Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2022 – 03:05 pm GMT

PDB ID	:	7OL9
Title	:	Crystal structure of C-terminally truncated Bacillus subtilis nucleoid occlusion protein (Noc) complexed to the Noc-binding site (NBS)
Authors	:	Jalal, A.S.B.; Lawson, D.M.; Le, T.B.K.
Deposited on	:	2021-05-19
Resolution	:	2.90 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 following versions of software and data (see references ⓘ) were used in the production of this report:

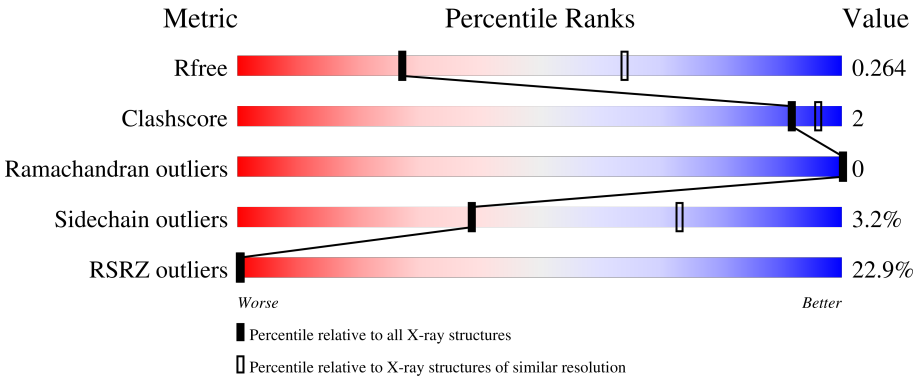
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.27
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 2.90 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	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	255	11% 73% 6% 21%
1	B	255	27% 73% 7% 20%
2	C	16	6% 100%
2	D	16	6% 94% 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3843 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 Nucleoid occlusion protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	201	Total	C	N	O	S	0	0	0
			1574	995	280	297	2			
1	B	203	Total	C	N	O	S	0	0	0
			1619	1020	287	309	3			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	243	LYS	-	expression tag	UNP P37524
A	244	LEU	-	expression tag	UNP P37524
A	245	ALA	-	expression tag	UNP P37524
A	246	ALA	-	expression tag	UNP P37524
A	247	ALA	-	expression tag	UNP P37524
A	248	LEU	-	expression tag	UNP P37524
A	249	GLU	-	expression tag	UNP P37524
A	250	HIS	-	expression tag	UNP P37524
A	251	HIS	-	expression tag	UNP P37524
A	252	HIS	-	expression tag	UNP P37524
A	253	HIS	-	expression tag	UNP P37524
A	254	HIS	-	expression tag	UNP P37524
A	255	HIS	-	expression tag	UNP P37524
B	243	LYS	-	expression tag	UNP P37524
B	244	LEU	-	expression tag	UNP P37524
B	245	ALA	-	expression tag	UNP P37524
B	246	ALA	-	expression tag	UNP P37524
B	247	ALA	-	expression tag	UNP P37524
B	248	LEU	-	expression tag	UNP P37524
B	249	GLU	-	expression tag	UNP P37524
B	250	HIS	-	expression tag	UNP P37524
B	251	HIS	-	expression tag	UNP P37524
B	252	HIS	-	expression tag	UNP P37524
B	253	HIS	-	expression tag	UNP P37524
B	254	HIS	-	expression tag	UNP P37524

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	255	HIS	-	expression tag	UNP P37524

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

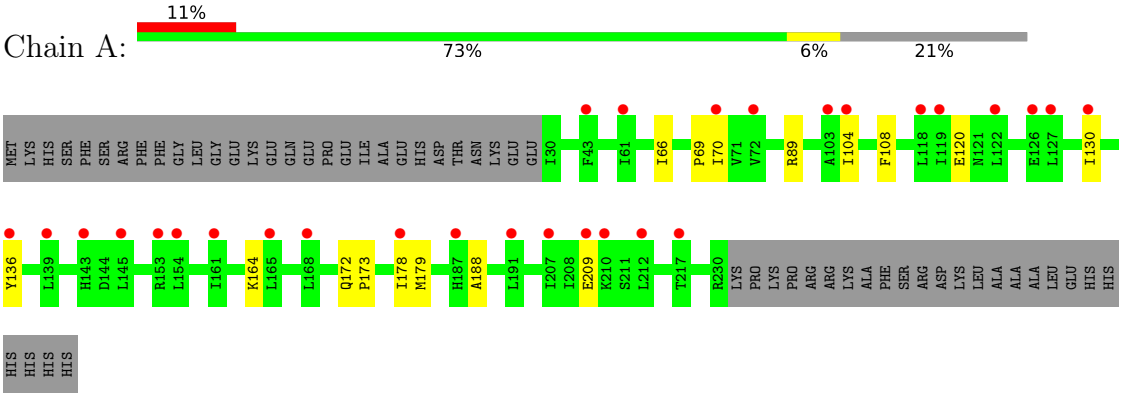
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	16	Total	C	N	O	P	0	0	0
			325	157	59	94	15			
2	D	16	Total	C	N	O	P	0	0	0
			325	157	59	94	15			

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.

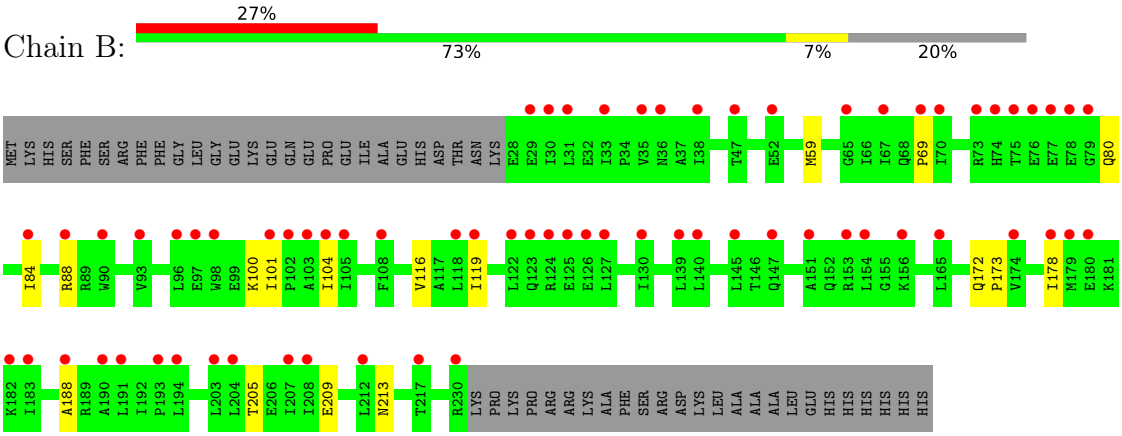
• Molecule 1: Nucleoid occlusion protein

Chain A:



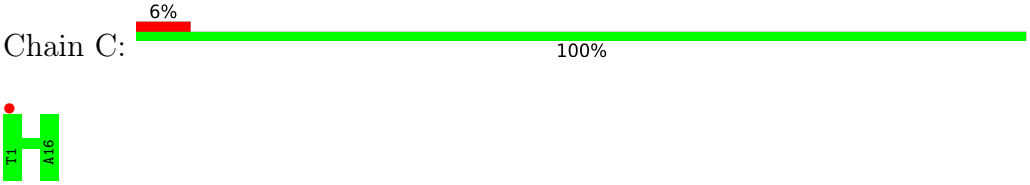
• Molecule 1: Nucleoid occlusion protein

Chain B:



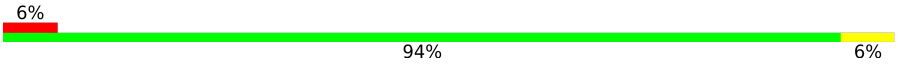
• Molecule 2: DNA (5'-D(*TP*AP*TP*TP*TP*CP*CP*CP*GP*GP*GP*AP*AP*AP*TP*A)-3')

Chain C:



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

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	70.50Å 99.33Å 99.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	70.36 – 2.90 70.27 – 2.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (70.36-2.90) 100.0 (70.27-2.90)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.230 , 0.278 0.231 , 0.264	Depositor DCC
R_{free} test set	828 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å^2)	88.2	Xtriage
Anisotropy	0.672	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.002 for -h,l,k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3843	wwPDB-VP
Average B, all atoms (Å^2)	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.55% 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	0.66	0/1596	0.71	0/2167
1	B	0.65	0/1641	0.72	0/2222
2	C	0.43	0/364	0.83	0/560
2	D	0.42	0/364	0.82	0/560
All	All	0.62	0/3965	0.74	0/5509

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

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	1574	0	1590	6	0
1	B	1619	0	1657	6	0
2	C	325	0	183	0	0
2	D	325	0	183	1	0
All	All	3843	0	3613	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 2.

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:130:ILE:HD11	1:A:179:MET:HB2	1.97	0.46
1:B:172:GLN:N	1:B:173:PRO:CD	2.79	0.46
1:B:178:ILE:HD11	1:B:188:ALA:CB	2.45	0.46
1:B:69:PRO:HG2	1:B:104:ILE:HD11	1.99	0.44
1:A:69:PRO:HG2	1:A:104:ILE:HD11	1.98	0.44
1:A:172:GLN:N	1:A:173:PRO:CD	2.80	0.44
1:B:213:ASN:HB2	2:D:9:DG:OP2	2.17	0.44
1:A:178:ILE:HD11	1:A:188:ALA:CB	2.48	0.44
1:A:70:ILE:HA	1:A:89:ARG:HH12	1.84	0.43
1:A:136:TYR:CE2	1:A:164:LYS:HD2	2.55	0.41
1:B:116:VAL:HA	1:B:119:ILE:HG12	2.02	0.41
1:B:84:ILE:HG22	1:B:116:VAL:HG11	2.03	0.41

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	199/255 (78%)	193 (97%)	6 (3%)	0	100	100
1	B	201/255 (79%)	193 (96%)	8 (4%)	0	100	100
All	All	400/510 (78%)	386 (96%)	14 (4%)	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	166/226 (74%)	162 (98%)	4 (2%)	49 79
1	B	176/226 (78%)	169 (96%)	7 (4%)	31 65
All	All	342/452 (76%)	331 (97%)	11 (3%)	39 73

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66	ILE
1	A	108	PHE
1	A	120	GLU
1	A	209	GLU
1	B	59	MET
1	B	80	GLN
1	B	88	ARG
1	B	100	LYS
1	B	101	ILE
1	B	205	THR
1	B	209	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	80	GLN
1	A	172	GLN
1	B	123	GLN

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 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, 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	201/255 (78%)	0.73	29 (14%) 2 2	82, 126, 164, 176	0
1	B	203/255 (79%)	1.42	69 (33%) 0 0	75, 114, 161, 194	0
2	C	16/16 (100%)	0.43	1 (6%) 20 16	75, 81, 93, 124	0
2	D	16/16 (100%)	0.35	1 (6%) 20 16	75, 82, 90, 150	0
All	All	436/542 (80%)	1.03	100 (22%) 0 0	75, 119, 163, 194	0

All (100) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	43	PHE	6.2
2	C	1	DT	4.9
1	A	139	LEU	4.5
2	D	16	DA	4.4
1	B	76	GLU	4.2
1	B	101	ILE	4.2
1	A	122	LEU	4.0
1	B	125	GLU	3.9
1	A	103	ALA	3.9
1	B	88	ARG	3.8
1	B	130	ILE	3.7
1	B	98	TRP	3.7
1	B	105	ILE	3.6
1	B	97	GLU	3.6
1	B	126	GLU	3.6
1	B	74	HIS	3.4
1	B	154	LEU	3.3
1	B	119	ILE	3.3
1	A	210	LYS	3.2
1	B	52	GLU	3.2
1	B	203	LEU	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	104	ILE	3.1
1	A	119	ILE	3.1
1	A	207	ILE	3.1
1	B	230	ARG	3.0
1	B	29	GLU	3.0
1	B	30	ILE	3.0
1	B	102	PRO	3.0
1	A	191	LEU	3.0
1	B	78	GLU	3.0
1	B	207	ILE	2.9
1	B	122	LEU	2.9
1	B	33	ILE	2.9
1	B	179	MET	2.9
1	B	127	LEU	2.8
1	B	65	GLY	2.8
1	B	191	LEU	2.8
1	B	75	THR	2.8
1	B	108	PHE	2.8
1	B	96	LEU	2.8
1	B	139	LEU	2.8
1	B	35	VAL	2.7
1	B	77	GLU	2.7
1	A	127	LEU	2.7
1	A	154	LEU	2.7
1	B	103	ALA	2.7
1	B	151	ALA	2.7
1	A	70	ILE	2.7
1	A	143	HIS	2.7
1	B	84	ILE	2.7
1	B	36	ASN	2.7
1	B	104	ILE	2.6
1	B	194	LEU	2.6
1	B	147	GLN	2.6
1	B	208	ILE	2.6
1	A	136	TYR	2.6
1	B	174	VAL	2.5
1	B	31	LEU	2.5
1	B	217	THR	2.5
1	A	153	ARG	2.5
1	A	165	LEU	2.5
1	B	165	LEU	2.5
1	B	212	LEU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	70	ILE	2.5
1	B	190	ALA	2.4
1	A	212	LEU	2.4
1	B	188	ALA	2.4
1	B	156	LYS	2.4
1	B	193	PRO	2.4
1	B	123	GLN	2.4
1	B	118	LEU	2.4
1	B	67	ILE	2.4
1	A	217	THR	2.3
1	B	178	ILE	2.3
1	B	183	ILE	2.3
1	A	72	VAL	2.3
1	A	168	LEU	2.3
1	B	79	GLY	2.3
1	A	209	GLU	2.3
1	A	61	ILE	2.3
1	A	161	ILE	2.2
1	B	124	ARG	2.2
1	B	38	ILE	2.2
1	B	69	PRO	2.2
1	A	145	LEU	2.2
1	B	90	TRP	2.2
1	A	118	LEU	2.1
1	B	153	ARG	2.1
1	A	178	ILE	2.1
1	A	126	GLU	2.1
1	B	204	LEU	2.1
1	B	140	LEU	2.1
1	B	180	GLU	2.1
1	B	182	LYS	2.1
1	A	130	ILE	2.1
1	B	73	ARG	2.1
1	B	47	THR	2.0
1	A	187	HIS	2.0
1	B	93	VAL	2.0
1	B	145	LEU	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 monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.