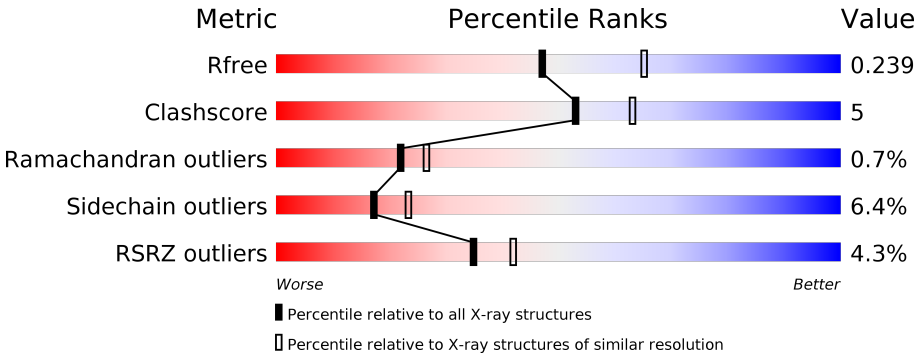
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	190	3% 59% 13% 27%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1141 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 FIBER KNOB DOMAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	138	Total	C	N	O	S	0	0	0
			1091	706	188	192	5			

There are 38 discrepancies between the modelled and reference sequences:

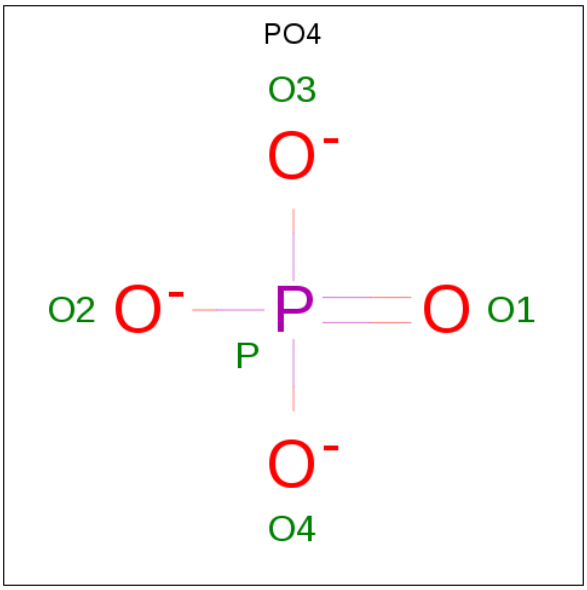
Chain	Residue	Modelled	Actual	Comment	Reference
A	265	MET	-	expression tag	UNP Q0GF90
A	266	GLY	-	expression tag	UNP Q0GF90
A	267	SER	-	expression tag	UNP Q0GF90
A	268	SER	-	expression tag	UNP Q0GF90
A	269	HIS	-	expression tag	UNP Q0GF90
A	270	HIS	-	expression tag	UNP Q0GF90
A	271	HIS	-	expression tag	UNP Q0GF90
A	272	HIS	-	expression tag	UNP Q0GF90
A	273	HIS	-	expression tag	UNP Q0GF90
A	274	HIS	-	expression tag	UNP Q0GF90
A	275	SER	-	expression tag	UNP Q0GF90
A	276	SER	-	expression tag	UNP Q0GF90
A	277	GLY	-	expression tag	UNP Q0GF90
A	278	LEU	-	expression tag	UNP Q0GF90
A	279	VAL	-	expression tag	UNP Q0GF90
A	280	PRO	-	expression tag	UNP Q0GF90
A	281	ARG	-	expression tag	UNP Q0GF90
A	282	GLY	-	expression tag	UNP Q0GF90
A	283	SER	-	expression tag	UNP Q0GF90
A	284	HIS	-	expression tag	UNP Q0GF90
A	285	MET	-	expression tag	UNP Q0GF90
A	286	ALA	-	expression tag	UNP Q0GF90
A	287	SER	-	expression tag	UNP Q0GF90
A	288	MET	-	expression tag	UNP Q0GF90
A	289	THR	-	expression tag	UNP Q0GF90
A	290	GLY	-	expression tag	UNP Q0GF90
A	291	GLY	-	expression tag	UNP Q0GF90

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	292	GLN	-	expression tag	UNP Q0GF90
A	293	GLN	-	expression tag	UNP Q0GF90
A	294	MET	-	expression tag	UNP Q0GF90
A	295	GLY	-	expression tag	UNP Q0GF90
A	296	ARG	-	expression tag	UNP Q0GF90
A	297	GLY	-	expression tag	UNP Q0GF90
A	298	SER	-	expression tag	UNP Q0GF90
A	299	GLU	-	expression tag	UNP Q0GF90
A	300	PHE	-	expression tag	UNP Q0GF90
A	354	ILE	MET	engineered mutation	UNP Q0GF90
A	376	THR	MET	engineered mutation	UNP Q0GF90

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0

- Molecule 3 is water.

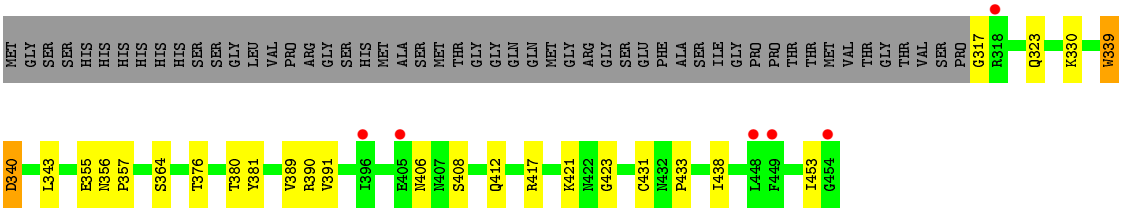
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	40	Total O 40 40	0	0

3 Residue-property plots (i)

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 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: FIBER KNOB DOMAIN

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	98.96Å 98.96Å 98.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.30 31.29 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (30.00-2.30) 100.0 (31.29-2.30)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.98 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.197 , 0.239 0.203 , 0.239	Depositor DCC
R_{free} test set	335 reflections (4.57%)	wwPDB-VP
Wilson B-factor (Å^2)	58.7	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 46.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.042 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1141	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.09% 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

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.65	0/1112	0.81	1/1508 (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	5

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	417	ARG	NE-CZ-NH1	5.69	123.14	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	317	GLY	Peptide
1	A	339	TRP	Peptide
1	A	340	ASP	Peptide
1	A	364	SER	Peptide
1	A	453	ILE	Peptide

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	1091	0	1142	12	0
2	A	10	0	0	0	0
3	A	40	0	0	0	0
All	All	1141	0	1142	12	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 5.

All (12) 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:343:LEU:HD23	1:A:391:VAL:HG11	1.65	0.77
1:A:390:ARG:NE	1:A:423:GLY:O	2.22	0.72
1:A:412:GLN:OE1	1:A:431:CYS:HB2	1.91	0.69
1:A:376:THR:HG23	1:A:380:THR:O	2.01	0.61
1:A:381:TYR:OH	1:A:438:ILE:HD12	2.06	0.56
1:A:406:ASN:OD1	1:A:408:SER:OG	2.24	0.51
1:A:343:LEU:HD23	1:A:391:VAL:CG1	2.38	0.48
1:A:376:THR:HG21	1:A:381:TYR:CE1	2.50	0.46
1:A:381:TYR:CZ	1:A:438:ILE:HD12	2.52	0.45
1:A:390:ARG:NH2	1:A:423:GLY:O	2.50	0.44
1:A:390:ARG:CZ	1:A:423:GLY:O	2.66	0.43
1:A:421:LYS:HD2	1:A:421:LYS:HA	1.93	0.40

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	136/190 (72%)	128 (94%)	7 (5%)	1 (1%)	22 26

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	357	PRO

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	125/166 (75%)	117 (94%)	8 (6%)	17 23

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	323	GLN
1	A	330	LYS
1	A	339	TRP
1	A	340	ASP
1	A	355	GLU
1	A	356	ASN
1	A	389	VAL
1	A	433	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

2 ligands are modelled in this entry.

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	$\#\|Z\| > 2$
2	PO4	A	1456	-	4,4,4	0.92	0	6,6,6	0.93	0
2	PO4	A	1455	-	4,4,4	0.94	0	6,6,6	0.55	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

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^{th} 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	138/190 (72%)	0.05	6 (4%) 35 42	42, 61, 99, 119	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	448	LEU	3.1
1	A	318	ARG	2.6
1	A	454	GLY	2.5
1	A	405	GLU	2.5
1	A	396	ILE	2.4
1	A	449	PHE	2.3

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

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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
2	PO4	A	1455	5/5	0.91	0.14	57,57,60,61	5
2	PO4	A	1456	5/5	0.92	0.33	54,55,62,63	5

6.5 Other polymers (i)

There are no such residues in this entry.