



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 05:14 PM GMT

PDB ID : 4HMG  
Title : REFINEMENT OF THE INFLUENZA VIRUS HEMAGGLUTININ BY SIMULATED ANNEALING  
Authors : Weis, W.I.; Bruenger, A.T.; Skehel, J.J.; Wiley, D.C.  
Deposited on : 1989-09-11  
Resolution : 3.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

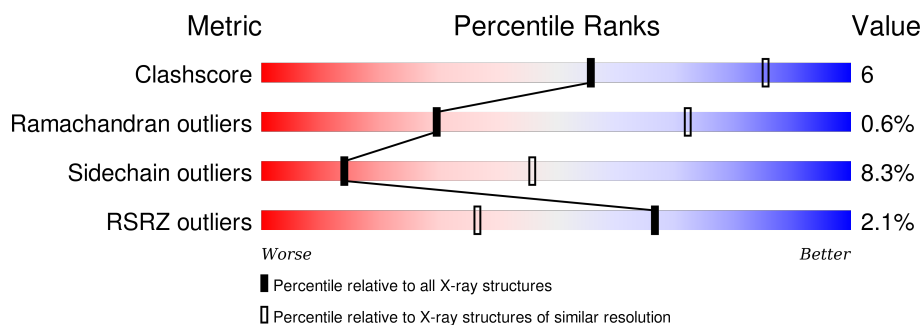
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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. 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	328	<div> <div>3%</div> <div> <div></div> <div>76%</div> <div>20%</div> <div>.</div> </div> </div>
1	C	328	<div> <div>2%</div> <div> <div></div> <div>75%</div> <div>21%</div> <div>.</div> </div> </div>
1	E	328	<div> <div>2%</div> <div> <div></div> <div>78%</div> <div>19%</div> <div>.</div> </div> </div>
2	B	175	<div> <div>%</div> <div> <div></div> <div>70%</div> <div>25%</div> <div>5%</div> </div> </div>
2	D	175	<div> <div>%</div> <div> <div></div> <div>71%</div> <div>24%</div> <div>5%</div> </div> </div>
2	F	175	<div> <div>%</div> <div> <div></div> <div>71%</div> <div>25%</div> <div>..</div> </div> </div>

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
3	SIA	C	329	-	-	-	X
4	NAG	A	330	-	-	-	X
4	NAG	A	348	-	-	-	X
4	NAG	B	401	-	-	-	X
4	NAG	C	330	-	-	-	X
4	NAG	C	348	-	-	-	X
4	NAG	D	401	-	-	-	X
4	NAG	E	330	-	-	-	X
4	NAG	E	348	-	-	-	X
5	MAN	A	341	X	-	-	-
5	NAG	C	340	-	-	-	X
5	MAN	C	341	X	-	-	-
5	MAN	E	341	X	-	-	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12243 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 HEMAGGLUTININ, CHAIN HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	328	Total	C	N	O	S	0	0	0
			2533	1580	446	494	13			
1	C	328	Total	C	N	O	S	0	0	0
			2533	1580	446	494	13			
1	E	328	Total	C	N	O	S	0	0	0
			2533	1580	446	494	13			

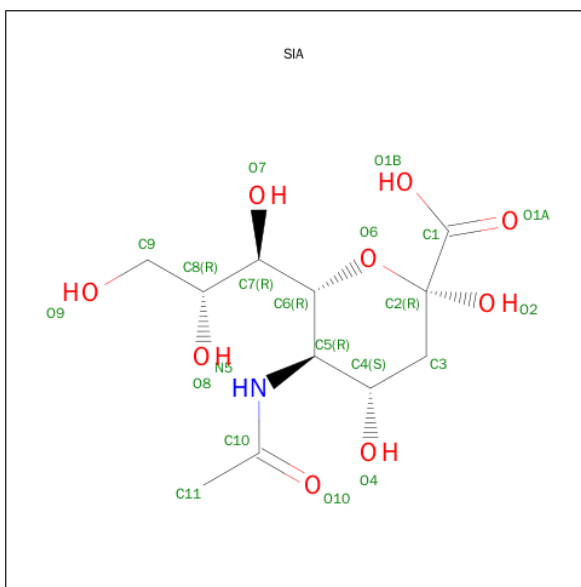
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	226	GLN	LEU	CONFLICT	UNP P03437
C	226	GLN	LEU	CONFLICT	UNP P03437
E	226	GLN	LEU	CONFLICT	UNP P03437

- Molecule 2 is a protein called HEMAGGLUTININ, CHAIN HA1.

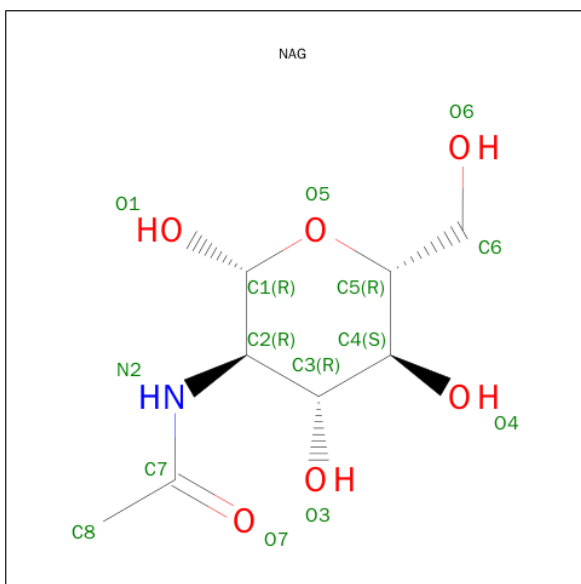
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	175	Total	C	N	O	S	0	0	0
			1421	882	250	283	6			
2	D	175	Total	C	N	O	S	0	0	0
			1421	882	250	283	6			
2	F	175	Total	C	N	O	S	0	0	0
			1421	882	250	283	6			

- Molecule 3 is SUGAR (O-SIALIC ACID) (three-letter code: SIA) (formula: C<sub>11</sub>H<sub>19</sub>NO<sub>9</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			21	11	1	9		
3	C	1	Total	C	N	O	0	0
			21	11	1	9		
3	E	1	Total	C	N	O	0	0
			21	11	1	9		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	B	1	Total C N O 14 8 1 5	0	0
4	C	1	Total C N O 14 8 1 5	0	0
4	C	1	Total C N O 14 8 1 5	0	0
4	C	1	Total C N O 14 8 1 5	0	0
4	D	1	Total C N O 14 8 1 5	0	0
4	E	1	Total C N O 14 8 1 5	0	0
4	E	1	Total C N O 14 8 1 5	0	0
4	E	1	Total C N O 14 8 1 5	0	0
4	F	1	Total C N O 14 8 1 5	0	0

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	3	Total C N O 39 22 2 15	0	0
5	C	3	Total C N O 39 22 2 15	0	0
5	E	3	Total C N O 39 22 2 15	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	6	Total O 6 6	0	0
6	B	5	Total O 5 5	0	0
6	C	7	Total O 7 7	0	0

*Continued on next page...*

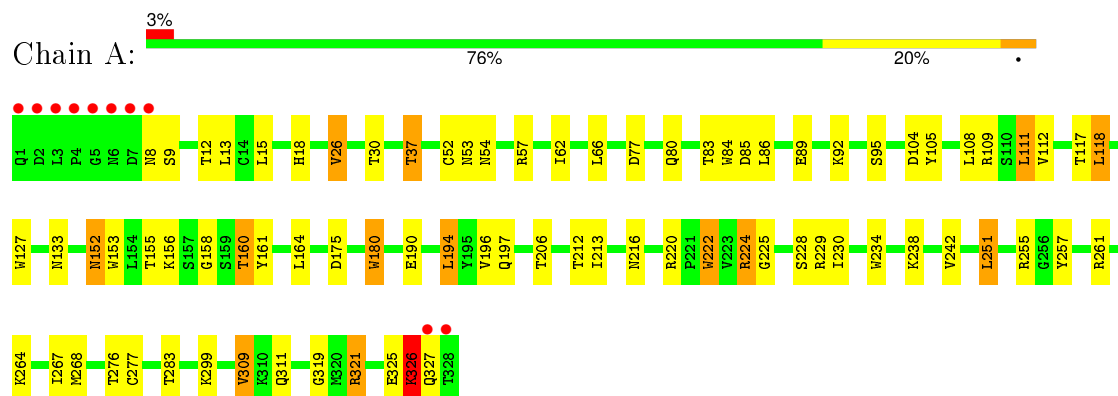
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	4	Total 4	O 4	0	0
6	E	6	Total 6	O 6	0	0
6	F	5	Total 5	O 5	0	0

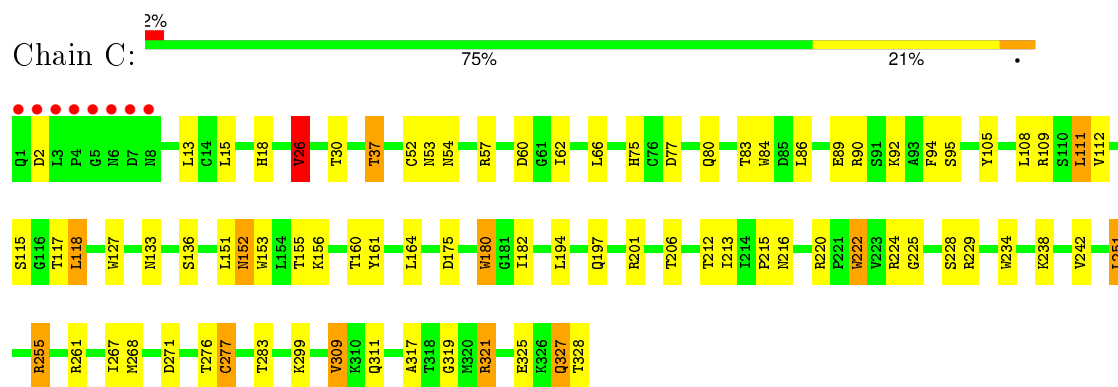
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

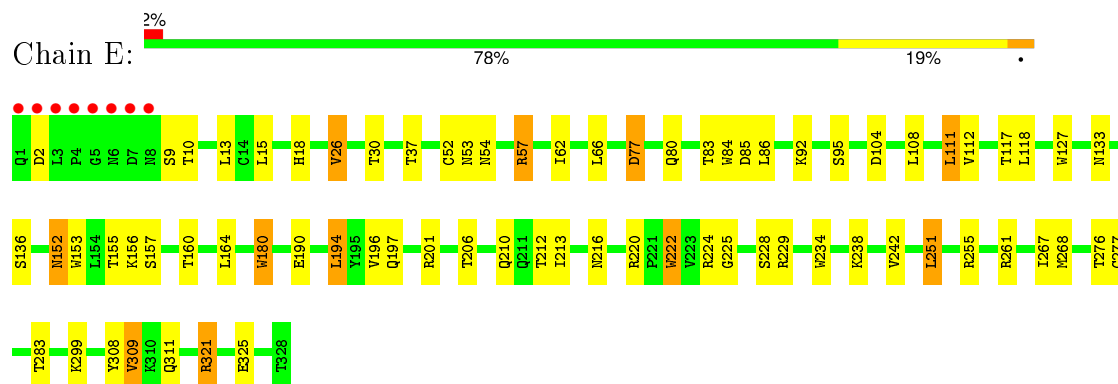
#### • Molecule 1: HEMAGGLUTININ, CHAIN HA1



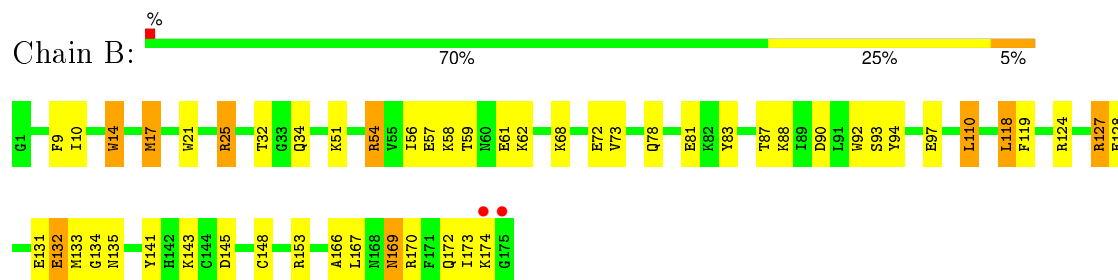
#### • Molecule 1: HEMAGGLUTININ, CHAIN HA1



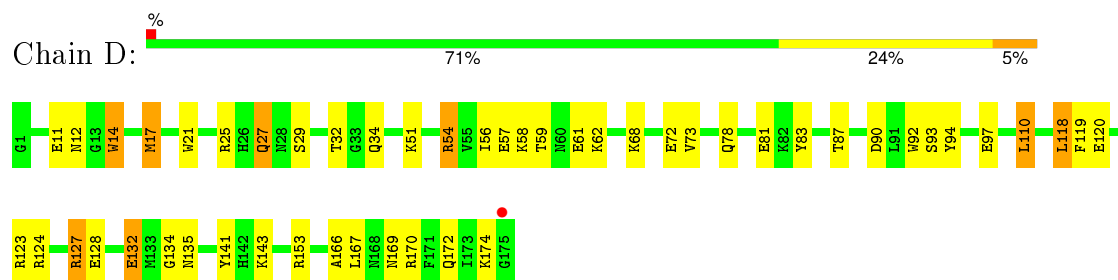
#### • Molecule 1: HEMAGGLUTININ, CHAIN HA1



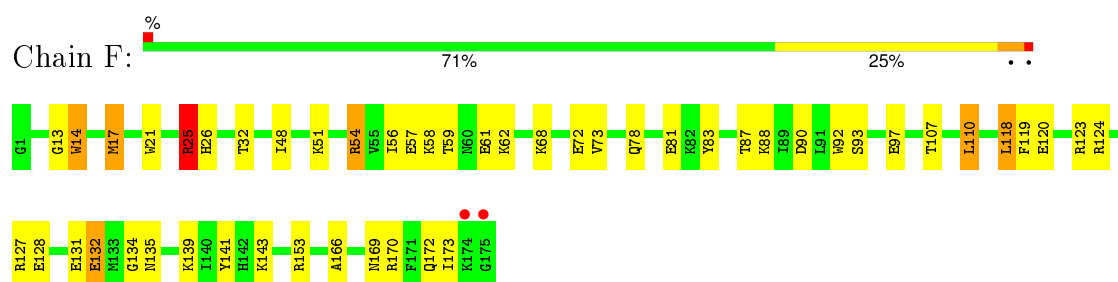
## ● Molecule 2: HEMAGGLUTININ, CHAIN HA1



## ● Molecule 2: HEMAGGLUTININ, CHAIN HA1



## ● Molecule 2: HEMAGGLUTININ, CHAIN HA1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	162.80Å 162.80Å 176.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 3.00 7.00 – 2.99	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-3.00) 75.1 (7.00-2.99)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$	-	Xtriage
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.231 , (Not available) 0.217 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	39.1	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 59.8	EDS
Estimated twinning fraction	0.039 for h,-k,-l	Xtriage
L-test for twinning <sup>1</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 64062 reflections	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	12243	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.80% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: SIA, NAG, MAN

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	0.87	1/2590 (0.0%)	1.44	33/3528 (0.9%)
1	C	0.86	1/2590 (0.0%)	1.45	32/3528 (0.9%)
1	E	0.89	1/2590 (0.0%)	1.45	34/3528 (1.0%)
2	B	0.94	0/1445	1.48	21/1939 (1.1%)
2	D	0.97	0/1445	1.49	20/1939 (1.0%)
2	F	0.97	0/1445	1.49	22/1939 (1.1%)
All	All	0.91	3/12105 (0.0%)	1.46	162/16401 (1.0%)

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
5	A	1	0
5	C	1	0
5	E	1	0
All	All	3	0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	234	TRP	CG-CD2	-5.71	1.33	1.43
1	C	234	TRP	CG-CD2	-5.16	1.34	1.43
1	A	234	TRP	CG-CD2	-5.15	1.34	1.43

All (162) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	321	ARG	NE-CZ-NH1	-11.94	114.33	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	321	ARG	NE-CZ-NH1	-11.71	114.44	120.30
1	A	321	ARG	NE-CZ-NH1	-11.45	114.58	120.30
2	D	170	ARG	NE-CZ-NH1	-11.34	114.63	120.30
2	F	170	ARG	NE-CZ-NH1	-10.75	114.92	120.30
2	F	124	ARG	NE-CZ-NH1	-10.46	115.07	120.30
2	B	170	ARG	NE-CZ-NH1	-10.44	115.08	120.30
2	D	124	ARG	NE-CZ-NH1	-10.41	115.09	120.30
2	B	124	ARG	NE-CZ-NH1	-9.77	115.42	120.30
1	A	180	TRP	CD1-CG-CD2	9.55	113.94	106.30
1	C	161	TYR	CB-CG-CD2	-9.34	115.40	121.00
1	E	180	TRP	CD1-CG-CD2	9.29	113.73	106.30
1	A	255	ARG	NE-CZ-NH1	-9.24	115.68	120.30
1	C	180	TRP	CD1-CG-CD2	9.14	113.61	106.30
1	C	321	ARG	NE-CZ-NH2	8.92	124.76	120.30
2	F	14	TRP	CD1-CG-CD2	8.56	113.15	106.30
2	B	92	TRP	CD1-CG-CD2	8.55	113.14	106.30
2	B	21	TRP	CD1-CG-CD2	8.51	113.11	106.30
1	A	180	TRP	CE2-CD2-CG	-8.35	100.62	107.30
2	B	14	TRP	CD1-CG-CD2	8.35	112.98	106.30
1	A	84	TRP	CD1-CG-CD2	8.34	112.97	106.30
2	F	21	TRP	CD1-CG-CD2	8.33	112.96	106.30
1	E	84	TRP	CD1-CG-CD2	8.27	112.91	106.30
2	D	14	TRP	CD1-CG-CD2	8.20	112.86	106.30
1	E	180	TRP	CE2-CD2-CG	-8.12	100.81	107.30
1	C	180	TRP	CE2-CD2-CG	-8.10	100.82	107.30
1	E	321	ARG	NE-CZ-NH2	8.06	124.33	120.30
2	D	90	ASP	CB-CG-OD2	8.03	125.52	118.30
2	D	21	TRP	CD1-CG-CD2	7.93	112.64	106.30
1	C	153	TRP	CD1-CG-CD2	7.85	112.58	106.30
1	E	222	TRP	CD1-CG-CD2	7.84	112.57	106.30
1	C	127	TRP	CD1-CG-CD2	7.83	112.56	106.30
1	E	153	TRP	CD1-CG-CD2	7.82	112.55	106.30
1	A	127	TRP	CD1-CG-CD2	7.76	112.51	106.30
1	A	321	ARG	NE-CZ-NH2	7.70	124.15	120.30
1	C	84	TRP	CD1-CG-CD2	7.69	112.45	106.30
1	C	222	TRP	CD1-CG-CD2	7.64	112.41	106.30
1	E	127	TRP	CD1-CG-CD2	7.63	112.41	106.30
2	F	92	TRP	CD1-CG-CD2	7.62	112.39	106.30
2	B	14	TRP	CE2-CD2-CG	-7.51	101.29	107.30
1	A	222	TRP	CD1-CG-CD2	7.51	112.31	106.30
1	E	84	TRP	CE2-CD2-CG	-7.47	101.32	107.30
2	F	21	TRP	CE2-CD2-CG	-7.44	101.35	107.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	92	TRP	CD1-CG-CD2	7.41	112.23	106.30
2	F	153	ARG	NE-CZ-NH1	-7.38	116.61	120.30
1	A	84	TRP	CE2-CD2-CG	-7.34	101.43	107.30
2	F	14	TRP	CE2-CD2-CG	-7.32	101.44	107.30
2	B	92	TRP	CE2-CD2-CG	-7.30	101.46	107.30
1	C	127	TRP	CE2-CD2-CG	-7.29	101.47	107.30
2	D	92	TRP	CE2-CD2-CG	-7.27	101.48	107.30
1	E	255	ARG	NE-CZ-NH2	7.20	123.90	120.30
2	D	14	TRP	CE2-CD2-CG	-7.19	101.55	107.30
2	B	21	TRP	CE2-CD2-CG	-7.18	101.56	107.30
1	A	105	TYR	CB-CG-CD2	-7.16	116.71	121.00
2	B	90	ASP	CB-CG-OD2	7.15	124.73	118.30
1	C	84	TRP	CE2-CD2-CG	-7.14	101.59	107.30
2	D	153	ARG	NE-CZ-NH1	-7.07	116.76	120.30
1	C	255	ARG	NE-CZ-NH1	-6.99	116.81	120.30
1	E	127	TRP	CE2-CD2-CG	-6.97	101.72	107.30
1	E	222	TRP	CE2-CD2-CG	-6.97	101.72	107.30
1	A	127	TRP	CE2-CD2-CG	-6.90	101.78	107.30
1	A	234	TRP	CD1-CG-CD2	6.89	111.81	106.30
2	F	92	TRP	CE2-CD2-CG	-6.85	101.82	107.30
2	D	54	ARG	NE-CZ-NH2	6.82	123.71	120.30
1	E	255	ARG	NE-CZ-NH1	-6.77	116.91	120.30
1	E	234	TRP	CD1-CG-CD2	6.71	111.67	106.30
1	C	57	ARG	NE-CZ-NH2	6.68	123.64	120.30
1	C	222	TRP	CE2-CD2-CG	-6.67	101.96	107.30
1	A	153	TRP	CD1-CG-CD2	6.65	111.62	106.30
1	E	180	TRP	CG-CD2-CE3	6.64	139.87	133.90
1	A	180	TRP	CG-CD1-NE1	-6.59	103.51	110.10
1	C	153	TRP	CE2-CD2-CG	-6.56	102.06	107.30
1	A	180	TRP	CG-CD2-CE3	6.54	139.79	133.90
1	A	234	TRP	CE2-CD2-CG	-6.53	102.08	107.30
1	A	222	TRP	CE2-CD2-CG	-6.47	102.12	107.30
2	F	132	GLU	CA-CB-CG	-6.40	99.32	113.40
1	C	180	TRP	CG-CD2-CE3	6.33	139.59	133.90
2	F	14	TRP	CG-CD2-CE3	6.29	139.56	133.90
1	E	153	TRP	CE2-CD2-CG	-6.29	102.27	107.30
1	E	180	TRP	CG-CD1-NE1	-6.28	103.82	110.10
2	F	25	ARG	CG-CD-NE	6.28	124.99	111.80
1	C	105	TYR	CB-CG-CD2	-6.23	117.26	121.00
1	E	234	TRP	CE2-CD2-CG	-6.21	102.33	107.30
2	B	25	ARG	NE-CZ-NH1	6.17	123.38	120.30
2	D	21	TRP	CE2-CD2-CG	-6.17	102.37	107.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	132	GLU	CA-CB-CG	-6.16	99.85	113.40
2	F	14	TRP	CG-CD1-NE1	-6.16	103.94	110.10
2	D	14	TRP	CG-CD2-CE3	6.16	139.44	133.90
1	C	234	TRP	CD1-CG-CD2	6.14	111.22	106.30
1	C	321	ARG	CB-CG-CD	-6.12	95.69	111.60
2	B	14	TRP	CG-CD2-CE3	6.09	139.38	133.90
2	B	132	GLU	CA-CB-CG	-6.07	100.04	113.40
1	A	255	ARG	NE-CZ-NH2	6.07	123.33	120.30
1	A	321	ARG	CB-CG-CD	-6.06	95.83	111.60
1	E	321	ARG	CB-CG-CD	-6.03	95.92	111.60
2	B	14	TRP	CG-CD1-NE1	-6.01	104.09	110.10
2	D	14	TRP	CG-CD1-NE1	-6.00	104.10	110.10
1	C	180	TRP	CG-CD1-NE1	-5.96	104.14	110.10
1	A	180	TRP	CB-CG-CD1	-5.92	119.30	127.00
2	B	21	TRP	CG-CD1-NE1	-5.90	104.20	110.10
2	F	90	ASP	CB-CG-OD2	5.84	123.56	118.30
2	D	21	TRP	CG-CD1-NE1	-5.76	104.33	110.10
1	E	180	TRP	CB-CG-CD1	-5.74	119.53	127.00
1	C	197	GLN	CA-CB-CG	-5.74	100.77	113.40
2	F	110	LEU	CA-CB-CG	5.72	128.46	115.30
2	B	92	TRP	CG-CD1-NE1	-5.72	104.38	110.10
1	E	197	GLN	CA-CB-CG	-5.71	100.83	113.40
1	A	197	GLN	CA-CB-CG	-5.69	100.89	113.40
1	C	180	TRP	CB-CG-CD1	-5.67	119.63	127.00
2	F	25	ARG	NE-CZ-NH2	5.67	123.13	120.30
2	F	54	ARG	NE-CZ-NH1	-5.65	117.47	120.30
2	B	153	ARG	NE-CZ-NH1	-5.65	117.48	120.30
1	A	26	VAL	CA-CB-CG2	-5.63	102.46	110.90
2	F	21	TRP	CG-CD1-NE1	-5.62	104.48	110.10
1	E	84	TRP	CG-CD1-NE1	-5.62	104.48	110.10
1	A	84	TRP	CG-CD1-NE1	-5.62	104.48	110.10
1	A	153	TRP	CE2-CD2-CG	-5.61	102.81	107.30
2	B	94	TYR	CB-CG-CD1	-5.58	117.65	121.00
1	C	234	TRP	CE2-CD2-CG	-5.55	102.86	107.30
1	C	175	ASP	CB-CG-OD1	5.55	123.29	118.30
1	A	257	TYR	CB-CG-CD2	-5.50	117.70	121.00
2	D	110	LEU	CA-CB-CG	5.50	127.95	115.30
1	E	57	ARG	NE-CZ-NH2	5.50	123.05	120.30
2	D	94	TYR	CB-CG-CD1	-5.46	117.73	121.00
2	D	170	ARG	NE-CZ-NH2	5.44	123.02	120.30
1	E	84	TRP	CG-CD2-CE3	5.37	138.74	133.90
1	C	201	ARG	NE-CZ-NH1	-5.36	117.62	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	26	VAL	CA-CB-CG2	-5.35	102.87	110.90
2	F	54	ARG	NE-CZ-NH2	5.32	122.96	120.30
1	E	104	ASP	CB-CG-OD1	5.30	123.07	118.30
1	E	153	TRP	CG-CD1-NE1	-5.29	104.81	110.10
1	A	161	TYR	CB-CG-CD2	-5.27	117.84	121.00
1	E	308	TYR	CB-CG-CD2	-5.26	117.84	121.00
1	A	175	ASP	CB-CG-OD1	5.26	123.03	118.30
2	B	153	ARG	CB-CG-CD	-5.24	97.97	111.60
1	A	84	TRP	CG-CD2-CE3	5.23	138.60	133.90
1	C	84	TRP	CG-CD1-NE1	-5.22	104.88	110.10
1	C	153	TRP	CG-CD1-NE1	-5.22	104.88	110.10
2	D	92	TRP	CG-CD1-NE1	-5.22	104.88	110.10
1	C	60	ASP	CB-CG-OD2	5.22	123.00	118.30
2	B	110	LEU	CA-CB-CG	5.20	127.25	115.30
1	E	84	TRP	CB-CG-CD1	-5.19	120.25	127.00
2	F	153	ARG	CB-CG-CD	-5.17	98.15	111.60
2	F	92	TRP	CG-CD1-NE1	-5.16	104.94	110.10
1	A	224	ARG	NE-CZ-NH1	-5.15	117.73	120.30
2	D	153	ARG	CB-CG-CD	-5.14	98.24	111.60
1	E	85	ASP	CB-CG-OD1	5.13	122.92	118.30
2	B	14	TRP	CB-CG-CD1	-5.11	120.36	127.00
1	C	127	TRP	CG-CD1-NE1	-5.10	105.00	110.10
1	A	111	LEU	CB-CA-C	-5.10	100.51	110.20
2	B	54	ARG	NE-CZ-NH1	-5.08	117.76	120.30
1	C	251	LEU	CA-CB-CG	5.07	126.95	115.30
1	A	127	TRP	CG-CD1-NE1	-5.06	105.04	110.10
1	A	251	LEU	CA-CB-CG	5.05	126.93	115.30
2	F	124	ARG	CG-CD-NE	-5.05	101.19	111.80
1	E	26	VAL	CA-CB-CG2	-5.04	103.34	110.90
1	C	111	LEU	CB-CA-C	-5.04	100.63	110.20
1	E	77	ASP	CB-CG-OD2	5.04	122.83	118.30
1	A	104	ASP	CB-CG-OD1	5.03	122.83	118.30
1	E	251	LEU	CA-CB-CG	5.02	126.86	115.30
1	E	111	LEU	CB-CA-C	-5.02	100.67	110.20
1	E	201	ARG	NE-CZ-NH1	-5.00	117.80	120.30

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	341	MAN	C1
5	C	341	MAN	C1
5	E	341	MAN	C1

There are no planarity outliers.

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2533	0	2470	38	0
1	C	2533	0	2470	35	0
1	E	2533	0	2470	32	0
2	B	1421	0	1345	29	0
2	D	1421	0	1345	25	0
2	F	1421	0	1345	26	0
3	A	21	0	18	0	0
3	C	21	0	18	0	0
3	E	21	0	18	0	0
4	A	42	0	39	0	0
4	B	14	0	13	0	0
4	C	42	0	39	0	0
4	D	14	0	13	0	0
4	E	42	0	39	0	0
4	F	14	0	13	0	0
5	A	39	0	34	0	0
5	C	39	0	34	0	0
5	E	39	0	34	0	0
6	A	6	0	0	0	0
6	B	5	0	0	0	0
6	C	7	0	0	0	0
6	D	4	0	0	0	0
6	E	6	0	0	0	0
6	F	5	0	0	0	0
All	All	12243	0	11757	149	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 6.

All (149) 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:C:216:ASN:HB3	1:E:212:THR:HG21	1.62	0.81

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:321:ARG:HH21	1:E:321:ARG:HG2	1.57	0.70
1:E:156:LYS:HD2	1:E:196:VAL:HG23	1.73	0.69
1:A:216:ASN:HB3	1:C:212:THR:HG21	1.75	0.69
1:A:212:THR:HG21	1:E:216:ASN:HB3	1.75	0.69
1:A:321:ARG:HH21	1:A:321:ARG:HG2	1.58	0.68
1:C:321:ARG:HH21	1:C:321:ARG:HG2	1.58	0.68
2:D:27:GLN:HG2	2:D:32:THR:HG22	1.76	0.68
2:B:51:LYS:HG3	1:E:30:THR:HG22	1.76	0.67
1:A:133:ASN:H	1:A:152:ASN:ND2	1.93	0.67
1:C:133:ASN:H	1:C:152:ASN:ND2	1.93	0.66
1:C:77:ASP:O	1:C:80:GLN:HG3	1.97	0.65
1:E:156:LYS:HD2	1:E:196:VAL:CG2	2.28	0.64
2:B:83:TYR:O	2:B:87:THR:HG23	1.98	0.64
1:A:133:ASN:H	1:A:152:ASN:HD21	1.45	0.63
1:E:133:ASN:H	1:E:152:ASN:ND2	1.95	0.63
1:E:77:ASP:O	1:E:80:GLN:HG3	1.99	0.63
2:D:83:TYR:O	2:D:87:THR:HG23	1.99	0.63
1:C:133:ASN:H	1:C:152:ASN:HD21	1.45	0.62
2:F:83:TYR:O	2:F:87:THR:HG23	1.99	0.62
1:E:133:ASN:H	1:E:152:ASN:HD21	1.48	0.62
1:A:30:THR:HG22	2:D:51:LYS:HG3	1.80	0.61
1:C:220:ARG:HD3	1:C:229:ARG:HD3	1.82	0.61
1:E:66:LEU:HD21	1:E:112:VAL:HG23	1.82	0.61
1:A:66:LEU:HD21	1:A:112:VAL:HG23	1.82	0.61
2:D:14:TRP:HE3	2:D:17:MET:HG3	1.66	0.61
1:A:160:THR:HA	1:A:196:VAL:HG21	1.84	0.59
1:A:77:ASP:O	1:A:80:GLN:HG3	2.02	0.59
1:C:66:LEU:HD21	1:C:112:VAL:HG23	1.83	0.59
1:C:30:THR:HG22	2:F:51:LYS:HG3	1.83	0.59
2:F:14:TRP:HE3	2:F:17:MET:HG3	1.67	0.59
2:B:14:TRP:HE3	2:B:17:MET:HG3	1.67	0.59
1:E:220:ARG:HD3	1:E:229:ARG:HD3	1.83	0.59
1:A:220:ARG:HD3	1:A:229:ARG:HD3	1.84	0.58
1:A:117:THR:HG21	1:A:261:ARG:HH21	1.69	0.57
2:D:54:ARG:O	2:D:57:GLU:HB2	2.06	0.56
1:E:117:THR:HG21	1:E:261:ARG:HH21	1.70	0.55
1:E:325:GLU:HG2	2:F:13:GLY:O	2.07	0.54
2:B:54:ARG:O	2:B:57:GLU:HB2	2.07	0.54
1:C:117:THR:HG21	1:C:261:ARG:HH21	1.71	0.54
1:C:228:SER:O	1:C:229:ARG:HD2	2.07	0.54
1:A:52:CYS:HB3	1:A:277:CYS:O	2.07	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:54:ARG:O	2:F:57:GLU:HB2	2.07	0.54
1:A:228:SER:O	1:A:229:ARG:HD2	2.08	0.54
1:A:309:VAL:HG22	2:B:93:SER:HA	1.90	0.54
1:A:53:ASN:OD1	1:A:276:THR:HA	2.08	0.54
1:E:53:ASN:OD1	1:E:276:THR:HA	2.07	0.54
2:F:78:GLN:HA	2:F:81:GLU:OE1	2.07	0.54
2:D:78:GLN:HA	2:D:81:GLU:OE1	2.08	0.53
1:C:53:ASN:OD1	1:C:276:THR:HA	2.08	0.53
1:E:52:CYS:HB3	1:E:277:CYS:O	2.09	0.52
1:E:228:SER:O	1:E:229:ARG:HD2	2.09	0.52
2:B:25:ARG:NH2	2:B:34:GLN:OE1	2.43	0.52
1:A:311:GLN:NE2	2:B:97:GLU:HB2	2.24	0.52
1:A:12:THR:HG23	2:B:133:MET:HE3	1.92	0.52
1:A:111:LEU:HD22	2:F:73:VAL:HG12	1.91	0.52
1:E:311:GLN:NE2	2:F:97:GLU:HB2	2.25	0.52
2:F:58:LYS:HE2	2:F:59:THR:O	2.11	0.51
2:B:58:LYS:HE2	2:B:59:THR:O	2.11	0.51
1:C:309:VAL:HG22	2:D:93:SER:HA	1.92	0.51
2:B:78:GLN:HA	2:B:81:GLU:OE1	2.11	0.51
2:D:58:LYS:HE2	2:D:59:THR:O	2.11	0.51
1:C:52:CYS:HB3	1:C:277:CYS:O	2.11	0.50
2:B:132:GLU:HG2	2:B:134:GLY:H	1.75	0.50
2:F:132:GLU:HG2	2:F:134:GLY:H	1.75	0.50
1:C:311:GLN:NE2	2:D:97:GLU:HB2	2.27	0.50
1:A:66:LEU:HD22	1:A:267:ILE:HD12	1.94	0.50
1:E:309:VAL:HG22	2:F:93:SER:HA	1.93	0.50
2:F:141:TYR:O	2:F:166:ALA:HA	2.12	0.50
2:B:141:TYR:O	2:B:166:ALA:HA	2.12	0.49
2:D:61:GLU:O	2:D:62:LYS:HD2	2.13	0.49
2:D:141:TYR:O	2:D:166:ALA:HA	2.12	0.49
2:D:132:GLU:HG2	2:D:134:GLY:H	1.78	0.49
1:A:222:TRP:CE2	1:A:225:GLY:HA2	2.48	0.49
1:C:66:LEU:HD22	1:C:267:ILE:HD12	1.95	0.48
1:E:222:TRP:CE2	1:E:225:GLY:HA2	2.49	0.48
2:D:25:ARG:HG2	2:D:34:GLN:HG3	1.96	0.48
2:D:25:ARG:CG	2:D:34:GLN:HG3	2.45	0.47
1:C:222:TRP:CE2	1:C:225:GLY:HA2	2.49	0.47
2:B:61:GLU:O	2:B:62:LYS:HD2	2.15	0.47
1:A:57:ARG:HH21	1:A:57:ARG:HG2	1.80	0.47
1:A:326:LYS:HG2	1:A:327:GLN:H	1.80	0.46
2:B:173:ILE:HG21	2:D:167:LEU:HB3	1.96	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:86:LEU:HD11	1:C:268:MET:HB2	1.98	0.46
1:E:299:LYS:HE2	2:F:68:LYS:HB3	1.98	0.46
1:E:15:LEU:HD22	2:F:119:PHE:HA	1.97	0.46
1:A:86:LEU:HD11	1:A:268:MET:HB2	1.98	0.45
2:B:88:LYS:HD2	2:F:87:THR:HB	1.98	0.45
1:A:299:LYS:HE2	2:B:68:LYS:HB3	1.98	0.45
1:E:66:LEU:HD22	1:E:267:ILE:HD12	1.97	0.45
1:C:299:LYS:HE2	2:D:68:LYS:HB3	1.99	0.45
1:C:15:LEU:HD23	2:D:118:LEU:HD13	1.98	0.45
2:B:167:LEU:HB3	2:F:173:ILE:HG21	1.99	0.45
2:F:61:GLU:O	2:F:62:LYS:HD2	2.16	0.45
1:A:15:LEU:HD22	2:B:119:PHE:HA	1.99	0.44
1:E:206:THR:HA	1:E:242:VAL:O	2.18	0.44
2:F:26:HIS:O	2:F:32:THR:HA	2.17	0.44
2:B:174:LYS:HD2	2:B:174:LYS:HA	1.71	0.44
1:A:8:ASN:HB2	2:B:169:ASN:OD1	2.18	0.44
1:C:325:GLU:HB2	2:D:12:ASN:ND2	2.34	0.43
1:C:118:LEU:HD12	1:C:118:LEU:HA	1.79	0.43
1:C:206:THR:HA	1:C:242:VAL:O	2.18	0.43
1:A:206:THR:HA	1:A:242:VAL:O	2.18	0.43
2:B:167:LEU:HA	2:B:167:LEU:HD12	1.84	0.43
1:C:37:THR:HG22	1:C:319:GLY:HA3	2.01	0.43
1:E:57:ARG:HG2	1:E:57:ARG:HH21	1.83	0.43
1:E:86:LEU:HD11	1:E:268:MET:HB2	2.01	0.43
2:D:87:THR:HB	2:F:88:LYS:HD2	2.01	0.43
1:A:15:LEU:HD23	2:B:118:LEU:HD13	1.99	0.43
1:A:37:THR:HG22	1:A:319:GLY:HA3	2.01	0.42
1:C:15:LEU:HD22	2:D:119:PHE:HA	2.00	0.42
2:D:73:VAL:HG12	1:E:111:LEU:HD22	2.02	0.42
1:A:180:TRP:CH2	1:A:213:ILE:HD13	2.55	0.42
1:A:109:ARG:HD2	1:A:109:ARG:HH21	1.66	0.42
1:E:15:LEU:HD23	2:F:118:LEU:HD13	2.02	0.42
1:C:109:ARG:HH21	1:C:109:ARG:HD2	1.63	0.42
2:B:14:TRP:CD2	2:B:25:ARG:NH2	2.87	0.42
1:E:180:TRP:HH2	1:E:213:ILE:HD13	1.85	0.42
2:B:145:ASP:O	2:B:148:CYS:HB3	2.20	0.42
1:C:182:ILE:HD11	1:C:215:PRO:HD3	2.02	0.41
2:D:174:LYS:HA	2:D:174:LYS:HD2	1.71	0.41
1:A:118:LEU:HD12	1:A:118:LEU:HA	1.81	0.41
1:C:89:GLU:OE2	1:C:109:ARG:NH2	2.53	0.41
2:D:167:LEU:HD12	2:D:167:LEU:HA	1.89	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:180:TRP:HH2	1:A:213:ILE:HD13	1.84	0.41
1:A:230:ILE:HG21	1:A:230:ILE:HD13	1.91	0.41
1:E:180:TRP:CH2	1:E:213:ILE:HD13	2.55	0.41
2:F:120:GLU:OE2	2:F:123:ARG:NH2	2.54	0.41
1:A:325:GLU:O	1:A:326:LYS:HD3	2.20	0.41
2:B:127:ARG:NH2	2:F:131:GLU:OE1	2.53	0.41
2:B:131:GLU:OE1	2:D:127:ARG:NH2	2.54	0.41
2:D:120:GLU:OE2	2:D:123:ARG:NH2	2.54	0.41
1:A:190:GLU:O	1:A:194:LEU:HB2	2.20	0.41
1:C:90:ARG:HH21	1:C:271:ASP:HA	1.86	0.41
1:E:190:GLU:O	1:E:194:LEU:HB2	2.21	0.41
1:A:85:ASP:CG	1:A:264:LYS:HZ1	2.24	0.41
1:C:115:SER:O	1:C:261:ARG:HB2	2.21	0.41
1:C:26:VAL:HG11	1:C:317:ALA:HB2	2.02	0.40
2:F:14:TRP:CH2	2:F:25:ARG:HB2	2.55	0.40
2:B:25:ARG:HH21	2:B:25:ARG:HG3	1.86	0.40
2:F:48:ILE:CD1	2:F:107:THR:HG23	2.51	0.40
1:C:151:LEU:O	1:C:255:ARG:NH2	2.54	0.40
1:E:10:THR:HG22	2:F:141:TYR:HA	2.03	0.40
1:E:10:THR:HG21	2:F:139:LYS:HE2	2.04	0.40
1:A:89:GLU:OE2	1:A:109:ARG:NH2	2.54	0.40
1:C:75:HIS:HE1	1:C:94:PHE:O	2.04	0.40
2:B:9:PHE:CD1	2:B:10:ILE:HG13	2.56	0.40
1:C:180:TRP:CH2	1:C:213:ILE:HD13	2.56	0.40
2:B:73:VAL:HG12	1:C:111:LEU:HD22	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	326/328 (99%)	309 (95%)	14 (4%)	3 (1%)	21 64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	326/328 (99%)	310 (95%)	14 (4%)	2 (1%)	30	72
1	E	326/328 (99%)	311 (95%)	14 (4%)	1 (0%)	46	84
2	B	173/175 (99%)	158 (91%)	14 (8%)	1 (1%)	30	72
2	D	173/175 (99%)	160 (92%)	12 (7%)	1 (1%)	30	72
2	F	173/175 (99%)	159 (92%)	13 (8%)	1 (1%)	30	72
All	All	1497/1509 (99%)	1407 (94%)	81 (5%)	9 (1%)	30	72

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	326	LYS
2	D	135	ASN
2	B	135	ASN
2	F	135	ASN
1	A	62	ILE
1	C	62	ILE
1	C	327	GLN
1	E	62	ILE
1	A	158	GLY

### 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	289/289 (100%)	266 (92%)	23 (8%)	15	47
1	C	289/289 (100%)	263 (91%)	26 (9%)	12	41
1	E	289/289 (100%)	264 (91%)	25 (9%)	13	43
2	B	149/149 (100%)	138 (93%)	11 (7%)	17	52
2	D	149/149 (100%)	136 (91%)	13 (9%)	13	43
2	F	149/149 (100%)	138 (93%)	11 (7%)	17	52
All	All	1314/1314 (100%)	1205 (92%)	109 (8%)	14	46

All (109) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	9	SER
1	A	13	LEU
1	A	18	HIS
1	A	26	VAL
1	A	37	THR
1	A	54	ASN
1	A	83	THR
1	A	92	LYS
1	A	95	SER
1	A	108	LEU
1	A	118	LEU
1	A	152	ASN
1	A	155	THR
1	A	156	LYS
1	A	160	THR
1	A	164	LEU
1	A	194	LEU
1	A	224	ARG
1	A	238	LYS
1	A	251	LEU
1	A	283	THR
1	A	309	VAL
1	A	326	LYS
2	B	17	MET
2	B	32	THR
2	B	56	ILE
2	B	72	GLU
2	B	110	LEU
2	B	118	LEU
2	B	127	ARG
2	B	128	GLU
2	B	143	LYS
2	B	169	ASN
2	B	172	GLN
1	C	2	ASP
1	C	13	LEU
1	C	18	HIS
1	C	26	VAL
1	C	37	THR
1	C	54	ASN
1	C	83	THR
1	C	92	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	95	SER
1	C	108	LEU
1	C	118	LEU
1	C	136	SER
1	C	152	ASN
1	C	155	THR
1	C	156	LYS
1	C	160	THR
1	C	164	LEU
1	C	194	LEU
1	C	224	ARG
1	C	238	LYS
1	C	251	LEU
1	C	277	CYS
1	C	283	THR
1	C	309	VAL
1	C	327	GLN
1	C	328	THR
2	D	11	GLU
2	D	17	MET
2	D	27	GLN
2	D	29	SER
2	D	56	ILE
2	D	72	GLU
2	D	110	LEU
2	D	118	LEU
2	D	127	ARG
2	D	128	GLU
2	D	143	LYS
2	D	169	ASN
2	D	172	GLN
1	E	2	ASP
1	E	9	SER
1	E	13	LEU
1	E	18	HIS
1	E	26	VAL
1	E	37	THR
1	E	54	ASN
1	E	83	THR
1	E	92	LYS
1	E	95	SER
1	E	108	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	118	LEU
1	E	136	SER
1	E	152	ASN
1	E	155	THR
1	E	157	SER
1	E	160	THR
1	E	164	LEU
1	E	194	LEU
1	E	210	GLN
1	E	224	ARG
1	E	238	LYS
1	E	251	LEU
1	E	283	THR
1	E	309	VAL
2	F	17	MET
2	F	25	ARG
2	F	56	ILE
2	F	72	GLU
2	F	110	LEU
2	F	118	LEU
2	F	127	ARG
2	F	128	GLU
2	F	143	LYS
2	F	169	ASN
2	F	172	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	132	GLN
1	A	152	ASN
1	A	171	ASN
1	C	132	GLN
1	C	152	ASN
1	C	171	ASN
1	C	327	GLN
2	D	12	ASN
1	E	132	GLN
1	E	152	ASN
1	E	171	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

9 carbohydrates 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$
5	NAG	A	339	1,5	14,14,15	0.53	0	15,19,21	0.90	1 (6%)
5	NAG	A	340	5	14,14,15	0.53	0	15,19,21	0.95	1 (6%)
5	MAN	A	341	5	11,11,12	0.78	0	14,15,17	1.61	2 (14%)
5	NAG	C	339	1,5	14,14,15	0.57	0	15,19,21	0.97	1 (6%)
5	NAG	C	340	5	14,14,15	0.45	0	15,19,21	0.86	1 (6%)
5	MAN	C	341	5	11,11,12	1.05	1 (9%)	14,15,17	1.42	3 (21%)
5	NAG	E	339	1,5	14,14,15	0.51	0	15,19,21	0.96	1 (6%)
5	NAG	E	340	5	14,14,15	0.39	0	15,19,21	0.98	1 (6%)
5	MAN	E	341	5	11,11,12	0.60	0	14,15,17	0.78	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
5	NAG	A	339	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	340	5	-	0/6/23/26	0/1/1/1
5	MAN	A	341	5	1/1/4/5	0/2/19/22	0/1/1/1
5	NAG	C	339	1,5	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	C	340	5	-	0/6/23/26	0/1/1/1
5	MAN	C	341	5	1/1/4/5	0/2/19/22	0/1/1/1
5	NAG	E	339	1,5	-	0/6/23/26	0/1/1/1
5	NAG	E	340	5	-	0/6/23/26	0/1/1/1
5	MAN	E	341	5	1/1/4/5	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	341	MAN	C2-C3	2.50	1.55	1.52

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	341	MAN	C1-C2-C3	-4.08	104.72	109.54
5	E	340	NAG	C4-C3-C2	-2.23	107.76	111.23
5	C	340	NAG	C4-C3-C2	-2.07	108.01	111.23
5	A	340	NAG	C4-C3-C2	-2.01	108.10	111.23
5	C	341	MAN	O5-C1-C2	-2.01	107.60	110.86
5	A	339	NAG	C1-O5-C5	2.03	114.82	112.25
5	C	339	NAG	C1-O5-C5	2.08	114.89	112.25
5	E	339	NAG	C1-O5-C5	2.38	115.27	112.25
5	A	341	MAN	C3-C4-C5	2.45	114.47	110.20
5	C	341	MAN	O3-C3-C2	2.77	115.00	110.00
5	C	341	MAN	C1-O5-C5	2.92	115.95	112.25

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	C	341	MAN	C1
5	A	341	MAN	C1
5	E	341	MAN	C1

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry [i](#)

15 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	SIA	A	329	-	17,21,21	0.82	0	19,31,31	1.33	4 (21%)
4	NAG	A	330	1	14,14,15	0.84	1 (7%)	15,19,21	1.07	1 (6%)
4	NAG	A	334	1	14,14,15	0.86	0	15,19,21	1.18	1 (6%)
4	NAG	A	348	1	14,14,15	0.61	0	15,19,21	0.96	1 (6%)
4	NAG	B	401	2	14,14,15	0.64	0	15,19,21	0.94	1 (6%)
3	SIA	C	329	-	17,21,21	0.78	1 (5%)	19,31,31	1.28	3 (15%)
4	NAG	C	330	1	14,14,15	0.97	1 (7%)	15,19,21	1.10	1 (6%)
4	NAG	C	334	1	14,14,15	0.76	0	15,19,21	1.30	1 (6%)
4	NAG	C	348	1	14,14,15	0.64	0	15,19,21	0.94	1 (6%)
4	NAG	D	401	2	14,14,15	0.55	0	15,19,21	0.91	1 (6%)
3	SIA	E	329	-	17,21,21	0.89	1 (5%)	19,31,31	1.28	4 (21%)
4	NAG	E	330	1	14,14,15	0.90	1 (7%)	15,19,21	1.05	1 (6%)
4	NAG	E	334	1	14,14,15	0.66	0	15,19,21	1.26	1 (6%)
4	NAG	E	348	1	14,14,15	0.65	0	15,19,21	0.92	1 (6%)
4	NAG	F	401	2	14,14,15	0.51	0	15,19,21	0.90	1 (6%)

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
3	SIA	A	329	-	-	0/14/38/38	0/1/1/1
4	NAG	A	330	1	-	0/6/23/26	0/1/1/1
4	NAG	A	334	1	-	0/6/23/26	0/1/1/1
4	NAG	A	348	1	-	0/6/23/26	0/1/1/1
4	NAG	B	401	2	-	0/6/23/26	0/1/1/1
3	SIA	C	329	-	-	0/14/38/38	0/1/1/1
4	NAG	C	330	1	-	0/6/23/26	0/1/1/1
4	NAG	C	334	1	-	0/6/23/26	0/1/1/1
4	NAG	C	348	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	D	401	2	-	0/6/23/26	0/1/1/1
3	SIA	E	329	-	-	0/14/38/38	0/1/1/1
4	NAG	E	330	1	-	0/6/23/26	0/1/1/1
4	NAG	E	334	1	-	0/6/23/26	0/1/1/1
4	NAG	E	348	1	-	0/6/23/26	0/1/1/1
4	NAG	F	401	2	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	329	SIA	C3-C2	2.03	1.54	1.51
3	E	329	SIA	C3-C2	2.45	1.55	1.51
4	A	330	NAG	C4-C5	2.47	1.58	1.53
4	E	330	NAG	C4-C5	2.54	1.58	1.53
4	C	330	NAG	C4-C5	2.99	1.59	1.53

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	329	SIA	C4-C5-N5	-2.39	105.22	110.41
3	C	329	SIA	O2-C2-O6	-2.38	106.33	110.22
3	A	329	SIA	C4-C5-N5	-2.35	105.31	110.41
3	E	329	SIA	C4-C5-N5	-2.31	105.39	110.41
3	A	329	SIA	O2-C2-O6	-2.19	106.64	110.22
3	E	329	SIA	O2-C2-O6	-2.09	106.81	110.22
4	C	348	NAG	C8-C7-N2	2.04	120.01	116.11
4	F	401	NAG	C1-O5-C5	2.05	114.85	112.25
3	E	329	SIA	C6-C5-N5	2.05	114.65	111.07
3	A	329	SIA	C7-C6-C5	2.06	117.44	114.32
3	A	329	SIA	C6-C5-N5	2.09	114.72	111.07
4	A	348	NAG	C8-C7-N2	2.10	120.12	116.11
4	B	401	NAG	C1-O5-C5	2.11	114.93	112.25
3	E	329	SIA	C7-C6-C5	2.11	117.52	114.32
4	D	401	NAG	C1-O5-C5	2.17	115.00	112.25
4	E	348	NAG	C8-C7-N2	2.19	120.29	116.11
3	C	329	SIA	C6-C5-N5	2.29	115.06	111.07
4	E	330	NAG	C1-O5-C5	2.65	115.61	112.25
4	C	330	NAG	C1-O5-C5	2.74	115.73	112.25
4	A	330	NAG	C1-O5-C5	2.88	115.91	112.25
4	A	334	NAG	C1-O5-C5	3.10	116.18	112.25
4	C	334	NAG	C1-O5-C5	3.24	116.36	112.25
4	E	334	NAG	C1-O5-C5	3.58	116.79	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	A	328/328 (100%)	-0.82	10 (3%) 54 25	4, 26, 56, 143	0
1	C	328/328 (100%)	-0.83	8 (2%) 62 32	4, 26, 54, 142	0
1	E	328/328 (100%)	-0.85	8 (2%) 62 32	4, 26, 54, 146	0
2	B	175/175 (100%)	-0.91	2 (1%) 82 58	3, 21, 53, 101	0
2	D	175/175 (100%)	-0.99	1 (0%) 90 73	2, 20, 50, 101	0
2	F	175/175 (100%)	-0.95	2 (1%) 82 58	2, 20, 51, 101	0
All	All	1509/1509 (100%)	-0.87	31 (2%) 67 36	2, 24, 54, 146	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	175	GLY	7.1
1	E	2	ASP	7.1
1	A	6	ASN	6.2
1	A	4	PRO	6.1
1	C	1	GLN	5.7
1	A	5	GLY	5.7
1	A	2	ASP	5.6
1	E	3	LEU	5.6
1	E	5	GLY	5.1
2	F	174	LYS	4.9
2	F	175	GLY	4.8
1	C	5	GLY	4.7
1	C	6	ASN	4.6
1	E	1	GLN	4.6
2	B	175	GLY	4.4
1	E	8	ASN	4.4
1	C	4	PRO	4.2
1	C	8	ASN	4.2
1	E	4	PRO	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	2	ASP	3.9
1	E	6	ASN	3.8
1	A	7	ASP	3.7
1	A	8	ASN	3.6
1	A	1	GLN	3.6
1	A	327	GLN	2.8
1	A	328	THR	2.8
1	A	3	LEU	2.7
2	B	174	LYS	2.4
1	E	7	ASP	2.3
1	C	7	ASP	2.2
1	C	3	LEU	2.0

## 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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	C	340	14/15	0.86	0.22	2.08	60,61,65,66	0
5	NAG	A	340	14/15	0.90	0.19	1.67	60,61,64,66	0
5	NAG	E	340	14/15	0.95	0.14	0.25	60,61,64,66	0
5	NAG	E	339	14/15	0.96	0.12	-0.57	51,53,55,57	0
5	MAN	A	341	11/12	0.87	0.28	-	68,69,70,71	0
5	NAG	A	339	14/15	0.93	0.19	-	51,55,56,58	0
5	MAN	E	341	11/12	0.89	0.24	-	68,70,71,71	0
5	MAN	C	341	11/12	0.83	0.31	-	69,71,71,71	0
5	NAG	C	339	14/15	0.90	0.20	-	51,54,57,57	0

## 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	C	330	14/15	0.89	0.23	10.57	42,45,46,46	0
4	NAG	E	330	14/15	0.89	0.27	8.01	42,46,47,47	0
3	SIA	C	329	21/21	0.81	0.31	5.61	35,36,37,38	21
4	NAG	E	348	14/15	0.89	0.25	4.54	40,43,46,46	0
4	NAG	C	348	14/15	0.88	0.25	4.47	40,44,45,46	0
4	NAG	A	330	14/15	0.94	0.19	4.16	41,45,46,46	0
4	NAG	A	348	14/15	0.92	0.18	4.08	39,43,45,46	0
4	NAG	D	401	14/15	0.87	0.28	3.39	55,58,59,60	0
4	NAG	B	401	14/15	0.85	0.27	3.15	55,58,59,60	0
4	NAG	F	401	14/15	0.87	0.27	1.71	54,57,59,59	0
3	SIA	A	329	21/21	0.91	0.22	1.48	35,36,36,37	21
3	SIA	E	329	21/21	0.89	0.23	0.85	34,35,36,36	21
4	NAG	C	334	14/15	0.93	0.21	-	39,42,43,45	0
4	NAG	A	334	14/15	0.92	0.18	-	39,42,43,44	0
4	NAG	E	334	14/15	0.94	0.16	-	39,41,42,44	0

## 6.5 Other polymers

There are no such residues in this entry.