wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:22 PM GMT

PDB ID	:	1OHF
Title	:	THE REFINED STRUCTURE OF NUDAURELIA CAPENSIS OMEGA VIRUS
Authors	:	Helgstrand, C.; Munshi, S.; Johnson, J.E.; Liljas, L.
Deposited on	:	2003-05-26
Resolution	:	2.80 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	FAILED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.80 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
Clashscore		19
Ramachandran outliers		1.1%
Sidechain outliers		6.8%

Worse *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
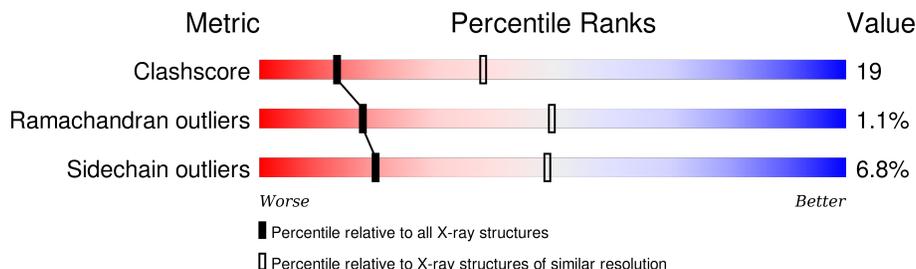
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)

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 <=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.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	644	51% 29% • 16%
1	B	644	54% 27% • 15%
1	C	644	64% 25% • 8%
1	D	644	62% 24% • 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 17851 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 NUDAURELIA CAPENSIS OMEGA VIRUS CAPSID PRO-TEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	539	Total	C	N	O	S	0	0	0
			4134	2632	683	806	13			
1	B	547	Total	C	N	O	S	0	0	0
			4187	2659	694	821	13			
1	C	590	Total	C	N	O	S	0	0	0
			4519	2861	770	874	14			
1	D	587	Total	C	N	O	S	0	0	0
			4491	2846	761	870	14			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	235	VAL	CYS	SEE REMARK 999	UNP Q90063
A	283	VAL	GLU	SEE REMARK 999	UNP Q90063
B	235	VAL	CYS	SEE REMARK 999	UNP Q90063
B	283	VAL	GLU	SEE REMARK 999	UNP Q90063
C	235	VAL	CYS	SEE REMARK 999	UNP Q90063
C	283	VAL	GLU	SEE REMARK 999	UNP Q90063
D	235	VAL	CYS	SEE REMARK 999	UNP Q90063
D	283	VAL	GLU	SEE REMARK 999	UNP Q90063

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		
2	D	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	97	Total 97	O 97	0	0
3	B	127	Total 127	O 127	0	0
3	C	148	Total 148	O 148	0	0
3	D	144	Total 144	O 144	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 errors 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.

Note EDS failed to run properly.

• Molecule 1: NUDAURELIA CAPENSIS OMEGA VIRUS CAPSID PROTEIN
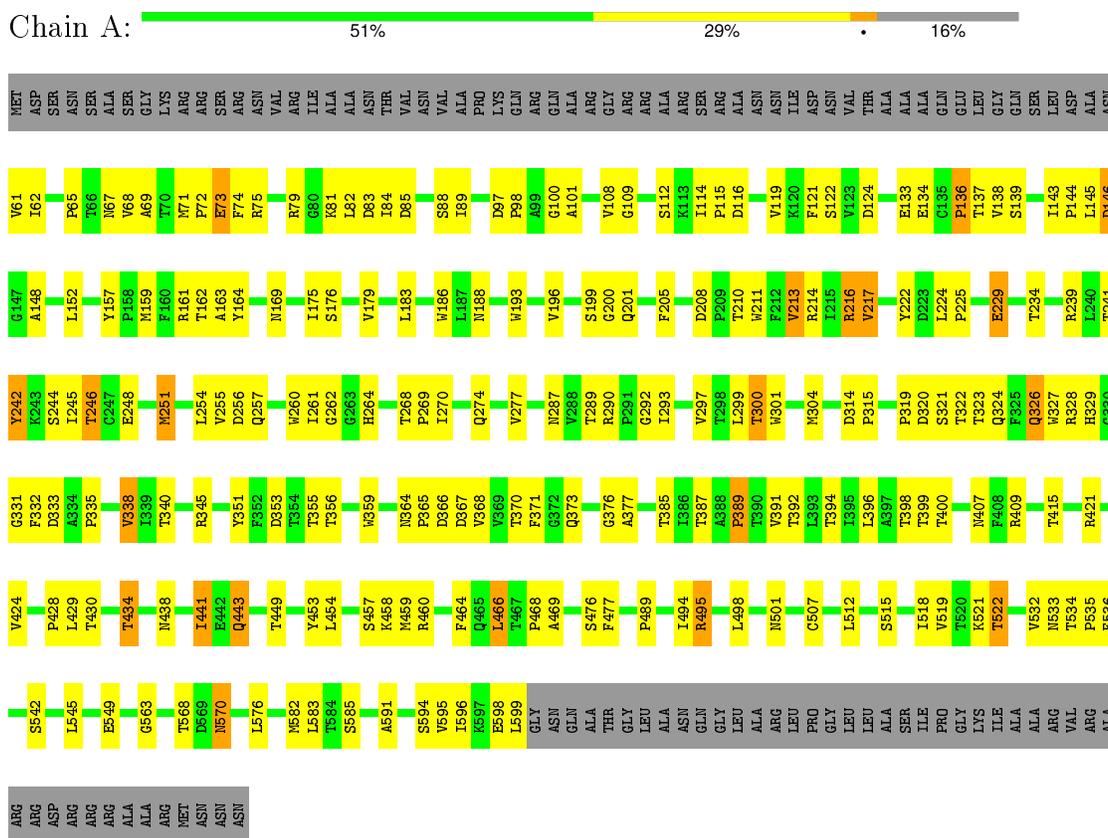
Chain A:

• Molecule 1: NUDAURELIA CAPENSIS OMEGA VIRUS CAPSID PROTEIN

Chain B:

● Molecule 1: NUDAURELIA CAPENSIS OMEGA VIRUS CAPSID PROTEIN

Chain C:

● Molecule 1: NUDAURELIA CAPENSIS OMEGA VIRUS CAPSID PROTEIN

Chain D:
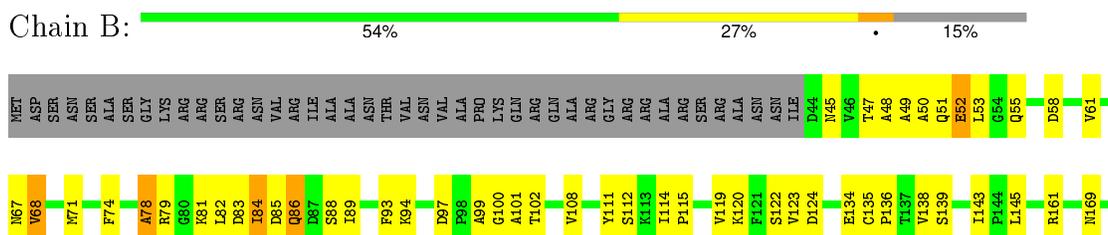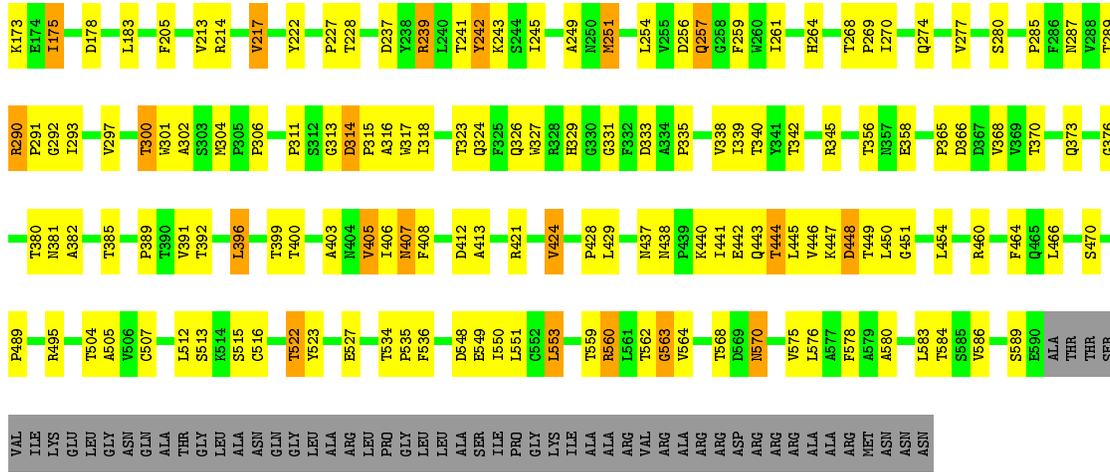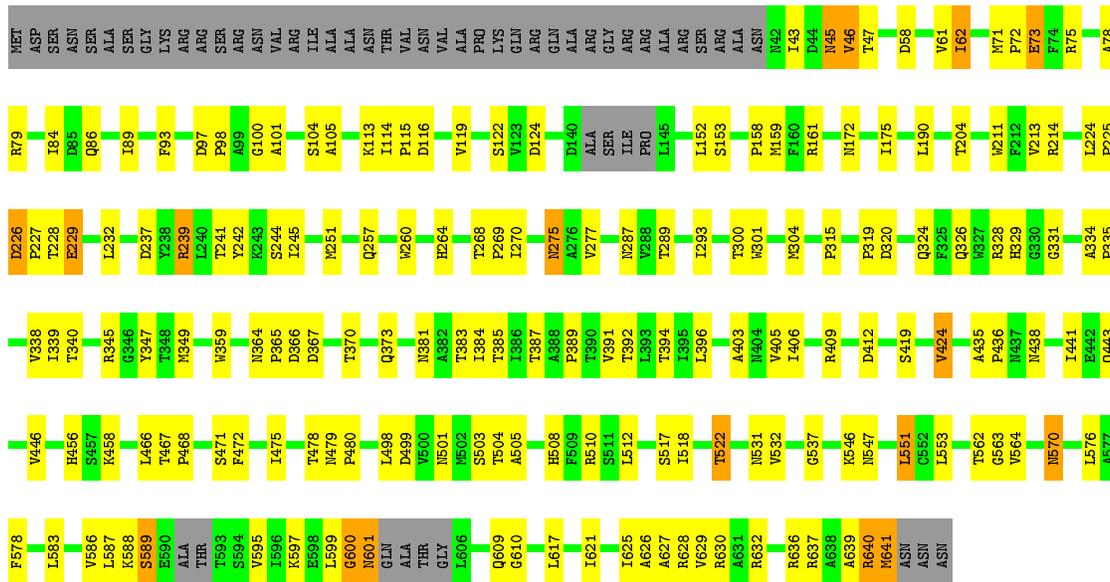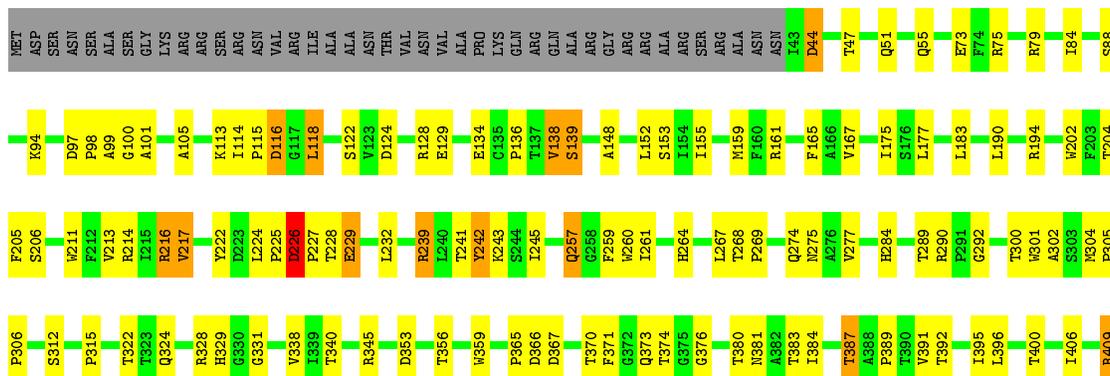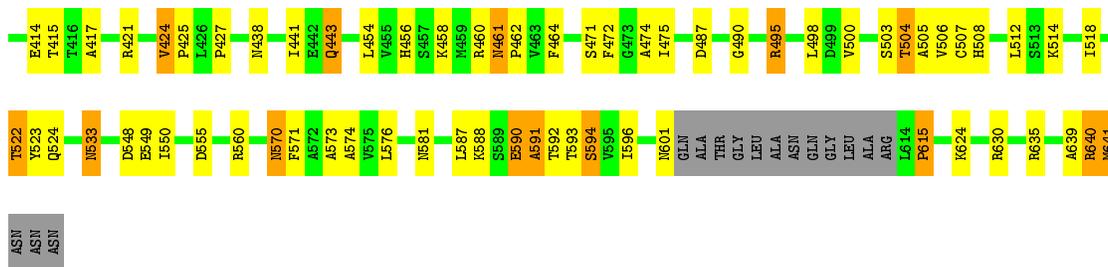
4 Data and refinement statistics (i)

EDS failed to run properly - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	413.55Å 410.22Å 419.67Å 59.13° 58.90° 64.04°	Depositor
Resolution (Å)	20.00 – 2.80	Depositor
% Data completeness (in resolution range)	50.9 (20.00-2.80)	Depositor
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	6.07 (at 2.79Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.219 , 0.221	Depositor
Wilson B-factor (Å^2)	36.1	Xtriage
Anisotropy	0.126	Xtriage
Estimated twinning fraction	0.045 for h-l,h,h-k 0.045 for k,k-l,-h+k 0.046 for -k,-h,-l 0.046 for -h+l,-k+l,l 0.046 for k-l,h-l,-l 0.046 for -h,-h+l,-h+k 0.046 for -k+l,-k,h-k	Xtriage
L-test for twinning[2]	$<\|L\|>=0.41$, $<L^2>=0.23$	Xtriage
Outliers	0 of 2465275 reflections	Xtriage
Total number of atoms	17851	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 5.23% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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

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	0.43	0/4244	0.67	0/5816
1	B	0.44	0/4297	0.67	0/5889
1	C	0.42	0/4627	0.65	0/6324
1	D	0.43	0/4602	0.68	0/6296
All	All	0.43	0/17770	0.67	0/24325

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 (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	4134	0	3987	189	0
1	B	4187	0	4021	175	0
1	C	4519	0	4389	169	0
1	D	4491	0	4363	154	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	97	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	127	0	0	5	0
3	C	148	0	0	2	0
3	D	144	0	0	4	0
All	All	17851	0	16760	643	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 19.

The worst 5 of 643 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:428:PRO:HB2	1:A:434:THR:HG22	1.34	1.08
1:A:175:ILE:HD11	1:A:429:LEU:HB3	1.37	1.06
1:B:342:THR:HG22	1:B:368:VAL:HG22	1.35	1.04
1:B:269:PRO:HB3	1:B:424:VAL:HG13	1.39	0.99
1:A:329:HIS:CE1	1:A:373:GLN:HE22	1.79	0.98

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	535/644 (83%)	495 (92%)	37 (7%)	3 (1%)	30	65
1	B	543/644 (84%)	500 (92%)	36 (7%)	7 (1%)	15	44
1	C	580/644 (90%)	552 (95%)	22 (4%)	6 (1%)	19	52
1	D	581/644 (90%)	540 (93%)	33 (6%)	8 (1%)	14	42
All	All	2239/2576 (87%)	2087 (93%)	128 (6%)	24 (1%)	17	50

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	589	SER
1	D	139	SER
1	D	615	PRO
1	D	640	ARG
1	A	200	GLY

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	450/527 (85%)	415 (92%)	35 (8%)	16	41
1	B	454/527 (86%)	421 (93%)	33 (7%)	17	44
1	C	486/527 (92%)	459 (94%)	27 (6%)	26	59
1	D	484/527 (92%)	451 (93%)	33 (7%)	20	49
All	All	1874/2108 (89%)	1746 (93%)	128 (7%)	20	49

5 of 128 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	424	VAL
1	C	86	GLN
1	D	495	ARG
1	B	504	THR
1	B	568	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 52 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	479	ASN
1	C	329	HIS
1	D	533	ASN
1	C	86	GLN
1	C	275	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)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS failed to run properly - this section will therefore be empty.

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

EDS failed to run properly - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS failed to run properly - this section will therefore be empty.

6.4 Ligands (i)

EDS failed to run properly - this section will therefore be empty.

6.5 Other polymers (i)

EDS failed to run properly - this section will therefore be empty.