Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2025 – 01:42 pm BST

PDB ID	:	7Z5A / pdb_00007z5a
Title	:	Crystal structure of the trapped complex of mouse Endonuclease VIII-LIKE 3 (mNEIL3) and hairpin DNA with 5'overhang
Authors	:	Silhan, J.; Huskova, A.
Deposited on	:	2022-03-08
Resolution	:	2.28 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 ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

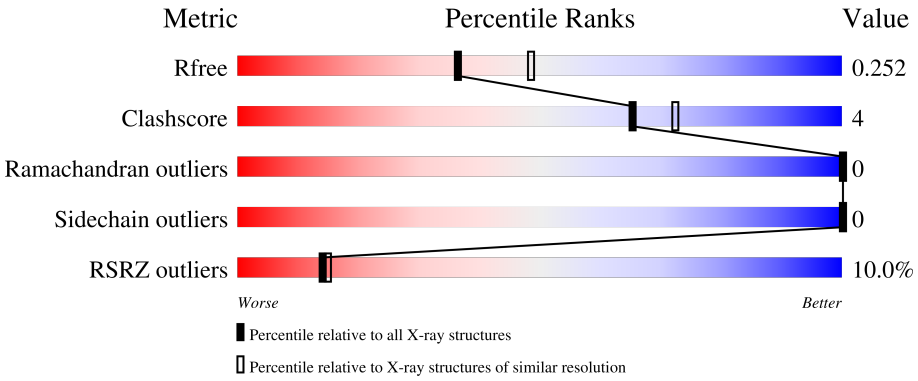
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.44

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 2.28 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}	164625	8487 (2.30-2.26)
Clashscore	180529	9437 (2.30-2.26)
Ramachandran outliers	177936	9341 (2.30-2.26)
Sidechain outliers	177891	9342 (2.30-2.26)
RSRZ outliers	164620	8487 (2.30-2.26)

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	303	9% 81% 6% 13%
2	E	19	11% 53% 32% 16%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2484 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 Endonuclease 8-like 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	264	Total	C	N	O	S	0	0	0
			2039	1282	370	366	21			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	46	PRO	LEU	conflict	UNP Q8K203
A	90	HIS	PRO	conflict	UNP Q8K203
A	114	GLY	ALA	conflict	UNP Q8K203
A	150	GLU	VAL	conflict	UNP Q8K203
A	220	ARG	CYS	conflict	UNP Q8K203
A	256	GLY	ASP	conflict	UNP Q8K203
A	283	ASP	-	expression tag	UNP Q8K203
A	284	GLY	-	expression tag	UNP Q8K203
A	285	LEU	-	expression tag	UNP Q8K203
A	286	GLU	-	expression tag	UNP Q8K203
A	287	VAL	-	expression tag	UNP Q8K203
A	288	LEU	-	expression tag	UNP Q8K203
A	289	PHE	-	expression tag	UNP Q8K203
A	290	GLN	-	expression tag	UNP Q8K203
A	291	GLY	-	expression tag	UNP Q8K203
A	292	PRO	-	expression tag	UNP Q8K203
A	293	GLY	-	expression tag	UNP Q8K203
A	294	SER	-	expression tag	UNP Q8K203
A	295	SER	-	expression tag	UNP Q8K203
A	296	HIS	-	expression tag	UNP Q8K203
A	297	HIS	-	expression tag	UNP Q8K203
A	298	HIS	-	expression tag	UNP Q8K203
A	299	HIS	-	expression tag	UNP Q8K203
A	300	HIS	-	expression tag	UNP Q8K203
A	301	HIS	-	expression tag	UNP Q8K203
A	302	HIS	-	expression tag	UNP Q8K203
A	303	HIS	-	expression tag	UNP Q8K203

- Molecule 2 is a DNA chain called DNA (5'-D(P*TP*TP*TP*(PED)P*AP*CP*GP*CP*G P*AP*AP*GP*CP*GP*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	16	Total	C	N	O	P	0	0	0
			321	152	57	96	16			

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		

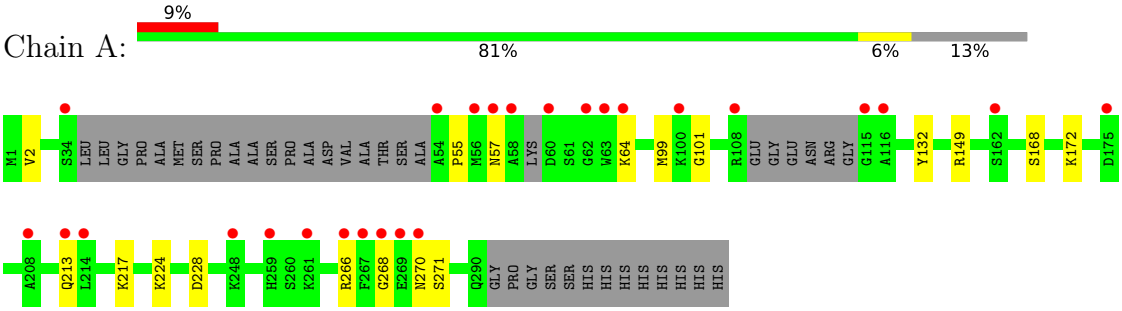
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	102	Total	O	0	0
			102	102		
4	E	21	Total	O	0	0
			21	21		

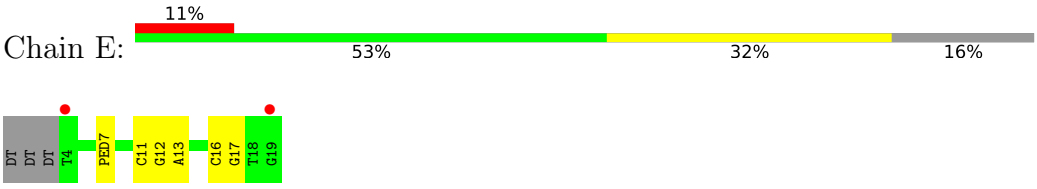
3 Residue-property plots (i)

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: Endonuclease 8-like 3



• Molecule 2: DNA (5'-D(P*TP*TP*TP*(PED)P*AP*CP*GP*CP*GP*AP*AP*GP*CP*GP*TP*G)-3')



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.12Å 71.53Å 94.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.13 – 2.28 36.13 – 2.28	Depositor EDS
% Data completeness (in resolution range)	98.9 (36.13-2.28) 99.1 (36.13-2.28)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.75 (at 2.01Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.237 , 0.248 0.234 , 0.252	Depositor DCC
R_{free} test set	897 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	11.0	Xtriage
Anisotropy	0.335	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 40.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	2484	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.93% 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: ZN, PED

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	0.42	1/2074 (0.0%)	0.65	5/2786 (0.2%)
2	E	0.27	0/346	0.55	0/530
All	All	0.41	1/2420 (0.0%)	0.64	5/3316 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	64	LYS	C-O	5.86	1.31	1.24

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	64	LYS	CA-C-N	-6.02	112.22	120.28
1	A	64	LYS	C-N-CA	-6.02	112.22	120.28
1	A	270	ASN	N-CA-C	-5.40	101.67	110.20
1	A	268	GLY	CA-C-O	-5.29	117.11	122.56
1	A	270	ASN	CA-C-O	-5.08	115.27	121.05

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	2039	0	2027	14	0
2	E	321	0	178	7	0
3	A	1	0	0	0	0
4	A	102	0	0	1	0
4	E	21	0	0	0	0
All	All	2484	0	2205	17	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 4.

All (17) 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:E:11:DC:H2"	2:E:12:DG:H5"	1.71	0.73
1:A:224:LYS:O	1:A:224:LYS:HD2	2.04	0.56
2:E:12:DG:H2"	2:E:13:DA:C8	2.41	0.56
1:A:149:ARG:HH11	1:A:149:ARG:HG2	1.71	0.56
1:A:213:GLN:NE2	4:A:503:HOH:O	2.38	0.55
1:A:55:PRO:HB2	1:A:57:ASN:OD1	2.07	0.54
1:A:2:VAL:N	2:E:7:PED:C1'	2.70	0.54
1:A:149:ARG:HG2	1:A:149:ARG:NH1	2.22	0.53
1:A:168:SER:O	1:A:172:LYS:HG3	2.09	0.53
1:A:101:GLY:HA2	1:A:132:TYR:O	2.10	0.52
1:A:2:VAL:H	2:E:7:PED:C1'	2.28	0.47
1:A:224:LYS:HE3	1:A:228:ASP:OD2	2.18	0.44
1:A:266:ARG:HE	1:A:271:SER:HB3	1.84	0.43
1:A:2:VAL:N	2:E:7:PED:H1'3	2.36	0.41
2:E:16:DC:H2"	2:E:17:DG:C8	2.55	0.41
1:A:217:LYS:HE3	1:A:217:LYS:HB3	1.85	0.41
1:A:99:MET:SD	2:E:7:PED:H2'1	2.61	0.40

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	256/303 (84%)	251 (98%)	5 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	220/255 (86%)	220 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	259	HIS
1	A	270	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 oligosaccharides in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is 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

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, 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	264/303 (87%)	0.78	26 (9%) 14 15	21, 36, 62, 99	0
2	E	15/19 (78%)	0.87	2 (13%) 8 9	29, 56, 84, 102	0
All	All	279/322 (86%)	0.78	28 (10%) 14 15	21, 37, 72, 102	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	269	GLU	5.5
1	A	108	ARG	5.2
1	A	54	ALA	5.2
1	A	115	GLY	5.0
1	A	34	SER	5.0
1	A	116	ALA	4.9
1	A	267	PHE	4.3
2	E	19	DG	3.7
1	A	64	LYS	3.6
1	A	266	ARG	3.6
1	A	58	ALA	3.5
1	A	62	GLY	3.4
1	A	268	GLY	3.2
2	E	4	DT	3.0
1	A	175	ASP	3.0
1	A	213	GLN	2.6
1	A	162	SER	2.6
1	A	208	ALA	2.6
1	A	56	MET	2.3
1	A	63	TRP	2.3
1	A	57	ASN	2.2
1	A	100	LYS	2.2
1	A	270	ASN	2.2
1	A	261	LYS	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	259	HIS	2.1
1	A	248	LYS	2.1
1	A	214	LEU	2.1
1	A	60	ASP	2.0

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 oligosaccharides in this entry.

6.4 Ligands (i)

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
3	ZN	A	401	1/1	0.93	0.11	53,53,53,53	0

6.5 Other polymers (i)

There are no such residues in this entry.