



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:47 PM GMT

PDB ID : 1M1G
Title : Crystal Structure of Aquifex aeolicus N-utilization substance G (NusG), Space Group P2(1)
Authors : Steiner, T.; Kaiser, J.T.; Marinkovic, S.; Huber, R.; Wahl, M.C.
Deposited on : 2002-06-19
Resolution : 2.00 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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
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

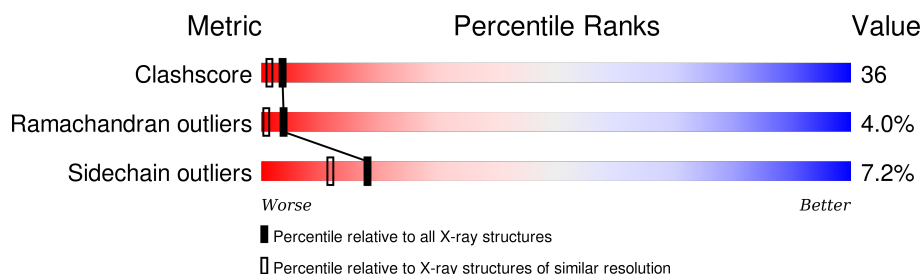
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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))
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)

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 $\leq 5\%$. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	248	
1	B	248	
1	C	248	
1	D	248	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8718 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 Transcription antitermination protein nusG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	240	Total	C	N	O	S	0	0	0
			1899	1214	323	354	8			
1	B	242	Total	C	N	O	S	0	0	0
			1917	1224	326	359	8			
1	C	244	Total	C	N	O	S	0	0	0
			1933	1234	329	362	8			
1	D	239	Total	C	N	O	S	0	0	0
			1891	1208	322	353	8			

- Molecule 2 is water.

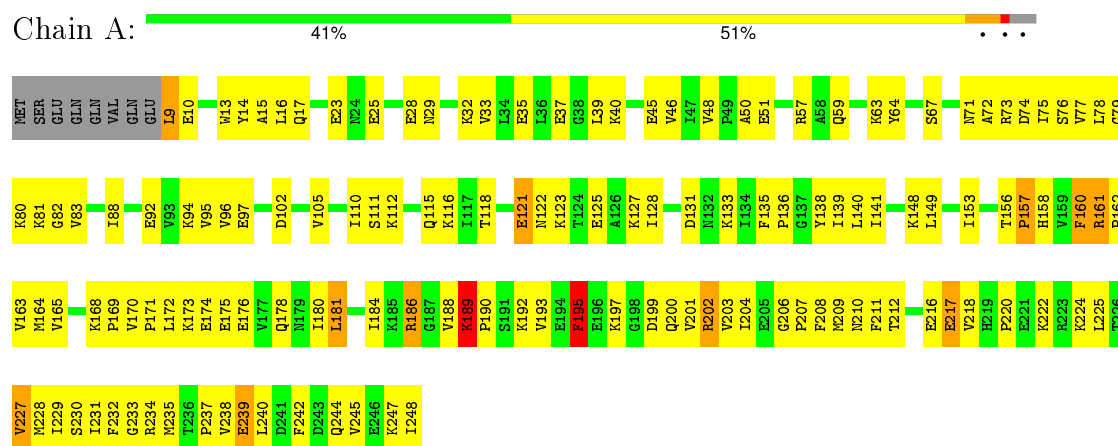
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	220	Total	O	0	0
			220	220		
2	B	338	Total	O	0	0
			338	338		
2	C	300	Total	O	0	0
			300	300		
2	D	220	Total	O	0	0
			220	220		

3 Residue-property plots

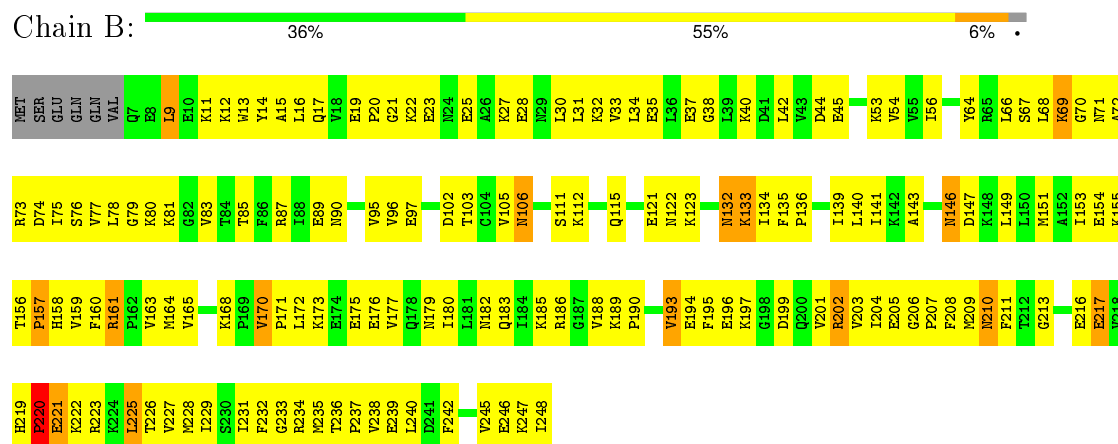
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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 was not executed.

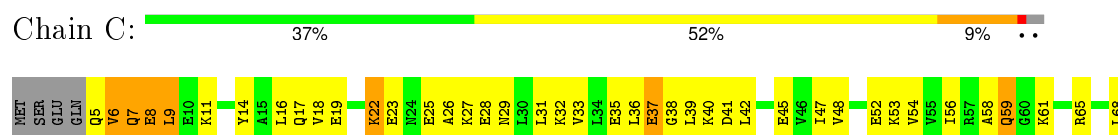
- Molecule 1: Transcription antitermination protein nusG

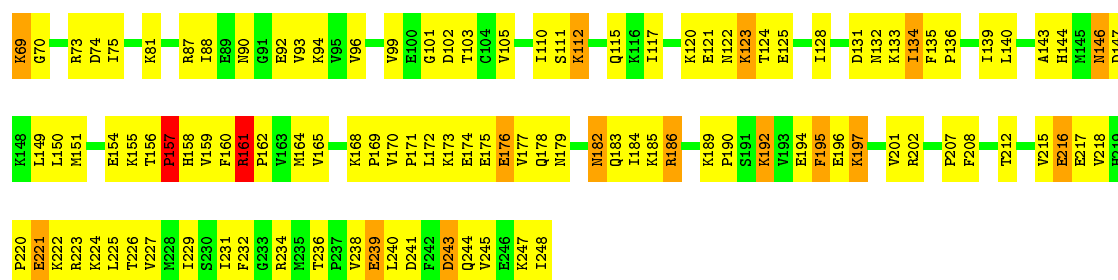


- Molecule 1: Transcription antitermination protein nusG



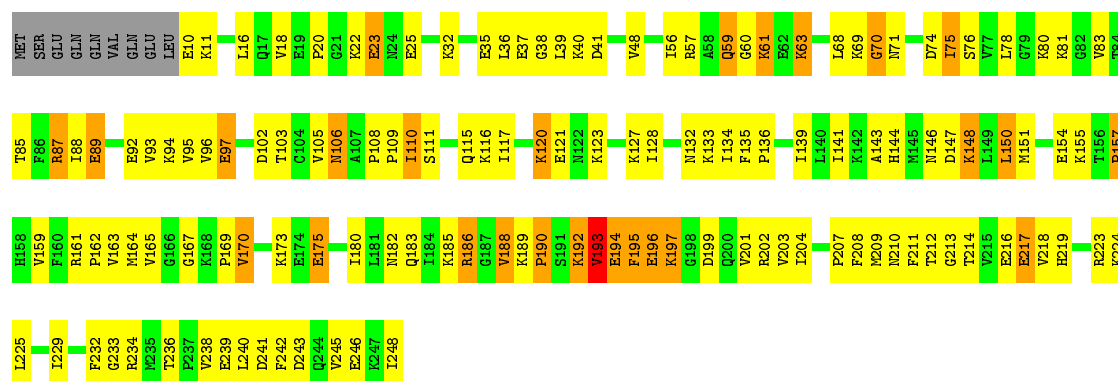
- Molecule 1: Transcription antitermination protein nusG





• Molecule 1: Transcription antitermination protein nusG

Chain D: 42% 44% 10%



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	87.50 Å 54.40 Å 113.20 Å 90.00° 89.30° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00	Depositor
% Data completeness (in resolution range)	94.9 (30.00-2.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.251 , 0.274	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8718	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.38	0/1928	0.64	0/2592
1	B	0.38	0/1946	0.65	1/2616 (0.0%)
1	C	0.38	0/1962	0.64	0/2638
1	D	0.38	0/1920	0.67	0/2581
All	All	0.38	0/7756	0.65	1/10427 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	225	LEU	CA-CB-CG	5.25	127.37	115.30

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	1899	0	1994	105	322
1	B	1917	0	2008	157	388
1	C	1933	0	2025	159	134
1	D	1891	0	1983	140	102
2	A	220	0	0	28	50
2	B	338	0	0	50	69
2	C	300	0	0	48	42
2	D	220	0	0	45	9

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	8718	0	8010	561	598

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 36.

The worst 5 of 561 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:207:PRO:CG	2:D:365:HOH:O	1.84	1.24
1:B:235:MET:HB3	2:B:267:HOH:O	1.09	1.23
1:D:212:THR:C	2:D:320:HOH:O	1.75	1.22
1:D:213:GLY:N	2:D:320:HOH:O	1.74	1.18
1:D:217:GLU:HB2	2:D:423:HOH:O	1.44	1.15

The worst 5 of 598 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:165:VAL:CB	1:B:238:VAL:CG1[2_757]	0.11	2.09
1:D:195:PHE:CD2	2:A:260:HOH:O[2_646]	0.20	2.00
1:B:207:PRO:C	2:A:258:HOH:O[2_747]	0.44	1.76
1:C:29:ASN:OD1	1:C:178:GLN:OE1[2_756]	0.45	1.75
1:B:161:ARG:NH1	2:B:293:HOH:O[2_747]	0.47	1.73

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	238/248 (96%)	214 (90%)	18 (8%)	6 (2%)	7 2
1	B	240/248 (97%)	206 (86%)	27 (11%)	7 (3%)	6 2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	242/248 (98%)	208 (86%)	22 (9%)	12 (5%)	3	0
1	D	237/248 (96%)	205 (86%)	19 (8%)	13 (6%)	2	0
All	All	957/992 (96%)	833 (87%)	86 (9%)	38 (4%)	4	1

5 of 38 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	189	LYS
1	B	193	VAL
1	B	220	PRO
1	C	8	GLU
1	C	157	PRO

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	214/222 (96%)	201 (94%)	13 (6%)	23	17
1	B	216/222 (97%)	205 (95%)	11 (5%)	29	23
1	C	218/222 (98%)	199 (91%)	19 (9%)	13	7
1	D	213/222 (96%)	194 (91%)	19 (9%)	12	7
All	All	861/888 (97%)	799 (93%)	62 (7%)	18	12

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	69	LYS
1	C	161	ARG
1	D	175	GLU
1	C	132	ASN
1	C	182	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	182	ASN
1	C	106	ASN
1	D	178	GLN
1	C	29	ASN
1	A	158	HIS

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](#)

There are no ligands in this entry.

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 was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.