



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 01:59 PM UTC

PDB ID : 5ITT / pdb_00005itt
Title : Crystal Structure of Human NEIL1 bound to duplex DNA containing THF
Authors : Zhu, C.; Lu, L.; Zhang, J.; Yue, Z.; Song, J.; Zong, S.; Liu, M.; Stovicek, O.;
Gao, Y.; Yi, C.
Deposited on : 2016-03-17
Resolution : 2.53 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

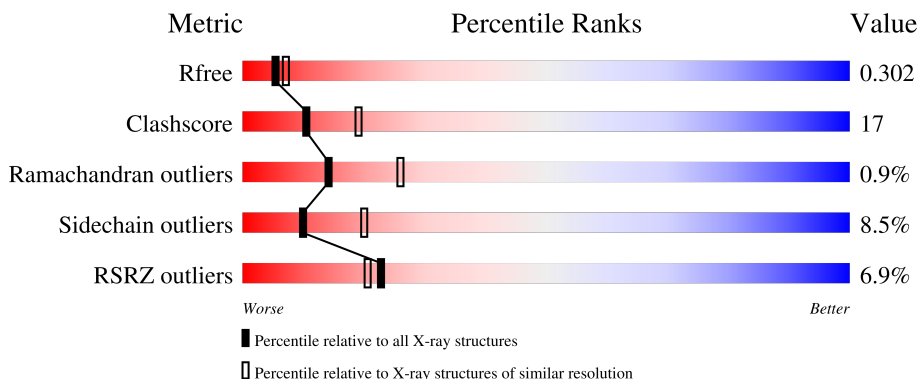
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.53 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}	180053	7383 (2.54-2.50)
Clashscore	190562	8079 (2.54-2.50)
Ramachandran outliers	187476	7944 (2.54-2.50)
Sidechain outliers	187428	7946 (2.54-2.50)
RSRZ outliers	180081	7387 (2.54-2.50)

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 $\leq 5\%$. 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	400	
1	B	400	
1	C	400	
2	D	26	
2	F	26	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	26	 50% 50%

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 8033 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 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	271	2143	1367	389	377	10	0	0	0
1	B	261	2076	1327	381	358	10	0	0	0
1	C	259	2074	1330	378	356	10	0	0	0

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	242	ARG	LYS	engineered mutation	UNP Q96FI4
A	391	ALA	-	expression tag	UNP Q96FI4
A	392	ALA	-	expression tag	UNP Q96FI4
A	393	LEU	-	expression tag	UNP Q96FI4
A	394	GLY	-	expression tag	UNP Q96FI4
A	395	HIS	-	expression tag	UNP Q96FI4
A	396	HIS	-	expression tag	UNP Q96FI4
A	397	HIS	-	expression tag	UNP Q96FI4
A	398	HIS	-	expression tag	UNP Q96FI4
A	399	HIS	-	expression tag	UNP Q96FI4
A	400	HIS	-	expression tag	UNP Q96FI4
B	242	ARG	LYS	engineered mutation	UNP Q96FI4
B	391	ALA	-	expression tag	UNP Q96FI4
B	392	ALA	-	expression tag	UNP Q96FI4
B	393	LEU	-	expression tag	UNP Q96FI4
B	394	GLY	-	expression tag	UNP Q96FI4
B	395	HIS	-	expression tag	UNP Q96FI4
B	396	HIS	-	expression tag	UNP Q96FI4
B	397	HIS	-	expression tag	UNP Q96FI4
B	398	HIS	-	expression tag	UNP Q96FI4
B	399	HIS	-	expression tag	UNP Q96FI4
B	400	HIS	-	expression tag	UNP Q96FI4
C	242	ARG	LYS	engineered mutation	UNP Q96FI4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	391	ALA	-	expression tag	UNP Q96FI4
C	392	ALA	-	expression tag	UNP Q96FI4
C	393	LEU	-	expression tag	UNP Q96FI4
C	394	GLY	-	expression tag	UNP Q96FI4
C	395	HIS	-	expression tag	UNP Q96FI4
C	396	HIS	-	expression tag	UNP Q96FI4
C	397	HIS	-	expression tag	UNP Q96FI4
C	398	HIS	-	expression tag	UNP Q96FI4
C	399	HIS	-	expression tag	UNP Q96FI4
C	400	HIS	-	expression tag	UNP Q96FI4

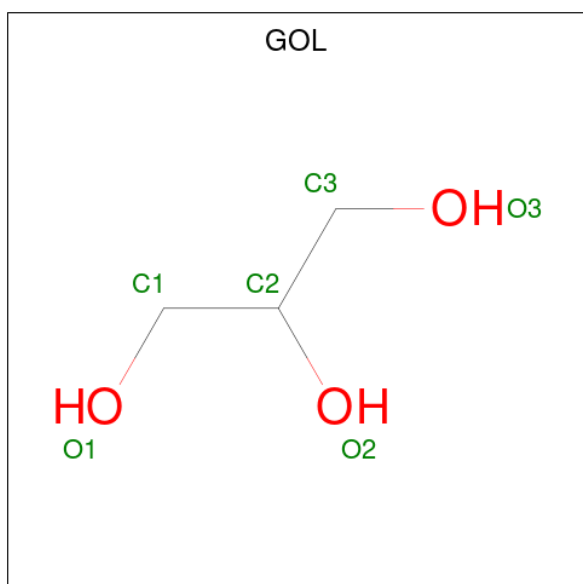
- Molecule 2 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
2	D	26	Total	C	N	O	P	0	0	0
			516	247	94	151	24			
2	F	26	Total	C	N	O	P	0	0	0
			516	247	94	151	24			

- Molecule 3 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
3	E	26	Total	C	N	O	P	0	0	0
			514	247	94	149	24			

- Molecule 4 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	108	Total	O	0	0
			108	108		
5	B	48	Total	O	0	0
			48	48		
5	C	4	Total	O	0	0
			4	4		
5	D	14	Total	O	0	0
			14	14		
5	E	6	Total	O	0	0
			6	6		
5	F	2	Total	O	0	0
			2	2		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	73.86Å 108.75Å 171.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	81.81 – 2.53 81.81 – 2.53	Depositor EDS
% Data completeness (in resolution range)	97.9 (81.81-2.53) 98.0 (81.81-2.53)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.19 (at 2.54Å)	Xtrriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.210 , 0.278 0.266 , 0.302	Depositor DCC
R_{free} test set	2318 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	56.2	Xtrriage
Anisotropy	0.418	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 21.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8033	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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: GOL

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	1.41	7/2202 (0.3%)	1.34	18/2979 (0.6%)
1	B	1.21	3/2133 (0.1%)	1.27	18/2885 (0.6%)
1	C	0.74	3/2129 (0.1%)	0.98	10/2878 (0.3%)
2	D	0.46	0/577	0.93	0/886
2	F	0.38	0/577	0.86	3/886 (0.3%)
3	E	0.51	0/575	1.03	2/882 (0.2%)
All	All	1.05	13/8193 (0.2%)	1.15	51/11396 (0.4%)

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	2
1	B	0	1
All	All	0	3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	113	CYS	CA-C	-6.39	1.44	1.52
1	B	121	GLY	C-O	-5.72	1.16	1.23
1	B	139	GLN	CA-C	-5.62	1.45	1.52
1	B	223	PRO	N-CD	5.59	1.55	1.47
1	C	61	SER	C-N	5.43	1.40	1.33

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	105	PRO	N-CA-C	12.88	126.42	110.70
1	A	250	GLU	CA-C-N	-8.78	107.56	120.38
1	A	250	GLU	C-N-CA	-8.78	107.56	120.38
1	A	67	GLN	C-N-CD	8.39	139.06	120.60
1	C	67	GLN	C-N-CD	8.36	138.99	120.60

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	249	GLY	Peptide
1	A	94	PRO	Peptide
1	B	67	GLN	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	2143	0	2103	85	0
1	B	2076	0	2061	53	0
1	C	2074	0	2064	88	0
2	D	516	0	290	7	0
2	F	516	0	290	23	0
3	E	514	0	290	8	0
4	A	6	0	8	1	0
4	B	6	0	8	0	0
5	A	108	0	0	4	0
5	B	48	0	0	2	0
5	C	4	0	0	0	0
5	D	14	0	0	0	0
5	E	6	0	0	0	0
5	F	2	0	0	0	0
All	All	8033	0	7114	252	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 17.

The worst 5 of 252 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:199:GLU:OE1	1:B:202:GLN:NE2	1.57	1.35
1:A:250:GLU:HG3	1:A:253:PHE:HD1	1.03	1.16
1:A:250:GLU:HG3	1:A:253:PHE:CD1	1.85	1.11
1:A:250:GLU:HB2	1:A:253:PHE:HB3	1.08	1.08
1:C:282:GLN:OE1	1:C:284:ASP:N	1.88	1.07

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	267/400 (67%)	254 (95%)	13 (5%)	0	100	100
1	B	255/400 (64%)	228 (89%)	25 (10%)	2 (1%)	16	29
1	C	249/400 (62%)	215 (86%)	29 (12%)	5 (2%)	6	9
All	All	771/1200 (64%)	697 (90%)	67 (9%)	7 (1%)	14	26

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	68	PRO
1	C	199	GLU
1	C	167	ASP
1	C	198	LEU
1	C	262	CYS

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	222/335 (66%)	206 (93%)	16 (7%)	13	26
1	B	217/335 (65%)	203 (94%)	14 (6%)	15	30
1	C	219/335 (65%)	193 (88%)	26 (12%)	5	9
All	All	658/1005 (66%)	602 (92%)	56 (8%)	10	20

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	270	SER
1	C	282	GLN
1	C	46	ARG
1	C	281	PHE
1	C	219	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	130	GLN
1	C	222	ASN
1	B	143	GLN
1	C	272	GLN
1	C	147	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

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
4	GOL	B	501	-	5,5,5	0.55	0	5,5,5	0.57	0
4	GOL	A	501	-	5,5,5	0.99	0	5,5,5	1.05	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	B	501	-	-	1/4/4/4	-
4	GOL	A	501	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	501	GOL	C1-C2-C3-O3
4	A	501	GOL	O2-C2-C3-O3
4	A	501	GOL	O1-C1-C2-O2
4	B	501	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	501	GOL	1	0

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	271/400 (67%)	-0.03	5 (1%) 67 64	2, 2, 17, 39	0
1	B	261/400 (65%)	0.44	10 (3%) 44 40	2, 12, 29, 42	0
1	C	259/400 (64%)	1.24	44 (16%) 4 3	29, 54, 74, 93	0
2	D	26/26 (100%)	0.00	1 (3%) 44 40	4, 25, 48, 53	0
2	F	26/26 (100%)	0.73	0 100 100	30, 55, 64, 68	0
3	E	26/26 (100%)	-0.04	0 100 100	8, 26, 36, 39	0
All	All	869/1278 (67%)	0.51	60 (6%) 23 20	2, 15, 65, 93	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	264	GLY	6.2
1	C	22	VAL	4.6
1	C	263	TYR	4.4
1	B	200	ALA	4.3
1	A	248	SER	3.9

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, 95th 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
4	GOL	B	501	6/6	0.80	0.38	64,73,79,79	0
4	GOL	A	501	6/6	0.82	0.31	58,61,63,78	0

6.5 Other polymers [i](#)

There are no such residues in this entry.