



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 21, 2022 – 02:38 PM EDT

PDB ID : 6UJI  
Title : Low resolution crystal structure (5.5 Å) of the anthrax toxin protective antigen heptamer prepore D425A mutant  
Authors : Lovell, S.; Mehzabeen, N.; Battaile, K.P.; Bann, J.G.  
Deposited on : 2019-10-03  
Resolution : 5.50 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.27
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)	:	2.27

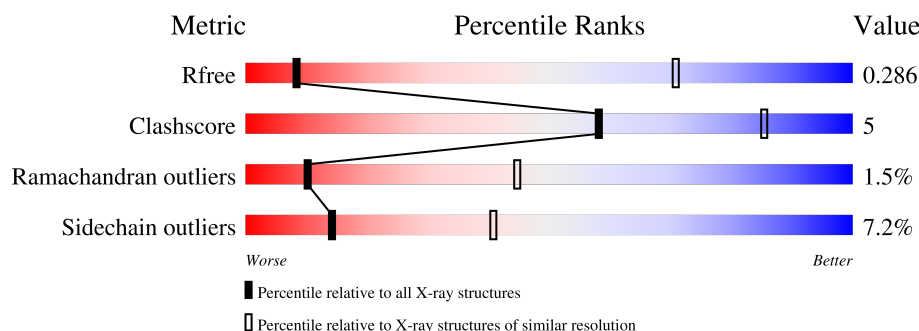
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 5.50 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	1019 (7.12-3.82)
Clashscore	141614	1010 (7.10-3.90)
Ramachandran outliers	138981	1014 (7.12-3.82)
Sidechain outliers	138945	1191 (7.20-3.80)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	568	67% 15% • 17%
1	B	568	72% 18% • 9%
1	C	568	57% 11% 32%
1	D	568	77% 16% • 5%
1	E	568	71% 19% • 9%
1	F	568	68% 13% • 18%
1	G	568	73% 17% • 8%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	568	<div><div></div><div>69%17%13%</div></div>
1	I	568	<div><div></div><div>74%18%7%</div></div>
1	J	568	<div><div></div><div>67%15%17%</div></div>
1	K	568	<div><div></div><div>68%16%14%</div></div>
1	L	568	<div><div></div><div>57%14%29%</div></div>
1	M	568	<div><div></div><div>65%18%17%</div></div>
1	N	568	<div><div></div><div>65%18%15%</div></div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 53490 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 Protective antigen PA-63.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	471	Total	C	N	O	S	0	0	0
			3712	2316	650	740	6			
1	B	519	Total	C	N	O	S	0	0	0
			4114	2578	710	820	6			
1	C	388	Total	C	N	O	S	0	0	0
			3083	1929	539	609	6			
1	D	539	Total	C	N	O	S	0	0	0
			4265	2672	735	852	6			
1	E	516	Total	C	N	O	S	0	0	0
			4089	2564	704	815	6			
1	F	467	Total	C	N	O	S	0	0	0
			3688	2316	641	725	6			
1	G	520	Total	C	N	O	S	0	0	0
			4125	2583	710	826	6			
1	H	493	Total	C	N	O	S	0	0	0
			3894	2434	678	776	6			
1	I	529	Total	C	N	O	S	0	0	0
			4189	2622	724	837	6			
1	J	470	Total	C	N	O	S	0	0	0
			3718	2322	649	741	6			
1	K	488	Total	C	N	O	S	0	0	0
			3867	2419	673	769	6			
1	L	404	Total	C	N	O	S	0	0	0
			3189	1992	555	637	5			
1	M	473	Total	C	N	O	S	0	0	0
			3749	2345	656	742	6			
1	N	480	Total	C	N	O	S	0	0	0
			3808	2375	663	764	6			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	425	ALA	ASP	engineered mutation	UNP P13423

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	425	ALA	ASP	engineered mutation	UNP P13423
C	425	ALA	ASP	engineered mutation	UNP P13423
D	425	ALA	ASP	engineered mutation	UNP P13423
E	425	ALA	ASP	engineered mutation	UNP P13423
F	425	ALA	ASP	engineered mutation	UNP P13423
G	425	ALA	ASP	engineered mutation	UNP P13423
H	425	ALA	ASP	engineered mutation	UNP P13423
I	425	ALA	ASP	engineered mutation	UNP P13423
J	425	ALA	ASP	engineered mutation	UNP P13423
K	425	ALA	ASP	engineered mutation	UNP P13423
L	425	ALA	ASP	engineered mutation	UNP P13423
M	425	ALA	ASP	engineered mutation	UNP P13423
N	425	ALA	ASP	engineered mutation	UNP P13423



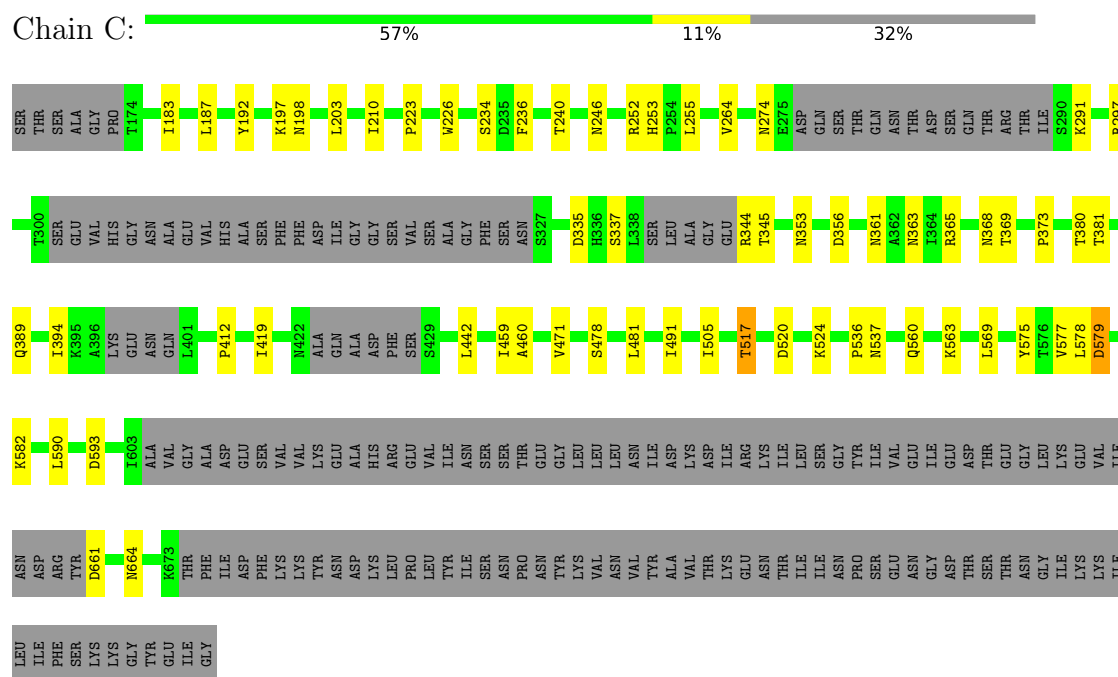
- Molecule 1: Protective antigen PA-63

LEU	TYR	R600	M422	V303	SER
ILE	ILE	I603	GLN	H304	THR
SER	SER	ALA	ALA	G305	ALA
ASN	ASN	K613	ASP	N306	GLY
PRO	PRO	E614	PHE	A307	PRO
ASN	ASN	V619	SER	E308	T174
TYR	TYR	ILE	S429	V309	
LYS	LYS	ASN	H310	H310	L187
VAL	VAL	ASN	L442	A311	I188
ASN	ASN	SER	D451	S312	V189
VAL	VAL	SER	THR	S319	E190
TYR	TYR	GLU	I459		G191
VAL	VAL	GLY	A460	F324	Y192
T702	T702	LEU	LEU		K197
		LEU	V471	D335	
I708	I708	LEU	LEU	H336	I210
		M630	M477	S337	
T716	T716	I631	S478	LEU	P223
		G639		SER	E224
T724	T724	SER	L481	LEU	D231
L725	L725	GLY	I484	ALA	P232
ILE	PHE	TYR	I491	GLY	
SER	SER	ILE	THR	GLU	D235
LYS	LYS	VAL	ILE	R344	P236
LYS	LYS	GLU	I505	T345	E237
GLY	GLY	ILE	ILE	N353	K238
TYR	TYR	GLU	P513		V239
GLU	GLU	ASP	THR	D356	T240
GLY	GLY	THR	T517		
		GLU	GLY	N361	N246
		GLY	D520	A362	
		LEU	LYS	N363	R252
		LYS	K524	I364	H253
		GLU	P536	R365	P254
		VAL	ILE	T369	L255
		ASN	N537		V264
		ASN	F554	P373	
		R659	S559	T380	N274
		I660	Q560	T381	E275
		D661			ASP
		N664	K563	Q389	GLN
					SER
					THR
					GLN
					ASN
					THR
					ASP
					LYS
					GLU
					ASN
					GLN
					THR
					ARG
					T288
					K291
					P007

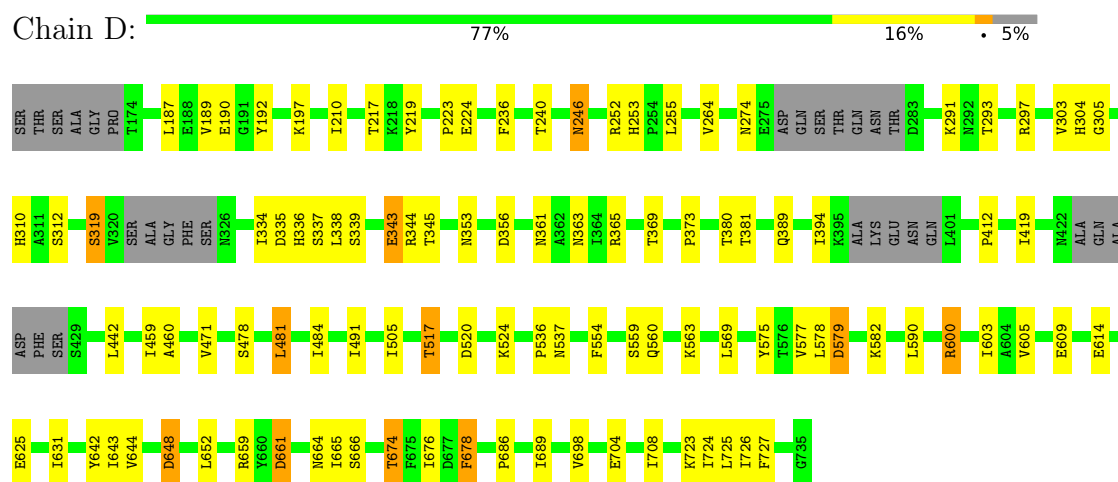
- Molecule 1: Protective antigen PA-63

[illegible]

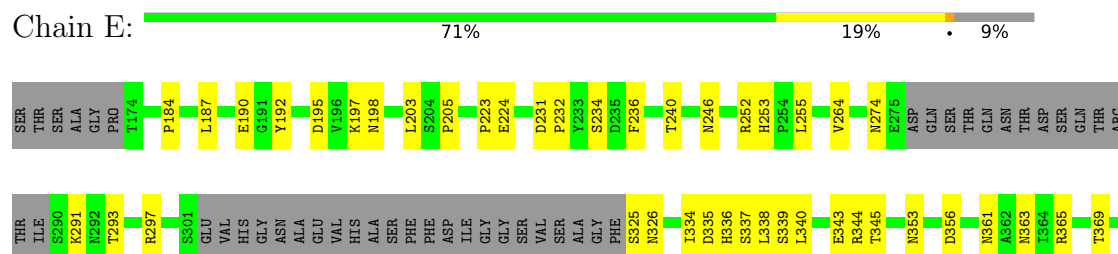
• Molecule 1: Protective antigen PA-63

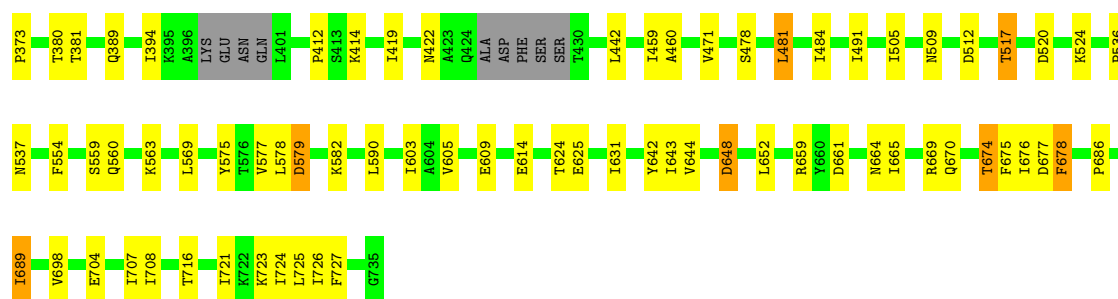


• Molecule 1: Protective antigen PA-63



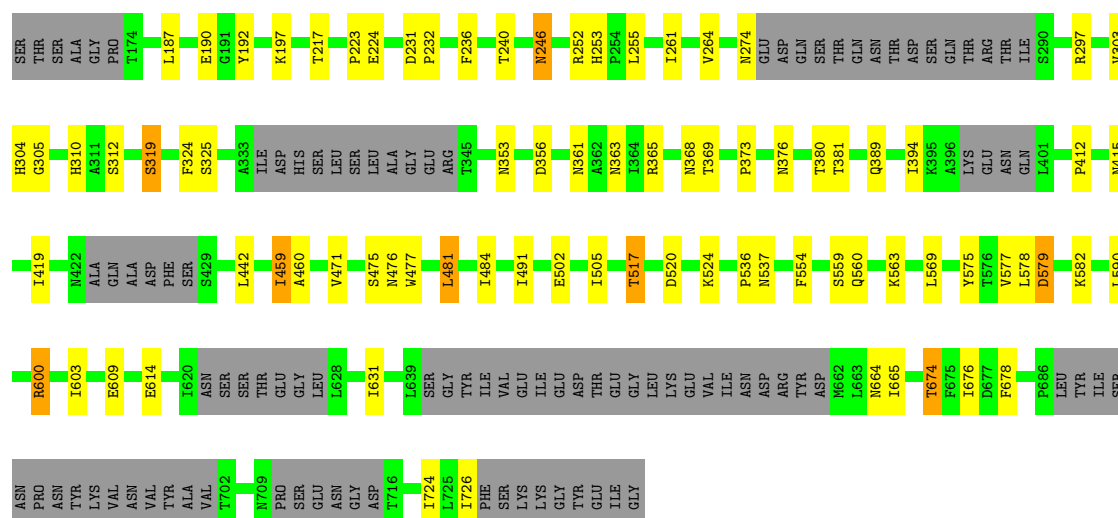
• Molecule 1: Protective antigen PA-63





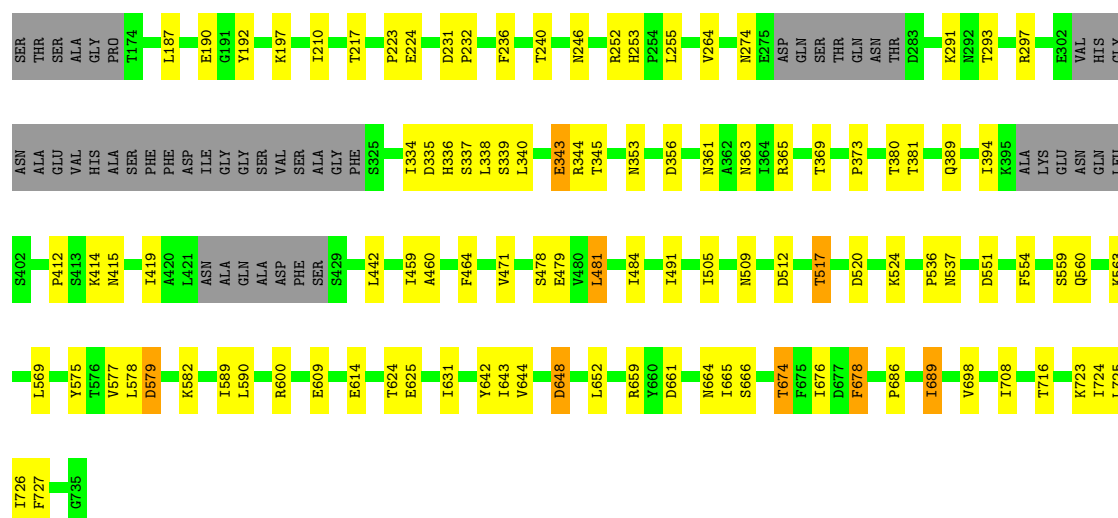
• Molecule 1: Protective antigen PA-63

Chain F: 68% 13% 18%



• Molecule 1: Protective antigen PA-63

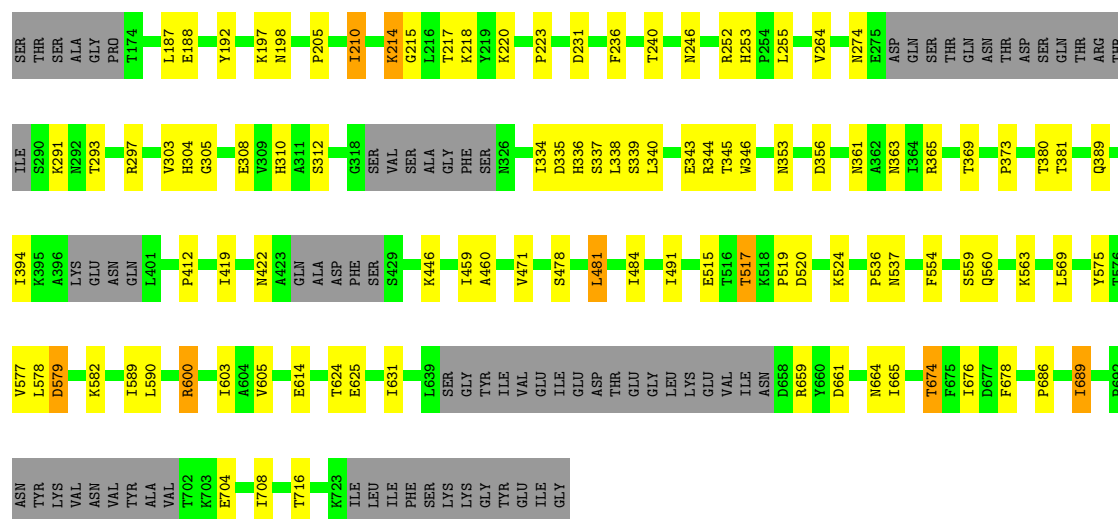
Chain G: 73% 17% 8%



• Molecule 1: Protective antigen PA-63

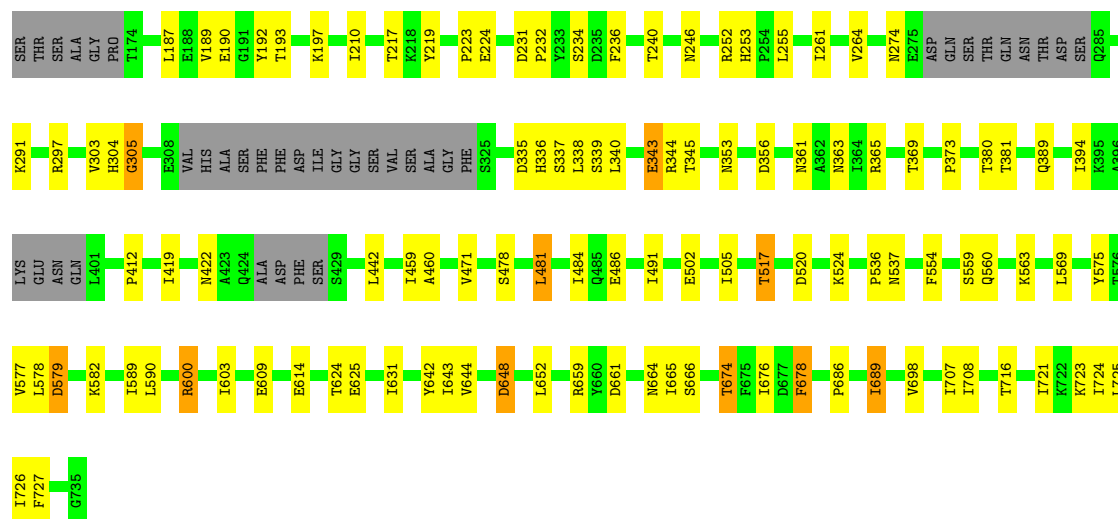


Chain H:  69% 17% 13%



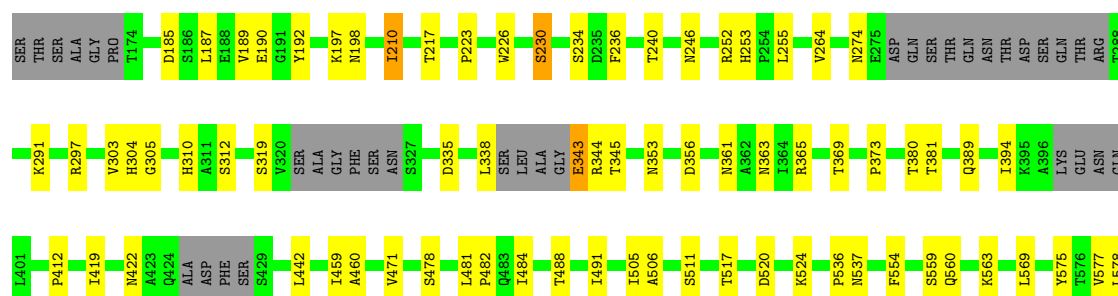
• Molecule 1: Protective antigen PA-63

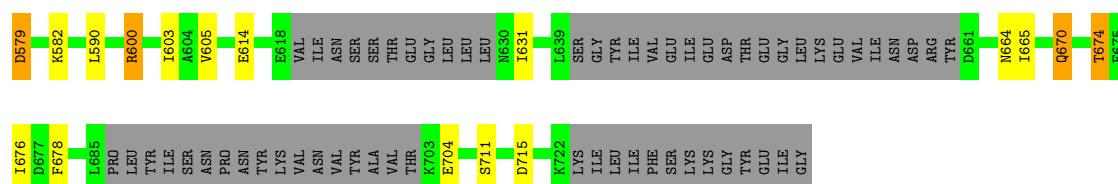
Chain I:  74% 18% 7%



• Molecule 1: Protective antigen PA-63

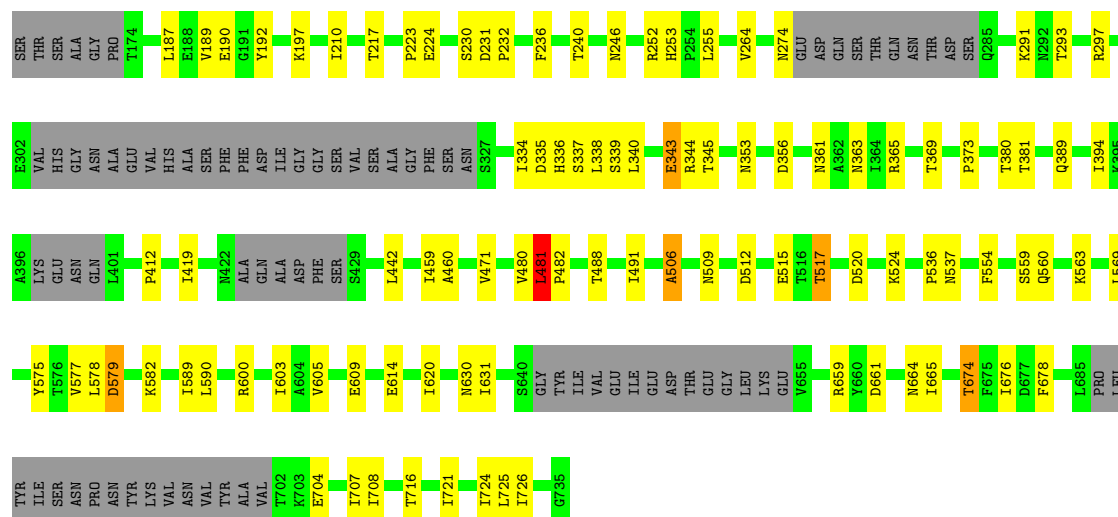
Chain J:  67% 15% 17%





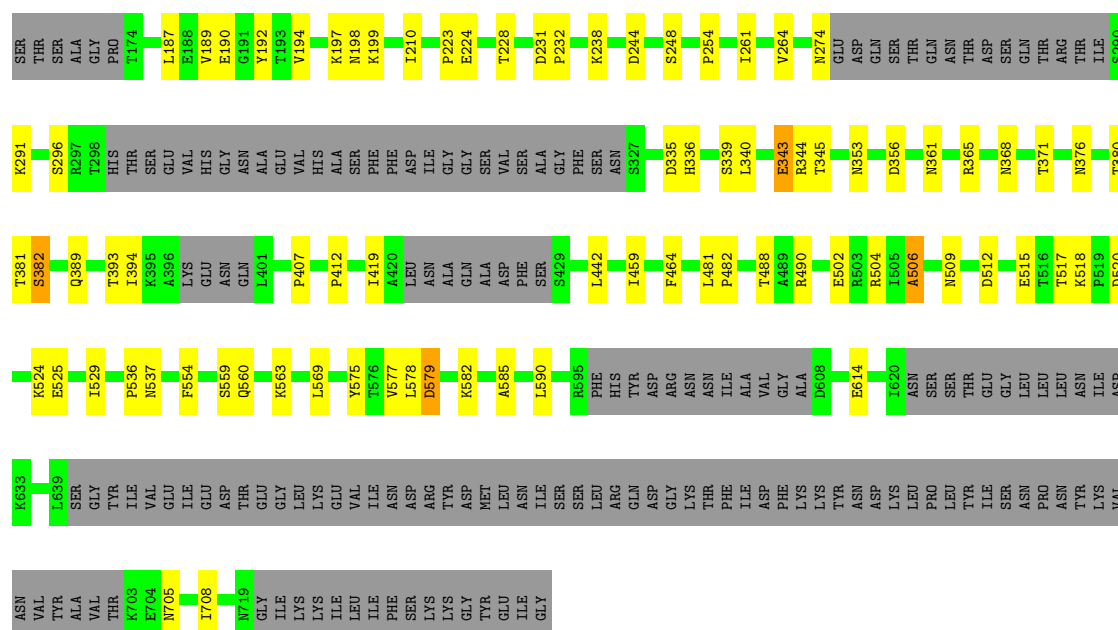
• Molecule 1: Protective antigen PA-63

Chain K: 68% 16% 14%

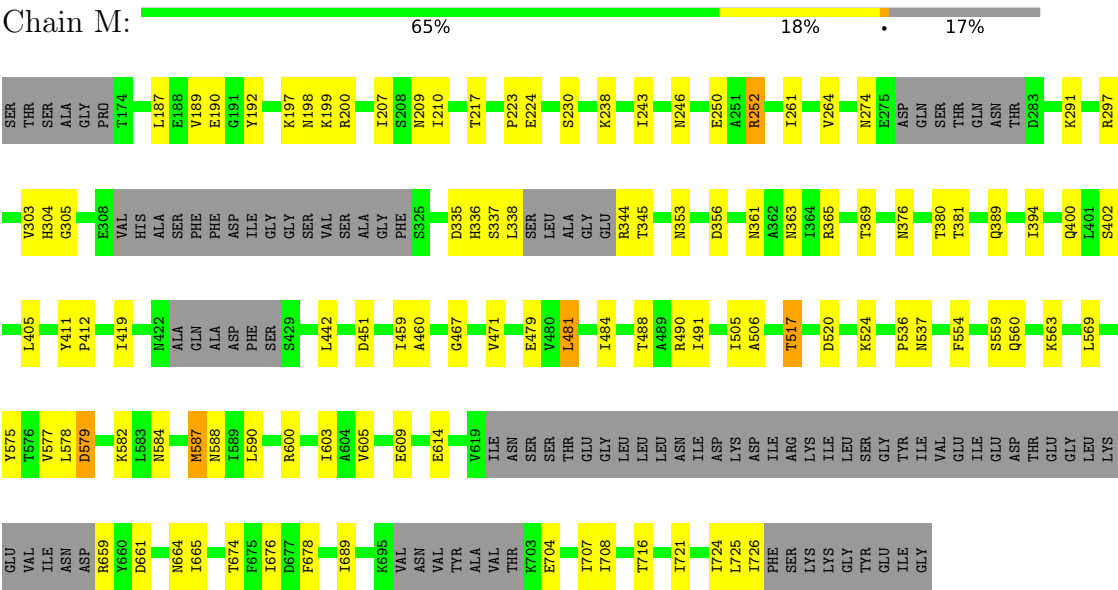


• Molecule 1: Protective antigen PA-63

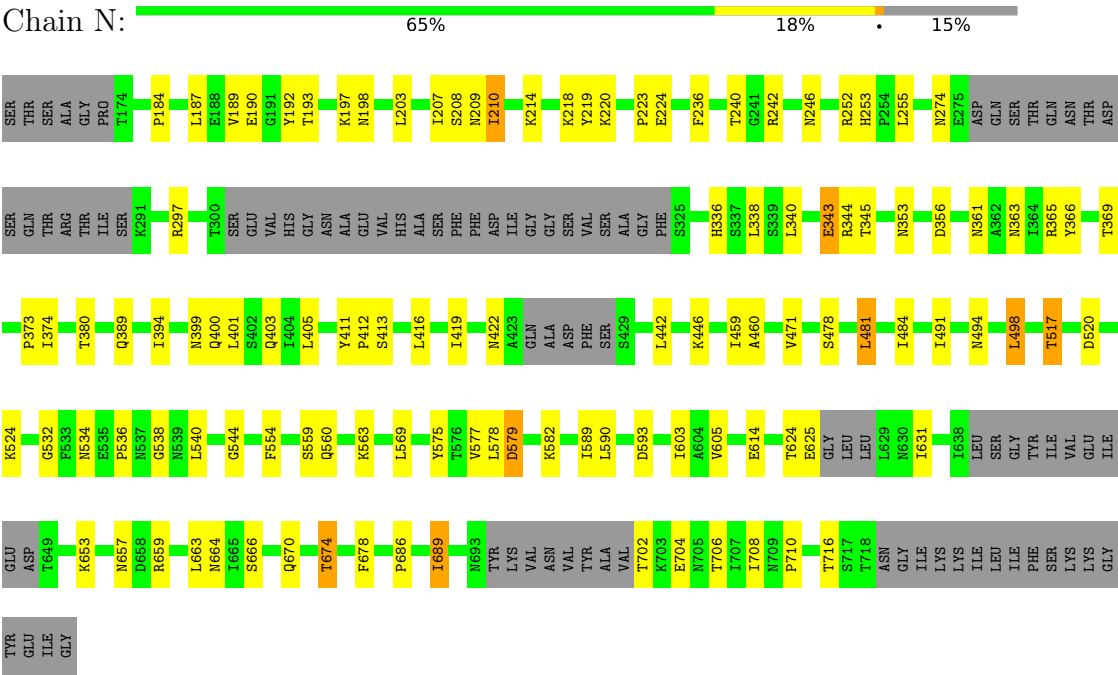
Chain L: 57% 14% 29%



• Molecule 1: Protective antigen PA-63



● Molecule 1: Protective antigen PA-63



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	166.18Å 144.25Å 304.82Å 90.00° 102.41° 90.00°	Depositor
Resolution (Å)	44.83 – 5.50 49.78 – 5.50	Depositor EDS
% Data completeness (in resolution range)	99.3 (44.83-5.50) 99.3 (49.78-5.50)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.34 (at 5.39Å)	Xtriage
Refinement program	BUSTER	Depositor
R, $R_{free}$	0.251 , 0.278 0.268 , 0.286	Depositor DCC
$R_{free}$ test set	2250 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	183.8	Xtriage
Anisotropy	0.481	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 204.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	53490	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	257.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.29% 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 [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	0.37	0/3771	0.58	0/5109
1	B	0.37	0/4181	0.58	0/5667
1	C	0.36	0/3134	0.58	0/4250
1	D	0.35	0/4336	0.58	0/5877
1	E	0.36	0/4156	0.58	0/5633
1	F	0.37	0/3746	0.58	0/5072
1	G	0.35	0/4192	0.57	0/5681
1	H	0.36	0/3958	0.57	0/5364
1	I	0.35	0/4257	0.58	0/5770
1	J	0.38	0/3776	0.60	0/5112
1	K	0.38	0/3926	0.58	0/5314
1	L	0.40	0/3237	0.60	0/4388
1	M	0.39	0/3811	0.60	0/5166
1	N	0.39	0/3868	0.60	0/5243
All	All	0.37	0/54349	0.58	0/73646

There are no bond length outliers.

There are no bond angle outliers.

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	3712	0	3666	39	0
1	B	4114	0	4098	48	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3083	0	3058	20	0
1	D	4265	0	4229	40	0
1	E	4089	0	4071	46	0
1	F	3688	0	3679	29	0
1	G	4125	0	4102	42	0
1	H	3894	0	3852	37	0
1	I	4189	0	4164	45	0
1	J	3718	0	3673	31	0
1	K	3867	0	3860	39	0
1	L	3189	0	3177	33	0
1	M	3749	0	3697	40	0
1	N	3808	0	3769	45	0
All	All	53490	0	53095	507	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 5.

The worst 5 of 507 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:642:TYR:HB2	1:G:665:ILE:HD11	1.57	0.87
1:I:642:TYR:HB2	1:I:665:ILE:HD11	1.58	0.83
1:B:642:TYR:HB2	1:B:665:ILE:HD11	1.60	0.81
1:N:297:ARG:HD3	1:N:369:THR:HG21	1.68	0.75
1:K:481:LEU:H	1:K:482:PRO:HD2	1.56	0.71

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	455/568 (80%)	407 (90%)	39 (9%)	9 (2%)	7	37
1	B	509/568 (90%)	454 (89%)	49 (10%)	6 (1%)	13	49
1	C	374/568 (66%)	335 (90%)	35 (9%)	4 (1%)	14	51
1	D	529/568 (93%)	473 (89%)	49 (9%)	7 (1%)	12	47
1	E	506/568 (89%)	457 (90%)	43 (8%)	6 (1%)	13	49
1	F	449/568 (79%)	402 (90%)	41 (9%)	6 (1%)	12	47
1	G	510/568 (90%)	460 (90%)	45 (9%)	5 (1%)	15	53
1	H	479/568 (84%)	424 (88%)	46 (10%)	9 (2%)	8	38
1	I	519/568 (91%)	460 (89%)	52 (10%)	7 (1%)	12	47
1	J	452/568 (80%)	401 (89%)	39 (9%)	12 (3%)	5	31
1	K	474/568 (84%)	424 (90%)	44 (9%)	6 (1%)	12	47
1	L	388/568 (68%)	347 (89%)	36 (9%)	5 (1%)	12	47
1	M	459/568 (81%)	403 (88%)	48 (10%)	8 (2%)	9	41
1	N	466/568 (82%)	415 (89%)	40 (9%)	11 (2%)	6	33
All	All	6569/7952 (83%)	5862 (89%)	606 (9%)	101 (2%)	10	45

5 of 101 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	587	MET
1	N	413	SER
1	B	422	ASN
1	I	304	HIS
1	J	715	ASP

### 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	418/504 (83%)	389 (93%)	29 (7%)	15	42
1	B	465/504 (92%)	434 (93%)	31 (7%)	16	42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	349/504 (69%)	327 (94%)	22 (6%)	18	44
1	D	482/504 (96%)	448 (93%)	34 (7%)	14	41
1	E	462/504 (92%)	429 (93%)	33 (7%)	14	41
1	F	416/504 (82%)	385 (92%)	31 (8%)	13	39
1	G	468/504 (93%)	435 (93%)	33 (7%)	14	41
1	H	439/504 (87%)	408 (93%)	31 (7%)	14	41
1	I	473/504 (94%)	440 (93%)	33 (7%)	15	41
1	J	419/504 (83%)	389 (93%)	30 (7%)	14	41
1	K	438/504 (87%)	406 (93%)	32 (7%)	14	40
1	L	362/504 (72%)	335 (92%)	27 (8%)	13	39
1	M	419/504 (83%)	383 (91%)	36 (9%)	10	34
1	N	432/504 (86%)	396 (92%)	36 (8%)	11	36
All	All	6042/7056 (86%)	5604 (93%)	438 (7%)	14	41

5 of 438 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	517	THR
1	J	361	ASN
1	M	689	ILE
1	H	603	ILE
1	I	481	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	485	GLN
1	N	454	GLN
1	J	476	ASN
1	N	485	GLN
1	M	483	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 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.