



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

Mar 2, 2017 – 01:08 pm GMT

PDB ID : 5WQ7  
EMDB ID: : EMD-6675  
Title : CryoEM structure of type II secretion system secretin GspD in E.coli K12  
Authors : Yan, Z.; Yin, M.; Li, X.  
Deposited on : 2016-11-23  
Resolution : 3.04 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

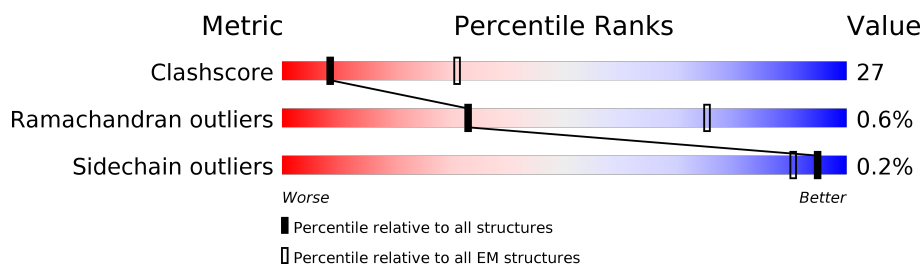
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.04 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	627	
1	B	627	
1	C	627	
1	D	627	
1	E	627	
1	F	627	
1	G	627	
1	H	627	
1	I	627	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	J	627	 44% 31% • 25%
1	K	627	 45% 30% • 25%
1	L	627	 45% 30% • 25%
1	M	627	 44% 30% • 25%
1	N	627	 44% 30% • 25%
1	O	627	 44% 30% • 25%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 54480 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 Putative type II secretion system protein D.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	B	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	C	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	D	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	E	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	F	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	G	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	H	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	I	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	J	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	K	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	L	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	M	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	N	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		
1	O	472	Total	C	N	O	S	0	0
			3632	2279	628	720	5		

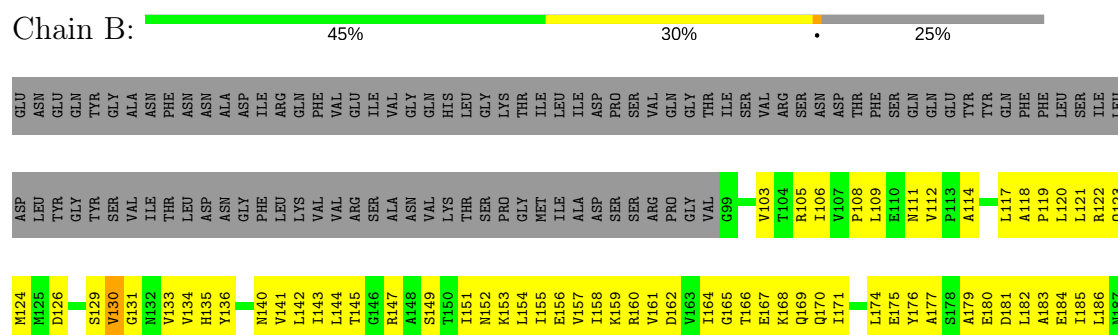
### 3 Residue-property plots

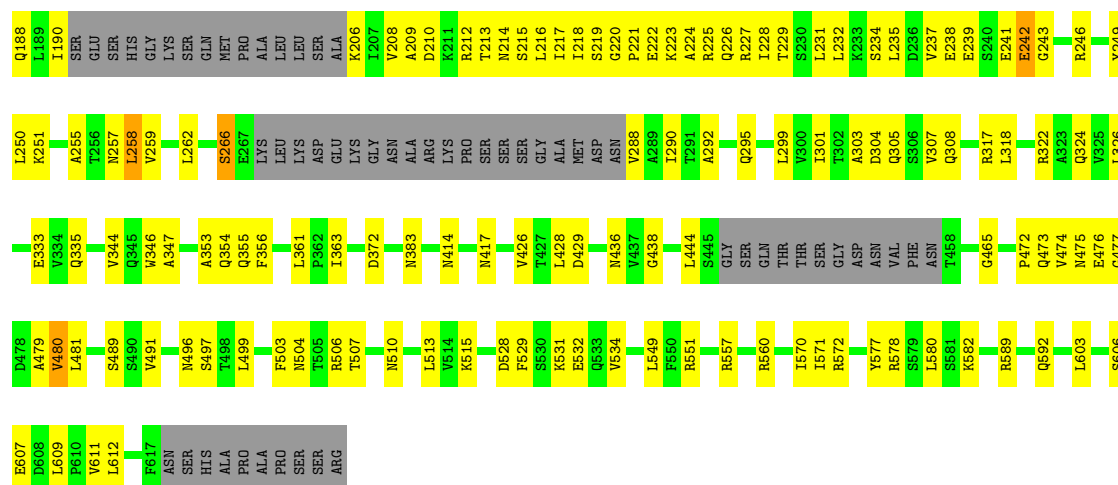
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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



#### • Molecule 1: Putative type II secretion system protein D

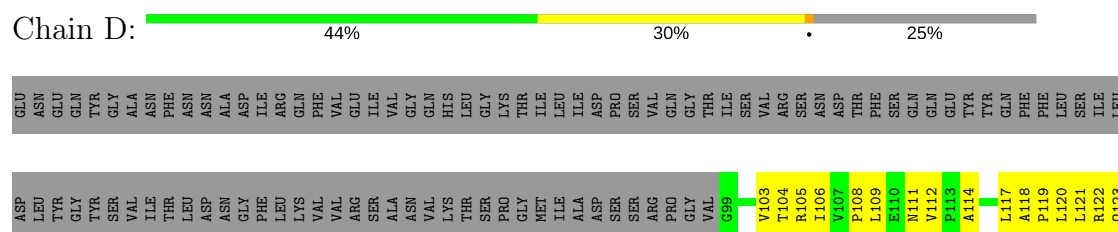


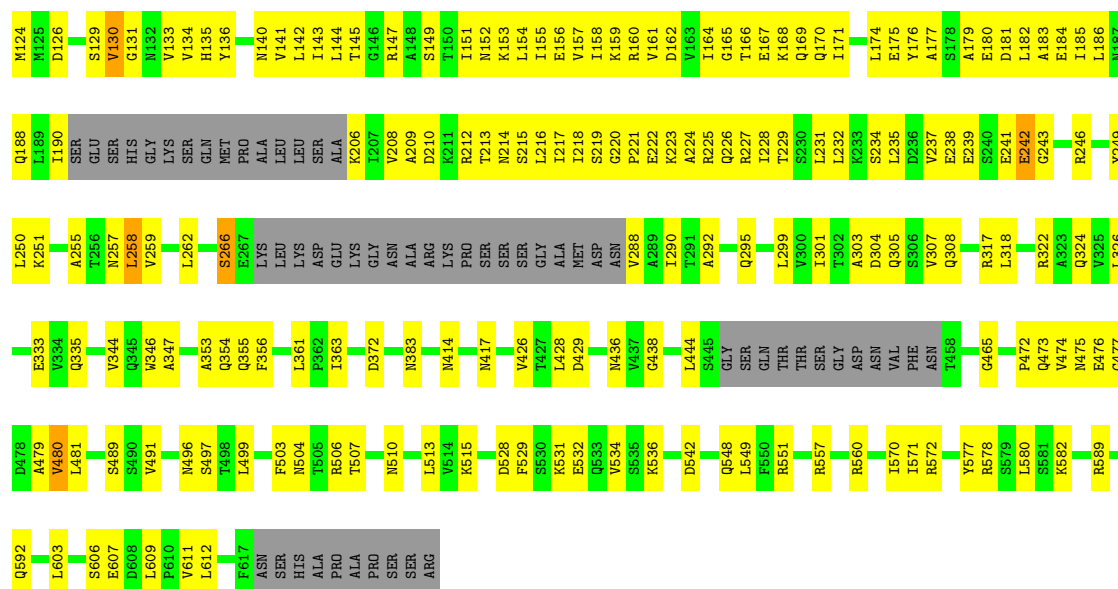


• Molecule 1: Putative type II secretion system protein D



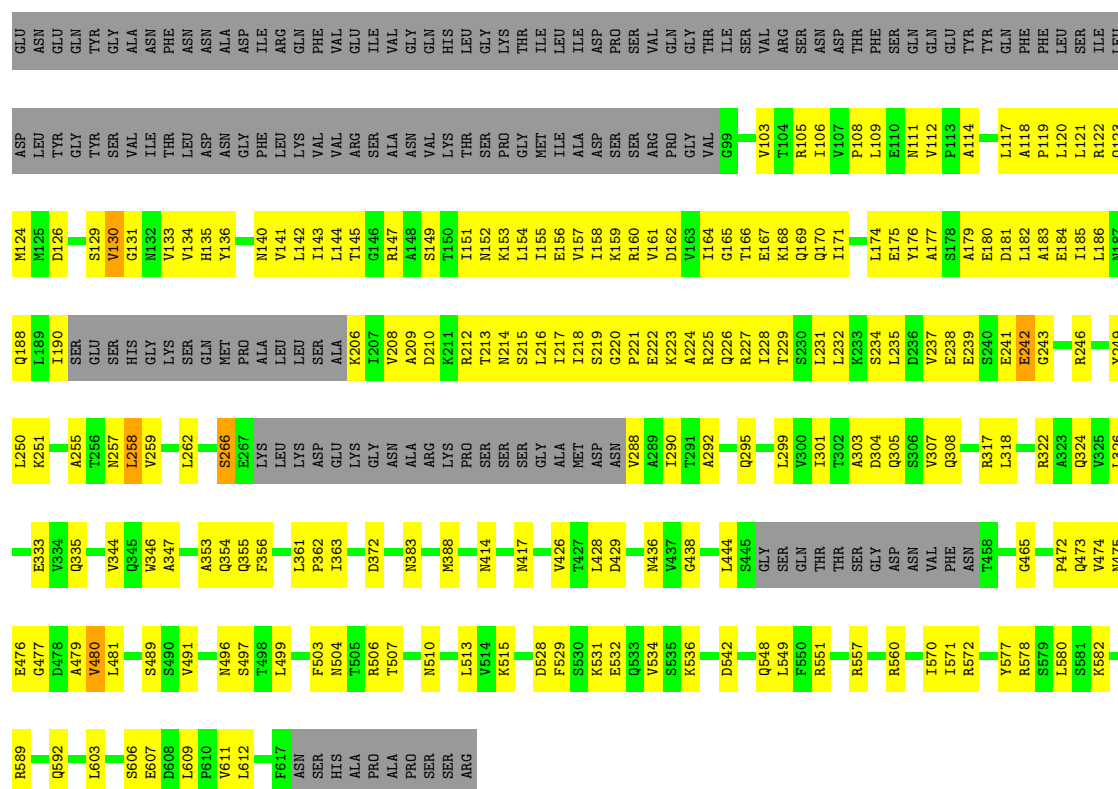
• Molecule 1: Putative type II secretion system protein D





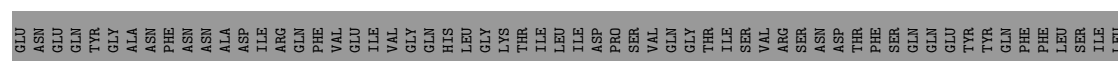
• Molecule 1: Putative type II secretion system protein D

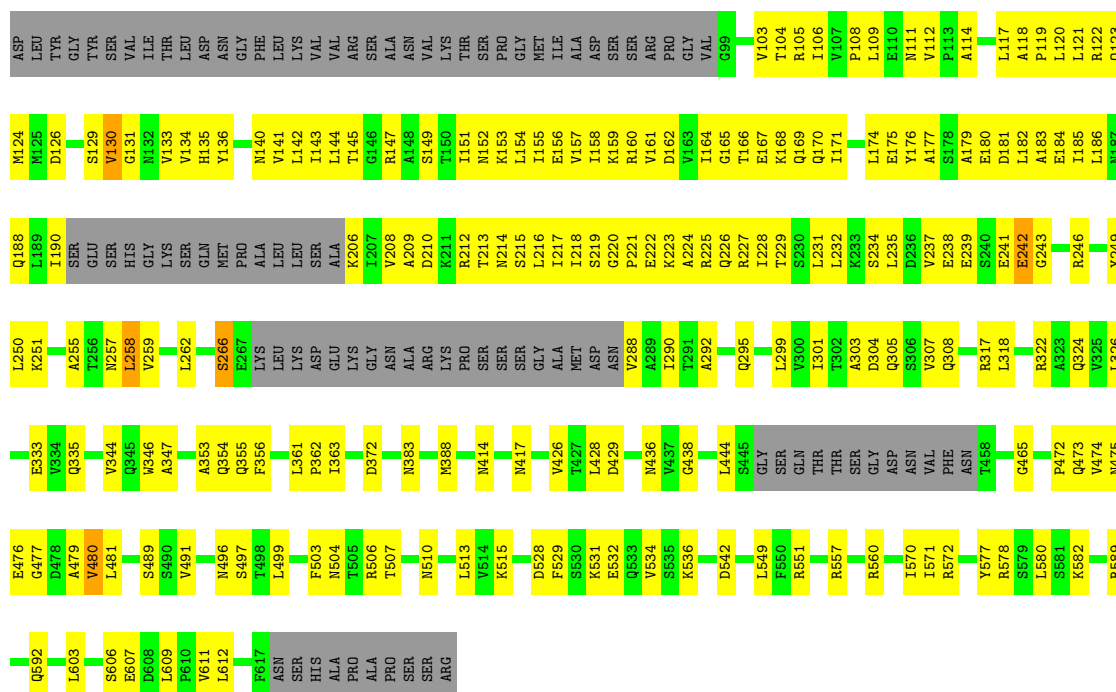
Chain E: 44% 30% 25%



• Molecule 1: Putative type II secretion system protein D

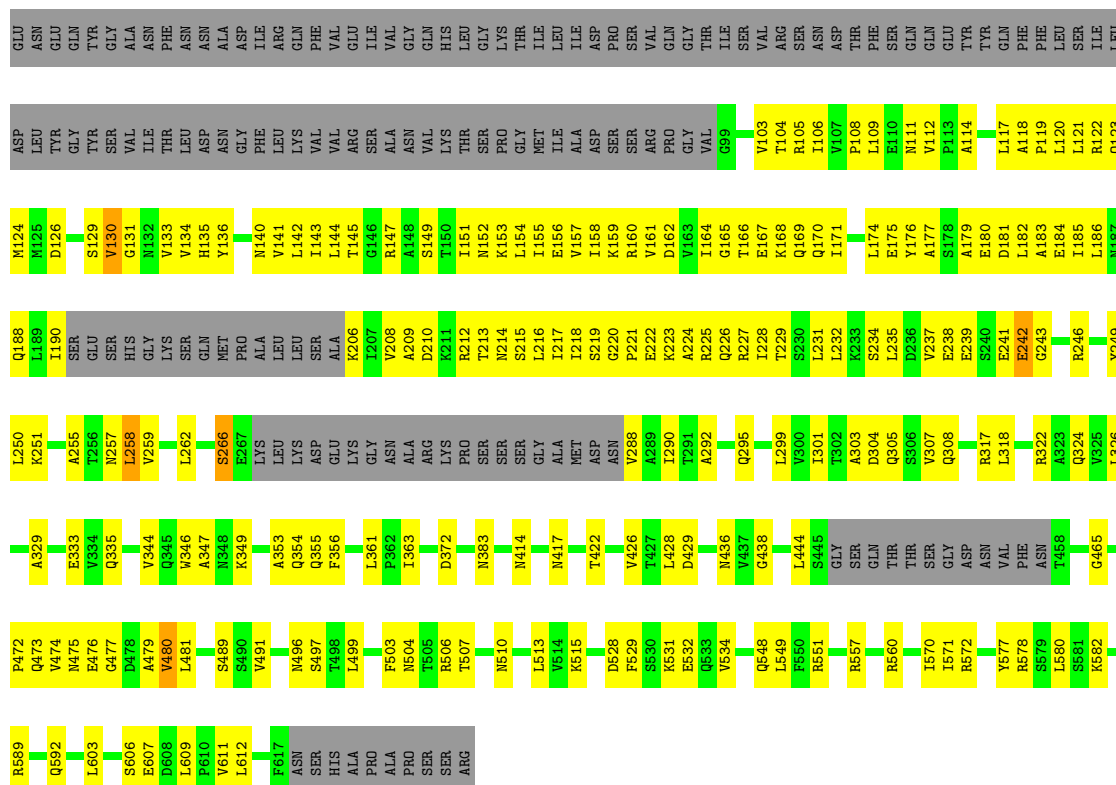
Chain F: 44% 30% 25%





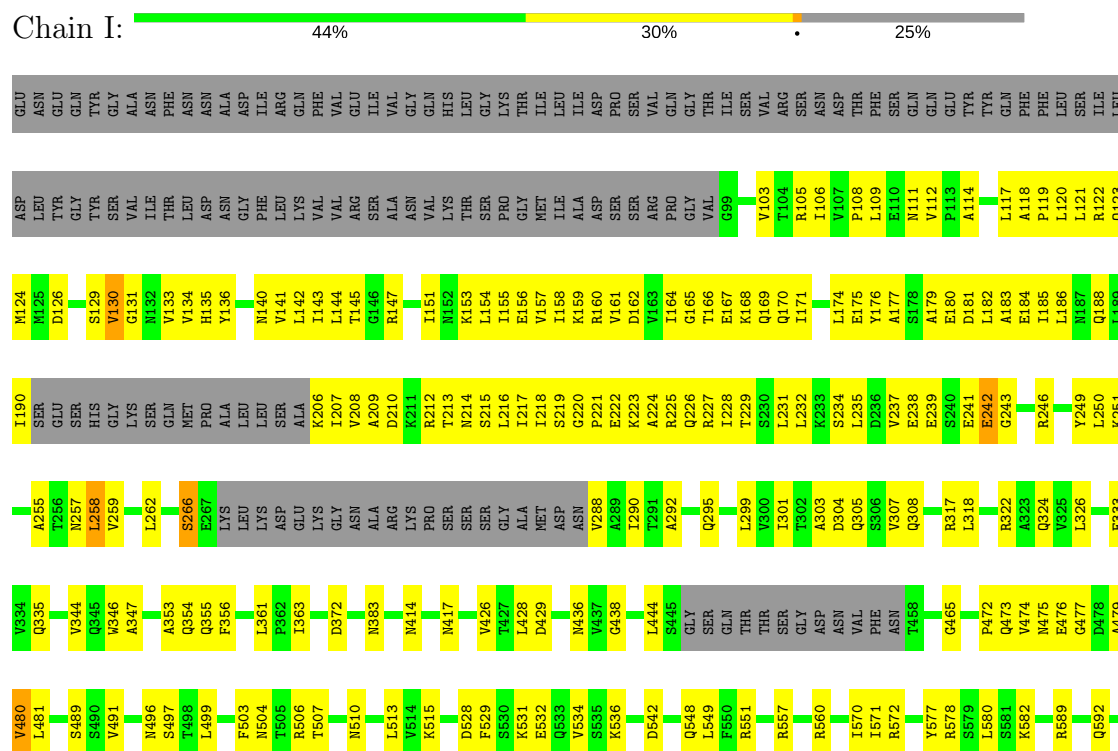
### • Molecule 1: Putative type II secretion system protein D

Chain G: 44% 30% 25%



### • Molecule 1: Putative type II secretion system protein D

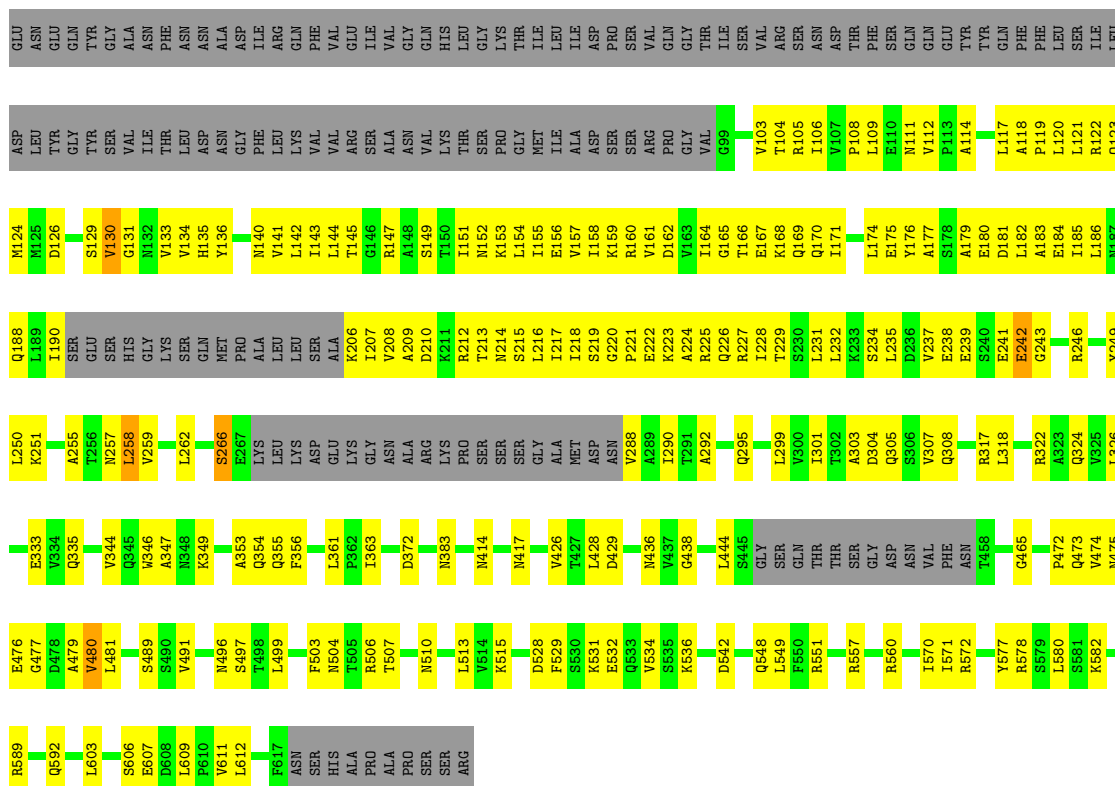
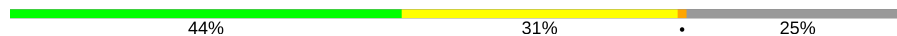






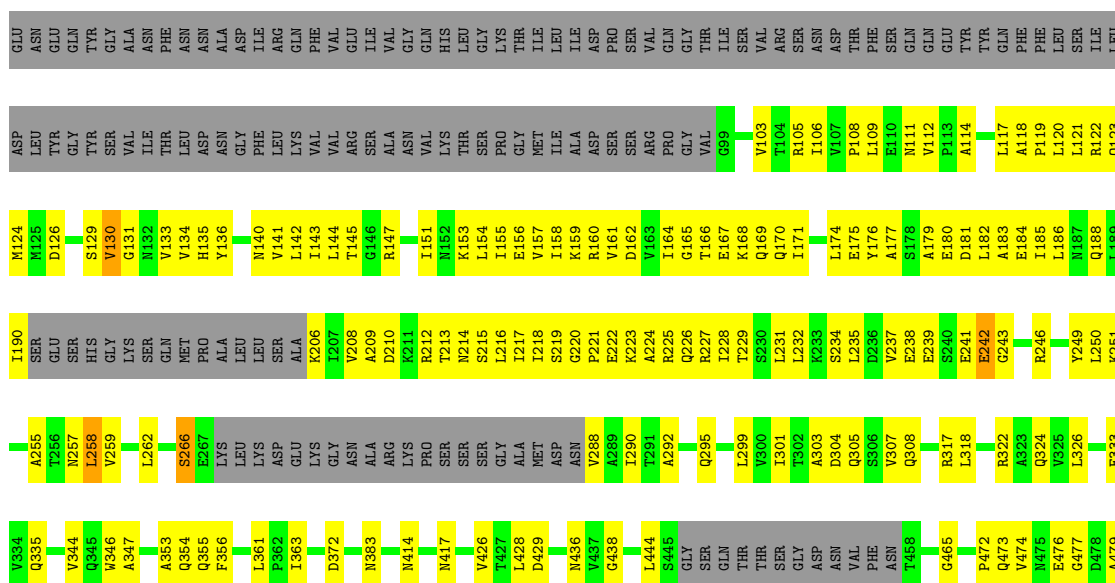
- Molecule 1: Putative type II secretion system protein D

Chain J:



- Molecule 1: Putative type II secretion system protein D

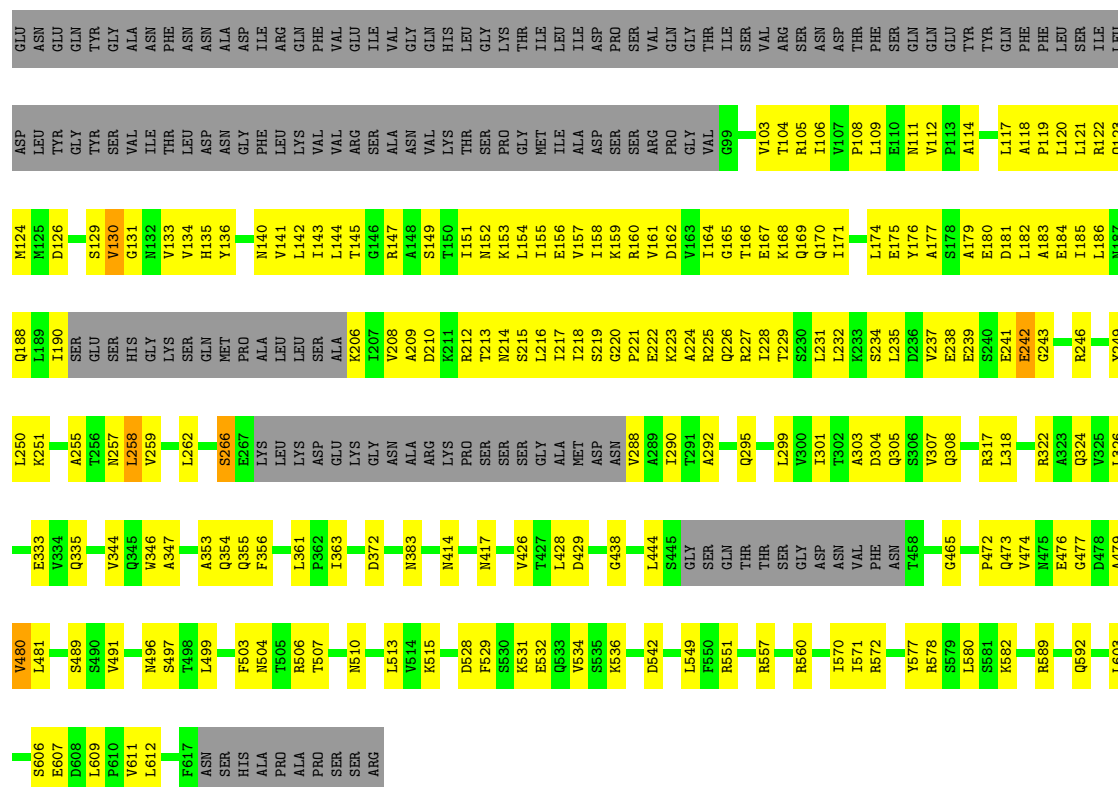
Chain K:





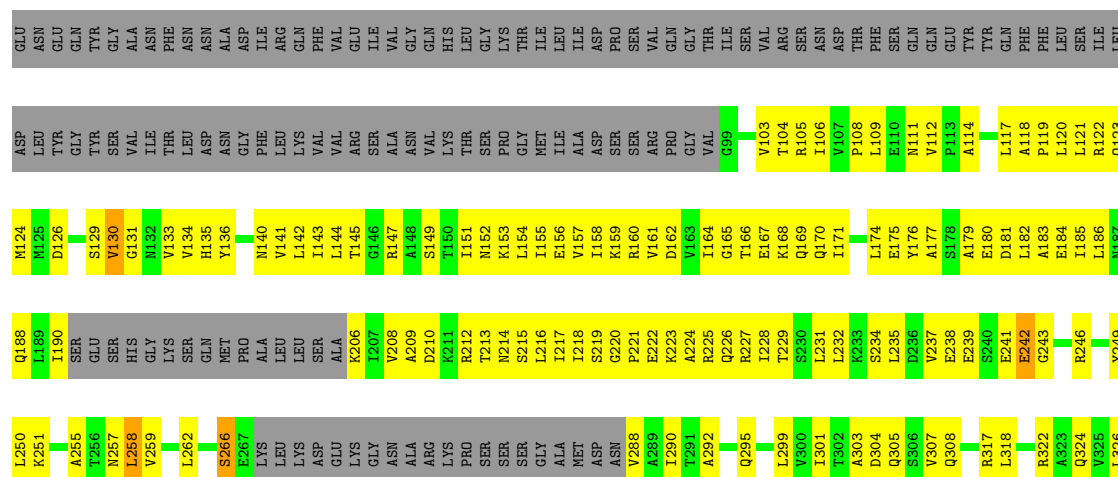
• Molecule 1: Putative type II secretion system protein D

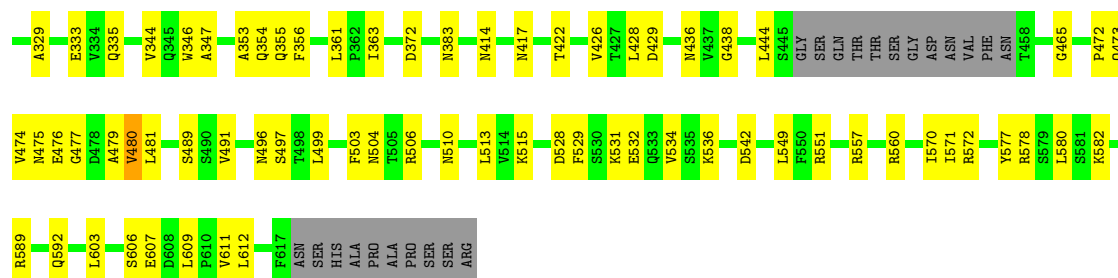
Chain L: 45% 30% 25%



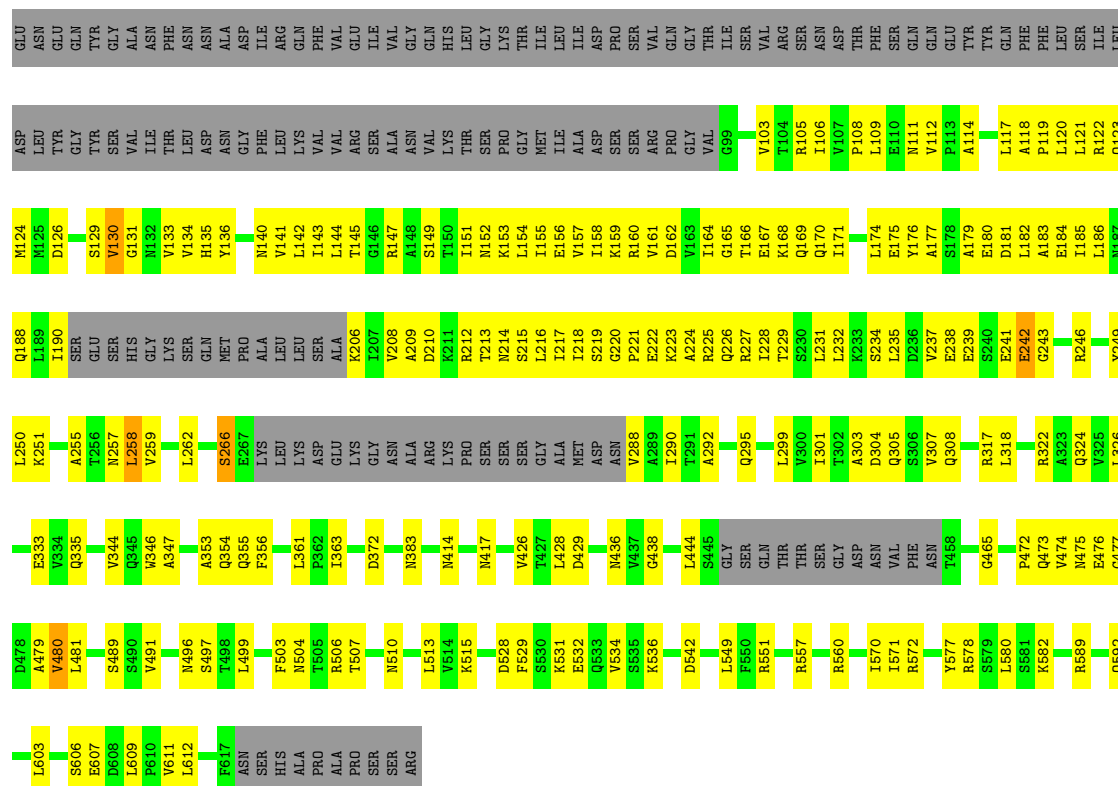
• Molecule 1: Putative type II secretion system protein D

Chain M: 44% 30% 25%

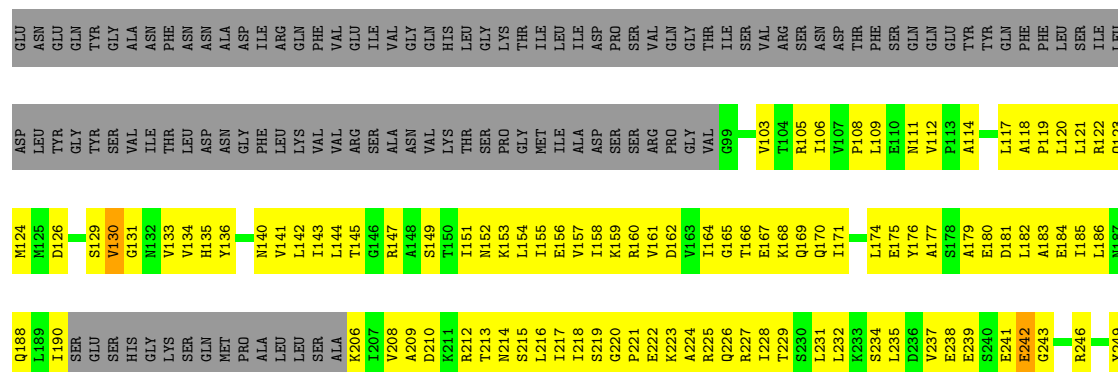




- Molecule 1: Putative type II secretion system protein D



- Molecule 1: Putative type II secretion system protein D



L250	E333	E476	R589
K251	V334	G477	Q592
A255	Q335	D478	L603
T256	V344	A479	S606
N257	Q345	V480	E607
L258	W346	L481	D608
V259	A347	S489	L609
L262	A353	S490	P610
S266	Q354	V491	V611
F267	Q355	N496	L612
LYS	F356	S497	F617
LEU	L361	T498	ASN
LYS	L361	L499	SER
ASP	P362	F503	HIS
GLU	I363	N504	ALA
LYS	D372	T505	PRO
ASN	N383	R506	ALA
ALA	M388	T507	PRO
ARG	N414	N510	SER
LYS	N417	L513	SER
PRO	V426	V514	ARG
SER	T427	K515	
SER	L428	D528	
GLY	D429	F529	
ALA	N436	S530	
MET	V437	K531	
ASP	G438	E532	
ASN	L444	Q533	
V288	S445	V534	
A289	GLY	S535	
I290	SER	K536	
T291	GLN	D542	
A292	L299	Q548	
Q295	V300	L549	
L299	I301	F550	
V300	T302	R551	
A303	THR	R557	
D304	SER	R560	
Q305	GLY	I570	
S306	ASP	I571	
V307	ASN	R572	
Q308	VAL	Y577	
R317	PHE	R578	
L318	ASN	S579	
R322	T458	L580	
A323	G465	S581	
Q324	P472	K582	
V325	Q473		
L326	V474		
	N475		

## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	30659	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

### 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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	A	0.42	0/3673	0.59	1/4983 (0.0%)
1	B	0.42	0/3673	0.59	1/4983 (0.0%)
1	C	0.42	0/3673	0.59	1/4983 (0.0%)
1	D	0.42	0/3673	0.59	1/4983 (0.0%)
1	E	0.42	0/3673	0.59	1/4983 (0.0%)
1	F	0.42	0/3673	0.59	1/4983 (0.0%)
1	G	0.42	0/3673	0.59	1/4983 (0.0%)
1	H	0.42	0/3673	0.59	1/4983 (0.0%)
1	I	0.42	0/3673	0.59	1/4983 (0.0%)
1	J	0.42	0/3673	0.59	1/4983 (0.0%)
1	K	0.42	0/3673	0.59	1/4983 (0.0%)
1	L	0.42	0/3673	0.59	1/4983 (0.0%)
1	M	0.42	0/3673	0.59	1/4983 (0.0%)
1	N	0.42	0/3673	0.59	1/4983 (0.0%)
1	O	0.42	0/3673	0.59	1/4983 (0.0%)
All	All	0.42	0/55095	0.59	15/74745 (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	B	0	1
1	C	0	1
1	D	0	1
1	E	0	1
1	F	0	1
1	G	0	1
1	H	0	1
1	I	0	1
1	J	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
1	K	0	1
1	L	0	1
1	M	0	1
1	N	0	1
1	O	0	1
All	All	0	15

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	266	SER	CB-CA-C	-5.47	99.70	110.10
1	L	266	SER	CB-CA-C	-5.47	99.70	110.10
1	B	266	SER	CB-CA-C	-5.46	99.73	110.10
1	D	266	SER	CB-CA-C	-5.46	99.73	110.10
1	O	266	SER	CB-CA-C	-5.45	99.74	110.10

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	166	THR	Peptide
1	B	166	THR	Peptide
1	C	166	THR	Peptide
1	D	166	THR	Peptide
1	E	166	THR	Peptide

## 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	3632	0	3702	225	0
1	B	3632	0	3702	226	0
1	C	3632	0	3702	227	0
1	D	3632	0	3702	226	0
1	E	3632	0	3702	224	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	3632	0	3702	225	0
1	G	3632	0	3702	227	0
1	H	3632	0	3702	224	0
1	I	3632	0	3702	228	0
1	J	3632	0	3702	233	0
1	K	3632	0	3702	224	0
1	L	3632	0	3702	223	0
1	M	3632	0	3702	227	0
1	N	3632	0	3702	227	0
1	O	3632	0	3702	227	0
All	All	54480	0	55530	2956	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 27.

The worst 5 of 2956 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:B:210:ASP:O	1:B:214:ASN:HA	1.50	1.11
1:E:210:ASP:O	1:E:214:ASN:HA	1.50	1.11
1:H:210:ASP:O	1:H:214:ASN:HA	1.50	1.11
1:I:210:ASP:O	1:I:214:ASN:HA	1.50	1.11
1:A:210:ASP:O	1:A:214:ASN:HA	1.50	1.11

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	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28 67
1	B	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28 67

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	D	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	E	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	F	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	G	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	H	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	I	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	J	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	K	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	L	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	M	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	N	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
1	O	464/627 (74%)	442 (95%)	19 (4%)	3 (1%)	28	67
All	All	6960/9405 (74%)	6630 (95%)	285 (4%)	45 (1%)	33	67

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	480	VAL
1	B	480	VAL
1	C	480	VAL
1	D	480	VAL
1	E	480	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	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	B	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	C	405/537 (75%)	404 (100%)	1 (0%)	94	98

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	E	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	F	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	G	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	H	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	I	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	J	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	K	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	L	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	M	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	N	405/537 (75%)	404 (100%)	1 (0%)	94	98
1	O	405/537 (75%)	404 (100%)	1 (0%)	94	98
All	All	6075/8055 (75%)	6060 (100%)	15 (0%)	95	98

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	258	LEU
1	H	258	LEU
1	M	258	LEU
1	F	258	LEU
1	L	258	LEU

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	368	GLN
1	I	305	GLN
1	N	473	GLN
1	G	475	ASN
1	H	324	GLN

### 5.3.3 RNA ⓘ

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.