



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2020 – 08:33 PM BST

PDB ID : 6Y61  
Title : Structure of apo Sheep Polyomavirus VP1  
Authors : Stroh, L.J.; Rustmeier, N.H.; Stehle, T.  
Deposited on : 2020-02-26  
Resolution : 2.45 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

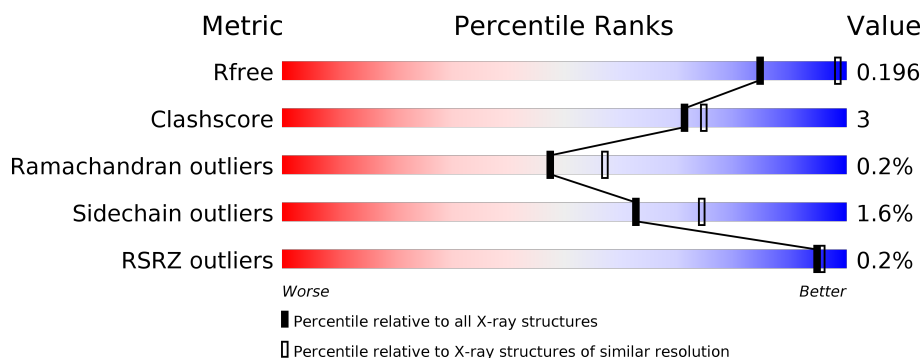
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.45 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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 on 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	AAA	293	<div> <div>82%</div> <div>8%</div> <div>10%</div> </div>
1	BBB	293	<div> <div>82%</div> <div>8%</div> <div>9%</div> </div>
1	CCC	293	<div> <div>85%</div> <div>6%</div> <div>8%</div> </div>
1	DDD	293	<div> <div>82%</div> <div>8%</div> <div>10%</div> </div>
1	EEE	293	<div> <div>83%</div> <div>9%</div> <div>8%</div> </div>
1	FFF	293	<div> <div>81%</div> <div>8%</div> <div>10%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	GGG	293	<div><div></div><div>80%</div><div>9%11%</div></div>
1	HHH	293	<div>%<div><div></div><div>83%</div><div>8%9%</div></div></div>
1	III	293	<div>%<div><div></div><div>83%</div><div>8%9%</div></div></div>
1	JJJ	293	<div><div></div><div>80%</div><div>12%8%</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 21437 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 VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	263	Total	C	N	O	S	0	0	0
			2026	1292	339	382	13			
1	BBB	266	Total	C	N	O	S	0	0	0
			2035	1297	347	379	12			
1	CCC	269	Total	C	N	O	S	0	0	0
			2065	1315	347	389	14			
1	DDD	265	Total	C	N	O	S	0	0	0
			2017	1288	341	376	12			
1	EEE	270	Total	C	N	O	S	0	0	0
			2070	1319	352	385	14			
1	FFF	264	Total	C	N	O	S	0	0	0
			2034	1295	345	382	12			
1	GGG	262	Total	C	N	O	S	0	0	0
			2010	1282	340	375	13			
1	HHH	268	Total	C	N	O	S	0	0	0
			2050	1304	346	386	14			
1	III	266	Total	C	N	O	S	0	0	0
			2026	1294	341	379	12			
1	JJJ	269	Total	C	N	O	S	0	0	0
			2062	1314	349	386	13			

There are 220 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
AAA	1	GLY	-	expression tag	UNP A0A0E3ZCF3
AAA	2	SER	-	expression tag	UNP A0A0E3ZCF3
AAA	3	SER	-	expression tag	UNP A0A0E3ZCF3
AAA	4	HIS	-	expression tag	UNP A0A0E3ZCF3
AAA	5	HIS	-	expression tag	UNP A0A0E3ZCF3
AAA	6	HIS	-	expression tag	UNP A0A0E3ZCF3
AAA	7	HIS	-	expression tag	UNP A0A0E3ZCF3
AAA	8	HIS	-	expression tag	UNP A0A0E3ZCF3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	9	HIS	-	expression tag	UNP A0A0E3ZCF3
AAA	10	SER	-	expression tag	UNP A0A0E3ZCF3
AAA	11	SER	-	expression tag	UNP A0A0E3ZCF3
AAA	12	GLY	-	expression tag	UNP A0A0E3ZCF3
AAA	13	LEU	-	expression tag	UNP A0A0E3ZCF3
AAA	14	VAL	-	expression tag	UNP A0A0E3ZCF3
AAA	15	PRO	-	expression tag	UNP A0A0E3ZCF3
AAA	16	ARG	-	expression tag	UNP A0A0E3ZCF3
AAA	17	GLY	-	expression tag	UNP A0A0E3ZCF3
AAA	18	SER	-	expression tag	UNP A0A0E3ZCF3
AAA	19	HIS	-	expression tag	UNP A0A0E3ZCF3
AAA	20	MET	-	expression tag	UNP A0A0E3ZCF3
AAA	96	SER	CYS	conflict	UNP A0A0E3ZCF3
BBB	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
BBB	1	GLY	-	expression tag	UNP A0A0E3ZCF3
BBB	2	SER	-	expression tag	UNP A0A0E3ZCF3
BBB	3	SER	-	expression tag	UNP A0A0E3ZCF3
BBB	4	HIS	-	expression tag	UNP A0A0E3ZCF3
BBB	5	HIS	-	expression tag	UNP A0A0E3ZCF3
BBB	6	HIS	-	expression tag	UNP A0A0E3ZCF3
BBB	7	HIS	-	expression tag	UNP A0A0E3ZCF3
BBB	8	HIS	-	expression tag	UNP A0A0E3ZCF3
BBB	9	HIS	-	expression tag	UNP A0A0E3ZCF3
BBB	10	SER	-	expression tag	UNP A0A0E3ZCF3
BBB	11	SER	-	expression tag	UNP A0A0E3ZCF3
BBB	12	GLY	-	expression tag	UNP A0A0E3ZCF3
BBB	13	LEU	-	expression tag	UNP A0A0E3ZCF3
BBB	14	VAL	-	expression tag	UNP A0A0E3ZCF3
BBB	15	PRO	-	expression tag	UNP A0A0E3ZCF3
BBB	16	ARG	-	expression tag	UNP A0A0E3ZCF3
BBB	17	GLY	-	expression tag	UNP A0A0E3ZCF3
BBB	18	SER	-	expression tag	UNP A0A0E3ZCF3
BBB	19	HIS	-	expression tag	UNP A0A0E3ZCF3
BBB	20	MET	-	expression tag	UNP A0A0E3ZCF3
BBB	96	SER	CYS	conflict	UNP A0A0E3ZCF3
CCC	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
CCC	1	GLY	-	expression tag	UNP A0A0E3ZCF3
CCC	2	SER	-	expression tag	UNP A0A0E3ZCF3
CCC	3	SER	-	expression tag	UNP A0A0E3ZCF3
CCC	4	HIS	-	expression tag	UNP A0A0E3ZCF3
CCC	5	HIS	-	expression tag	UNP A0A0E3ZCF3
CCC	6	HIS	-	expression tag	UNP A0A0E3ZCF3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	7	HIS	-	expression tag	UNP A0A0E3ZCF3
CCC	8	HIS	-	expression tag	UNP A0A0E3ZCF3
CCC	9	HIS	-	expression tag	UNP A0A0E3ZCF3
CCC	10	SER	-	expression tag	UNP A0A0E3ZCF3
CCC	11	SER	-	expression tag	UNP A0A0E3ZCF3
CCC	12	GLY	-	expression tag	UNP A0A0E3ZCF3
CCC	13	LEU	-	expression tag	UNP A0A0E3ZCF3
CCC	14	VAL	-	expression tag	UNP A0A0E3ZCF3
CCC	15	PRO	-	expression tag	UNP A0A0E3ZCF3
CCC	16	ARG	-	expression tag	UNP A0A0E3ZCF3
CCC	17	GLY	-	expression tag	UNP A0A0E3ZCF3
CCC	18	SER	-	expression tag	UNP A0A0E3ZCF3
CCC	19	HIS	-	expression tag	UNP A0A0E3ZCF3
CCC	20	MET	-	expression tag	UNP A0A0E3ZCF3
CCC	96	SER	CYS	conflict	UNP A0A0E3ZCF3
DDD	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
DDD	1	GLY	-	expression tag	UNP A0A0E3ZCF3
DDD	2	SER	-	expression tag	UNP A0A0E3ZCF3
DDD	3	SER	-	expression tag	UNP A0A0E3ZCF3
DDD	4	HIS	-	expression tag	UNP A0A0E3ZCF3
DDD	5	HIS	-	expression tag	UNP A0A0E3ZCF3
DDD	6	HIS	-	expression tag	UNP A0A0E3ZCF3
DDD	7	HIS	-	expression tag	UNP A0A0E3ZCF3
DDD	8	HIS	-	expression tag	UNP A0A0E3ZCF3
DDD	9	HIS	-	expression tag	UNP A0A0E3ZCF3
DDD	10	SER	-	expression tag	UNP A0A0E3ZCF3
DDD	11	SER	-	expression tag	UNP A0A0E3ZCF3
DDD	12	GLY	-	expression tag	UNP A0A0E3ZCF3
DDD	13	LEU	-	expression tag	UNP A0A0E3ZCF3
DDD	14	VAL	-	expression tag	UNP A0A0E3ZCF3
DDD	15	PRO	-	expression tag	UNP A0A0E3ZCF3
DDD	16	ARG	-	expression tag	UNP A0A0E3ZCF3
DDD	17	GLY	-	expression tag	UNP A0A0E3ZCF3
DDD	18	SER	-	expression tag	UNP A0A0E3ZCF3
DDD	19	HIS	-	expression tag	UNP A0A0E3ZCF3
DDD	20	MET	-	expression tag	UNP A0A0E3ZCF3
DDD	96	SER	CYS	conflict	UNP A0A0E3ZCF3
EEE	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
EEE	1	GLY	-	expression tag	UNP A0A0E3ZCF3
EEE	2	SER	-	expression tag	UNP A0A0E3ZCF3
EEE	3	SER	-	expression tag	UNP A0A0E3ZCF3
EEE	4	HIS	-	expression tag	UNP A0A0E3ZCF3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
EEE	5	HIS	-	expression tag	UNP A0A0E3ZCF3
EEE	6	HIS	-	expression tag	UNP A0A0E3ZCF3
EEE	7	HIS	-	expression tag	UNP A0A0E3ZCF3
EEE	8	HIS	-	expression tag	UNP A0A0E3ZCF3
EEE	9	HIS	-	expression tag	UNP A0A0E3ZCF3
EEE	10	SER	-	expression tag	UNP A0A0E3ZCF3
EEE	11	SER	-	expression tag	UNP A0A0E3ZCF3
EEE	12	GLY	-	expression tag	UNP A0A0E3ZCF3
EEE	13	LEU	-	expression tag	UNP A0A0E3ZCF3
EEE	14	VAL	-	expression tag	UNP A0A0E3ZCF3
EEE	15	PRO	-	expression tag	UNP A0A0E3ZCF3
EEE	16	ARG	-	expression tag	UNP A0A0E3ZCF3
EEE	17	GLY	-	expression tag	UNP A0A0E3ZCF3
EEE	18	SER	-	expression tag	UNP A0A0E3ZCF3
EEE	19	HIS	-	expression tag	UNP A0A0E3ZCF3
EEE	20	MET	-	expression tag	UNP A0A0E3ZCF3
EEE	96	SER	CYS	conflict	UNP A0A0E3ZCF3
FFF	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
FFF	1	GLY	-	expression tag	UNP A0A0E3ZCF3
FFF	2	SER	-	expression tag	UNP A0A0E3ZCF3
FFF	3	SER	-	expression tag	UNP A0A0E3ZCF3
FFF	4	HIS	-	expression tag	UNP A0A0E3ZCF3
FFF	5	HIS	-	expression tag	UNP A0A0E3ZCF3
FFF	6	HIS	-	expression tag	UNP A0A0E3ZCF3
FFF	7	HIS	-	expression tag	UNP A0A0E3ZCF3
FFF	8	HIS	-	expression tag	UNP A0A0E3ZCF3
FFF	9	HIS	-	expression tag	UNP A0A0E3ZCF3
FFF	10	SER	-	expression tag	UNP A0A0E3ZCF3
FFF	11	SER	-	expression tag	UNP A0A0E3ZCF3
FFF	12	GLY	-	expression tag	UNP A0A0E3ZCF3
FFF	13	LEU	-	expression tag	UNP A0A0E3ZCF3
FFF	14	VAL	-	expression tag	UNP A0A0E3ZCF3
FFF	15	PRO	-	expression tag	UNP A0A0E3ZCF3
FFF	16	ARG	-	expression tag	UNP A0A0E3ZCF3
FFF	17	GLY	-	expression tag	UNP A0A0E3ZCF3
FFF	18	SER	-	expression tag	UNP A0A0E3ZCF3
FFF	19	HIS	-	expression tag	UNP A0A0E3ZCF3
FFF	20	MET	-	expression tag	UNP A0A0E3ZCF3
FFF	96	SER	CYS	conflict	UNP A0A0E3ZCF3
GGG	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
GGG	1	GLY	-	expression tag	UNP A0A0E3ZCF3
GGG	2	SER	-	expression tag	UNP A0A0E3ZCF3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
GGG	3	SER	-	expression tag	UNP A0A0E3ZCF3
GGG	4	HIS	-	expression tag	UNP A0A0E3ZCF3
GGG	5	HIS	-	expression tag	UNP A0A0E3ZCF3
GGG	6	HIS	-	expression tag	UNP A0A0E3ZCF3
GGG	7	HIS	-	expression tag	UNP A0A0E3ZCF3
GGG	8	HIS	-	expression tag	UNP A0A0E3ZCF3
GGG	9	HIS	-	expression tag	UNP A0A0E3ZCF3
GGG	10	SER	-	expression tag	UNP A0A0E3ZCF3
GGG	11	SER	-	expression tag	UNP A0A0E3ZCF3
GGG	12	GLY	-	expression tag	UNP A0A0E3ZCF3
GGG	13	LEU	-	expression tag	UNP A0A0E3ZCF3
GGG	14	VAL	-	expression tag	UNP A0A0E3ZCF3
GGG	15	PRO	-	expression tag	UNP A0A0E3ZCF3
GGG	16	ARG	-	expression tag	UNP A0A0E3ZCF3
GGG	17	GLY	-	expression tag	UNP A0A0E3ZCF3
GGG	18	SER	-	expression tag	UNP A0A0E3ZCF3
GGG	19	HIS	-	expression tag	UNP A0A0E3ZCF3
GGG	20	MET	-	expression tag	UNP A0A0E3ZCF3
GGG	96	SER	CYS	conflict	UNP A0A0E3ZCF3
HHH	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
HHH	1	GLY	-	expression tag	UNP A0A0E3ZCF3
HHH	2	SER	-	expression tag	UNP A0A0E3ZCF3
HHH	3	SER	-	expression tag	UNP A0A0E3ZCF3
HHH	4	HIS	-	expression tag	UNP A0A0E3ZCF3
HHH	5	HIS	-	expression tag	UNP A0A0E3ZCF3
HHH	6	HIS	-	expression tag	UNP A0A0E3ZCF3
HHH	7	HIS	-	expression tag	UNP A0A0E3ZCF3
HHH	8	HIS	-	expression tag	UNP A0A0E3ZCF3
HHH	9	HIS	-	expression tag	UNP A0A0E3ZCF3
HHH	10	SER	-	expression tag	UNP A0A0E3ZCF3
HHH	11	SER	-	expression tag	UNP A0A0E3ZCF3
HHH	12	GLY	-	expression tag	UNP A0A0E3ZCF3
HHH	13	LEU	-	expression tag	UNP A0A0E3ZCF3
HHH	14	VAL	-	expression tag	UNP A0A0E3ZCF3
HHH	15	PRO	-	expression tag	UNP A0A0E3ZCF3
HHH	16	ARG	-	expression tag	UNP A0A0E3ZCF3
HHH	17	GLY	-	expression tag	UNP A0A0E3ZCF3
HHH	18	SER	-	expression tag	UNP A0A0E3ZCF3
HHH	19	HIS	-	expression tag	UNP A0A0E3ZCF3
HHH	20	MET	-	expression tag	UNP A0A0E3ZCF3
HHH	96	SER	CYS	conflict	UNP A0A0E3ZCF3
III	0	MET	-	initiating methionine	UNP A0A0E3ZCF3

*Continued on next page...*



*Continued from previous page...*

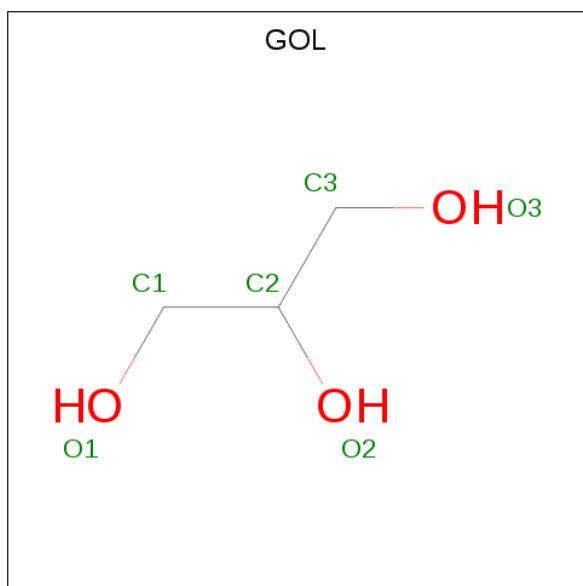
Chain	Residue	Modelled	Actual	Comment	Reference
III	1	GLY	-	expression tag	UNP A0A0E3ZCF3
III	2	SER	-	expression tag	UNP A0A0E3ZCF3
III	3	SER	-	expression tag	UNP A0A0E3ZCF3
III	4	HIS	-	expression tag	UNP A0A0E3ZCF3
III	5	HIS	-	expression tag	UNP A0A0E3ZCF3
III	6	HIS	-	expression tag	UNP A0A0E3ZCF3
III	7	HIS	-	expression tag	UNP A0A0E3ZCF3
III	8	HIS	-	expression tag	UNP A0A0E3ZCF3
III	9	HIS	-	expression tag	UNP A0A0E3ZCF3
III	10	SER	-	expression tag	UNP A0A0E3ZCF3
III	11	SER	-	expression tag	UNP A0A0E3ZCF3
III	12	GLY	-	expression tag	UNP A0A0E3ZCF3
III	13	LEU	-	expression tag	UNP A0A0E3ZCF3
III	14	VAL	-	expression tag	UNP A0A0E3ZCF3
III	15	PRO	-	expression tag	UNP A0A0E3ZCF3
III	16	ARG	-	expression tag	UNP A0A0E3ZCF3
III	17	GLY	-	expression tag	UNP A0A0E3ZCF3
III	18	SER	-	expression tag	UNP A0A0E3ZCF3
III	19	HIS	-	expression tag	UNP A0A0E3ZCF3
III	20	MET	-	expression tag	UNP A0A0E3ZCF3
III	96	SER	CYS	conflict	UNP A0A0E3ZCF3
JJJ	0	MET	-	initiating methionine	UNP A0A0E3ZCF3
JJJ	1	GLY	-	expression tag	UNP A0A0E3ZCF3
JJJ	2	SER	-	expression tag	UNP A0A0E3ZCF3
JJJ	3	SER	-	expression tag	UNP A0A0E3ZCF3
JJJ	4	HIS	-	expression tag	UNP A0A0E3ZCF3
JJJ	5	HIS	-	expression tag	UNP A0A0E3ZCF3
JJJ	6	HIS	-	expression tag	UNP A0A0E3ZCF3
JJJ	7	HIS	-	expression tag	UNP A0A0E3ZCF3
JJJ	8	HIS	-	expression tag	UNP A0A0E3ZCF3
JJJ	9	HIS	-	expression tag	UNP A0A0E3ZCF3
JJJ	10	SER	-	expression tag	UNP A0A0E3ZCF3
JJJ	11	SER	-	expression tag	UNP A0A0E3ZCF3
JJJ	12	GLY	-	expression tag	UNP A0A0E3ZCF3
JJJ	13	LEU	-	expression tag	UNP A0A0E3ZCF3
JJJ	14	VAL	-	expression tag	UNP A0A0E3ZCF3
JJJ	15	PRO	-	expression tag	UNP A0A0E3ZCF3
JJJ	16	ARG	-	expression tag	UNP A0A0E3ZCF3
JJJ	17	GLY	-	expression tag	UNP A0A0E3ZCF3
JJJ	18	SER	-	expression tag	UNP A0A0E3ZCF3
JJJ	19	HIS	-	expression tag	UNP A0A0E3ZCF3
JJJ	20	MET	-	expression tag	UNP A0A0E3ZCF3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
JJJ	96	SER	CYS	conflict	UNP A0A0E3ZCF3

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	AAA	1	Total	C	O	0	0
			6	3	3		
2	AAA	1	Total	C	O	0	0
			6	3	3		
2	BBB	1	Total	C	O	0	0
			6	3	3		
2	CCC	1	Total	C	O	0	0
			6	3	3		
2	DDD	1	Total	C	O	0	0
			6	3	3		
2	EEE	1	Total	C	O	0	0
			6	3	3		
2	FFF	1	Total	C	O	0	0
			6	3	3		
2	GGG	1	Total	C	O	0	0
			6	3	3		
2	HHH	1	Total	C	O	0	0
			6	3	3		
2	III	1	Total	C	O	0	0
			6	3	3		
2	JJJ	1	Total	C	O	0	0
			6	3	3		

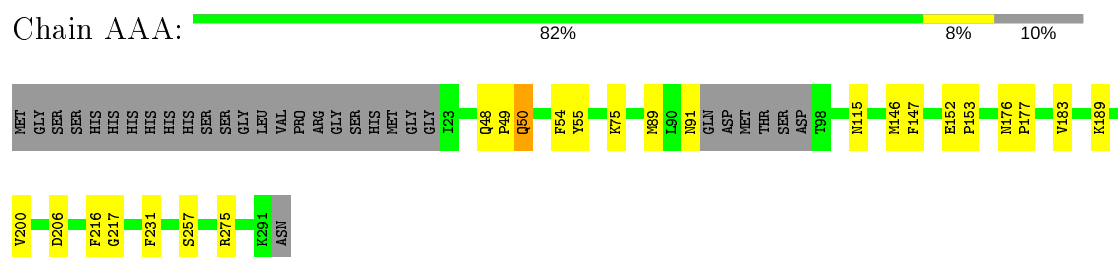
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	108	Total 109	O 109	0	1
3	BBB	97	Total 97	O 97	0	0
3	CCC	101	Total 101	O 101	0	0
3	DDD	110	Total 110	O 110	0	0
3	EEE	92	Total 92	O 92	0	0
3	FFF	85	Total 85	O 85	0	0
3	GGG	87	Total 88	O 88	0	1
3	HHH	110	Total 111	O 111	0	1
3	III	90	Total 92	O 92	0	2
3	JJJ	90	Total 91	O 91	0	1

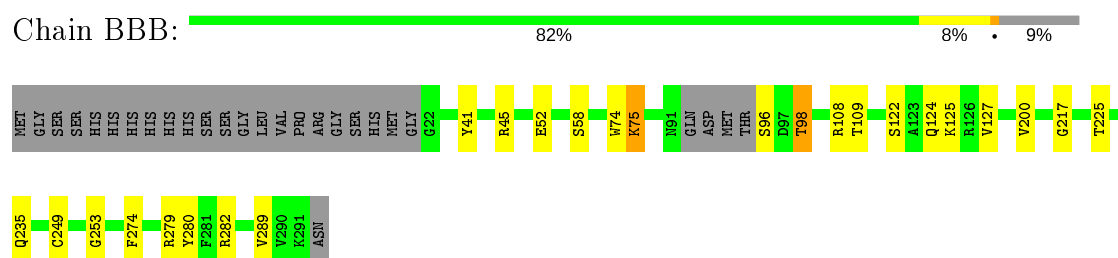
### 3 Residue-property plots [i](#)

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.

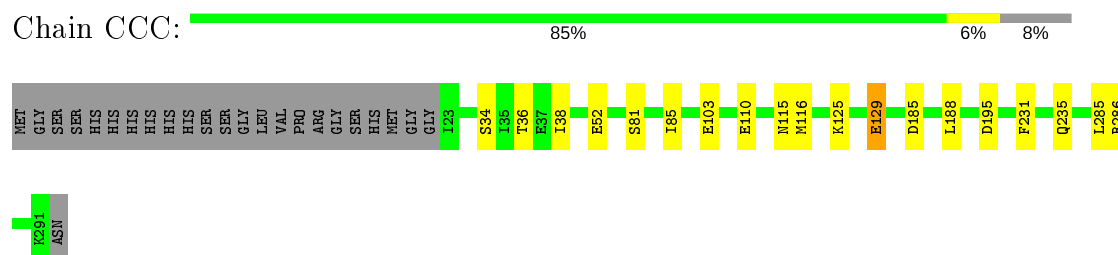
- Molecule 1: Capsid protein VP1



- Molecule 1: Capsid protein VP1



- Molecule 1: Capsid protein VP1





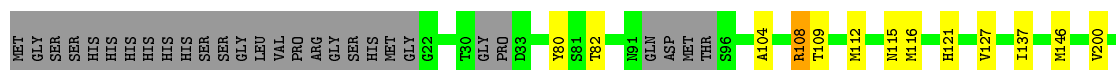
- Molecule 1: Capsid protein VP1

Chain EEE: 83% 9% 8%



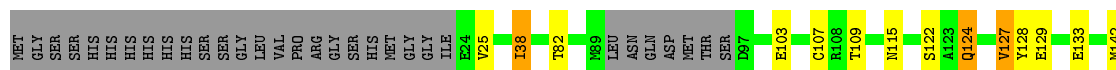
- Molecule 1: Capsid protein VP1

Chain FFF: 81% 8% 10%



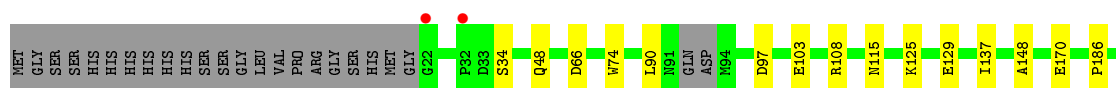
- Molecule 1: Capsid protein VP1

Chain GGG: 80% 9% 11%



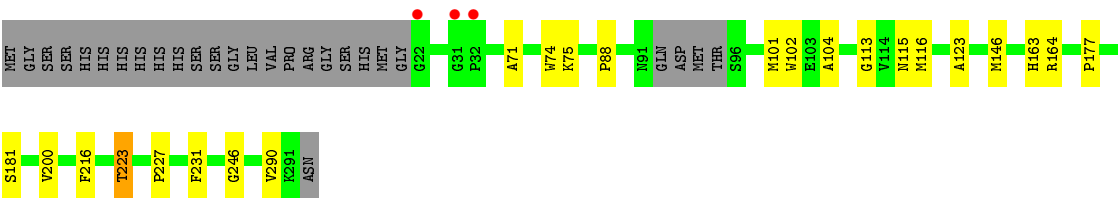
- Molecule 1: Capsid protein VP1

Chain HHH: 83% 8% 9%

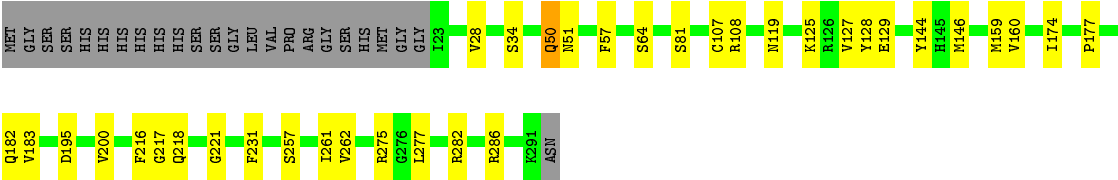
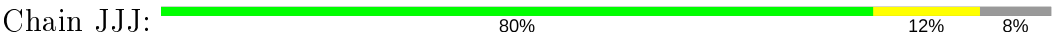


- Molecule 1: Capsid protein VP1

Chain III: 83% 8% 9%



● Molecule 1: Capsid protein VP1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	130.43Å 130.43Å 221.77Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.77 – 2.45 49.77 – 2.45	Depositor EDS
% Data completeness (in resolution range)	94.8 (49.77-2.45) 94.8 (49.77-2.45)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.65 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
R, $R_{free}$	0.161 , 0.197 0.165 , 0.196	Depositor DCC
$R_{free}$ test set	7336 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.2	Xtriage
Anisotropy	0.779	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 11.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.069 for -h,-k,l 0.246 for h,-h-k,-l 0.070 for -k,-h,-l	Xtriage
Reported twinning fraction	0.771 for H, K, L 0.229 for K, H, -L	Depositor
Outliers	0 of 147077 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	21437	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% 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](#)

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

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	AAA	0.69	0/2078	0.88	0/2827
1	BBB	0.68	0/2087	0.86	0/2839
1	CCC	0.67	0/2118	0.85	0/2881
1	DDD	0.68	0/2069	0.84	1/2814 (0.0%)
1	EEE	0.69	0/2123	0.86	0/2887
1	FFF	0.69	0/2084	0.83	0/2832
1	GGG	0.67	0/2062	0.85	0/2804
1	HHH	0.69	0/2102	0.85	0/2858
1	III	0.68	0/2078	0.85	0/2828
1	JJJ	0.69	0/2115	0.84	0/2879
All	All	0.68	0/20916	0.85	1/28449 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	DDD	164	ARG	NE-CZ-NH1	5.36	122.98	120.30

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	AAA	2026	0	1938	15	0
1	BBB	2035	0	1948	14	0
1	CCC	2065	0	1971	10	0
1	DDD	2017	0	1916	15	0
1	EEE	2070	0	1993	17	0
1	FFF	2034	0	1943	18	0
1	GGG	2010	0	1922	19	0
1	HHH	2050	0	1957	14	0
1	III	2026	0	1931	19	0
1	JJJ	2062	0	1975	25	0
2	AAA	12	0	16	0	0
2	BBB	6	0	8	0	0
2	CCC	6	0	8	0	0
2	DDD	6	0	8	0	0
2	EEE	6	0	8	0	0
2	FFF	6	0	8	0	0
2	GGG	6	0	8	0	0
2	HHH	6	0	8	0	0
2	III	6	0	8	1	0
2	JJJ	6	0	8	0	0
3	AAA	109	0	0	1	0
3	BBB	97	0	0	2	0
3	CCC	101	0	0	0	0
3	DDD	110	0	0	1	0
3	EEE	92	0	0	0	0
3	FFF	85	0	0	0	0
3	GGG	88	0	0	1	0
3	HHH	111	0	0	0	0
3	III	92	0	0	0	0
3	JJJ	91	0	0	1	0
All	All	21437	0	19582	139	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 3.

The worst 5 of 139 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:AAA:48:GLN:HG2	1:AAA:54:PHE:HB2	1.69	0.74
1:BBB:108:ARG:HG2	1:BBB:235:GLN:O	1.96	0.65
1:EEE:153:PRO:HB2	1:EEE:191:THR:CG2	2.27	0.65
1:JJJ:127:VAL:HG23	1:JJJ:128:TYR:CD2	2.33	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:FFF:280:TYR:OH	1:FFF:282:ARG:NH1	2.32	0.63

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	AAA	259/293 (88%)	247 (95%)	11 (4%)	1 (0%)	34	41
1	BBB	262/293 (89%)	243 (93%)	18 (7%)	1 (0%)	34	41
1	CCC	267/293 (91%)	253 (95%)	14 (5%)	0	100	100
1	DDD	261/293 (89%)	246 (94%)	15 (6%)	0	100	100
1	EEE	268/293 (92%)	256 (96%)	12 (4%)	0	100	100
1	FFF	258/293 (88%)	244 (95%)	14 (5%)	0	100	100
1	GGG	258/293 (88%)	241 (93%)	16 (6%)	1 (0%)	34	41
1	HHH	264/293 (90%)	249 (94%)	14 (5%)	1 (0%)	34	41
1	III	262/293 (89%)	247 (94%)	15 (6%)	0	100	100
1	JJJ	267/293 (91%)	257 (96%)	10 (4%)	0	100	100
All	All	2626/2930 (90%)	2483 (95%)	139 (5%)	4 (0%)	47	57

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	HHH	90	LEU
1	BBB	75	LYS
1	AAA	183	VAL
1	GGG	183	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 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	AAA	216/249 (87%)	212 (98%)	4 (2%)	57	69
1	BBB	214/249 (86%)	210 (98%)	4 (2%)	57	69
1	CCC	218/249 (88%)	214 (98%)	4 (2%)	59	71
1	DDD	209/249 (84%)	207 (99%)	2 (1%)	76	84
1	EEE	219/249 (88%)	215 (98%)	4 (2%)	59	71
1	FFF	215/249 (86%)	213 (99%)	2 (1%)	78	86
1	GGG	212/249 (85%)	205 (97%)	7 (3%)	38	49
1	HHH	217/249 (87%)	214 (99%)	3 (1%)	67	77
1	III	212/249 (85%)	211 (100%)	1 (0%)	88	93
1	JJJ	218/249 (88%)	214 (98%)	4 (2%)	59	71
All	All	2150/2490 (86%)	2115 (98%)	35 (2%)	62	74

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	EEE	66	ASP
1	FFF	282	ARG
1	JJJ	50	GLN
1	EEE	127	VAL
1	EEE	146	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

11 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	$\# Z  > 2$
2	GOL	GGG	301	-	5,5,5	0.14	0	5,5,5	0.46	0
2	GOL	JJJ	301	-	5,5,5	0.20	0	5,5,5	0.52	0
2	GOL	EEE	301	-	5,5,5	0.20	0	5,5,5	0.47	0
2	GOL	FFF	301	-	5,5,5	0.15	0	5,5,5	0.43	0
2	GOL	III	301	-	5,5,5	0.09	0	5,5,5	0.41	0
2	GOL	HHH	301	-	5,5,5	0.16	0	5,5,5	0.63	0
2	GOL	CCC	301	-	5,5,5	0.12	0	5,5,5	0.41	0
2	GOL	DDD	301	-	5,5,5	0.16	0	5,5,5	0.48	0
2	GOL	AAA	301	-	5,5,5	0.12	0	5,5,5	0.48	0
2	GOL	AAA	302	-	5,5,5	0.13	0	5,5,5	0.38	0
2	GOL	BBB	301	-	5,5,5	0.11	0	5,5,5	0.48	0

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. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	GGG	301	-	-	2/4/4/4	-
2	GOL	JJJ	301	-	-	2/4/4/4	-
2	GOL	EEE	301	-	-	2/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	FFF	301	-	-	2/4/4/4	-
2	GOL	III	301	-	-	2/4/4/4	-
2	GOL	HHH	301	-	-	2/4/4/4	-
2	GOL	CCC	301	-	-	0/4/4/4	-
2	GOL	DDD	301	-	-	4/4/4/4	-
2	GOL	AAA	301	-	-	2/4/4/4	-
2	GOL	AAA	302	-	-	2/4/4/4	-
2	GOL	BBB	301	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 23 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	GGG	301	GOL	O1-C1-C2-C3
2	EEE	301	GOL	O1-C1-C2-C3
2	FFF	301	GOL	O1-C1-C2-C3
2	DDD	301	GOL	O1-C1-C2-C3
2	DDD	301	GOL	C1-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	III	301	GOL	1	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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	AAA	263/293 (89%)	-0.46	0	100	100	18, 27, 49, 76	0
1	BBB	266/293 (90%)	-0.49	0	100	100	19, 27, 51, 78	0
1	CCC	269/293 (91%)	-0.45	0	100	100	19, 28, 50, 74	0
1	DDD	265/293 (90%)	-0.50	0	100	100	20, 28, 52, 72	0
1	EEE	270/293 (92%)	-0.53	0	100	100	19, 27, 42, 61	0
1	FFF	264/293 (90%)	-0.44	0	100	100	20, 30, 54, 71	0
1	GGG	262/293 (89%)	-0.51	0	100	100	19, 28, 51, 79	0
1	HHH	268/293 (91%)	-0.43	2 (0%)	87	88	19, 29, 50, 85	0
1	III	266/293 (90%)	-0.46	3 (1%)	80	80	20, 30, 59, 83	0
1	JJJ	269/293 (91%)	-0.46	0	100	100	21, 31, 46, 74	0
All	All	2662/2930 (90%)	-0.47	5 (0%)	95	95	18, 29, 51, 85	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	III	22	GLY	2.6
1	HHH	32	PRO	2.4
1	HHH	22	GLY	2.4
1	III	31	GLY	2.3
1	III	32	PRO	2.2

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

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( $\text{\AA}^2$ )	Q<0.9
2	GOL	BBB	301	6/6	0.94	0.17	34,37,39,39	0
2	GOL	EEE	301	6/6	0.95	0.11	26,29,32,38	0
2	GOL	GGG	301	6/6	0.95	0.15	37,37,40,44	0
2	GOL	FFF	301	6/6	0.96	0.18	32,35,36,36	0
2	GOL	III	301	6/6	0.96	0.16	38,41,44,51	0
2	GOL	AAA	301	6/6	0.96	0.14	29,33,34,36	0
2	GOL	AAA	302	6/6	0.96	0.10	40,42,45,46	0
2	GOL	JJJ	301	6/6	0.96	0.22	35,37,39,46	0
2	GOL	HHH	301	6/6	0.97	0.15	30,32,34,34	0
2	GOL	CCC	301	6/6	0.97	0.13	29,32,33,33	0
2	GOL	DDD	301	6/6	0.97	0.15	32,33,33,40	0

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

There are no such residues in this entry.