wwPDB X-ray Structure Validation Summary Report (i)

Aug 21, 2023 – 04:22 PM EDT

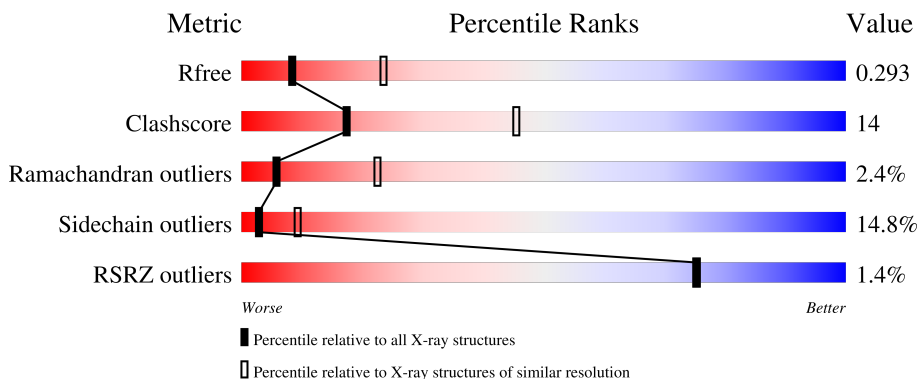|		
PDB ID	:	2OZK
Title	:	Structure of an N-Terminal Truncated Form of Nendou (NSP15) From SARS-CORONAVIRUS
Authors	:	Saikatendu, K.; Joseph, J.; Subramanian, V.; Neuman, B.; Buchmeier, M.; Stevens, R.C.; Kuhn, P.
Deposited on	:	2007-02-26
Resolution	:	2.90 Å(reported)

1   Overall quality at a glance ⓘ

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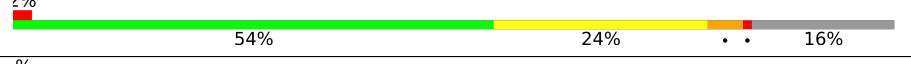*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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		0.293
Clashscore		14
Ramachandran outliers		2.4%
Sidechain outliers		14.8%
RSRZ outliers		1.4%

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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	346	61% 23% • 12%
1	B	346	61% 23% • • 12%
1	C	346	54% 24% • • 16%
1	D	346	51% 30% 5% 14%

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 9402 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 Uridylate-specific endoribonuclease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	306	Total	C	N	O	S	0	0	0
			2402	1538	391	463	10			
1	B	305	Total	C	N	O	S	0	0	0
			2394	1532	390	462	10			
1	C	291	Total	C	N	O	S	0	0	0
			2294	1473	371	440	10			
1	D	296	Total	C	N	O	S	0	0	0
			2312	1485	374	443	10			

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: Uridylate-specific endoribonuclease

Chain A:

• Molecule 1: Uridylate-specific endoribonuclease

Chain B:

• Molecule 1: Uridylate-specific endoribonuclease

Chain C:

SER LEU GLU ASN VAL ALA TYR ASN VAL VAL ASN LYS GLY HIS PHE ASP GLY HIS ALA GLY GLU ALA PRO VAL SER ILE ILE ASN N29 A30 V31 K34 V35 D36 G37 I38 I42 F43 E44 N45 K46 T47 T48 V51 N52 V53 W58 A59 K60 R61 N62 P65 I69 K70

I71 L72 L75 G76 V77 D78 I79 T83 V84 I85 H95 V96 S97 C102 T103 M104 T112 E113 S114 A115 C116 L122 G125 R126 V127 Q130 F134 R138 L142 I143 T144 P157 A158 Q159 V162 T166 L167 E170 F179 I185 I186 T192

Y193 F194 T195 L200 E201 T211 D212 E215 L216 A217 M218 D219 I222 Y230 A231 F232 E233 H234 I235 G238 ASP PHE SER HIS GLY GLN LEU GLY G247 L248 M251 I252 K256 R257 D260 L263 K264 L265 E266 D267 F268 I269 P270 M271 D272 S273 K276 N277 Y278

F279 Q284 THR GLY SER SER LYS C290 V291 C292 I295 D300 D301 F302 I306 Q309 S312 V313 I314 S315 K316 V320 T321 I322 D323 Y324 W332 CYS LYS ASP GLY HIS VAL GLU THR PHE TYR PRO LYS LEU GLN

• Molecule 1: Uridylate-specific endoribonuclease

%

Chain D: 51% 30% 5% 14%

SER LEU GLU ASN VAL ALA TYR ASN VAL VAL ASN LYS GLY HIS PHE ASP GLY HIS ALA GLY GLU ALA PRO VAL SER ILE ILE ASN N29 A30 V31 G37 I38 I42 F43 E44 N45 K46 T47 T48 V51 N52 W58 A59 K60 R61 K64 P65 V66 P67 I71 L72 N73 N74

L75 G76 V84 I85 W86 D87 A92 P93 A94 H95 V96 S97 T98 I99 G100 V101 M104 T105 D106 P111 V121 G125 R126 V127 E128 V131 D132 L133 R138 N139 G140 T144 E145 K149 S154 K155 Q159 V162 E170 S171 V172 K173 T174 Q175 V182

D183 G184 I185 I186 Q187 Q188 L189 T192 Y193 Q196 L200 R206 E210 L214 E215 L216 E220 K226 A231 F232 E233 H234 I235 V236 D239 F240 S241 HIS GLY GLN LEU G246 L250 R257 S258 Q259 D260 S261 P262 L263 K264 L265 E266 M271 N277 Y278

F279 I280 T281 D282 ALA GLN THR GLY S287 S288 K289 C290 V291 C292 I295 D296 L297 L298 L299 D300 D301 F302 V303 I306 Q309 D310 L311 S312 V313 I314 S315 K316 V320 T321 Y324 I327 S328 F329 M330 L331 W332 CYS LYS ASP GLY HIS VAL GLU THR PHE TYR PRO LYS LEU

GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	98.97Å 98.97Å 214.93Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	85.75 – 2.90 48.23 – 2.90	Depositor EDS
% Data completeness (in resolution range)	97.8 (85.75-2.90) 97.8 (48.23-2.90)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$<I/\sigma(I)>$ [1]	1.98 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.243 , 0.301 0.237 , 0.293	Depositor DCC
R_{free} test set	2601 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	73.6	Xtriage
Anisotropy	0.061	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 20.4	EDS
L-test for twinning[2]	$<|L|> = 0.51, <L^2> = 0.34$	Xtriage
Estimated twinning fraction	0.014 for -h,-k,l 0.448 for h,-h-k,-l 0.018 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9402	wwPDB-VP
Average B, all atoms (Å^2)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.79% 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	0.55	0/2447	0.71	1/3316 (0.0%)
1	B	0.57	0/2439	0.75	2/3305 (0.1%)
1	C	0.56	0/2335	0.68	0/3163
1	D	0.56	0/2354	0.71	0/3190
All	All	0.56	0/9575	0.71	3/12974 (0.0%)

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	C	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	167	LEU	CA-CB-CG	6.84	131.04	115.30
1	A	245	LEU	CA-CB-CG	6.80	130.95	115.30
1	B	245	LEU	CA-CB-CG	6.26	129.71	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	35	VAL	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	2402	0	2410	57	0
1	B	2394	0	2401	53	0
1	C	2294	0	2308	64	0
1	D	2312	0	2311	92	0
All	All	9402	0	9430	260	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 14.

The worst 5 of 260 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:D:281:THR:HA	1:D:282:ASP:CB	1.80	1.10
1:D:281:THR:HA	1:D:282:ASP:HB2	1.12	1.05
1:D:281:THR:CA	1:D:282:ASP:HB2	1.94	0.96
1:D:192:THR:HG21	1:D:324:TYR:H	1.31	0.96
1:B:192:THR:HG21	1:B:324:TYR:H	1.31	0.92

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	304/346 (88%)	267 (88%)	34 (11%)	3 (1%)	15	45
1	B	303/346 (88%)	271 (89%)	27 (9%)	5 (2%)	9	31

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	285/346 (82%)	240 (84%)	32 (11%)	13 (5%)	2 9
1	D	290/346 (84%)	247 (85%)	36 (12%)	7 (2%)	6 22
All	All	1182/1384 (85%)	1025 (87%)	129 (11%)	28 (2%)	6 22

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	46	LYS
1	C	115	ALA
1	C	232	PHE
1	D	52	ASN
1	D	232	PHE

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	270/303 (89%)	230 (85%)	40 (15%)	3 9
1	B	269/303 (89%)	229 (85%)	40 (15%)	3 9
1	C	258/303 (85%)	212 (82%)	46 (18%)	2 5
1	D	258/303 (85%)	228 (88%)	30 (12%)	5 16
All	All	1055/1212 (87%)	899 (85%)	156 (15%)	3 9

5 of 156 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	271	MET
1	D	192	THR
1	C	284	GLN
1	D	60	LYS
1	D	300	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	130	GLN
1	D	309	GLN
1	C	259	GLN
1	D	196	GLN
1	C	249	HIS

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(Å²)	Q<0.9
1	A	306/346 (88%)	0.28	4 (1%)	77	77	48, 63, 79, 86	0
1	B	305/346 (88%)	0.29	2 (0%)	87	87	47, 63, 79, 82	0
1	C	291/346 (84%)	0.39	6 (2%)	63	61	50, 78, 95, 103	0
1	D	296/346 (85%)	0.36	5 (1%)	70	69	55, 77, 95, 104	0
All	All	1198/1384 (86%)	0.33	17 (1%)	75	75	47, 70, 92, 104	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	112	THR	3.4
1	D	311	LEU	2.9
1	C	134	PHE	2.9
1	D	331	LEU	2.3
1	C	96	VAL	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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.