



## wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Mar 2, 2017 – 01:08 pm GMT

PDB ID : 5WQ9  
EMDB ID: : EMD-6677  
Title : CryoEM structure of type II secretion system secretin GspD G453A mutant  
in *Vibrio cholerae*  
Authors : Yan, Z.; Yin, M.; Li, X.  
Deposited on : 2016-11-23  
Resolution : 4.22 Å(reported)

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report  
for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc29047

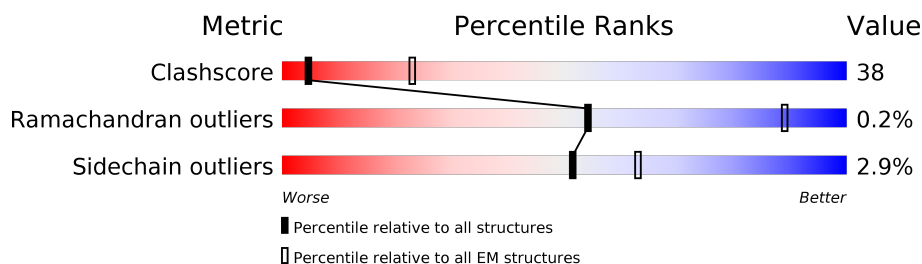
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.22 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)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	650	
1	B	650	
1	C	650	
1	D	650	
1	E	650	
1	F	650	
1	G	650	
1	H	650	
1	I	650	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	J	650	<div><div></div><div>41%33%24%</div></div>
1	K	650	<div><div></div><div>41%33%24%</div></div>
1	L	650	<div><div></div><div>40%33%24%</div></div>
1	M	650	<div><div></div><div>40%33%24%</div></div>
1	N	650	<div><div></div><div>40%33%24%</div></div>
1	O	650	<div><div></div><div>41%33%24%</div></div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 56340 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Type II secretion system protein D.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	B	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	C	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	D	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	E	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	F	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	G	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	H	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	I	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	J	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	K	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	L	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	M	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	N	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		
1	O	493	Total	C	N	O	S	0	0
			3756	2347	656	740	13		

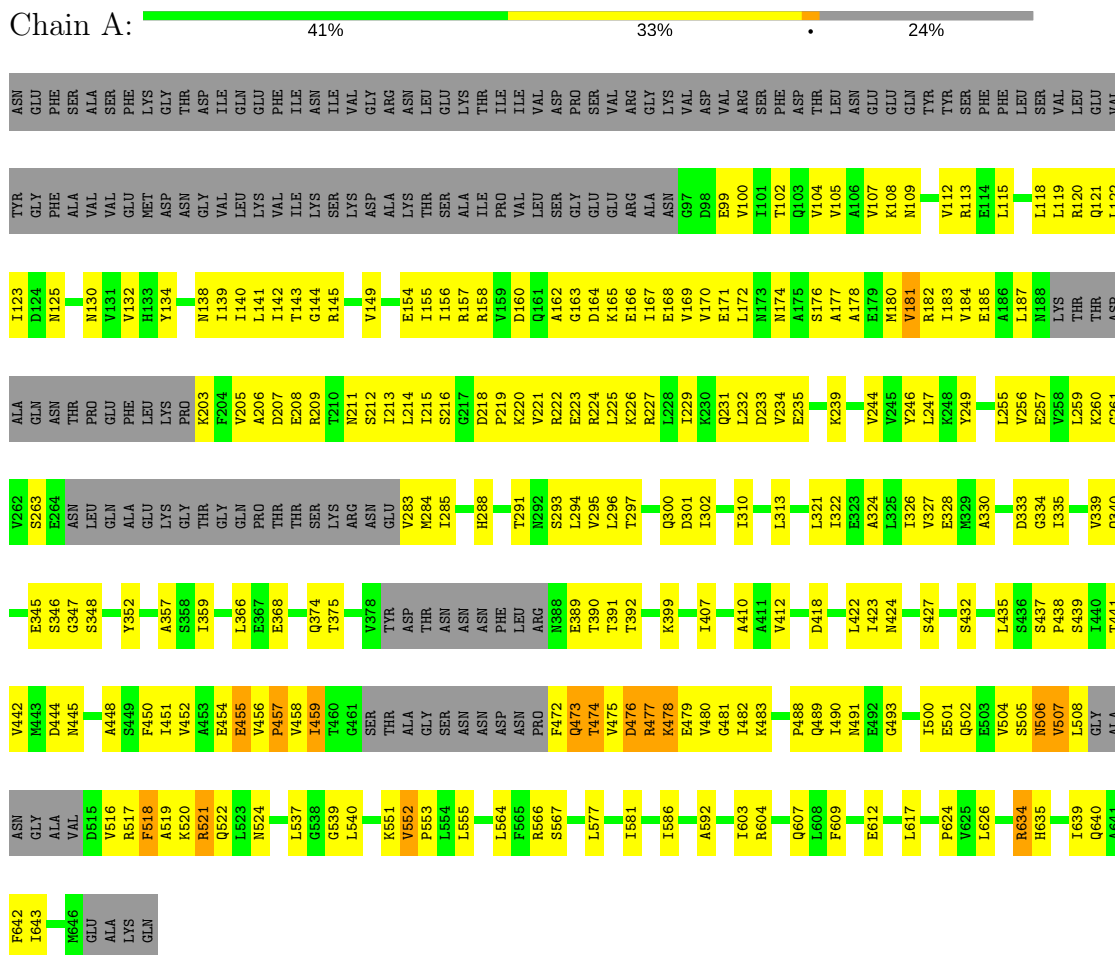
There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	453	ALA	GLY	engineered mutation	UNP P45779
B	453	ALA	GLY	engineered mutation	UNP P45779
C	453	ALA	GLY	engineered mutation	UNP P45779
D	453	ALA	GLY	engineered mutation	UNP P45779
E	453	ALA	GLY	engineered mutation	UNP P45779
F	453	ALA	GLY	engineered mutation	UNP P45779
G	453	ALA	GLY	engineered mutation	UNP P45779
H	453	ALA	GLY	engineered mutation	UNP P45779
I	453	ALA	GLY	engineered mutation	UNP P45779
J	453	ALA	GLY	engineered mutation	UNP P45779
K	453	ALA	GLY	engineered mutation	UNP P45779
L	453	ALA	GLY	engineered mutation	UNP P45779
M	453	ALA	GLY	engineered mutation	UNP P45779
N	453	ALA	GLY	engineered mutation	UNP P45779
O	453	ALA	GLY	engineered mutation	UNP P45779

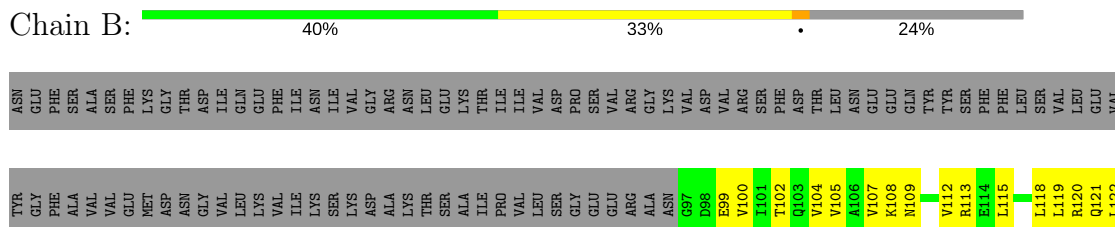
### 3 Residue-property plots

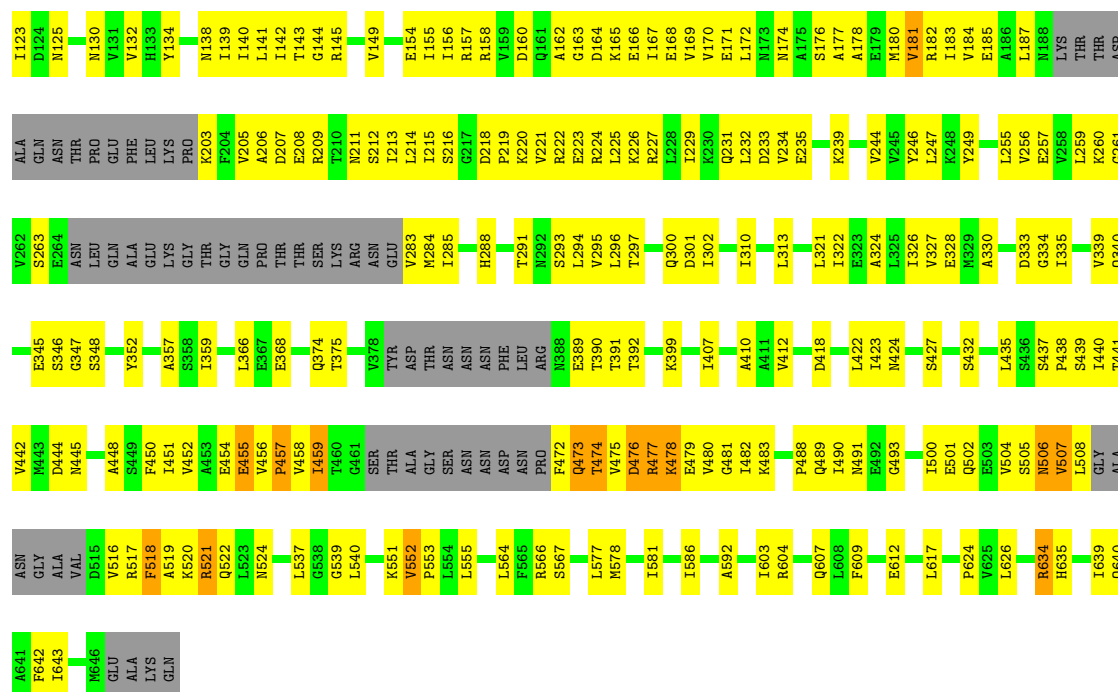
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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. 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: Type II secretion system protein D



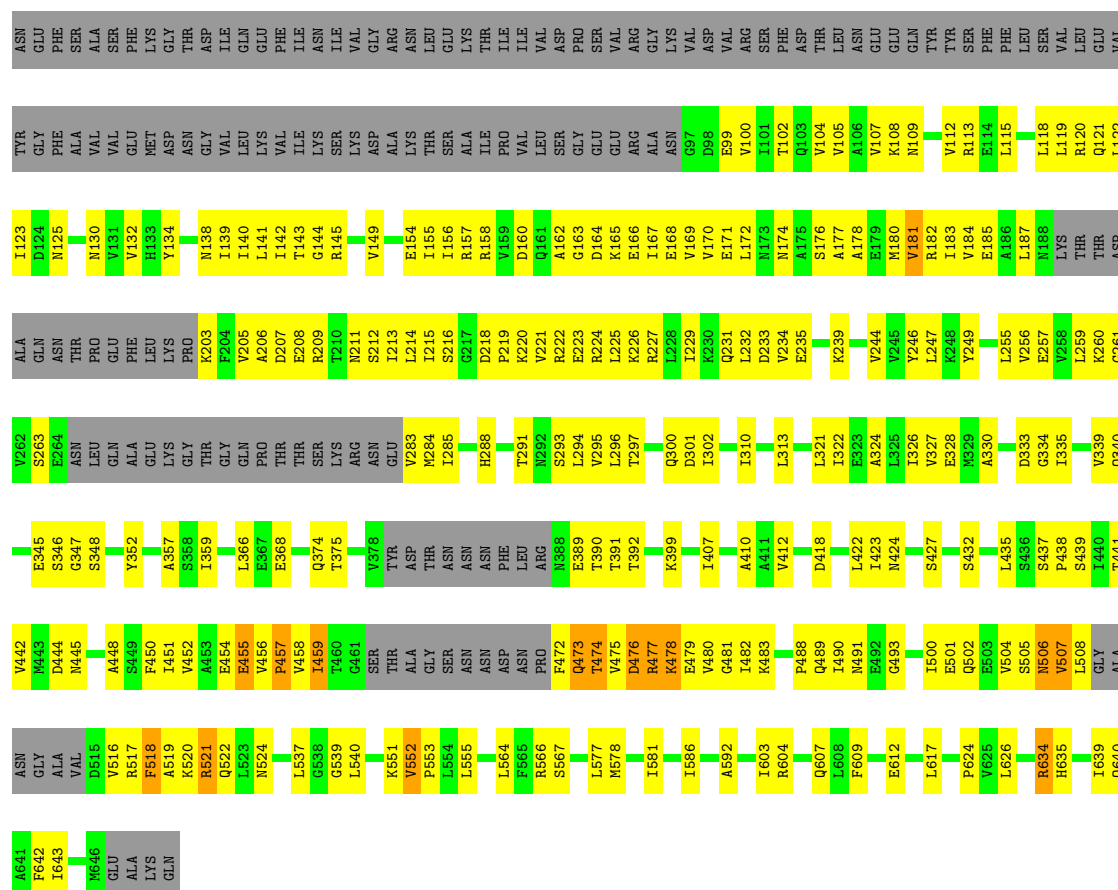
- Molecule 1: Type II secretion system protein D





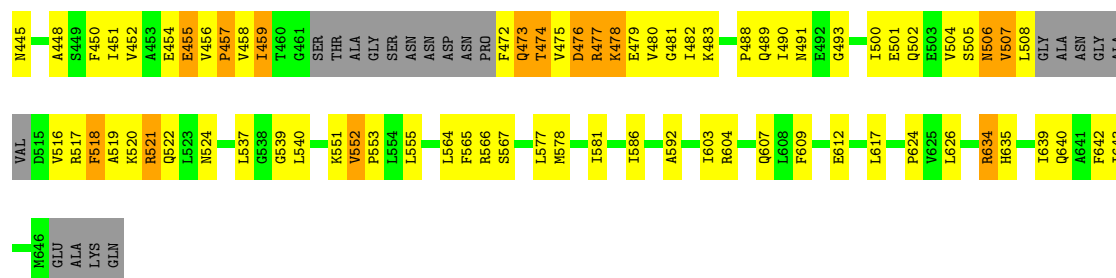
• Molecule 1: Type II secretion system protein D

Chain C: 40% 33% 24%



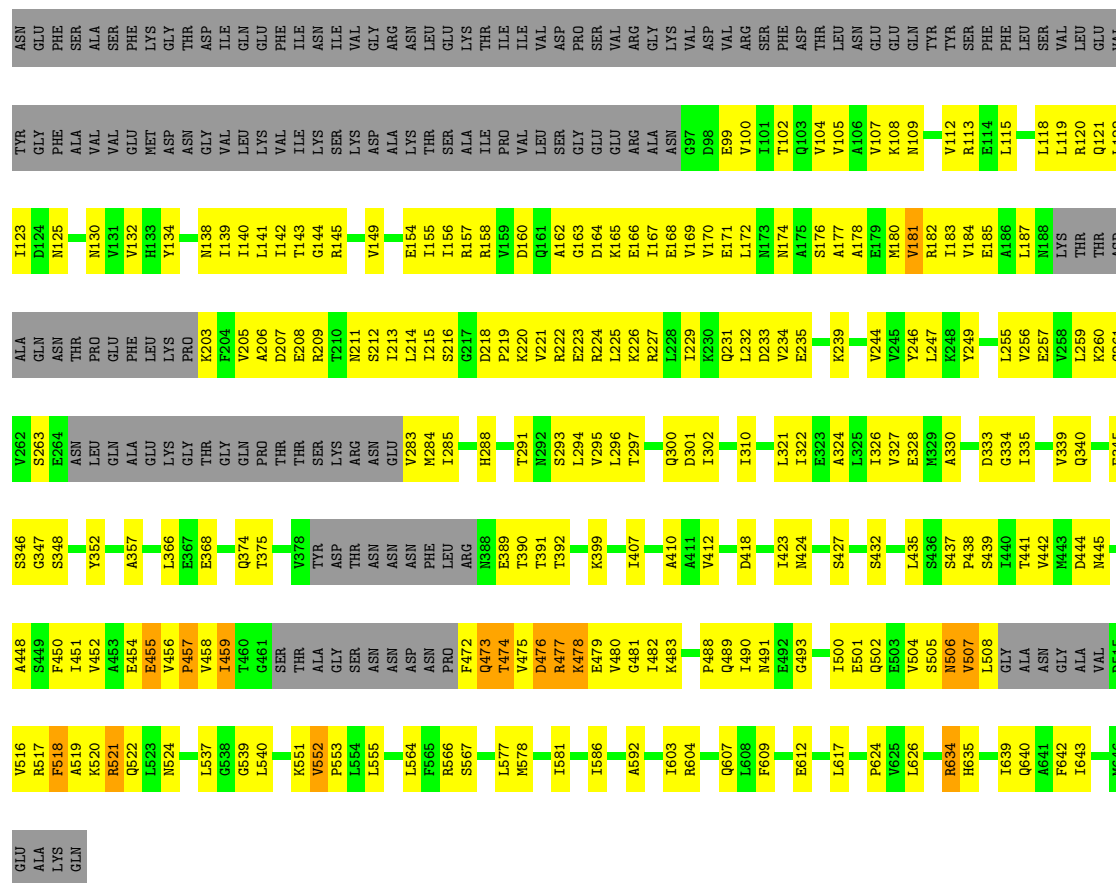






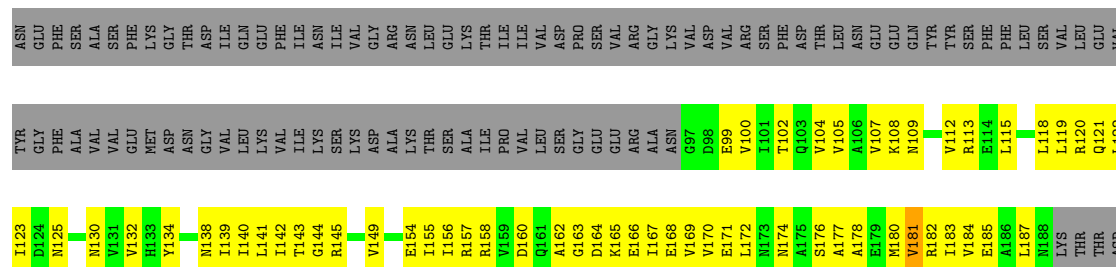
# • Molecule 1: Type II secretion system protein D

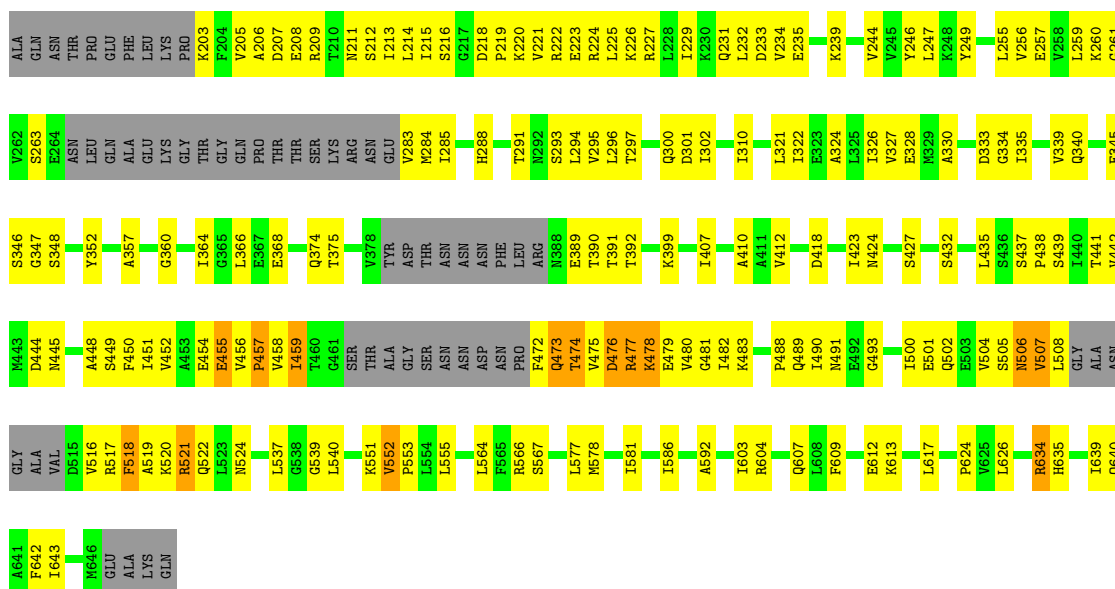
Chain F:



# • Molecule 1: Type II secretion system protein D

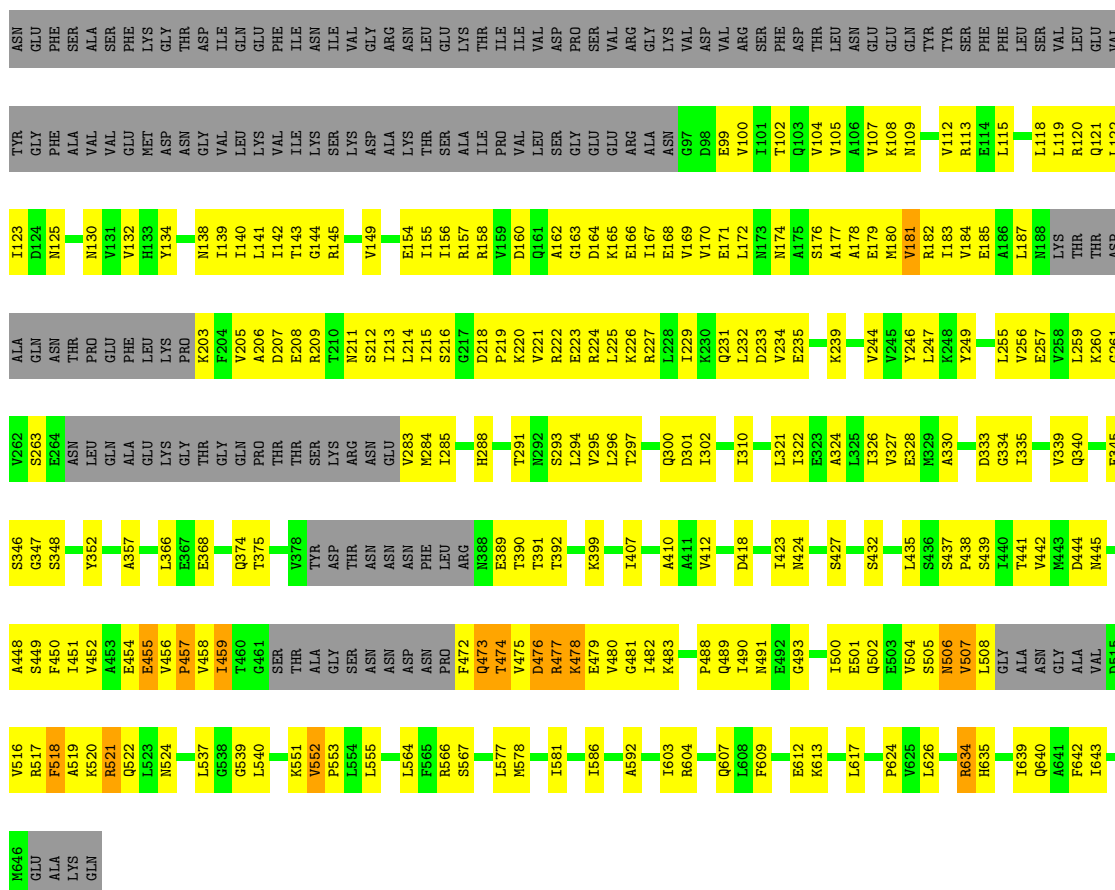
Chain G:





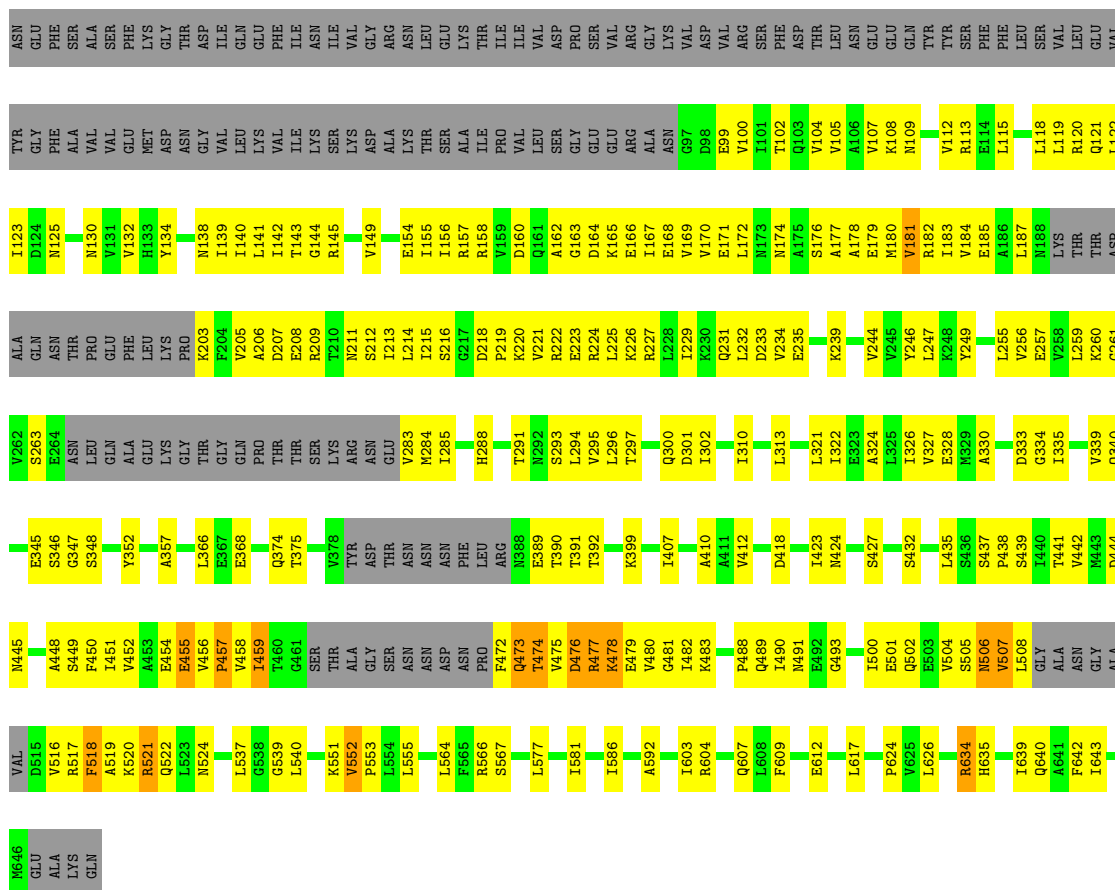
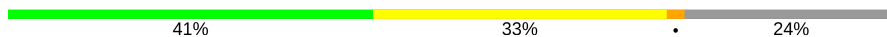
- Molecule 1: Type II secretion system protein D

Chain H:



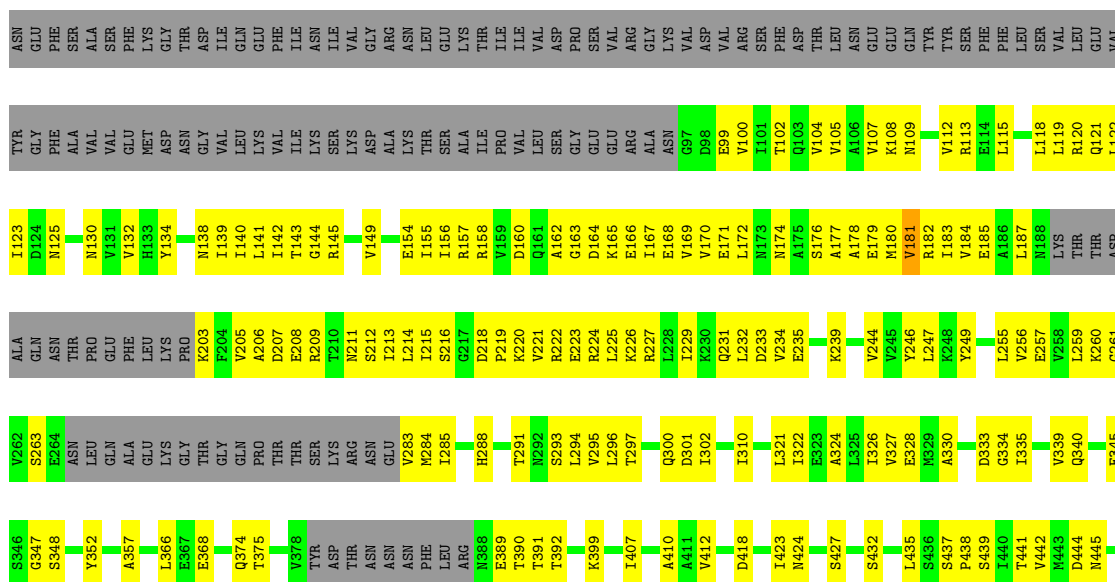
- Molecule 1: Type II secretion system protein D

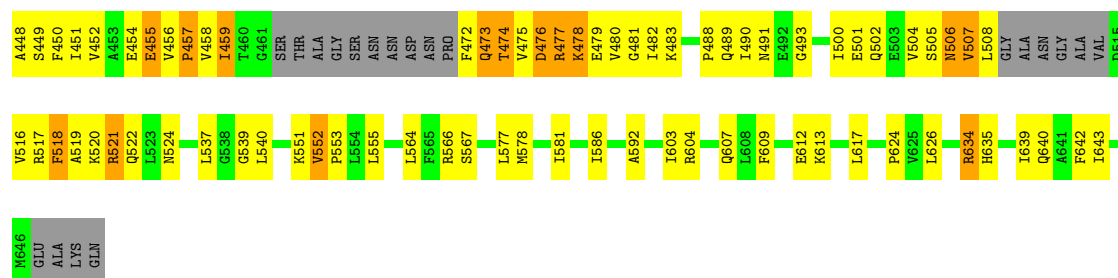
Chain I:



- Molecule 1: Type II secretion system protein D

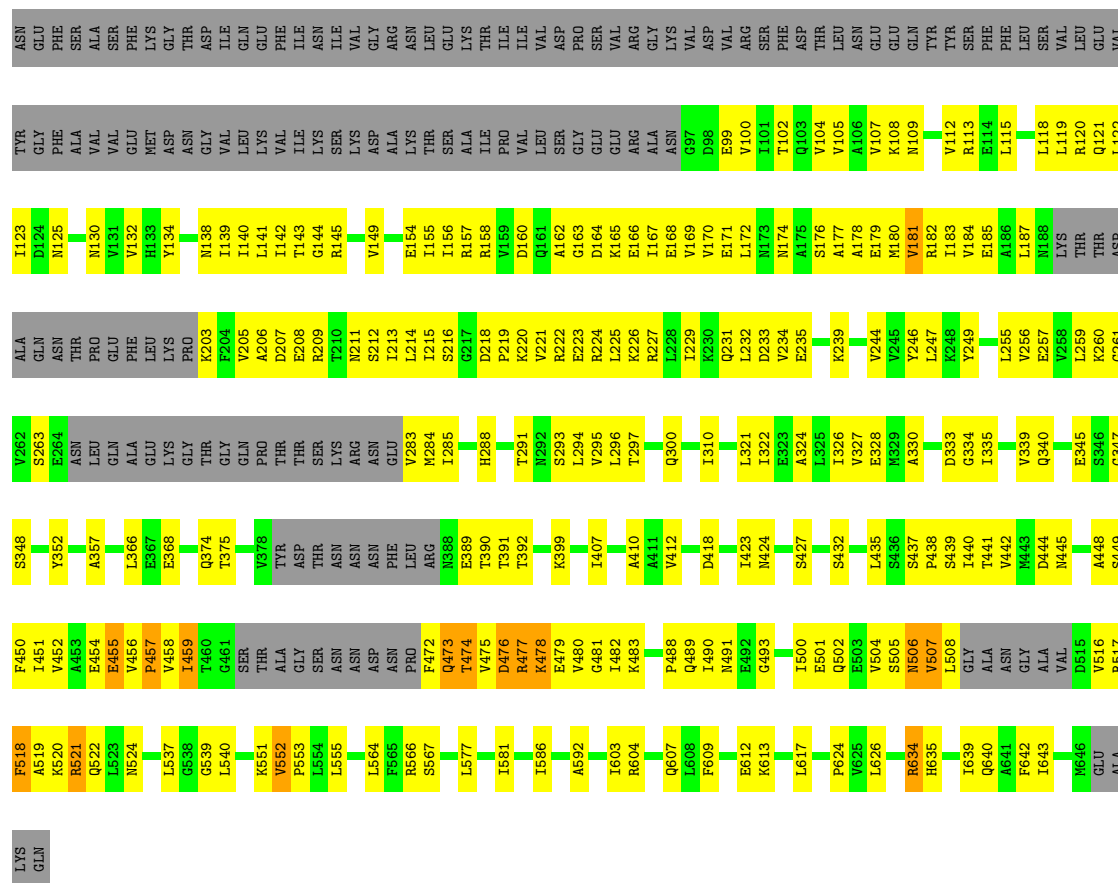
## Chain J:





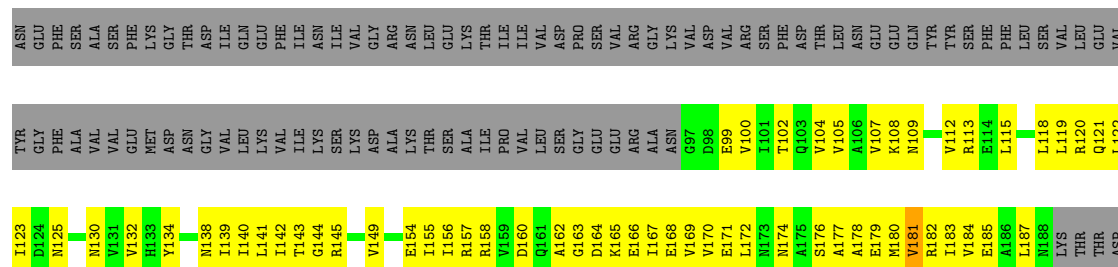
• Molecule 1: Type II secretion system protein D

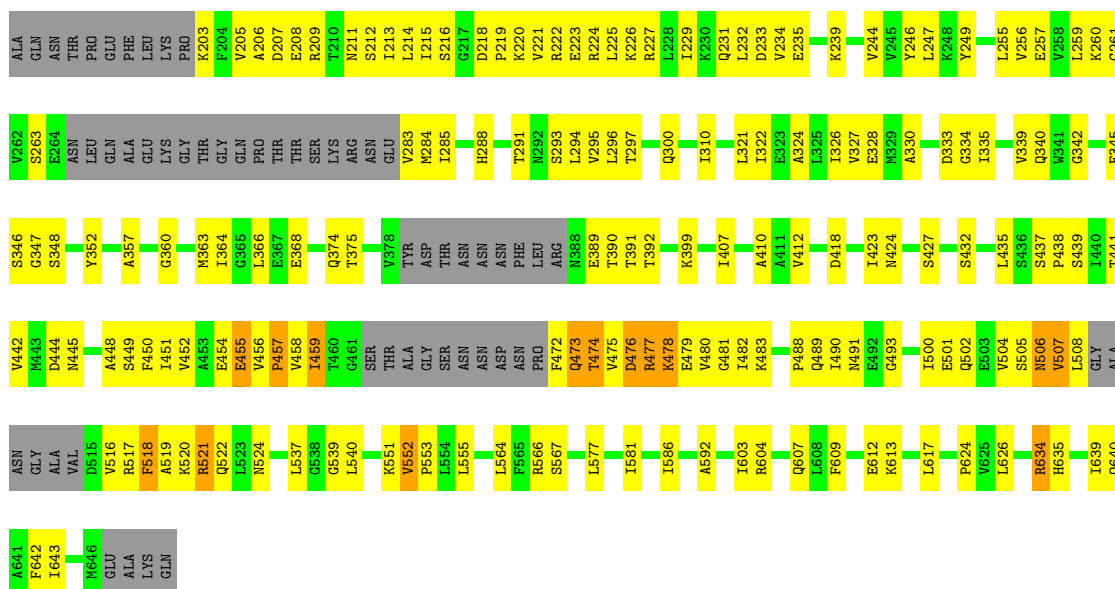
Chain K: 41% 33% 24%



• Molecule 1: Type II secretion system protein D

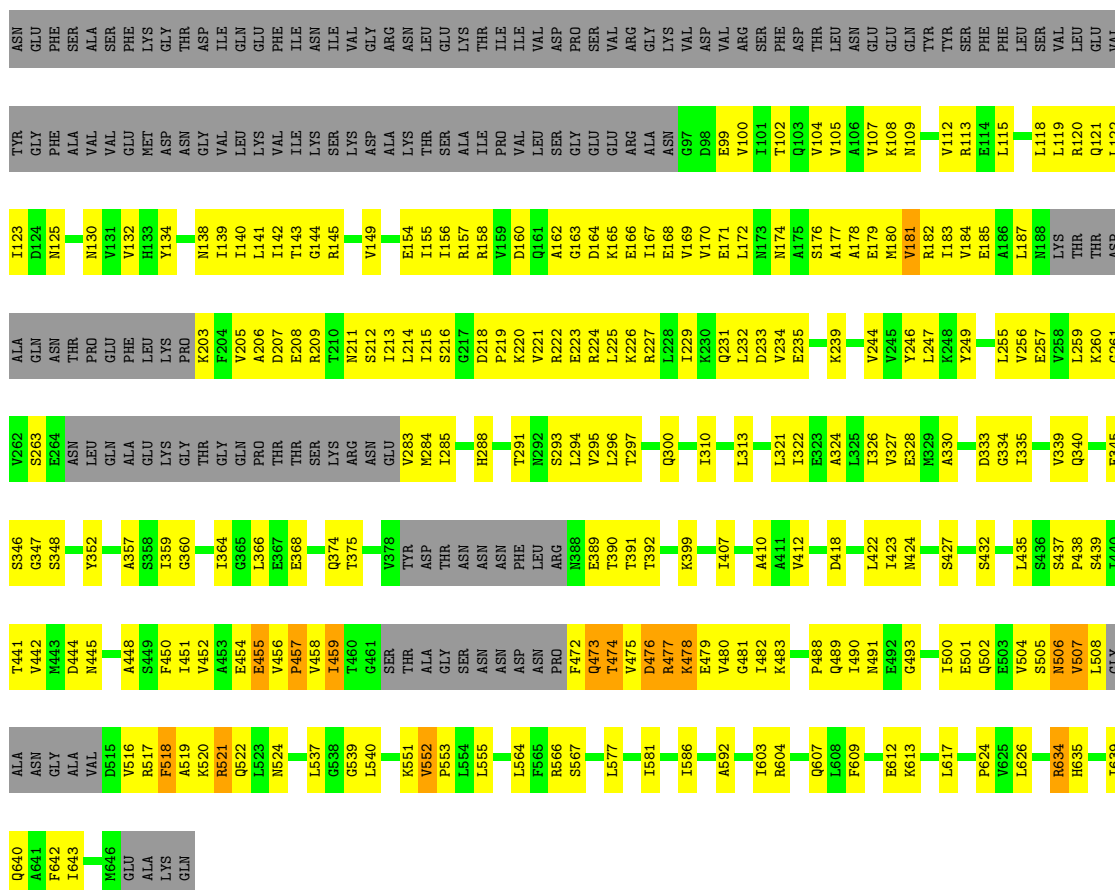
Chain L: 40% 33% 24%





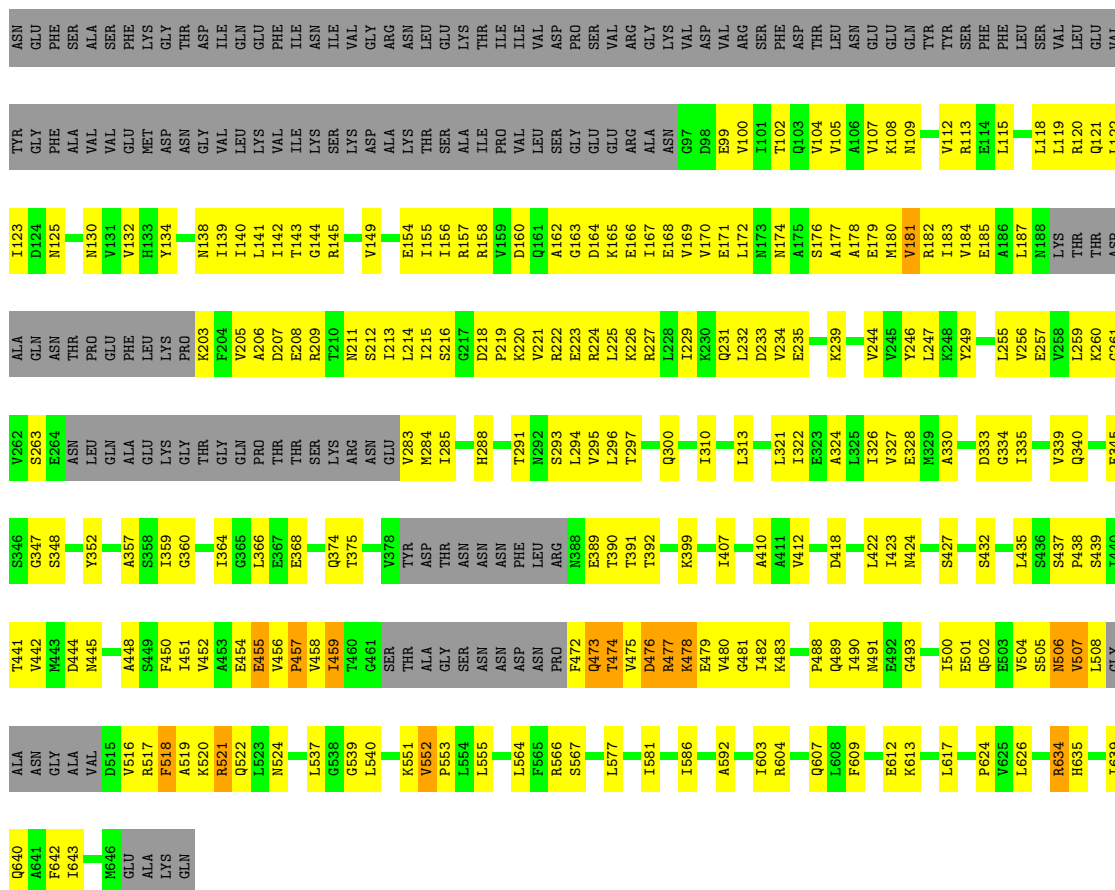
- Molecule 1: Type II secretion system protein D

Chain M:



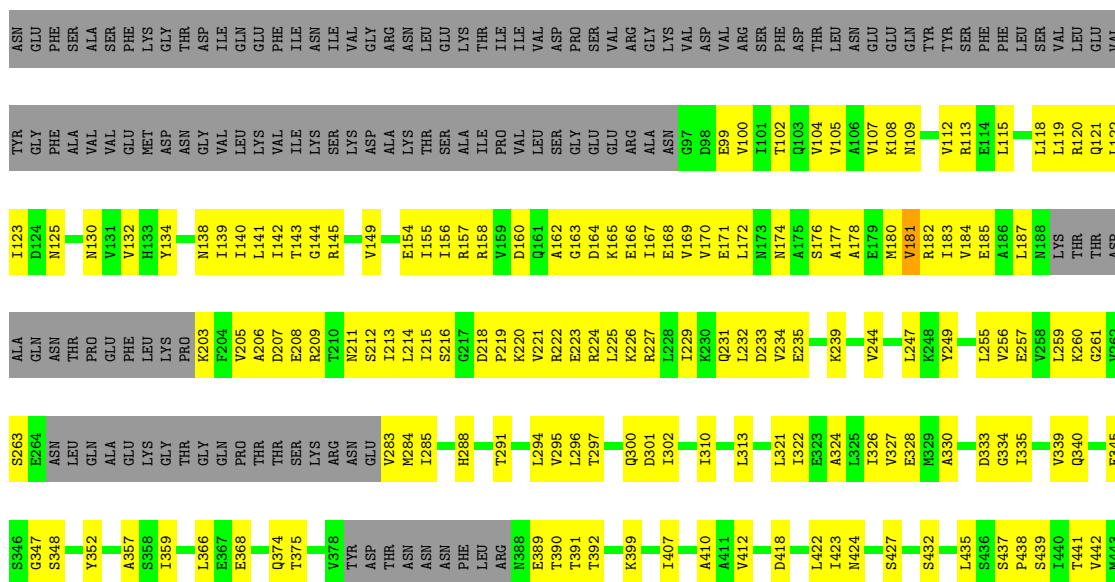
- Molecule 1: Type II secretion system protein D

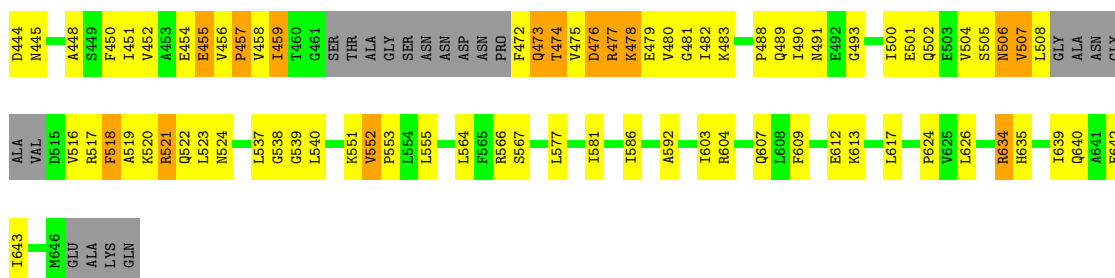
Chain N:  40% 33% • 24%



- Molecule 1: Type II secretion system protein D

Chain O: 





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	26955	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor



## 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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	A	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	B	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	C	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	D	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	E	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	F	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	G	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	H	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	I	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	J	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	K	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	L	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	M	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	N	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
1	O	0.41	1/3786 (0.0%)	0.64	4/5117 (0.1%)
All	All	0.41	15/56790 (0.0%)	0.64	60/76755 (0.1%)

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	M	457	PRO	N-CD	5.24	1.55	1.47
1	E	457	PRO	N-CD	5.23	1.55	1.47
1	J	457	PRO	N-CD	5.23	1.55	1.47
1	C	457	PRO	N-CD	5.21	1.55	1.47
1	B	457	PRO	N-CD	5.21	1.55	1.47

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	507	VAL	CB-CA-C	-10.70	91.07	111.40
1	E	507	VAL	CB-CA-C	-10.70	91.07	111.40
1	C	507	VAL	CB-CA-C	-10.70	91.07	111.40
1	I	507	VAL	CB-CA-C	-10.70	91.08	111.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	M	507	VAL	CB-CA-C	-10.70	91.08	111.40

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	3756	0	3858	323	0
1	B	3756	0	3858	327	0
1	C	3756	0	3858	320	0
1	D	3756	0	3858	317	0
1	E	3756	0	3858	319	0
1	F	3756	0	3858	319	0
1	G	3756	0	3858	318	0
1	H	3756	0	3858	319	0
1	I	3756	0	3858	318	0
1	J	3756	0	3858	324	0
1	K	3756	0	3858	322	0
1	L	3756	0	3858	320	0
1	M	3756	0	3858	317	0
1	N	3756	0	3858	321	0
1	O	3756	0	3858	324	0
All	All	56340	0	57870	4309	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 38.

The worst 5 of 4309 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:481:GLY:H	1:G:505:SER:CB	1.22	1.53
1:I:481:GLY:H	1:I:505:SER:CB	1.22	1.53
1:L:481:GLY:H	1:L:505:SER:CB	1.22	1.52
1:N:481:GLY:H	1:N:505:SER:CB	1.22	1.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:K:481:GLY:H	1:K:505:SER:CB	1.22	1.52

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 PDB entries followed by that with respect to all EM entries.

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	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	B	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	C	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	D	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	E	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	F	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	G	481/650 (74%)	458 (95%)	22 (5%)	1 (0%)	51	85
1	H	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	I	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	J	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	K	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	L	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	M	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	N	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
1	O	481/650 (74%)	459 (95%)	21 (4%)	1 (0%)	51	85
All	All	7215/9750 (74%)	6884 (95%)	316 (4%)	15 (0%)	54	85

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	181	VAL
1	B	181	VAL
1	C	181	VAL
1	E	181	VAL
1	F	181	VAL

### 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 PDB entries followed by that with respect to all EM entries.

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	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	B	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	C	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	D	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	E	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	F	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	G	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	H	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	I	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	J	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	K	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	L	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	M	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	N	415/550 (76%)	403 (97%)	12 (3%)	48	74
1	O	415/550 (76%)	403 (97%)	12 (3%)	48	74
All	All	6225/8250 (76%)	6045 (97%)	180 (3%)	51	74

5 of 180 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	518	PHE
1	I	476	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	N	520	LYS
1	G	521	ARG
1	H	506	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 165 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	502	GLN
1	I	340	GLN
1	N	506	ASN
1	G	640	GLN
1	H	424	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 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.