

# wwPDB X-ray Structure Validation Summary Report (i)

### Jan 13, 2024 - 06:48 pm GMT

PDB ID	:	6RUT
Title	:	Mycoplasma Genitalium Heterodimer Nap Complex (P140-P110 globular)
Authors	:	Fita, I.; Aparicio, D.
Deposited on	:	2019-05-29
Resolution	:	2.65  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 (i) 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 (1)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber $(2001)$
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 2.65 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))				
Rfree	130704	1332 (2.68-2.64)				
Clashscore	141614	1374 (2.68-2.64)				
Ramachandran outliers	138981	1349 (2.68-2.64)				
Sidechain outliers	138945	1349 (2.68-2.64)				
RSRZ outliers	127900	1318 (2.68-2.64)				

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	А	802	% • 79%	16%	
1	C	200	% •	1070	
	C	802	77%	17%	••
1	Е	802	79%	16%	••
1	G	802	78%	17%	••
2	В	1334	69%	22%	• 5%

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain		
2	D	1334	% 69%	22%	• 5%
2	F	1334	3% 67%	22%	• 6%
2	Н	1334	3% 67%	23%	• 7%



# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 64181 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.

Mol	Chain	Residues		A	toms			ZeroOcc	AltConf	Trace
1	Λ	781	Total	С	Ν	Ο	$\mathbf{S}$	0	0	0
1	Л	101	6013	3758	1009	1240	6	0	0	0
1	С	770	Total	С	Ν	Ο	S	0	0	0
1		119	5999	3751	1006	1236	6	0	0	0
1	F	792	Total	С	Ν	Ο	S	0	0	0
	103	6017	3760	1010	1241	6	0	0	0	
1	1 C	794	Total	С	Ν	Ο	S	0	0	0
I G	104	6023	3763	1011	1243	6	0	0	0	

• Molecule 1 is a protein called Mgp-operon protein 3.

• Molecule 2 is a protein called Adhesin P1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	В	1968	Total	С	Ν	Ο	$\mathbf{S}$	0	0	0
	D	1208	9938	6303	1679	1942	14	0	0	
9	Л	1965	Total	С	Ν	Ο	S	0	0	0
		1205	9918	6293	1676	1935	14	0	0	
0	Б	1950	Total	С	Ν	Ο	S	0	0	0
	1230	9822	6242	1658	1909	13	0	0		
9	о п	1947	Total	С	Ν	Ο	S	0	0	0
	1247	9798	6228	1654	1903	13		0	U	

• Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	62	$\begin{array}{cc} \text{Total} & \text{O} \\ 62 & 62 \end{array}$	0	0
3	В	121	Total O   121 121	0	0
3	С	87	Total O 87 87	0	0
3	D	137	Total O 137 137	0	0

Continued on next page...



Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	Е	69	Total O 69 69	0	0
3	F	70	TotalO7070	0	0
3	G	49	TotalO4949	0	0
3	Н	58	Total O 58 58	0	0



# 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: Mgp-operon protein 3



# P681 5416 M591 K421 M591 K421 M591 K440 M599 K440 M599 K440 M599 K440 M599 K440 M639 M454 M639 M454 M639 M454 M679 M454 M679 M454 M679 M454 M674 M454 M674 M454 M674 M454 M674 M454 M674 M456 M674 M456 M674 M456 M674 M456 M674 M436 M674</t

# 





### HIS HIS HIS

• Molecule 2: Adhesin P1

Chain	B:				6	9%							22%		•	5%	
GLY VAL ILE THR	GLY VAL GLY GLY	PHE LEU PHE ASN	GLN ASN LYS GLN	ARG SER SER	VAL SER ASN	ALA ALA TYR	GLN PRO LYS	GLN LEU SER	VAL LYS HIS	GLN Q59	E6 <mark>3</mark> T64 L65	E83	F88 Vq1	L98 Mgg	K105	SER SER S108 D109	D110 N111 L112 K113
Y114 L115 N116 A117	Q125 Q126 N127	1130 R131 <mark>R132</mark> Y133	R138 D142	1143 M148	E149 N150 P151	T153 V154	L172	W180 K184	<mark>G185</mark> K186 V187	S194	N1 <mark>99</mark> L200 Y201	F202 V203 L204	L205	q219	E223	LZZ V227 LYS ALA	THR GLN S232
N235 P236 Q238 Q238	R239 L240 Q241 K242 D243	S244 L256	T259 S262 S263	M264 S265 S266	G267 M268 A269	1270 8271 T272	A276 L277	K278 V279 E280	V281 E282 R283	<mark>գշ86</mark>	S2 <mark>89</mark> L290 L291	K292 N293	Р2 <mark>99</mark> L300 КЗ01	T318	L324	1328 D328	K334 V341
Y344 D345 E350	N351 H352 V358	D362 K365 M366	V367 1384	Y387 N391	L393	1390 E411	N418 T419	S420 P421 G422	K431 K432	K437 N438	S439 E448	N453	M457 V458	F465 I466 F467	K475	n470 F477 T478 T479	L523 P524
T527 N528 R529 W530	R536 V542	L549 M557	T562 R565 L566	K567 Y568 D569	ub/0 L571 E572	H574	L584 L585 R586	E587	(12) (12) (12) (12) (12) (12) (12) (12)	L609 Q610 D611	G614	G618 P619 H620	Y621	V632	E638	S643	V653 Y654
P655	N668 W669 L673	L678 S679 A680 N681	E685 N686	L693 F694 A695	A696 1697 L698	D701	5707 D708 K709	I710 R725	<mark>Q728</mark>	L732 N733 <mark>P734</mark>	N735 T738	N739 W740 A741	R742 R749	F750 T751	L756 D757 e750	D766	P769 W770 I771 G772
N773 <mark>G774</mark> K775 P776	F777 S782 PRO SFR	THR SER ALA SER	S7 89 S7 90	F796 S797	M806	180/ L811	R817 W818	1821	1828 Y834	R835 V836 Q837	0841 K842	N843 G844 I845	P846 F847 F848	U850 V850	P852 S853 Nee4	N00 <del>4</del> N855 S856 T857	D864
<mark>P870</mark> S871 L883	D892 N895	T900 N903	q906 R907	V921 L922 1923	N924 K925	r 933 N934	E938 Q939 K940	K943 T944	E945 T946 N947	E948 <mark>G949</mark> N950	L951 P952	E956 L965	L966	6970 F971	T974 N975	Т979 F985	K986 A987 D988
2993	L1002 N1003 V1004 T1005 S1006	q1007 D1008 L1012	D1017	W1 027 F1 028	R1029	T1039 G1040	Y1041 L1042 G1043	11044 T1045 L1046	Q1052	Q1059 P1060 W1061	T1062 K1076	S1077 L1078	K1085 M1089	SER SER THR	THR TYR	THR N1097	P1102 Q1105 L1106
Y1107 Q1108 P1109	V1112 N1122 K1123	P1127 N1136 M1137	L1143	T1157 A1158	SER SER GLN	GLY ASN N1164	N1165	Q1171 T1172	S1183 K1187	E1188	T1196	D1202 S1203 K1204	I1213	L1225	V1234 N1235 N1235	A1242	S1246 D1247 T1248 E1249
K1250 P1255 G1256	N1257 Q1258 11259 D1260 F1261	N1262 R1263 L1264 F1265	V1269 T1270 E1271	L1272 F1273 D1274	T1277	11281 Y1284	L1288	R1301	<mark>\$1308 V1309</mark>	E1310	R1317 L1318	T1325	11329 P1330 V1331	S1335	T1341 V1342	r 1344 01344 11349	ALA HIS HIS HIS
HIS NH NIH																	
• Mol	ecule 2	2: Adl	nesin	P1													
Chain	D:				6	9%							22%		•	5%	















# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	151.00Å 157.28Å 192.37Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $93.81^{\circ}$ $90.00^{\circ}$	Depositor
Bosolution (Å)	121.66 - 2.65	Depositor
Resolution (A)	150.67 - 2.65	EDS
% Data completeness	73.2 (121.66-2.65)	Depositor
(in resolution range)	$73.3\ (150.67-2.65)$	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.81 (at 2.65 \text{\AA})$	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
P. P.	0.187 , $0.224$	Depositor
$n, n_{free}$	0.199 , $0.229$	DCC
$R_{free}$ test set	9457 reflections $(4.99\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	33.5	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.33, 26.2	EDS
L-test for $twinning^2$	$ < L >=0.44, < L^2>=0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	64181	wwPDB-VP
Average B, all atoms $(Å^2)$	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 8.87% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical 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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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	Bo	nd lengths	Bond angles			
	Ullalli	RMSZ	# Z  > 5	RMSZ	# Z  > 5		
1	А	0.48	0/6136	0.74	0/8340		
1	С	0.48	0/6122	0.74	0/8321		
1	Ε	0.47	0/6141	0.74	0/8347		
1	G	0.48	0/6147	0.74	1/8355~(0.0%)		
2	В	0.50	0/10187	0.76	1/13863~(0.0%)		
2	D	0.49	0/10167	0.76	4/13836~(0.0%)		
2	F	0.46	0/10069	0.73	2/13701~(0.0%)		
2	Н	0.48	1/10044~(0.0%)	0.72	2/13668~(0.0%)		
All	All	0.48	1/65013~(0.0%)	0.74	10/88431~(0.0%)		

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	$\mathrm{Ideal}(\mathrm{\AA})$
2	Н	1115	TYR	CA-C	7.32	1.72	1.52

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms		$Observed(^{o})$	$Ideal(^{o})$
2	D	1165	ASN	C-N-CA	7.25	137.53	122.30
2	Н	989	SER	CA-CB-OG	-5.95	95.13	111.20
2	D	530	TRP	N-CA-C	-5.77	95.43	111.00
1	G	367	SER	C-N-CA	5.75	136.09	121.70
2	В	856	SER	C-N-CA	5.64	135.79	121.70

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



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	А	6013	0	5805	61	0
1	С	5999	0	5794	71	0
1	Ε	6017	0	5809	62	0
1	G	6023	0	5814	90	0
2	В	9938	0	9645	145	0
2	D	9918	0	9631	163	0
2	F	9822	0	9548	155	0
2	Н	9798	0	9529	199	0
3	А	62	0	0	0	0
3	В	121	0	0	2	0
3	С	87	0	0	0	0
3	D	137	0	0	5	0
3	Ε	69	0	0	0	0
3	F	70	0	0	1	0
3	G	49	0	0	0	0
3	Н	58	0	0	2	0
All	All	64181	0	61575	915	0

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.

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 7.

The worst 5 of 915 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:G:812:ASP:OD2	1:G:815:LEU:CD1	1.72	1.35
1:G:255:VAL:HG23	1:G:726:GLN:HB3	1.19	1.11
1:G:812:ASP:OD2	1:G:815:LEU:HD13	1.36	1.09
1:G:812:ASP:OD2	1:G:815:LEU:HD12	1.54	1.06
1:G:460:GLN:CA	2:H:812:ASN:HD21	1.68	1.05

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.



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	entiles
1	А	773/802~(96%)	717 (93%)	50~(6%)	6 (1%)	19	29
1	С	771/802~(96%)	715 (93%)	48 (6%)	8 (1%)	15	23
1	Е	777/802~(97%)	710 (91%)	60 (8%)	7 (1%)	17	26
1	G	778/802~(97%)	719 (92%)	50 (6%)	9 (1%)	13	19
2	В	1256/1334~(94%)	1162 (92%)	87 (7%)	7 (1%)	25	37
2	D	1253/1334~(94%)	1157 (92%)	84 (7%)	12 (1%)	15	23
2	F	1234/1334~(92%)	1139 (92%)	80 (6%)	15 (1%)	13	19
2	Н	1231/1334~(92%)	1125 (91%)	91 (7%)	15 (1%)	13	19
All	All	8073/8544 (94%)	7444 (92%)	550 (7%)	79 (1%)	15	23

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

5 of 79 Ramachandran outliers are listed below:

Mol	Chain	$\mathbf{Res}$	Type
1	А	366	THR
1	А	367	SER
1	С	368	THR
2	D	618	GLY
2	D	1157	THR

### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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	Pe	erce	entiles
1	А	685/701~(98%)	615~(90%)	70 (10%)		7	10
1	С	683/701~(97%)	616~(90%)	67~(10%)		8	11
1	Ε	684/701~(98%)	612~(90%)	72 (10%)		7	10
1	G	685/701~(98%)	624 (91%)	61 (9%)		9	14
2	В	$1112/1171 \ (95\%)$	955~(86%)	157 (14%)		3	4
2	D	$1109/1171 \ (95\%)$	949 (86%)	160 (14%)		3	3

Continued on next page...



Mol	Chain	Analysed	Rotameric	Outliers	Pe	rce	ntil	es
2	F	1097/1171~(94%)	938~(86%)	159 (14%)		3	3	
2	Н	1093/1171~(93%)	951 (87%)	142 (13%)		4	5	
All	All	7148/7488~(96%)	6260 (88%)	888 (12%)		4	6	

Continued from previous page...

5 of 888 residues with a non-rotameric sidechain are listed below:

Mol	Chain	$\mathbf{Res}$	Type
1	Е	103	SER
2	Н	1346	PHE
2	F	437	LYS
2	Н	1301	ARG
2	Н	629	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 188 such sidechains are listed below:

Mol	Chain	$\mathbf{Res}$	Type
1	Е	804	ASN
2	F	1311	ASN
2	F	352	HIS
2	F	684	ASN
1	G	183	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 (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)

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	$\mathbf{OWAB}(\mathbf{\mathring{A}}^2)$	Q<0.9
1	А	781/802~(97%)	-0.55	6 (0%) 86	85	8, 29, 71, 112	0
1	С	779/802~(97%)	-0.55	6 (0%) 86	85	9, 28, 73, 99	0
1	E	783/802~(97%)	-0.58	6 (0%) 86	85	7, 29, 68, 109	0
1	G	784/802~(97%)	-0.53	7 (0%) 84	83	8, 32, 69, 104	0
2	В	1268/1334 (95%)	-0.51	5 (0%) 92	93	7, 29, 76, 130	0
2	D	1265/1334~(94%)	-0.57	7 (0%) 89	89	6, 27, 73, 118	0
2	F	1250/1334~(93%)	-0.19	42 (3%) 45	41	9, 44, 103, 136	0
2	Н	1247/1334 (93%)	-0.03	46 (3%) 41	38	14, 55, 115, 149	0
All	All	8157/8544 (95%)	-0.41	125 (1%) 73	71	6, 34, 91, 149	0

The worst 5 of 125 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	Н	856	SER	7.1
1	А	367	SER	6.8
2	F	1265	PHE	6.1
2	Н	855	ASN	6.0
2	Н	1294	GLY	6.0

# 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.

