



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2023 – 02:58 AM EDT

PDB ID : 3NG9  
Title : Structure to Function Correlations for Adeno-associated Virus Serotype 1  
Authors : Govindasamy, L.; Miller, E.B.; Gurda, B.; McKenna, R.; Zolotukhin, S.; Muzy-  
czka, N.; Agbandje-McKenna, M.  
Deposited on : 2010-06-11  
Resolution : 2.50 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 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

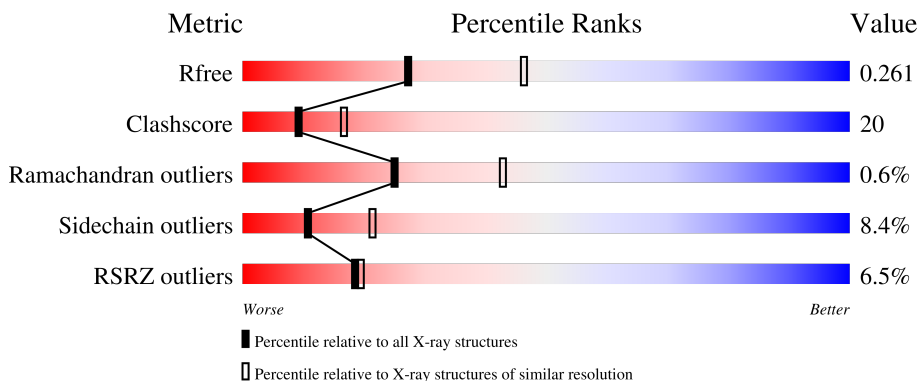
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	736	 2% 47% 20% 29%
1	B	736	 4% 46% 21% 29%
1	C	736	 4% 48% 20% 29%
1	D	736	 2% 46% 21% 29%
1	E	736	 5% 47% 21% 29%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	736	 3% 46% 22% 29%
1	G	736	 5% 46% 22% 29%
1	H	736	 3% 47% 20% 29%
1	I	736	 5% 46% 21% 29%
1	J	736	 11% 49% 18% 29%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ADE	F	737	-	-	-	X
3	CYT	C	738	-	-	-	X
3	CYT	E	738	-	-	-	X
3	CYT	H	738	-	-	-	X

## 2 Entry composition [i](#)

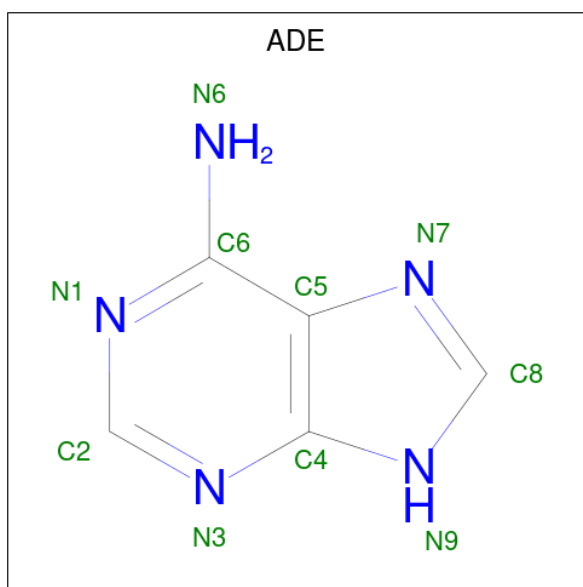
There are 4 unique types of molecules in this entry. The entry contains 42680 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Capsid protein.

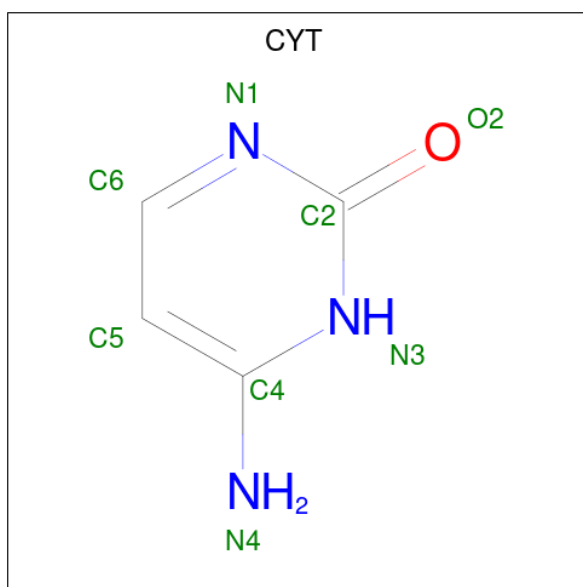
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	520	4120	2606	710	788	16	0	0	0
1	B	520	4120	2606	710	788	16	0	0	0
1	C	520	4120	2606	710	788	16	0	0	0
1	D	520	4120	2606	710	788	16	0	0	0
1	E	520	4120	2606	710	788	16	0	0	0
1	F	520	4120	2606	710	788	16	0	0	0
1	G	520	4120	2606	710	788	16	0	0	0
1	H	520	4120	2606	710	788	16	0	0	0
1	I	520	4120	2606	710	788	16	0	0	0
1	J	520	4120	2606	710	788	16	0	0	0

- Molecule 2 is ADENINE (three-letter code: ADE) (formula: C<sub>5</sub>H<sub>5</sub>N<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	N	0	0
			10	5	5		
2	B	1	Total	C	N	0	0
			10	5	5		
2	C	1	Total	C	N	0	0
			10	5	5		
2	D	1	Total	C	N	0	0
			10	5	5		
2	E	1	Total	C	N	0	0
			10	5	5		
2	F	1	Total	C	N	0	0
			10	5	5		
2	G	1	Total	C	N	0	0
			10	5	5		
2	H	1	Total	C	N	0	0
			10	5	5		
2	I	1	Total	C	N	0	0
			10	5	5		
2	J	1	Total	C	N	0	0
			10	5	5		

- Molecule 3 is 6-AMINOPYRIMIDIN-2(1H)-ONE (three-letter code: CYT) (formula: C<sub>4</sub>H<sub>5</sub>N<sub>3</sub>O).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			8	4	3	1		
3	B	1	Total	C	N	O	0	0
			8	4	3	1		
3	B	1	Total	C	N	O	0	0
			8	4	3	1		
3	C	1	Total	C	N	O	0	0
			8	4	3	1		
3	D	1	Total	C	N	O	0	0
			8	4	3	1		
3	E	1	Total	C	N	O	0	0
			8	4	3	1		
3	F	1	Total	C	N	O	0	0
			8	4	3	1		
3	H	1	Total	C	N	O	0	0
			8	4	3	1		
3	I	1	Total	C	N	O	0	0
			8	4	3	1		
3	J	1	Total	C	N	O	0	0
			8	4	3	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	131	Total	O	0	0
			131	131		
4	B	130	Total	O	0	0
			130	130		

Continued on next page...

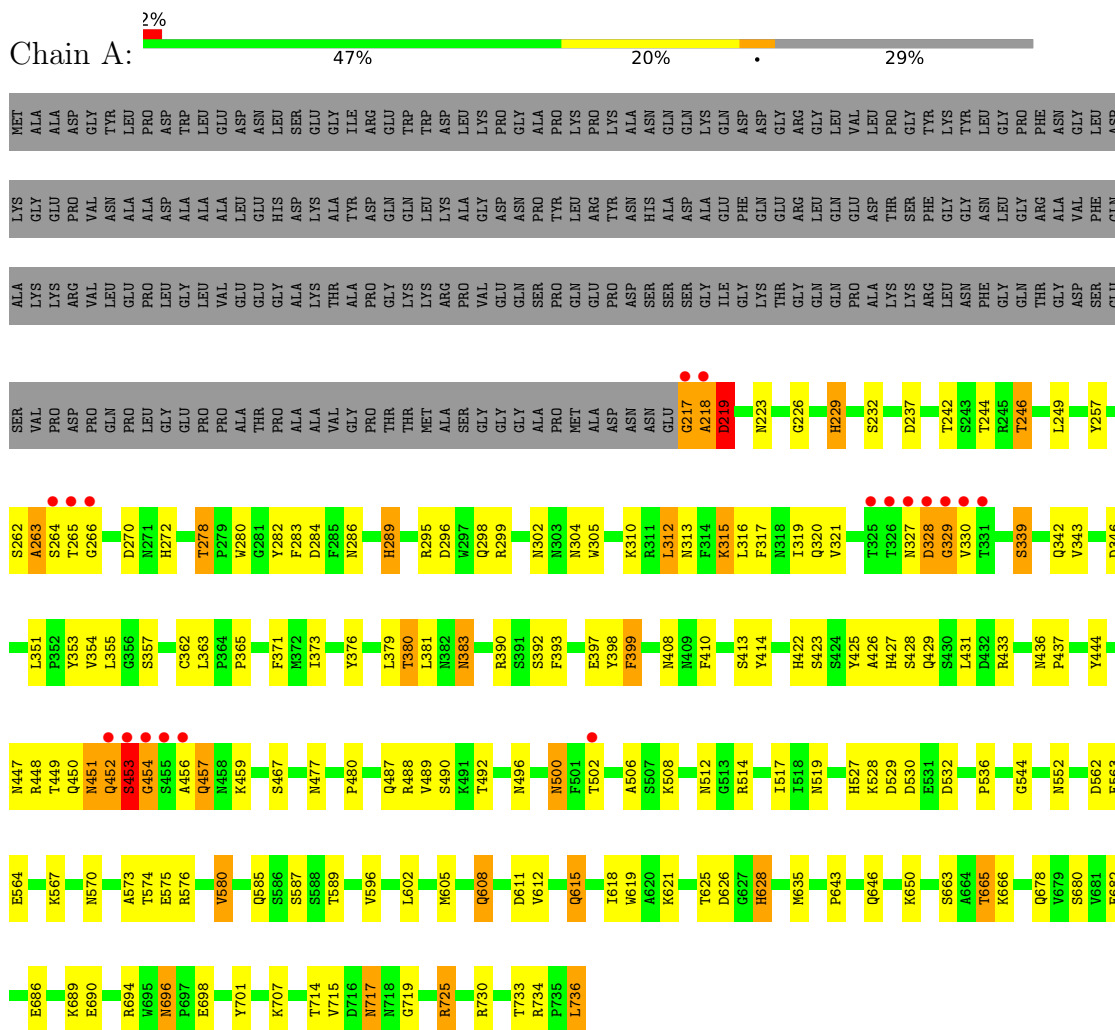
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	C	128	Total 128	O 128	0	0
4	D	131	Total 131	O 131	0	0
4	E	133	Total 133	O 133	0	0
4	F	131	Total 131	O 131	0	0
4	G	127	Total 127	O 127	0	0
4	H	130	Total 130	O 130	0	0
4	I	130	Total 130	O 130	0	0
4	J	129	Total 129	O 129	0	0

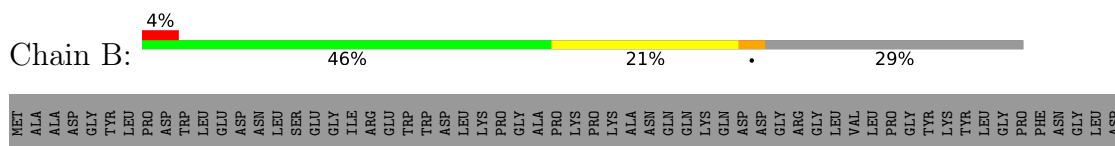
### 3 Residue-property plots [i](#)

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.

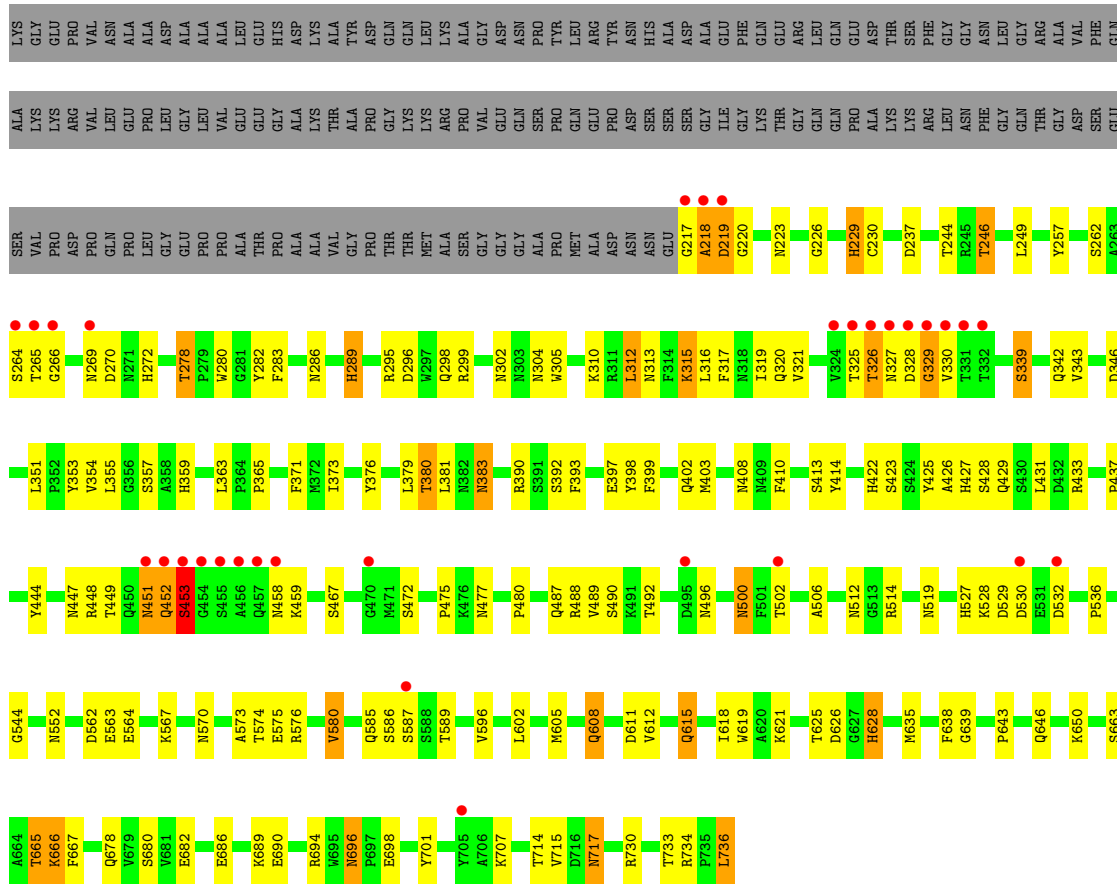
- Molecule 1: Capsid protein



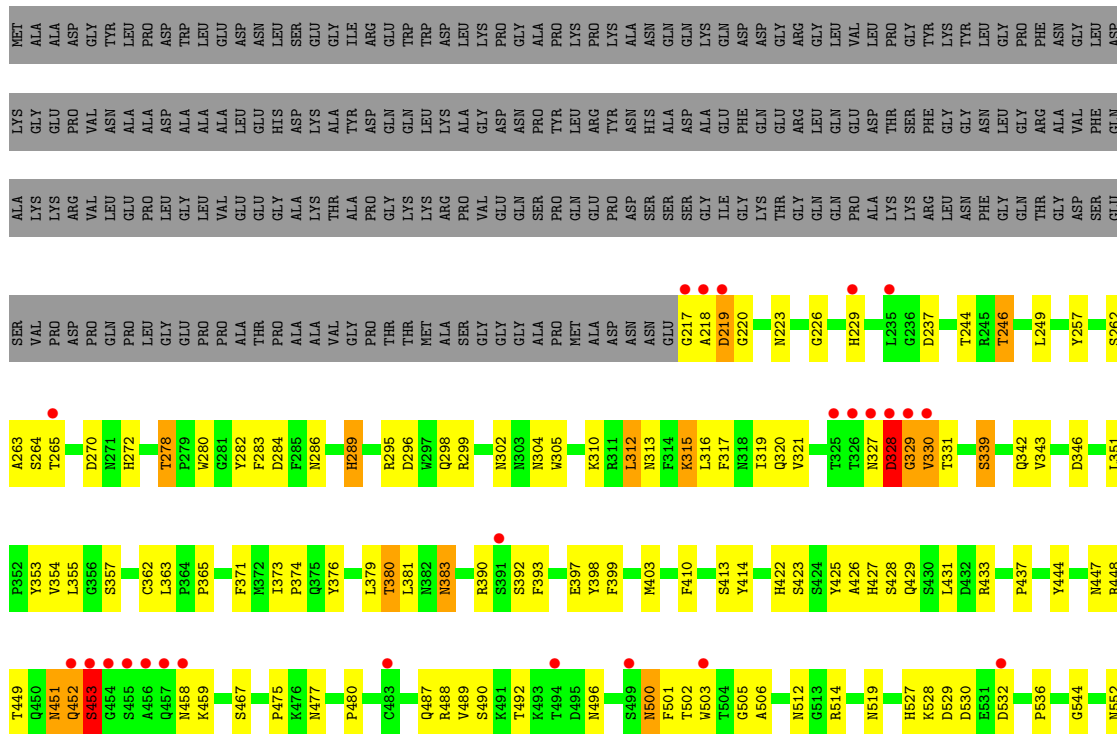
- Molecule 1: Capsid protein

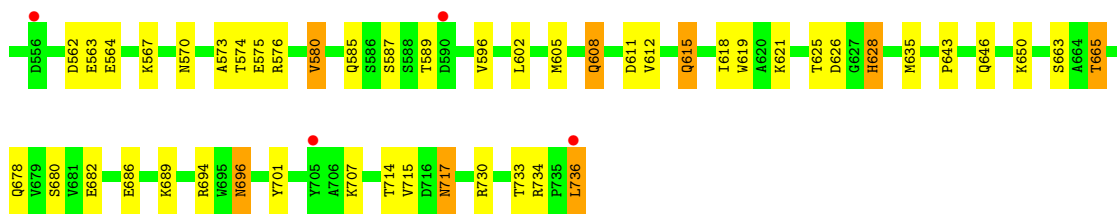




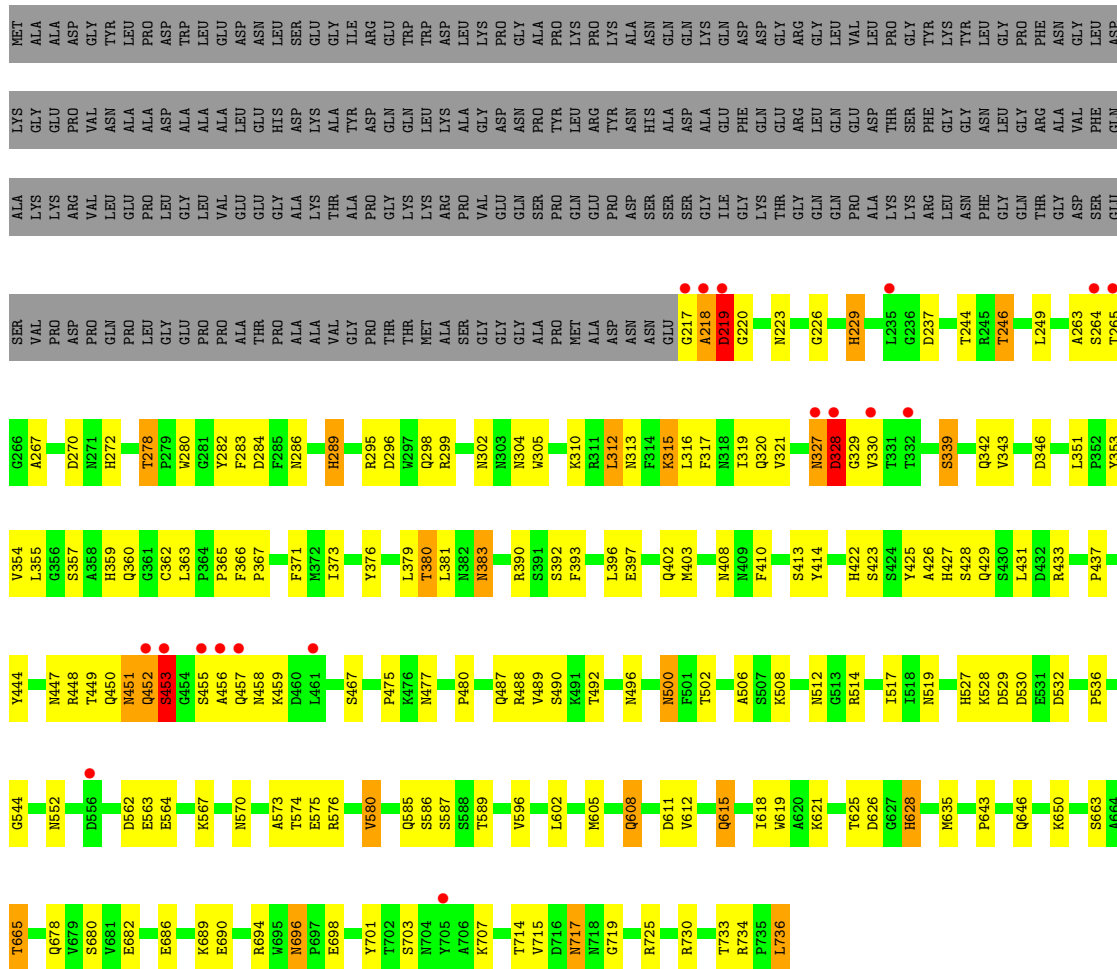


● Molecule 1: Capsid protein

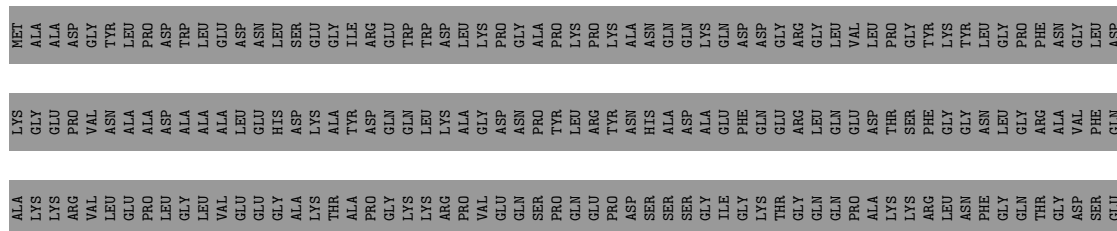


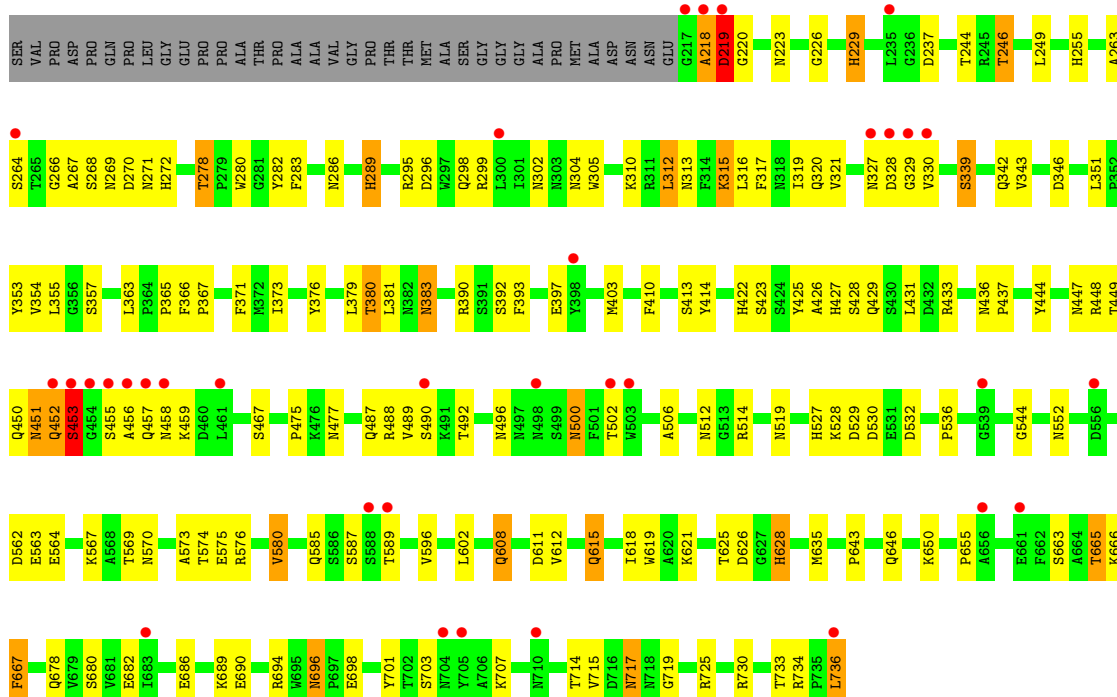


● Molecule 1: Capsid protein

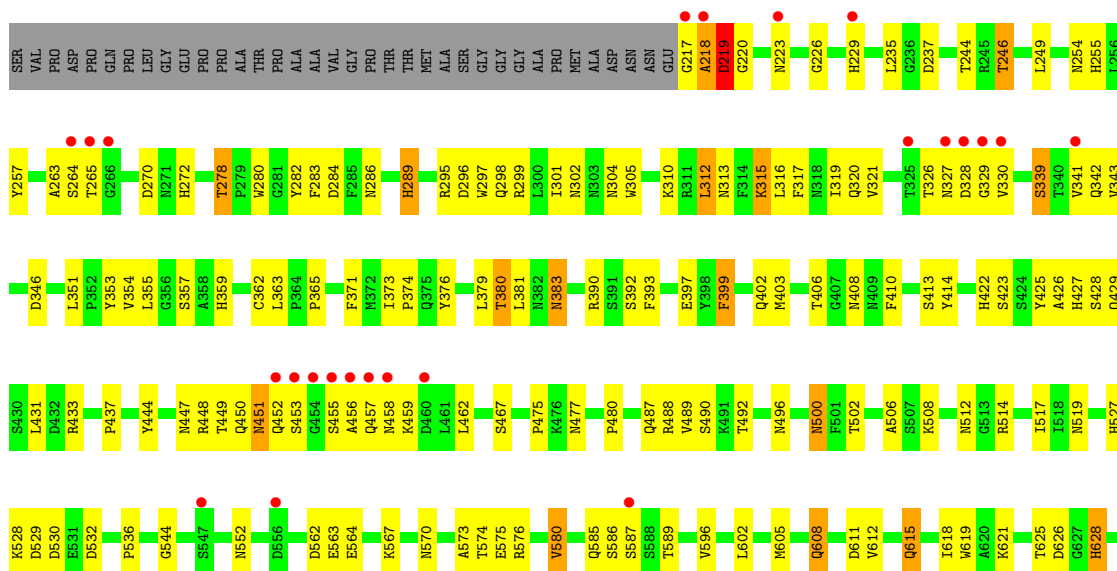
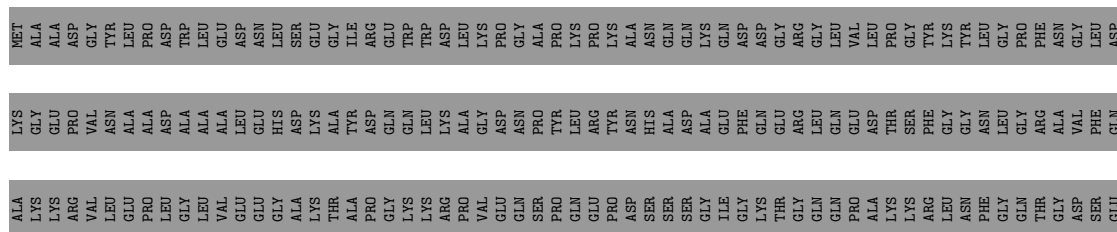


● Molecule 1: Capsid protein

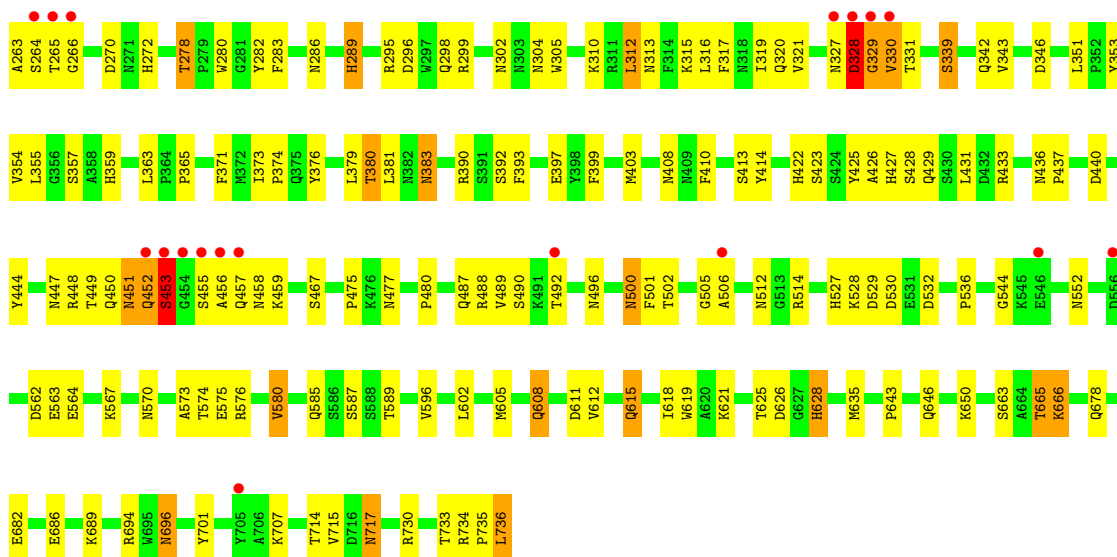




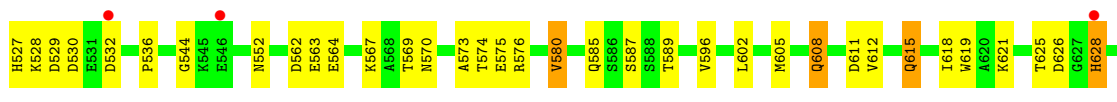
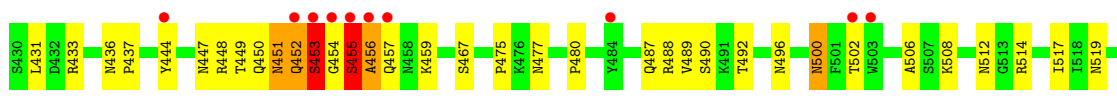
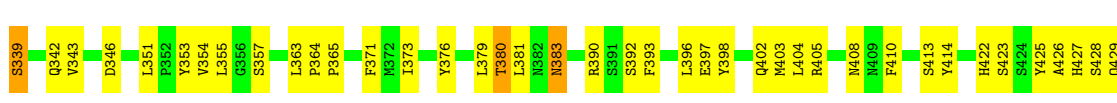
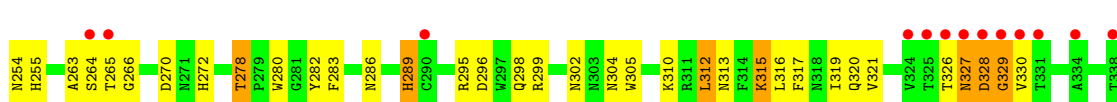
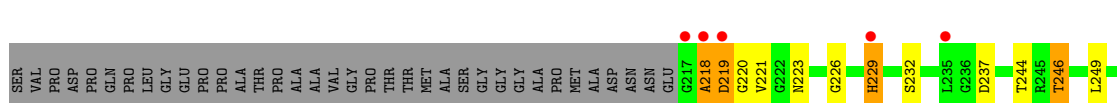
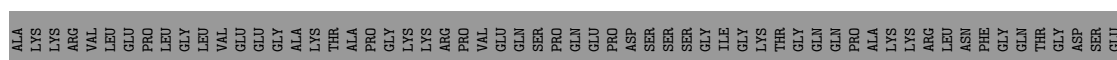
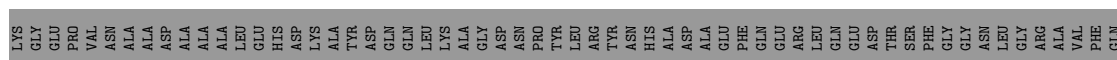
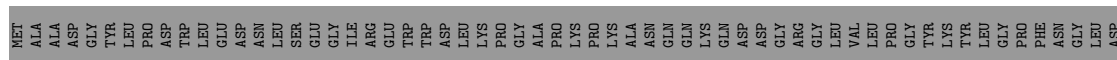
● Molecule 1: Capsid protein



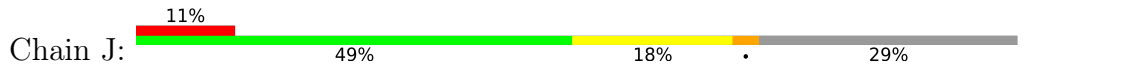




• Molecule 1: Capsid protein



• Molecule 1: Capsid protein



MET	ALA	ALA	ALA	GLY	GLY	LEU	PRO	PRO	ASP	TRP	LEU	GLU	ASP	ASN	LEU	SER	SER	GLY	GLY	ILE	ARG	GLU	TRP	TRP	LEU	TRP	ASP	LEU	LYS	PRO	LYS	PRO	LEU	GLN	GLY	GLN	GLY	GLY	ASP	ASP	GLN	LEU	VAL	LEU	LEU	PRO	GLY	PRO	PHE	ASN	GLY	LEU	ASP			
LYS	GLY	GLU	PRO	ASP	VAL	ASN	ALA	ALA	ASP	ALA	ALA	VAL	VAL	GLU	GLU	HIS	LEU	LYS	LYS	ALA	TYR	ALA	GLY	GLN	GLN	LEU	LEU	LYS	ALA	GLY	GLY	ASP	ASN	GLY	ASP	GLY	PHE	GLN	GLN	GLY	ASP	GLY	VAL	LEU	VAL	GLY	PRO	GLY	GLY	ASN	PHE	ARG	ALA	VAL	PHE	GLN
ALA	LYS	ARG	VAL	LEU	PRO	PRO	LEU	LEU	GLY	GLY	THR	THR	ALA	VAL	PRO	GLY	ALA	THR	ALA	ALA	PRO	VAL	VAL	GLY	GLY	LEU	LEU	ARG	PRO	VAL	GLY	GLY	GLY	LEU	ILE	GLY	GLY	PHE	GLN	GLN	GLY	THR	THR	THR	LEU	VAL	LEU	GLY	GLY	GLN	GLY	GLY	ASP	GLU		
SER	VAL	ASP	GLN	PRO	LEU	GLY	GLU	PRO	PRO	ALA	THR	R295	T278	P279	W280	G281	Y282	F283	M286	V287	H288	R295	D296	W297	R299	N302	N303	N304	W305	K310	R311	N313	F314	K315	L316	F317	N318	L319	Q320	Y321	K322	E323	N327	D237	S243	T244	E245	T246	L249	R253	H254	A263	S264			
T266	G266	A267	S268	N269	D270	Y273	T278	P279	W280	G281	Y282	F283	M286	V287	H288	R295	D296	W297	R299	N302	N303	N304	W305	K310	R311	N313	F314	K315	L316	F317	N318	L319	Q320	Y321	K322	E323	N327	D237	S243	T244	E245	T246	L249	R253	H254	A263	S264									
L351	P352	Y353	V354	L355	G356	S357	A358	H359	L363	Q364	P365	F366	F371	M372	L373	Y376	T380	L381	N382	R383	G384	S385	Q386	A387	R390	S391	S392	C395	L396	E397	Y398	M403	G407	M408	M409	F410	S413	Q329	Y330	F331	T332	H422	S423	S424	A426	H427	S428	Q429								
S430	L431	R432	R433	L438	I439	T449	Q450	M451	Q452	S453	G454	S455	A456	Q457	M458	K459	D460	L461	L462	F463	S467	P480	Q487	R488	V489	S490	K491	T492	K493	T494	D495	M496	R497	M498	S499	N500	F501	T502	W503	T504	G505	W519	A620	K621	H628	P643	Q646	K650	P655	A656						
D532	K533	F534	F535	P536	E546	D562	E563	E564	K567	N570	A573	T574	E575	R576	V580	F584	Q585	S586	S587	S588	T589	V596	A601	L602	M605	Q608	D611	V612	W613	L614	Q615	I618	W619	A620	K621	H628	P643	Q646	K650	P655	A656															
A660	E661	F662	S663	A664	T665	K666	F667	A668	S669	F670	Q678	V679	S680	V681	E682	E686	K689	E690	K693	M696	P697	E698	Y701	T702	S703	W704	Y705	A706	K707	S708	A709	N710	F713	T714	V715	D716	N717	N718	G719	L720	Y721	P726	R730	T733	R734	P735	L736									

## 4 Data and refinement statistics i

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	262.70Å 262.70Å 612.30Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.50 49.88 – 2.50	Depositor EDS
% Data completeness (in resolution range)	89.0 (50.00-2.50) 89.0 (49.88-2.50)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.74 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.230 , 0.250 0.257 , 0.261	Depositor DCC
$R_{free}$ test set	12421 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.3	Xtriage
Anisotropy	0.686	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 29.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.004 for -h,1/3*h-1/3*k-1/3*l,-4/3*h-8/3*k+1/3*l 0.009 for -1/3*h+1/3*k+1/3*l,-k,8/3*h+4/3*k+1/3*l 0.017 for -2/3*h-1/3*k-1/3*l,-1/3*h-2/3*k+1/3*l,-4/3*h+4/3*k+1/3*l	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	42680	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.69% 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

Bond lengths and bond angles in the following residue types are not validated in this section: CYT, ADE

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	0.79	1/4246 (0.0%)	0.78	0/5790
1	B	0.79	1/4246 (0.0%)	0.79	0/5790
1	C	0.79	1/4246 (0.0%)	0.78	0/5790
1	D	0.79	1/4246 (0.0%)	0.78	0/5790
1	E	0.79	1/4246 (0.0%)	0.79	0/5790
1	F	0.78	1/4246 (0.0%)	0.77	0/5790
1	G	0.78	1/4246 (0.0%)	0.78	0/5790
1	H	0.79	1/4246 (0.0%)	0.79	0/5790
1	I	0.78	1/4246 (0.0%)	0.77	0/5790
1	J	0.78	1/4246 (0.0%)	0.78	0/5790
All	All	0.79	10/42460 (0.0%)	0.78	0/57900

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	8
1	B	0	8
1	C	0	8
1	D	0	7
1	E	0	5
1	F	0	8
1	G	0	6
1	H	0	6
1	I	0	10
1	J	0	6
All	All	0	72

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



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	580	VAL	CB-CG2	-5.46	1.41	1.52
1	C	580	VAL	CB-CG2	-5.46	1.41	1.52
1	J	580	VAL	CB-CG2	-5.46	1.41	1.52
1	D	580	VAL	CB-CG2	-5.45	1.41	1.52
1	H	580	VAL	CB-CG2	-5.45	1.41	1.52

There are no bond angle outliers.

There are no chirality outliers.

5 of 72 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	217	GLY	Peptide
1	A	219	ASP	Peptide
1	A	263	ALA	Peptide
1	A	328	ASP	Peptide
1	A	329	GLY	Peptide

## 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	4120	0	3884	206	0
1	B	4120	0	3884	194	0
1	C	4120	0	3884	175	0
1	D	4120	0	3884	192	0
1	E	4120	0	3884	186	0
1	F	4120	0	3884	193	0
1	G	4120	0	3884	194	0
1	H	4120	0	3884	187	0
1	I	4120	0	3884	209	0
1	J	4120	0	3884	127	0
2	A	10	0	4	0	0
2	B	10	0	4	3	0
2	C	10	0	4	0	0
2	D	10	0	4	0	0
2	E	10	0	4	0	0
2	F	10	0	4	0	0
2	G	10	0	4	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	10	0	4	0	0
2	I	10	0	4	0	0
2	J	10	0	4	0	0
3	A	8	0	4	3	0
3	B	16	0	8	1	0
3	C	8	0	4	3	0
3	D	8	0	4	3	0
3	E	8	0	4	1	0
3	F	8	0	4	2	0
3	H	8	0	4	3	0
3	I	8	0	4	2	0
3	J	8	0	4	1	0
4	A	131	0	0	6	0
4	B	130	0	0	7	0
4	C	128	0	0	4	0
4	D	131	0	0	9	0
4	E	133	0	0	9	0
4	F	131	0	0	4	0
4	G	127	0	0	7	0
4	H	130	0	0	5	0
4	I	130	0	0	6	0
4	J	129	0	0	5	0
All	All	42680	0	38920	1595	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 20.

The worst 5 of 1595 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:328:ASP:HB3	1:G:329:GLY:C	1.41	1.40
1:C:328:ASP:HB3	1:C:329:GLY:C	1.41	1.39
1:H:328:ASP:HB3	1:H:329:GLY:C	1.41	1.37
1:D:328:ASP:CB	1:D:329:GLY:HA3	1.42	1.29
1:B:328:ASP:OD1	1:B:329:GLY:HA2	1.28	1.28

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	518/736 (70%)	499 (96%)	15 (3%)	4 (1%)	19	35
1	B	518/736 (70%)	500 (96%)	17 (3%)	1 (0%)	47	68
1	C	518/736 (70%)	498 (96%)	18 (4%)	2 (0%)	34	54
1	D	518/736 (70%)	498 (96%)	16 (3%)	4 (1%)	19	35
1	E	518/736 (70%)	500 (96%)	16 (3%)	2 (0%)	34	54
1	F	518/736 (70%)	501 (97%)	14 (3%)	3 (1%)	25	43
1	G	518/736 (70%)	498 (96%)	16 (3%)	4 (1%)	19	35
1	H	518/736 (70%)	498 (96%)	17 (3%)	3 (1%)	25	43
1	I	518/736 (70%)	501 (97%)	12 (2%)	5 (1%)	15	28
1	J	518/736 (70%)	497 (96%)	18 (4%)	3 (1%)	25	43
All	All	5180/7360 (70%)	4990 (96%)	159 (3%)	31 (1%)	25	43

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	218	ALA
1	A	219	ASP
1	C	328	ASP
1	E	219	ASP
1	F	218	ALA

### 5.3.2 Protein sidechains [i](#)

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	451/617 (73%)	413 (92%)	38 (8%)	11	21
1	B	451/617 (73%)	411 (91%)	40 (9%)	9	19
1	C	451/617 (73%)	413 (92%)	38 (8%)	11	21
1	D	451/617 (73%)	415 (92%)	36 (8%)	12	23
1	E	451/617 (73%)	413 (92%)	38 (8%)	11	21
1	F	451/617 (73%)	413 (92%)	38 (8%)	11	21
1	G	451/617 (73%)	414 (92%)	37 (8%)	11	22
1	H	451/617 (73%)	413 (92%)	38 (8%)	11	21
1	I	451/617 (73%)	414 (92%)	37 (8%)	11	22
1	J	451/617 (73%)	413 (92%)	38 (8%)	11	21
All	All	4510/6170 (73%)	4132 (92%)	378 (8%)	11	21

5 of 378 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	278	THR
1	H	489	VAL
1	G	328	ASP
1	G	667	PHE
1	H	707	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 318 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	375	GLN
1	J	223	ASN
1	H	457	GLN
1	I	302	ASN
1	J	422	HIS

### 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](#)

20 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	CYT	E	738	-	7,8,8	1.68	1 (14%)	8,10,10	9.06	6 (75%)
3	CYT	I	738	-	7,8,8	1.69	1 (14%)	8,10,10	9.04	6 (75%)
3	CYT	J	738	-	7,8,8	1.68	1 (14%)	8,10,10	9.09	6 (75%)
2	ADE	H	737	-	9,11,11	0.90	0	7,15,15	2.55	3 (42%)
2	ADE	E	737	-	9,11,11	0.91	0	7,15,15	2.55	3 (42%)
3	CYT	B	739	-	7,8,8	1.69	1 (14%)	8,10,10	9.06	6 (75%)
2	ADE	J	737	-	9,11,11	0.90	0	7,15,15	2.54	2 (28%)
2	ADE	A	737	-	9,11,11	0.91	0	7,15,15	2.55	3 (42%)
3	CYT	D	738	-	7,8,8	1.69	1 (14%)	8,10,10	9.04	6 (75%)
2	ADE	D	737	-	9,11,11	0.91	0	7,15,15	2.55	3 (42%)
3	CYT	B	738	-	7,8,8	1.69	1 (14%)	8,10,10	9.07	6 (75%)
2	ADE	F	737	-	9,11,11	0.92	0	7,15,15	2.54	2 (28%)
3	CYT	A	738	-	7,8,8	1.69	1 (14%)	8,10,10	9.05	6 (75%)
2	ADE	B	737	-	9,11,11	0.91	0	7,15,15	2.54	2 (28%)
2	ADE	G	737	-	9,11,11	0.91	0	7,15,15	2.56	3 (42%)
2	ADE	I	737	-	9,11,11	0.92	0	7,15,15	2.56	2 (28%)
3	CYT	H	738	-	7,8,8	1.69	1 (14%)	8,10,10	9.02	6 (75%)
3	CYT	F	738	-	7,8,8	1.70	1 (14%)	8,10,10	8.99	6 (75%)
3	CYT	C	738	-	7,8,8	1.68	1 (14%)	8,10,10	9.06	6 (75%)
2	ADE	C	737	-	9,11,11	0.92	0	7,15,15	2.55	2 (28%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	CYT	E	738	-	-	-	0/1/1/1
3	CYT	I	738	-	-	-	0/1/1/1
3	CYT	J	738	-	-	-	0/1/1/1
2	ADE	H	737	-	-	-	0/2/2/2
2	ADE	E	737	-	-	-	0/2/2/2
3	CYT	B	739	-	-	-	0/1/1/1
2	ADE	J	737	-	-	-	0/2/2/2
2	ADE	A	737	-	-	-	0/2/2/2
3	CYT	D	738	-	-	-	0/1/1/1
2	ADE	D	737	-	-	-	0/2/2/2
3	CYT	B	738	-	-	-	0/1/1/1
2	ADE	F	737	-	-	-	0/2/2/2
3	CYT	A	738	-	-	-	0/1/1/1
2	ADE	B	737	-	-	-	0/2/2/2
2	ADE	G	737	-	-	-	0/2/2/2
2	ADE	I	737	-	-	-	0/2/2/2
3	CYT	H	738	-	-	-	0/1/1/1
3	CYT	F	738	-	-	-	0/1/1/1
3	CYT	C	738	-	-	-	0/1/1/1
2	ADE	C	737	-	-	-	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	738	CYT	C5-C6	-2.63	1.33	1.38
3	I	738	CYT	C5-C6	-2.63	1.33	1.38
3	D	738	CYT	C5-C6	-2.61	1.33	1.38
3	A	738	CYT	C5-C6	-2.61	1.33	1.38
3	C	738	CYT	C5-C6	-2.59	1.33	1.38

The worst 5 of 85 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	J	738	CYT	C6-N1-C2	18.60	123.59	114.42
3	C	738	CYT	C6-N1-C2	18.55	123.57	114.42
3	B	738	CYT	C6-N1-C2	18.54	123.56	114.42
3	E	738	CYT	C6-N1-C2	18.52	123.55	114.42
3	B	739	CYT	C6-N1-C2	18.51	123.55	114.42

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	738	CYT	1	0
3	I	738	CYT	2	0
3	J	738	CYT	1	0
3	B	739	CYT	1	0
3	D	738	CYT	3	0
3	A	738	CYT	3	0
2	B	737	ADE	3	0
3	H	738	CYT	3	0
3	F	738	CYT	2	0
3	C	738	CYT	3	0

## 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	520/736 (70%)	0.26	18 (3%) 44 47	25, 34, 55, 92	0
1	B	520/736 (70%)	0.58	31 (5%) 21 22	25, 34, 55, 92	0
1	C	520/736 (70%)	0.83	29 (5%) 24 25	25, 34, 55, 92	0
1	D	520/736 (70%)	0.45	18 (3%) 44 47	25, 34, 55, 92	0
1	E	520/736 (70%)	0.79	34 (6%) 18 19	25, 34, 55, 92	0
1	F	520/736 (70%)	0.59	25 (4%) 30 32	25, 34, 55, 92	0
1	G	520/736 (70%)	0.77	39 (7%) 14 14	25, 34, 55, 92	0
1	H	520/736 (70%)	0.41	23 (4%) 34 37	25, 34, 55, 92	0
1	I	520/736 (70%)	0.90	36 (6%) 16 17	25, 34, 55, 92	0
1	J	520/736 (70%)	1.25	83 (15%) 1 1	25, 34, 55, 92	0
All	All	5200/7360 (70%)	0.68	336 (6%) 18 19	25, 34, 55, 92	0

The worst 5 of 336 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	453	SER	13.2
1	G	453	SER	11.1
1	F	453	SER	11.0
1	F	456	ALA	10.5
1	I	453	SER	9.7

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	CYT	E	738	8/8	0.42	0.47	63,67,68,70	0
3	CYT	D	738	8/8	0.52	0.40	63,67,68,70	0
3	CYT	F	738	8/8	0.54	0.38	63,67,68,70	0
3	CYT	H	738	8/8	0.54	0.43	63,67,68,70	0
3	CYT	B	739	8/8	0.58	0.38	63,67,68,70	0
3	CYT	C	738	8/8	0.60	0.42	63,67,68,70	0
2	ADE	F	737	10/10	0.63	0.46	78,81,82,83	0
2	ADE	I	737	10/10	0.72	0.34	78,81,82,83	0
3	CYT	B	738	8/8	0.73	0.33	63,67,68,70	0
2	ADE	E	737	10/10	0.74	0.40	78,81,82,83	0
3	CYT	A	738	8/8	0.76	0.34	63,67,68,70	0
3	CYT	I	738	8/8	0.76	0.33	63,67,68,70	0
2	ADE	A	737	10/10	0.77	0.33	78,81,82,83	0
2	ADE	D	737	10/10	0.78	0.30	78,81,82,83	0
3	CYT	J	738	8/8	0.79	0.36	63,67,68,70	0
2	ADE	G	737	10/10	0.80	0.36	78,81,82,83	0
2	ADE	C	737	10/10	0.82	0.30	78,81,82,83	0
2	ADE	J	737	10/10	0.82	0.36	78,81,82,83	0
2	ADE	H	737	10/10	0.82	0.31	78,81,82,83	0
2	ADE	B	737	10/10	0.83	0.25	78,81,82,83	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.