



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 22, 2024 – 11:46 PM EDT

PDB ID : 6S3U  
Title : Adhesin P140 from Mycoplasma Genitalium  
Authors : Fita, I.; Aparicio, D.  
Deposited on : 2019-06-26  
Resolution : 3.24 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](mailto: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	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

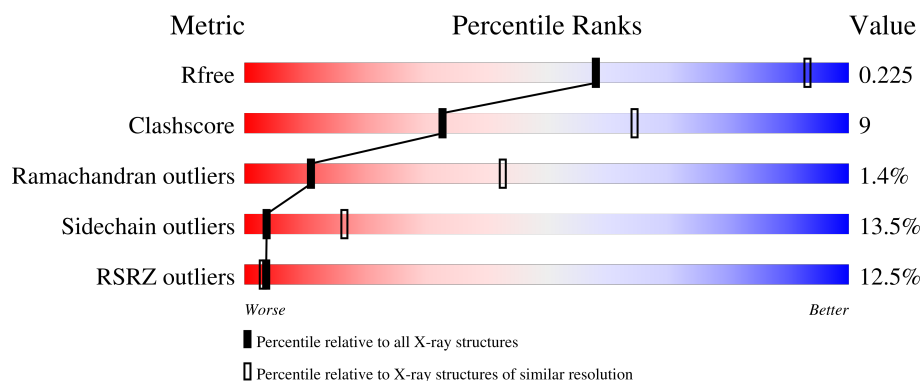
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.24 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	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)
RSRZ outliers	127900	1567 (3.28-3.20)

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  $\geq 3$ , 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	1304	<div> <div style="width: 2%;"></div> <div style="width: 68%; background-color: green;"></div> <div style="width: 26%; background-color: yellow;"></div> <div style="width: 4%; background-color: orange;"></div> <div style="width: 2%; background-color: red;"></div> <div style="width: 2%; background-color: grey;"></div> </div> <div>68% 26% . .</div>
1	B	1304	<div> <div style="width: 2%;"></div> <div style="width: 68%; background-color: green;"></div> <div style="width: 26%; background-color: yellow;"></div> <div style="width: 4%; background-color: orange;"></div> <div style="width: 2%; background-color: red;"></div> <div style="width: 2%; background-color: grey;"></div> </div> <div>68% 26% . .</div>
1	C	1304	<div> <div style="width: 2%;"></div> <div style="width: 70%; background-color: green;"></div> <div style="width: 25%; background-color: yellow;"></div> <div style="width: 4%; background-color: orange;"></div> <div style="width: 2%; background-color: red;"></div> <div style="width: 2%; background-color: grey;"></div> </div> <div>70% 25% . .</div>
1	D	1304	<div> <div style="width: 2%;"></div> <div style="width: 69%; background-color: green;"></div> <div style="width: 24%; background-color: yellow;"></div> <div style="width: 4%; background-color: orange;"></div> <div style="width: 2%; background-color: red;"></div> <div style="width: 2%; background-color: grey;"></div> </div> <div>69% 24% . .</div>
1	E	1304	<div> <div style="width: 34%; background-color: red;"></div> <div style="width: 69%; background-color: green;"></div> <div style="width: 24%; background-color: yellow;"></div> <div style="width: 4%; background-color: orange;"></div> <div style="width: 2%; background-color: red;"></div> <div style="width: 2%; background-color: grey;"></div> </div> <div>34% 69% 24% . .</div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	1304	<div> <div>33%</div> <div>69%</div> <div>24%</div> <div>• •</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 59767 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 Adhesin P1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1278	Total	C	N	O	S	0	0	0
			10016	6351	1692	1960	13			
1	B	1275	Total	C	N	O	S	0	0	0
			9999	6340	1689	1957	13			
1	C	1274	Total	C	N	O	S	0	0	0
			9993	6339	1688	1953	13			
1	D	1271	Total	C	N	O	S	0	0	0
			9976	6328	1685	1950	13			
1	E	1262	Total	C	N	O	S	0	0	0
			9886	6272	1666	1935	13			
1	F	1262	Total	C	N	O	S	0	0	0
			9897	6281	1668	1935	13			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1352	HIS	-	expression tag	UNP P20796
A	1353	HIS	-	expression tag	UNP P20796
A	1354	HIS	-	expression tag	UNP P20796
A	1355	HIS	-	expression tag	UNP P20796
A	1356	HIS	-	expression tag	UNP P20796
A	1357	HIS	-	expression tag	UNP P20796
B	1352	HIS	-	expression tag	UNP P20796
B	1353	HIS	-	expression tag	UNP P20796
B	1354	HIS	-	expression tag	UNP P20796
B	1355	HIS	-	expression tag	UNP P20796
B	1356	HIS	-	expression tag	UNP P20796
B	1357	HIS	-	expression tag	UNP P20796
C	1352	HIS	-	expression tag	UNP P20796
C	1353	HIS	-	expression tag	UNP P20796
C	1354	HIS	-	expression tag	UNP P20796
C	1355	HIS	-	expression tag	UNP P20796
C	1356	HIS	-	expression tag	UNP P20796

*Continued on next page...*

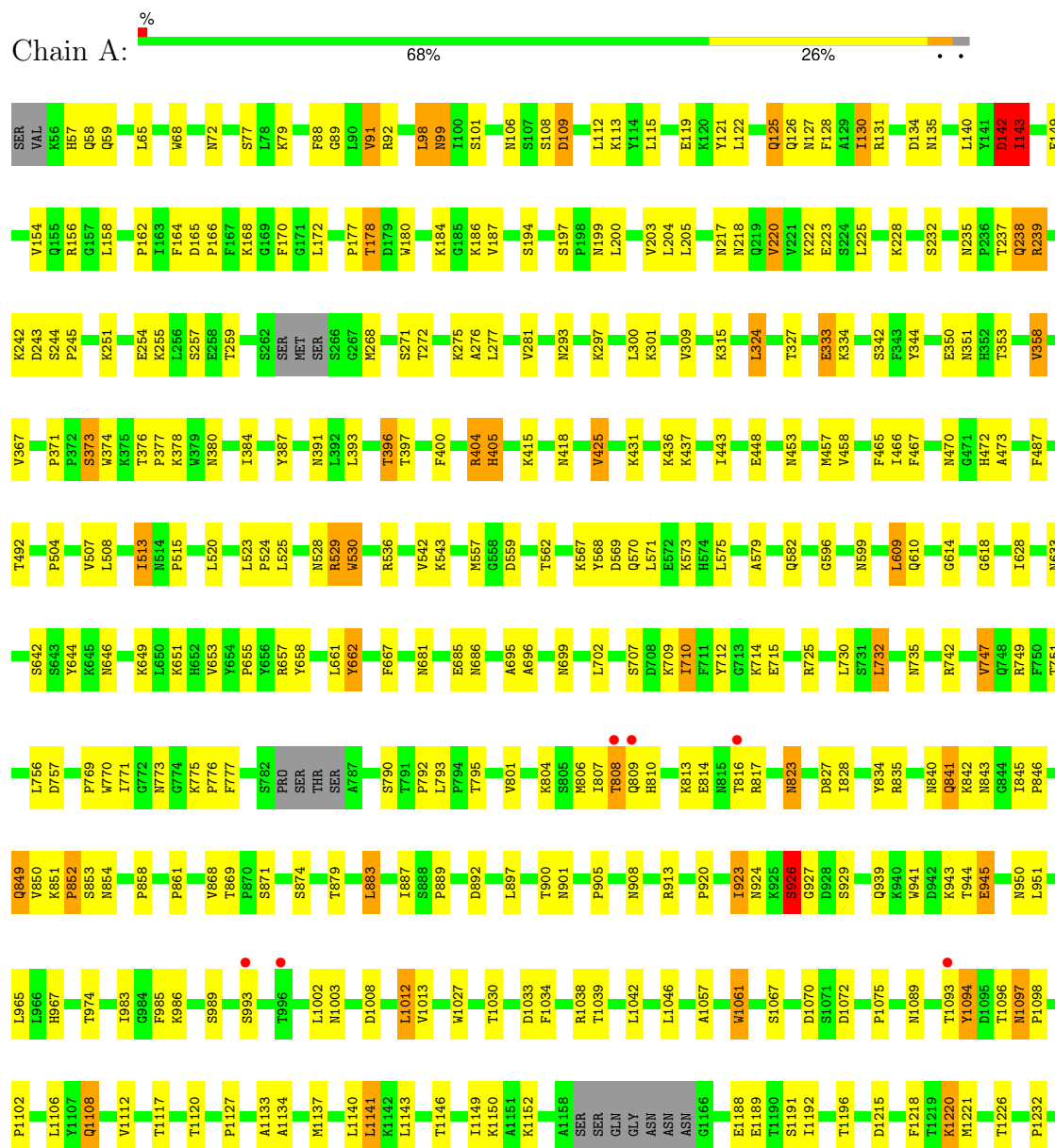
*Continued from previous page...*

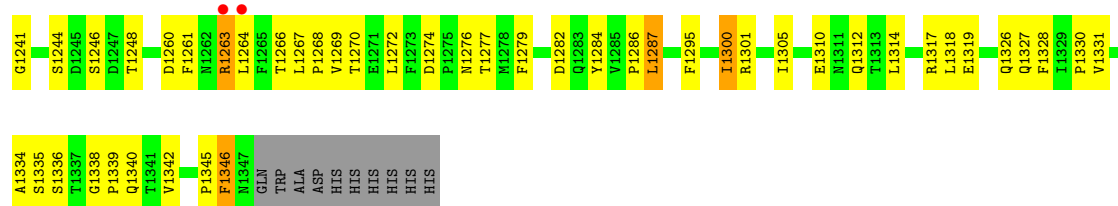
Chain	Residue	Modelled	Actual	Comment	Reference
C	1357	HIS	-	expression tag	UNP P20796
D	1352	HIS	-	expression tag	UNP P20796
D	1353	HIS	-	expression tag	UNP P20796
D	1354	HIS	-	expression tag	UNP P20796
D	1355	HIS	-	expression tag	UNP P20796
D	1356	HIS	-	expression tag	UNP P20796
D	1357	HIS	-	expression tag	UNP P20796
E	1352	HIS	-	expression tag	UNP P20796
E	1353	HIS	-	expression tag	UNP P20796
E	1354	HIS	-	expression tag	UNP P20796
E	1355	HIS	-	expression tag	UNP P20796
E	1356	HIS	-	expression tag	UNP P20796
E	1357	HIS	-	expression tag	UNP P20796
F	1352	HIS	-	expression tag	UNP P20796
F	1353	HIS	-	expression tag	UNP P20796
F	1354	HIS	-	expression tag	UNP P20796
F	1355	HIS	-	expression tag	UNP P20796
F	1356	HIS	-	expression tag	UNP P20796
F	1357	HIS	-	expression tag	UNP P20796

### 3 Residue-property plots

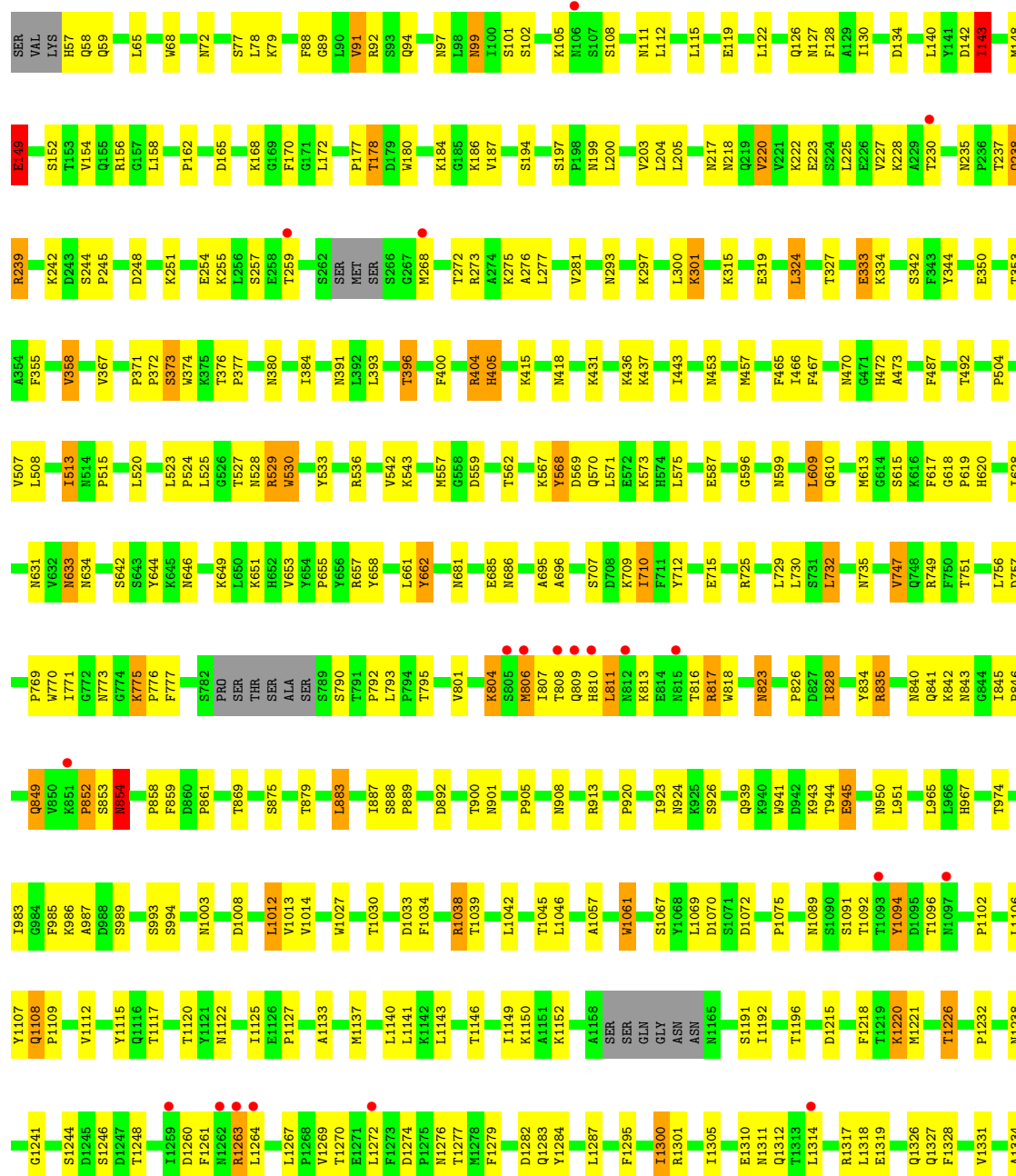
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Adhesin P1





• Molecule 1: Adhesin P1

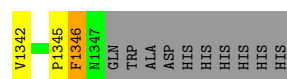




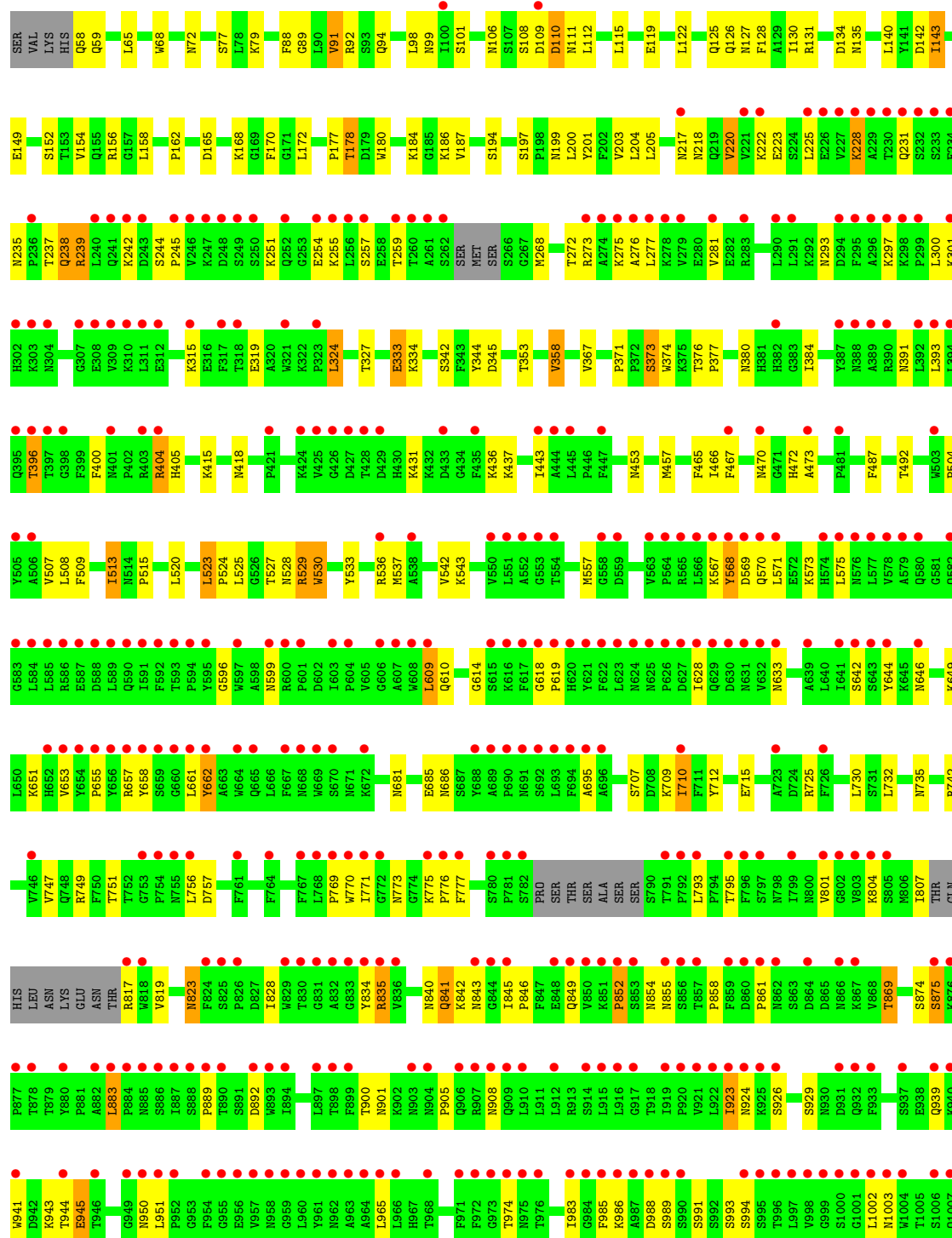


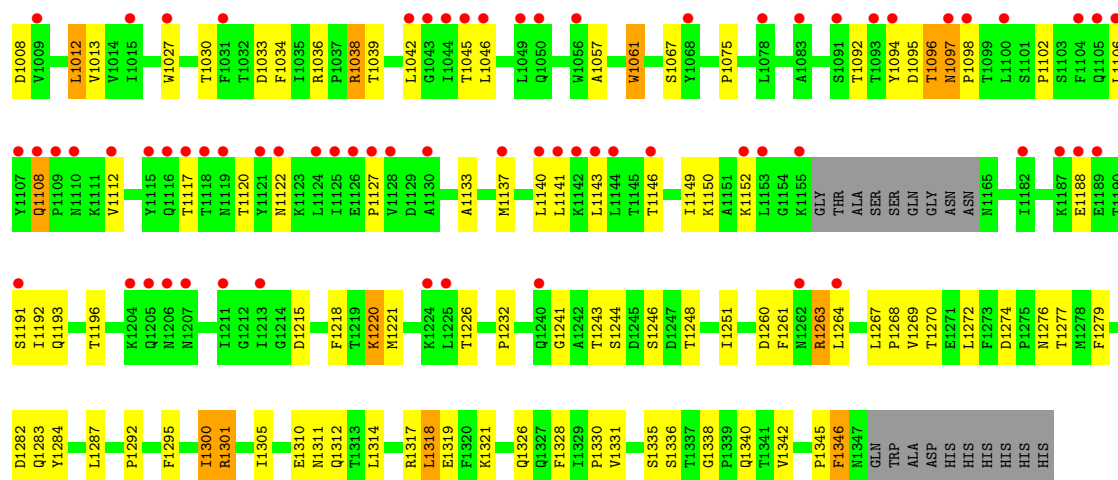


T1248	L1140	W1061	I983	L912	E848	SER	L702	K567	V387	P299	S233
D1260	L1141	K1065	G984	G916	Q849	THR	L707	Y568	N388	L300	F234
F1261	L1142	K1066	F985	G917	V850	SER	S707	D630	A389	K301	N235
T1145	L1143	S1067	K851	T918	K851	ALA	D708	Q570	H472	H302	P236
T1146	L1144	S1067	P852	T919	P852	SER	K708	V632	N391	K303	T237
		P1075	S853	P920	S853	SER	K709	E572	L392	N304	Q238
		L1078	N854	P921	N854		L710	K573	L393	R239	R239
		P1081	N855	P922	N855		F711	H574	G307	G307	L240
		S1090	P858	N923	P858		F712	L575	E308	Q241	Q241
		S1091	F859	N924	F859		G713	N576	V309	K243	K243
		T1092	P792	N925	P792		K714	L577	K310	S244	S244
		T1093	T795	K926	T795		E715	V578	L311	P245	P245
		D1098	F796	G927	F796		A719	Q580	F400	V246	V246
		S1099	N862	D928	N862		E720	Q581	K401	K247	K247
		S1100	S863	S929	S863		R725	Q582	P402	K248	K248
		T1099	D864	N930	D864		F726	G583	R403	S249	S249
		L1100	N866	D931	N866		K727	L584	R404	S250	S250
		P1098	K867	K935	K867		Q728	L585	F317	K251	K251
		S1101	V868	Q939	V868		L732	R586	W408	Q252	Q252
		P1099	T869	K940	T869		R731	E587	K415	G253	G253
		L1101	S874	D941	S874		L733	D588	D417	E254	E254
		P1102	K876	N942	K876		W735	L591	M418	L256	L256
		S1103	P877	K943	P877		R742	F592	F421	S257	S257
		F1104	T878	E945	T878		L751	F593	K330	E258	E258
		L1105	T879	N946	T879		F752	P594	E333	T259	T259
		T1106	Y880	N947	Y880		G753	G595	K334		
		Q1107	P881	G948	P881		L756	W597	S342		
		P1108	A882	G949	A882		D757	N528	F344		
		S1109	L883	N950	L883		F761	R529	M268		
		T1110	P884	L951	P884		L765	W530	S271		
		L1111	N885	D952	N885		F764	F531	T272		
		P1112	S886	G953	S886		L768	E532	A273		
		K4113	I887	F954	I887		G769	R536	A274		
		Y1114	S888	G955	S888		W770	M537	K275		
		T1115	T890	E956	T890		F773	V542	A276		
		Q1116	S891	N957	S891		N773	K543	L277		
		T1117	D892	G959	D892		G774	G553	K278		
		L1118	W893	N960	W893		N775	M557	V279		
		T1119	I894	Y961	I894		K775	T560	E280		
		P1120	N895	N962	N895		F777	A448	V281		
		S1121	A896	G963	A896		S780	E449	R282		
		T1122	L965	L966	L965		F781	V376	R283		
		K1123	T900	H967	T900		PR0	P377	G284		
		L1124	N901	T968	N901			G376	L290		
		E1125	K902	T969	K902			P377	L291		
		P1127	N903	Y969	N903			N380	K292		
		T1128	N904	G970	N904			H381	N293		
		D1129	P905	F971	P905			H382	D294		
		A1130	Q906	F972	Q906			G383	F295		
		L1131	K907	G973	K907			I384	A296		
		T1132	N908	T974	N908			W385	K297		
		S1133	Q909	N975	Q909			D386	K298		
		A1134	L945	T976	L945						
		M1137	P946		P946						
			F947		F947						



● Molecule 1: Adhesin P1





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	459.19Å 116.66Å 285.64Å 90.00° 124.20° 90.00°	Depositor
Resolution (Å)	37.36 – 3.24 236.25 – 3.24	Depositor EDS
% Data completeness (in resolution range)	61.7 (37.36-3.24) 61.8 (236.25-3.24)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.62 (at 3.26Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, $R_{free}$	0.184 , 0.205 0.202 , 0.225	Depositor DCC
$R_{free}$ test set	6155 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.0	Xtriage
Anisotropy	0.038	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 96.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	59767	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.81% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<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.

## 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	0.52	0/10269	0.78	6/13978 (0.0%)
1	B	0.49	0/10252	0.76	3/13956 (0.0%)
1	C	0.50	0/10246	0.76	1/13947 (0.0%)
1	D	0.50	0/10229	0.76	3/13925 (0.0%)
1	E	0.45	0/10135	0.73	2/13795 (0.0%)
1	F	0.45	0/10147	0.73	2/13812 (0.0%)
All	All	0.49	0/61278	0.75	17/83413 (0.0%)

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	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	476	MET	CB-CG-SD	7.17	133.92	112.40
1	E	861	PRO	N-CA-C	-6.71	94.65	112.10
1	F	861	PRO	N-CA-C	-6.65	94.81	112.10
1	A	109	ASP	C-N-CA	6.20	137.21	121.70
1	A	806	MET	C-N-CA	6.17	137.12	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	926	SER	Mainchain
1	C	926	SER	Mainchain

## 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	10016	0	9726	182	0
1	B	9999	0	9707	183	0
1	C	9993	0	9706	182	0
1	D	9976	0	9687	205	0
1	E	9886	0	9597	156	0
1	F	9897	0	9610	163	0
All	All	59767	0	58033	1029	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 9.

The worst 5 of 1029 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:100:ILE:CG1	1:D:100:ILE:CD1	1.76	1.57
1:D:812:ASN:ND2	1:D:815:ASN:HD22	1.28	1.31
1:F:1279:PHE:CE2	1:F:1301:ARG:HD2	1.78	1.18
1:A:1266:THR:HG21	1:C:1264:LEU:HD11	1.19	1.15
1:A:1264:LEU:HG	1:C:1283:GLN:HB2	1.30	1.09

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	1270/1304 (97%)	1128 (89%)	122 (10%)	20 (2%)	9	40
1	B	1267/1304 (97%)	1136 (90%)	114 (9%)	17 (1%)	12	44
1	C	1266/1304 (97%)	1128 (89%)	123 (10%)	15 (1%)	13	46
1	D	1263/1304 (97%)	1127 (89%)	117 (9%)	19 (2%)	10	41
1	E	1252/1304 (96%)	1118 (89%)	117 (9%)	17 (1%)	11	43
1	F	1252/1304 (96%)	1114 (89%)	119 (10%)	19 (2%)	10	41
All	All	7570/7824 (97%)	6751 (89%)	712 (9%)	107 (1%)	11	43

5 of 107 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	108	SER
1	A	1094	TYR
1	A	1097	ASN
1	A	1241	GLY
1	B	811	LEU

### 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	1121/1146 (98%)	964 (86%)	157 (14%)	3	15
1	B	1120/1146 (98%)	965 (86%)	155 (14%)	3	15
1	C	1118/1146 (98%)	973 (87%)	145 (13%)	4	18
1	D	1117/1146 (98%)	969 (87%)	148 (13%)	4	17
1	E	1107/1146 (97%)	958 (86%)	149 (14%)	4	16
1	F	1108/1146 (97%)	958 (86%)	150 (14%)	4	16
All	All	6691/6876 (97%)	5787 (86%)	904 (14%)	4	16

5 of 904 residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	D	127	ASN
1	F	1263	ARG
1	D	1108	GLN
1	F	1192	ILE
1	F	353	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 174 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	947	ASN
1	E	939	GLN
1	D	1327	GLN
1	E	610	GLN
1	F	217	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 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	1278/1304 (98%)	0.15	8 (0%) 89 85	13, 49, 106, 165	0
1	B	1275/1304 (97%)	0.28	20 (1%) 72 63	24, 60, 127, 238	0
1	C	1274/1304 (97%)	0.24	21 (1%) 72 63	11, 50, 128, 234	0
1	D	1271/1304 (97%)	0.32	27 (2%) 63 53	18, 58, 132, 228	0
1	E	1262/1304 (96%)	1.91	445 (35%) 0 0	25, 156, 238, 252	0
1	F	1262/1304 (96%)	1.75	431 (34%) 0 0	21, 149, 229, 244	0
All	All	7622/7824 (97%)	0.77	952 (12%) 3 3	11, 68, 224, 252	0

The worst 5 of 952 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	920	PRO	20.2
1	F	836	VAL	14.6
1	F	920	PRO	14.3
1	E	1125	ILE	14.1
1	F	961	TYR	13.9

### 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 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.