wwPDB X-ray Structure Validation Summary Report (i)

Feb 28, 2024 – 12:33 pm GMT

PDB ID	:	8PN5
Title	:	Crystal structure of the HC7-Glu200Ala mutant complexed to a triglycopeptide
Authors	:	Taleb, V.; Hurtado-Guerrero, R.
Deposited on	:	2023-06-29
Resolution	:	1.72 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

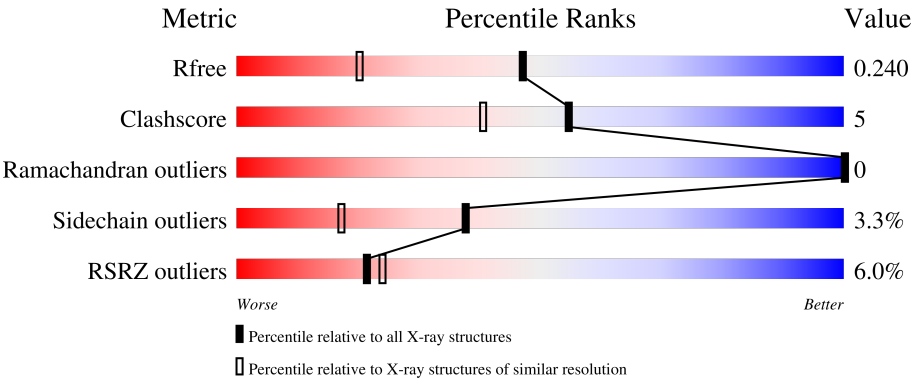
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.72 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	5722 (1.74-1.70)
Clashscore	141614	6152 (1.74-1.70)
Ramachandran outliers	138981	6051 (1.74-1.70)
Sidechain outliers	138945	6051 (1.74-1.70)
RSRZ outliers	127900	5629 (1.74-1.70)

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 >=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 <=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	627	50% · 45%
1	B	627	2% 48% 6% · 45%
1	C	627	4% 49% 5% · 45%
1	D	627	4% 49% 5% 45%
1	E	627	4% 49% 5% · 45%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	627	3% 49% 5% 45%
1	G	627	5% 49% 5% 46%
1	H	627	2% 50% 45%
2	I	15	13% 67% 7% 7% 20%
2	J	15	7% 53% 27% 20%
2	K	15	7% 53% 20% 7% 20%
2	L	15	7% 53% 27% 20%
2	M	15	7% 73% 7% 20%
2	N	15	7% 53% 13% 7% 27%
2	O	15	7% 60% 20% 20%
2	P	15	7% 67% 7% 7% 20%

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	GOL	G	701	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 25304 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 DUF3472 domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	342	Total	C	N	O	S	0	2	0
			2745	1740	463	538	4			
1	B	343	Total	C	N	O	S	0	2	0
			2748	1743	463	538	4			
1	C	342	Total	C	N	O	S	0	3	0
			2741	1740	462	535	4			
1	D	342	Total	C	N	O	S	0	0	0
			2735	1734	462	535	4			
1	E	344	Total	C	N	O	S	0	1	0
			2746	1740	464	538	4			
1	F	344	Total	C	N	O	S	0	2	0
			2752	1745	464	539	4			
1	G	341	Total	C	N	O	S	0	1	0
			2728	1732	461	531	4			
1	H	342	Total	C	N	O	S	0	0	0
			2730	1731	462	533	4			

- Molecule 2 is a protein called Triglycopeptide.

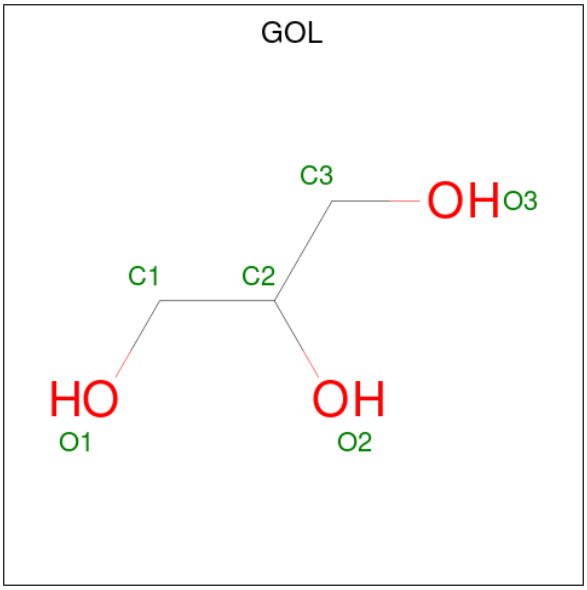
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	I	12	Total	C	N	O	0	0	1
			72	44	13	15			
2	J	12	Total	C	N	O	0	0	1
			72	44	13	15			
2	K	12	Total	C	N	O	0	0	1
			72	44	13	15			
2	L	12	Total	C	N	O	0	0	1
			72	44	13	15			
2	M	12	Total	C	N	O	0	0	1
			72	44	13	15			
2	N	11	Total	C	N	O	0	0	1
			67	41	12	14			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	O	12	Total C N O 72 44 13 15	0	0	1
2	P	12	Total C N O 72 44 13 15	0	0	1

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	D	1	Total C O 6 3 3	0	0
3	E	1	Total C O 6 3 3	0	0
3	F	1	Total C O 6 3 3	0	0
3	G	1	Total C O 6 3 3	0	0

Continued on next page...

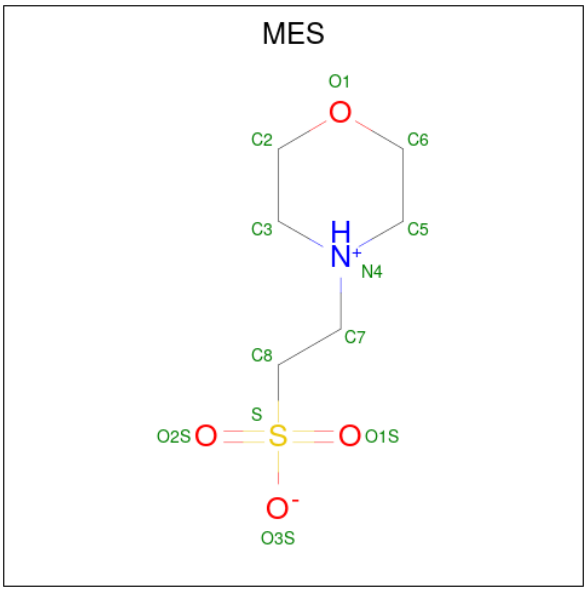
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	H	1	Total C O 6 3 3	0	0
3	H	1	Total C O 6 3 3	0	0
3	H	1	Total C O 6 3 3	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

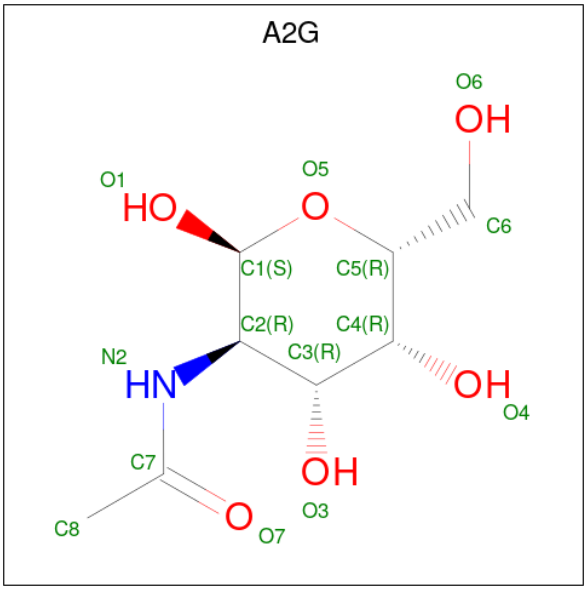
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Zn 1 1	0	0
4	B	1	Total Zn 1 1	0	0
4	E	1	Total Zn 1 1	0	0
4	F	2	Total Zn 2 2	0	0
4	G	2	Total Zn 2 2	0	0
4	H	1	Total Zn 1 1	0	0

- Molecule 5 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	F	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 6 is 2-acetamido-2-deoxy-alpha-D-galactopyranose (three-letter code: A2G) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	I	1	Total	C	N	O		0	0
			14	8	1	5			
6	I	1	Total	C	N	O		0	0
			14	8	1	5			
6	I	1	Total	C	N	O		0	0
			14	8	1	5			
6	J	1	Total	C	N	O		0	0
			14	8	1	5			
6	J	1	Total	C	N	O		0	0
			14	8	1	5			
6	J	1	Total	C	N	O		0	0
			14	8	1	5			
6	K	1	Total	C	N	O		0	0
			14	8	1	5			
6	K	1	Total	C	N	O		0	0
			14	8	1	5			
6	K	1	Total	C	N	O		0	0
			14	8	1	5			
6	L	1	Total	C	N	O		0	0
			14	8	1	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	L	1	Total C N O 14 8 1 5	0	0
6	L	1	Total C N O 14 8 1 5	0	0
6	M	1	Total C N O 14 8 1 5	0	0
6	M	1	Total C N O 14 8 1 5	0	0
6	M	1	Total C N O 14 8 1 5	0	0
6	N	1	Total C N O 14 8 1 5	0	0
6	N	1	Total C N O 14 8 1 5	0	0
6	N	1	Total C N O 14 8 1 5	0	0
6	O	1	Total C N O 14 8 1 5	0	0
6	O	1	Total C N O 14 8 1 5	0	0
6	O	1	Total C N O 14 8 1 5	0	0
6	P	1	Total C N O 14 8 1 5	0	0
6	P	1	Total C N O 14 8 1 5	0	0
6	P	1	Total C N O 14 8 1 5	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	344	Total O 344 344	0	0
7	B	332	Total O 332 332	0	0
7	C	312	Total O 312 312	0	0
7	D	221	Total O 221 221	0	0
7	E	217	Total O 217 217	0	0

Continued on next page...

Continued from previous page...

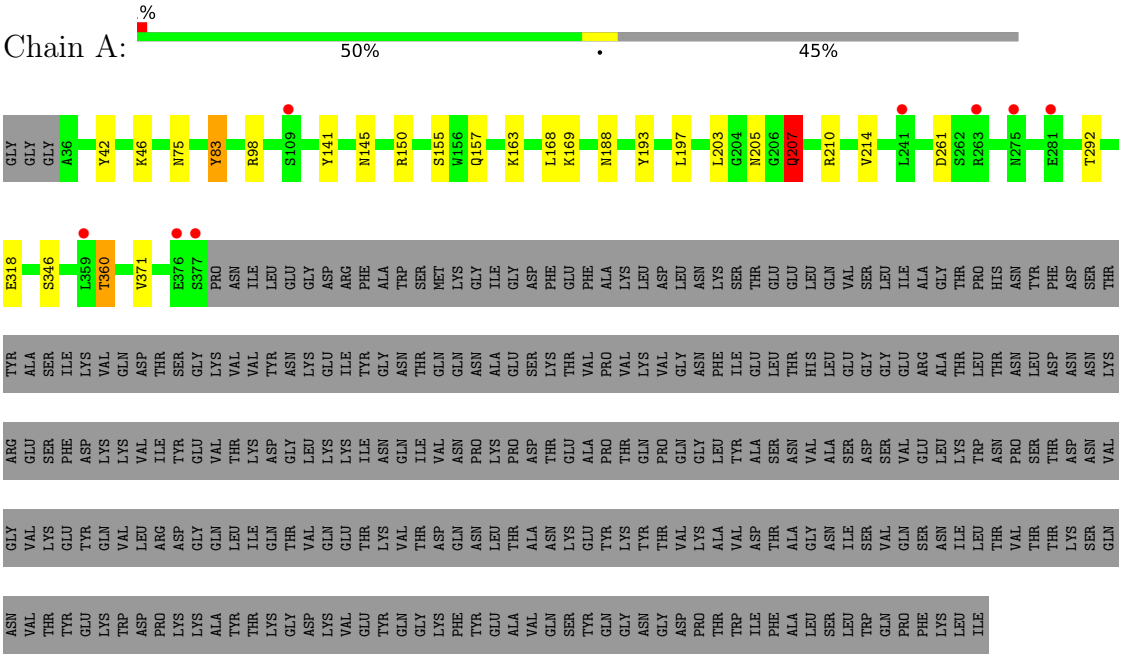
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	F	251	Total O 251 251	0	0
7	G	191	Total O 191 191	0	0
7	H	277	Total O 277 277	0	0
7	I	31	Total O 31 31	0	0
7	J	36	Total O 36 36	0	0
7	K	38	Total O 38 38	0	0
7	L	29	Total O 29 29	0	0
7	M	21	Total O 21 21	0	0
7	N	32	Total O 32 32	0	0
7	O	21	Total O 21 21	0	0
7	P	27	Total O 27 27	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.

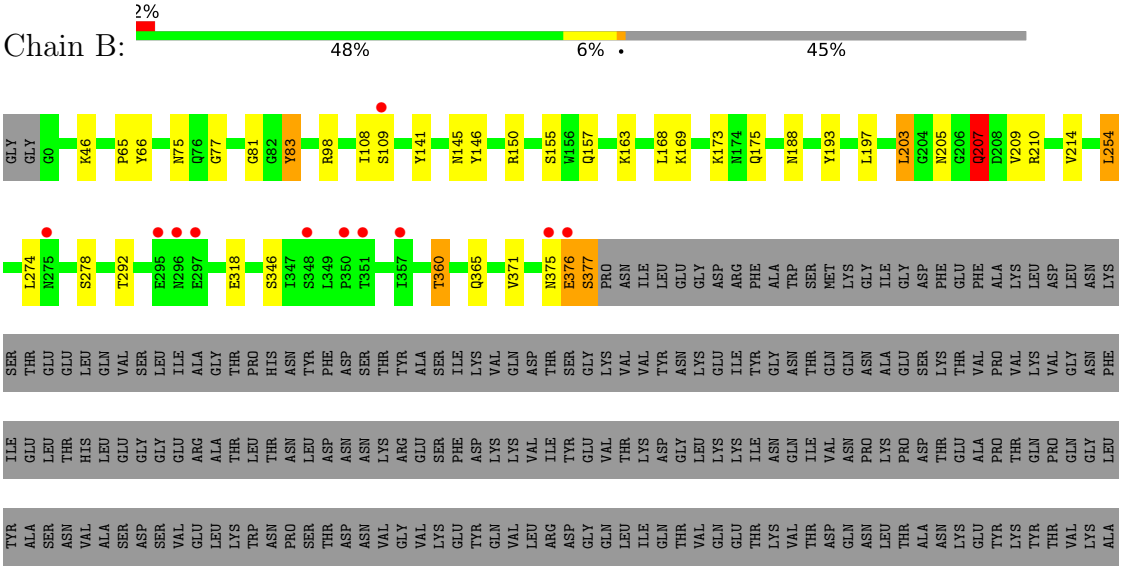
• Molecule 1: DUF3472 domain-containing protein

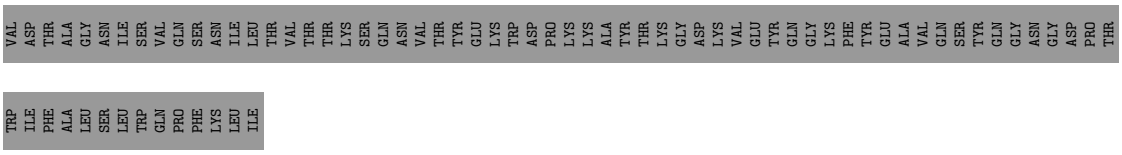
Chain A:



• Molecule 1: DUF3472 domain-containing protein

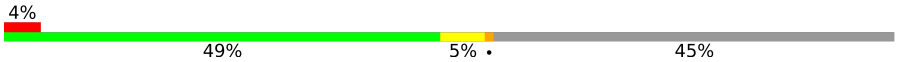
Chain B:

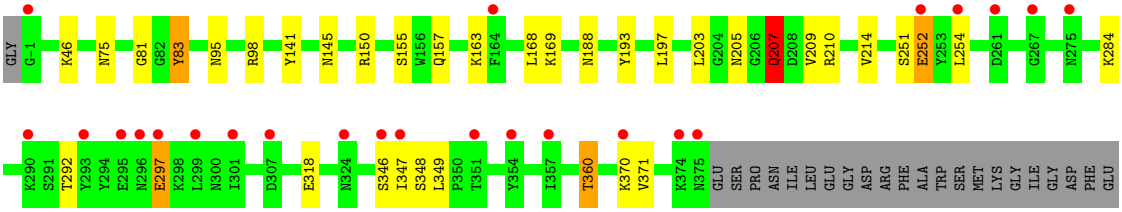


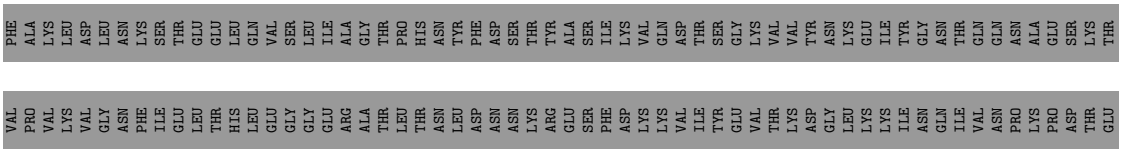


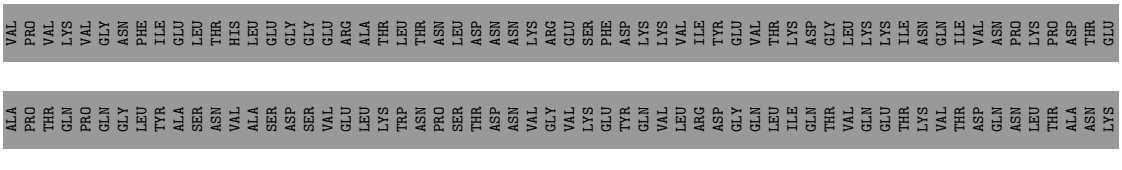
• Molecule 1: DUF3472 domain-containing protein

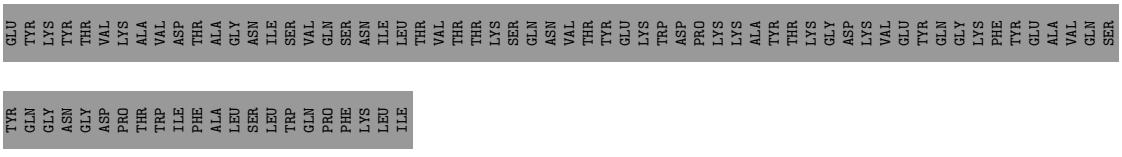
Chain C:

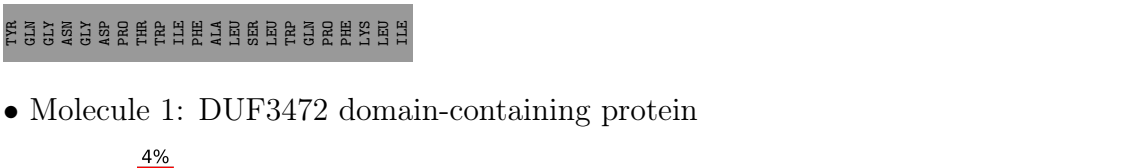


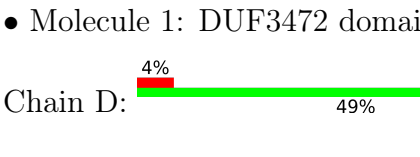






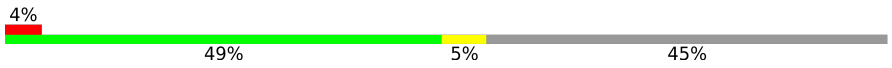


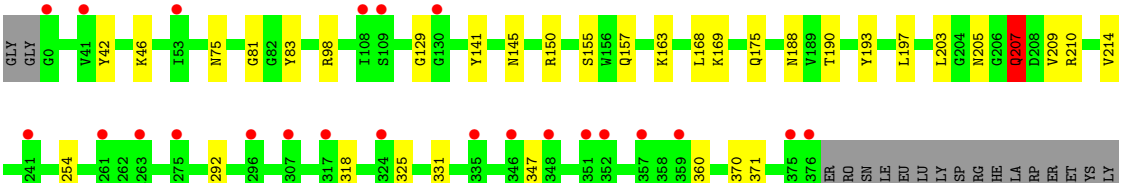


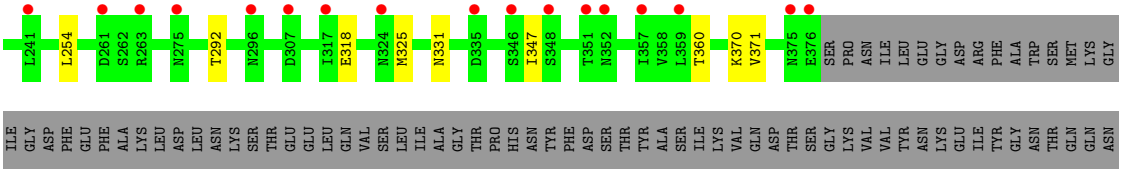


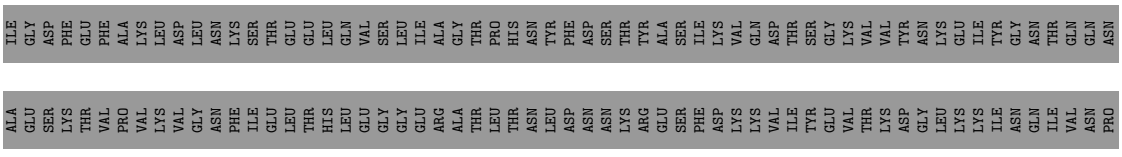
• Molecule 1: DUF3472 domain-containing protein

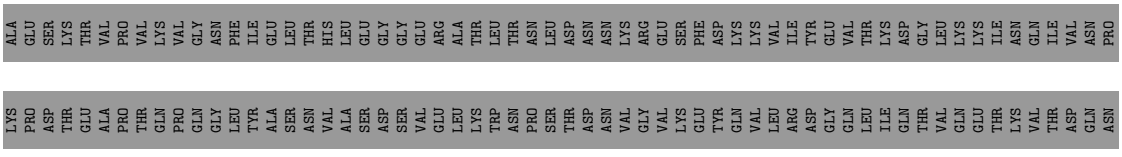
Chain D:

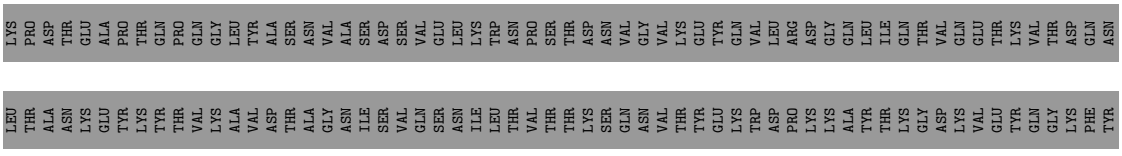


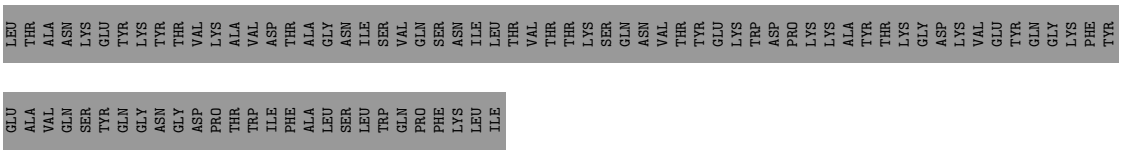






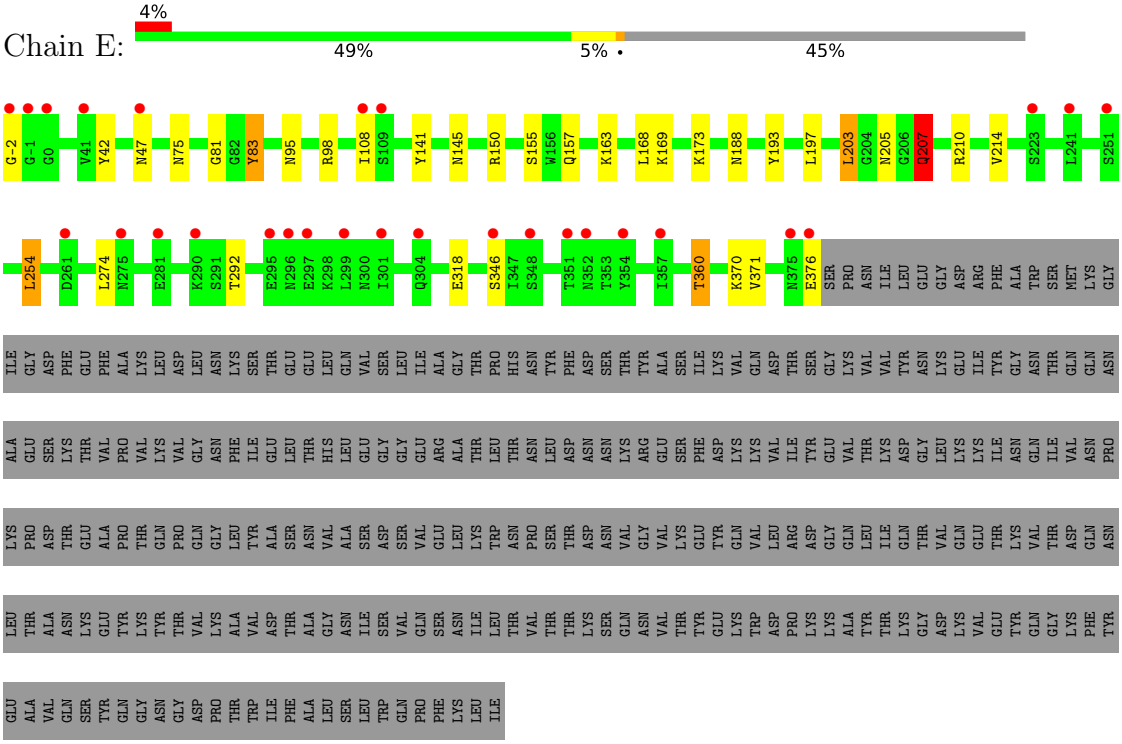






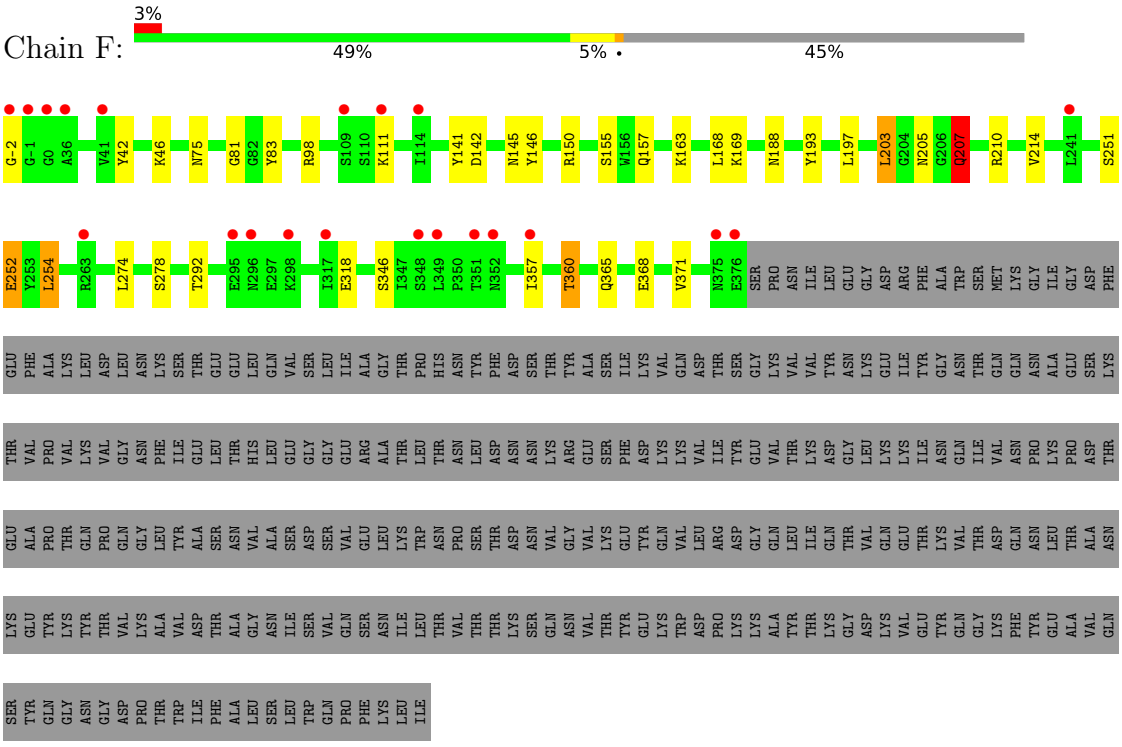
• Molecule 1: DUF3472 domain-containing protein

Chain E:



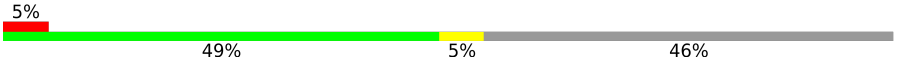
• Molecule 1: DUF3472 domain-containing protein

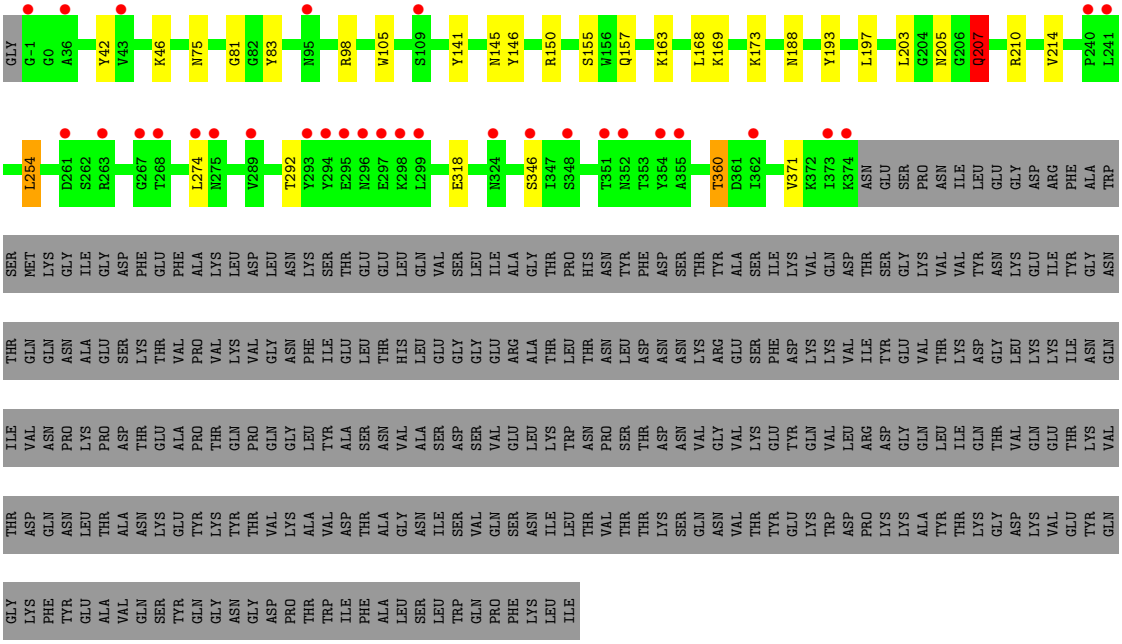
Chain F:



• Molecule 1: DUF3472 domain-containing protein

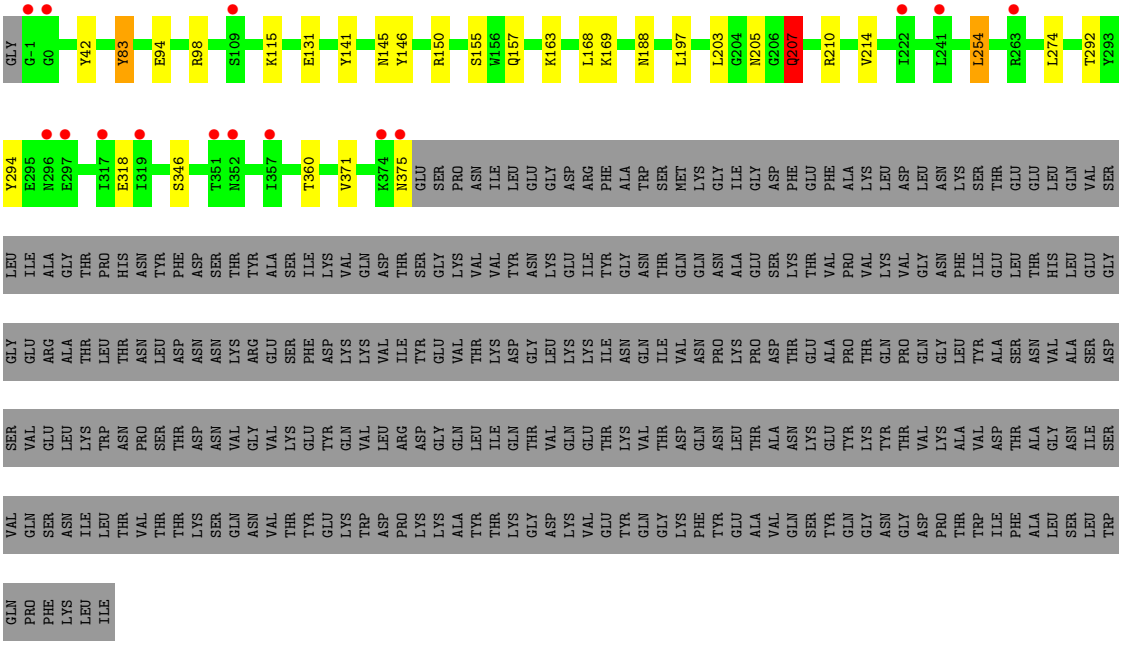
Chain G:





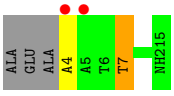
- Molecule 1: DUF3472 domain-containing protein

Chain H:



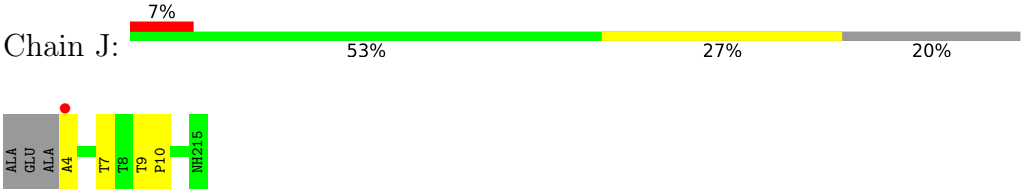
- Molecule 2: Triglycopeptide

Chain I:



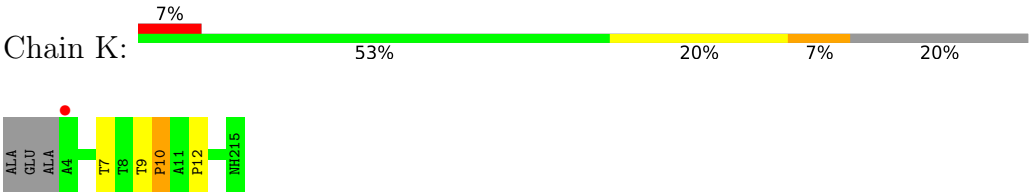
• Molecule 2: Triglycopeptide

Chain J:



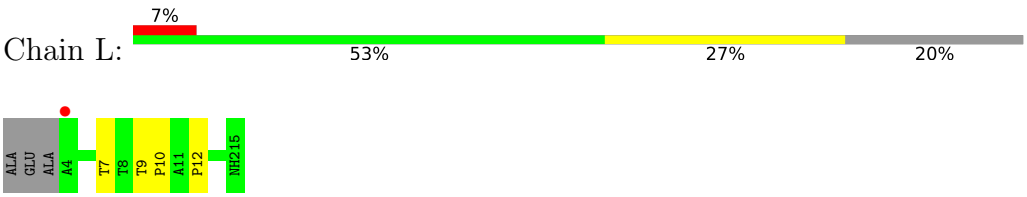
• Molecule 2: Triglycopeptide

Chain K:



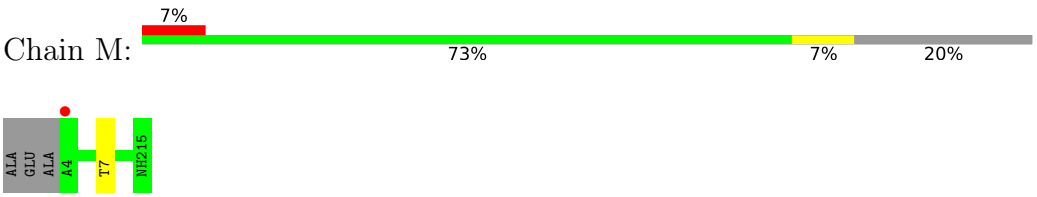
• Molecule 2: Triglycopeptide

Chain L:



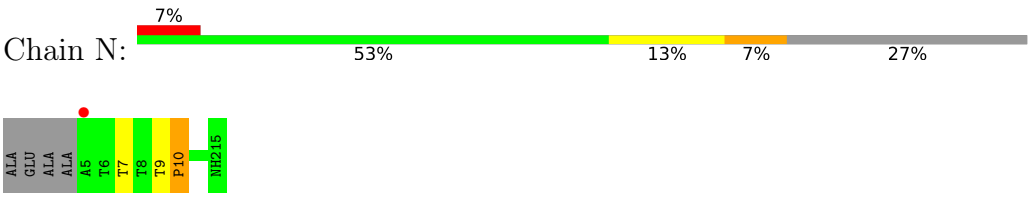
• Molecule 2: Triglycopeptide

Chain M:



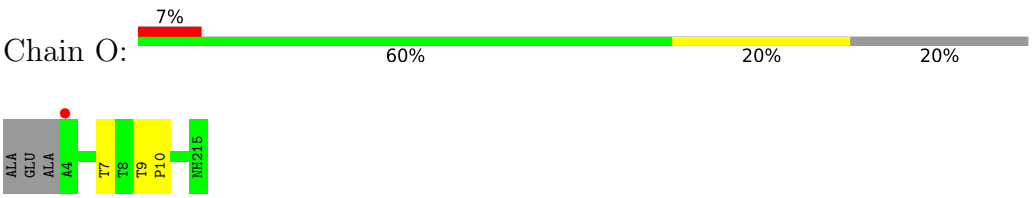
• Molecule 2: Triglycopeptide

Chain N:



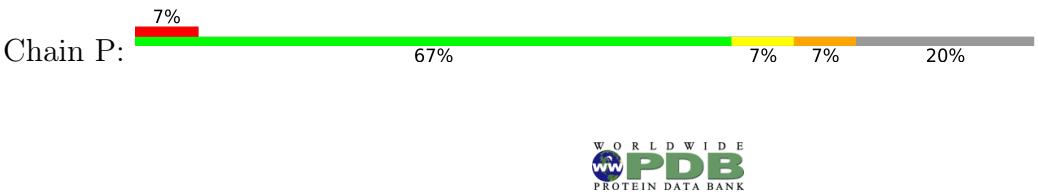
• Molecule 2: Triglycopeptide

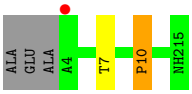
Chain O:



• Molecule 2: Triglycopeptide

Chain P:




ALA
GLU
ALA
A4
T7
P10
NH215

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	163.33Å 260.64Å 82.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.97 – 1.72 19.97 – 1.72	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.97-1.72) 100.0 (19.97-1.72)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.50 (at 1.72Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.208 , 0.235 0.215 , 0.240	Depositor DCC
R_{free} test set	14960 reflections (4.03%)	wwPDB-VP
Wilson B-factor (Å^2)	24.7	Xtriage
Anisotropy	0.637	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 39.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	25304	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.44% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: NH2, MES, ZN, A2G, 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	A	0.70	1/2826 (0.0%)	0.83	2/3841 (0.1%)
1	B	0.73	1/2829 (0.0%)	0.83	2/3845 (0.1%)
1	C	0.69	1/2825 (0.0%)	0.81	2/3840 (0.1%)
1	D	0.67	1/2810 (0.0%)	0.80	1/3819 (0.0%)
1	E	0.65	1/2824 (0.0%)	0.80	2/3837 (0.1%)
1	F	0.70	1/2833 (0.0%)	0.82	1/3850 (0.0%)
1	G	0.66	1/2806 (0.0%)	0.79	1/3812 (0.0%)
1	H	0.67	1/2805 (0.0%)	0.81	2/3812 (0.1%)
2	I	1.40	0/72	1.54	2/100 (2.0%)
2	J	1.48	0/72	1.30	0/100
2	K	1.56	2/72 (2.8%)	1.48	0/100
2	L	1.33	1/72 (1.4%)	1.36	0/100
2	M	1.33	0/72	1.12	0/100
2	N	1.33	0/67	1.42	0/93
2	O	1.17	0/72	1.20	0/100
2	P	1.62	1/72 (1.4%)	1.48	0/100
All	All	0.71	12/23129 (0.1%)	0.83	15/31449 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	H	0	1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	318	GLU	CD-OE2	10.21	1.36	1.25
1	F	318	GLU	CD-OE2	8.98	1.35	1.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	318	GLU	CD-OE1	8.21	1.34	1.25
1	E	318	GLU	CD-OE2	7.49	1.33	1.25
1	H	318	GLU	CD-OE2	7.43	1.33	1.25

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	4	ALA	N-CA-CB	-6.46	101.06	110.10
1	A	207	GLN	CB-CA-C	-5.57	99.27	110.40
1	E	83	TYR	CB-CG-CD1	5.41	124.24	121.00
1	H	207	GLN	CB-CA-C	-5.39	99.62	110.40
1	E	207	GLN	CB-CA-C	-5.37	99.67	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	131	GLU	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	2745	0	2613	21	0
1	B	2748	0	2622	33	1
1	C	2741	0	2621	34	0
1	D	2735	0	2603	30	0
1	E	2746	0	2614	29	0
1	F	2752	0	2622	31	1
1	G	2728	0	2607	24	0
1	H	2730	0	2600	22	0
2	I	72	0	71	0	0
2	J	72	0	71	2	0
2	K	72	0	71	1	0
2	L	72	0	71	1	0
2	M	72	0	71	0	0
2	N	67	0	66	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	O	72	0	71	1	0
2	P	72	0	71	2	0
3	A	12	0	16	1	0
3	B	6	0	8	2	0
3	C	12	0	16	0	0
3	D	6	0	8	0	0
3	E	6	0	8	1	0
3	F	6	0	8	0	0
3	G	6	0	8	5	0
3	H	18	0	24	3	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	E	1	0	0	0	0
4	F	2	0	0	0	0
4	G	2	0	0	0	0
4	H	1	0	0	0	0
5	F	12	0	13	4	0
6	I	42	0	36	2	0
6	J	42	0	35	0	0
6	K	42	0	35	0	0
6	L	42	0	36	1	0
6	M	42	0	36	1	0
6	N	42	0	36	1	0
6	O	42	0	36	4	0
6	P	42	0	36	1	0
7	A	344	0	0	3	0
7	B	332	0	0	4	0
7	C	312	0	0	14	0
7	D	221	0	0	4	0
7	E	217	0	0	5	0
7	F	251	0	0	7	0
7	G	191	0	0	1	0
7	H	277	0	0	0	0
7	I	31	0	0	0	0
7	J	36	0	0	1	0
7	K	38	0	0	0	0
7	L	29	0	0	0	0
7	M	21	0	0	0	0
7	N	32	0	0	0	0
7	O	21	0	0	0	0
7	P	27	0	0	0	0
All	All	25304	0	21860	232	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 232 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:175:GLN:HG2	7:B:1370:HOH:O	1.31	1.28
1:C:349:LEU:HG	7:C:806:HOH:O	1.36	1.21
1:B:109:SER:HB3	7:B:1350:HOH:O	1.39	1.19
1:D:175:GLN:HG2	7:D:987:HOH:O	1.37	1.19
1:E:47:ASN:HB2	7:E:974:HOH:O	1.53	1.08

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:278:SER:OG	1:F:278:SER:OG[3_554]	1.87	0.33

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	342/627 (54%)	334 (98%)	8 (2%)	0	100	100
1	B	343/627 (55%)	332 (97%)	11 (3%)	0	100	100
1	C	343/627 (55%)	333 (97%)	10 (3%)	0	100	100
1	D	340/627 (54%)	330 (97%)	10 (3%)	0	100	100
1	E	343/627 (55%)	333 (97%)	10 (3%)	0	100	100
1	F	344/627 (55%)	335 (97%)	9 (3%)	0	100	100
1	G	340/627 (54%)	331 (97%)	9 (3%)	0	100	100
1	H	340/627 (54%)	330 (97%)	10 (3%)	0	100	100
2	I	10/15 (67%)	10 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	J	10/15 (67%)	9 (90%)	1 (10%)	0	100	100
2	K	10/15 (67%)	10 (100%)	0	0	100	100
2	L	10/15 (67%)	10 (100%)	0	0	100	100
2	M	10/15 (67%)	9 (90%)	1 (10%)	0	100	100
2	N	9/15 (60%)	9 (100%)	0	0	100	100
2	O	10/15 (67%)	10 (100%)	0	0	100	100
2	P	10/15 (67%)	10 (100%)	0	0	100	100
All	All	2814/5136 (55%)	2735 (97%)	79 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	298/545 (55%)	291 (98%)	7 (2%)	50	31
1	B	298/545 (55%)	287 (96%)	11 (4%)	34	14
1	C	297/545 (54%)	287 (97%)	10 (3%)	37	16
1	D	295/545 (54%)	288 (98%)	7 (2%)	49	29
1	E	296/545 (54%)	288 (97%)	8 (3%)	44	25
1	F	297/545 (54%)	287 (97%)	10 (3%)	37	16
1	G	294/545 (54%)	286 (97%)	8 (3%)	44	25
1	H	294/545 (54%)	285 (97%)	9 (3%)	40	19
2	I	7/8 (88%)	6 (86%)	1 (14%)	3	0
2	J	7/8 (88%)	6 (86%)	1 (14%)	3	0
2	K	7/8 (88%)	6 (86%)	1 (14%)	3	0
2	L	7/8 (88%)	6 (86%)	1 (14%)	3	0
2	M	7/8 (88%)	6 (86%)	1 (14%)	3	0
2	N	7/8 (88%)	5 (71%)	2 (29%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	O	7/8 (88%)	6 (86%)	1 (14%)	3 0
2	P	7/8 (88%)	6 (86%)	1 (14%)	3 0
All	All	2425/4424 (55%)	2346 (97%)	79 (3%)	38 17

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	210	ARG
2	J	7	THR
1	G	346	SER
1	H	207	GLN
2	N	7	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	375	ASN
1	H	188	ASN
1	E	365	GLN
1	H	157	GLN
1	H	365	GLN

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)

Of 45 ligands modelled in this entry, 8 are monoatomic - leaving 37 for Mogul analysis.

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	$\#\|Z\| > 2$
6	A2G	P	103	2	14,14,15	1.77	5 (35%)	17,19,21	1.82	6 (35%)
6	A2G	K	103	2	14,14,15	2.74	5 (35%)	17,19,21	1.85	6 (35%)
6	A2G	M	102	2	14,14,15	1.03	0	17,19,21	1.36	2 (11%)
3	GOL	C	701	-	5,5,5	0.18	0	5,5,5	0.58	0
3	GOL	C	702	-	5,5,5	0.16	0	5,5,5	0.37	0
6	A2G	O	102	2	14,14,15	1.25	1 (7%)	17,19,21	1.19	1 (5%)
6	A2G	J	101	2	14,14,15	2.38	6 (42%)	17,19,21	1.51	4 (23%)
3	GOL	A	1002	-	5,5,5	0.20	0	5,5,5	0.61	0
6	A2G	I	103	2	14,14,15	2.33	4 (28%)	17,19,21	1.48	2 (11%)
6	A2G	I	101	2	14,14,15	1.82	3 (21%)	17,19,21	1.08	1 (5%)
3	GOL	B	1001	-	5,5,5	0.20	0	5,5,5	0.46	0
6	A2G	K	101	2	14,14,15	1.67	4 (28%)	17,19,21	1.65	5 (29%)
3	GOL	G	701	-	5,5,5	0.10	0	5,5,5	0.33	0
6	A2G	I	102	2	14,14,15	1.61	2 (14%)	17,19,21	1.43	3 (17%)
5	MES	F	702	-	12,12,12	0.97	0	14,16,16	1.36	2 (14%)
6	A2G	O	103	2	14,14,15	1.71	4 (28%)	17,19,21	1.26	2 (11%)
6	A2G	J	102	2	14,14,15	1.25	1 (7%)	17,19,21	1.32	2 (11%)
3	GOL	F	701	-	5,5,5	0.16	0	5,5,5	0.38	0
6	A2G	P	101	2	14,14,15	1.79	5 (35%)	17,19,21	1.92	5 (29%)
6	A2G	N	103	2	14,14,15	1.73	4 (28%)	17,19,21	1.59	3 (17%)
3	GOL	A	1001	-	5,5,5	0.14	0	5,5,5	0.38	0
6	A2G	N	101	2	14,14,15	2.15	4 (28%)	17,19,21	1.32	2 (11%)
3	GOL	H	702	-	5,5,5	0.12	0	5,5,5	0.24	0
6	A2G	M	103	2	14,14,15	1.93	3 (21%)	17,19,21	1.76	2 (11%)
6	A2G	N	102	2	14,14,15	1.42	2 (14%)	17,19,21	1.42	1 (5%)
6	A2G	K	102	2	14,14,15	1.14	1 (7%)	17,19,21	2.04	5 (29%)
3	GOL	D	701	-	5,5,5	0.20	0	5,5,5	0.32	0
6	A2G	M	101	2	14,14,15	1.76	3 (21%)	17,19,21	1.06	0
6	A2G	L	101	2	14,14,15	2.30	3 (21%)	17,19,21	1.24	1 (5%)
3	GOL	E	701	-	5,5,5	0.10	0	5,5,5	0.46	0
6	A2G	L	103	2	14,14,15	2.26	2 (14%)	17,19,21	1.10	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
6	A2G	O	101	2	14,14,15	0.94	1 (7%)	17,19,21	1.25	1 (5%)
6	A2G	P	102	2	14,14,15	1.26	2 (14%)	17,19,21	1.32	3 (17%)
6	A2G	J	103	2	14,14,15	1.15	1 (7%)	17,19,21	1.38	2 (11%)
3	GOL	H	703	-	5,5,5	0.19	0	5,5,5	0.26	0
3	GOL	H	701	-	5,5,5	0.20	0	5,5,5	0.39	0
6	A2G	L	102	2	14,14,15	1.16	2 (14%)	17,19,21	1.51	5 (29%)

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
6	A2G	P	103	2	-	0/6/23/26	0/1/1/1
6	A2G	K	103	2	-	0/6/23/26	0/1/1/1
6	A2G	M	102	2	-	0/6/23/26	0/1/1/1
3	GOL	C	701	-	-	2/4/4/4	-
3	GOL	C	702	-	-	0/4/4/4	-
6	A2G	O	102	2	-	0/6/23/26	0/1/1/1
6	A2G	J	101	2	-	0/6/23/26	0/1/1/1
3	GOL	A	1002	-	-	2/4/4/4	-
6	A2G	I	103	2	-	0/6/23/26	0/1/1/1
6	A2G	I	101	2	-	0/6/23/26	0/1/1/1
3	GOL	B	1001	-	-	0/4/4/4	-
6	A2G	K	101	2	-	0/6/23/26	0/1/1/1
3	GOL	G	701	-	-	2/4/4/4	-
6	A2G	I	102	2	-	0/6/23/26	0/1/1/1
5	MES	F	702	-	-	4/6/14/14	0/1/1/1
6	A2G	O	103	2	-	0/6/23/26	0/1/1/1
6	A2G	J	102	2	-	0/6/23/26	0/1/1/1
3	GOL	F	701	-	-	2/4/4/4	-
6	A2G	P	101	2	-	0/6/23/26	0/1/1/1
6	A2G	N	103	2	-	0/6/23/26	0/1/1/1
3	GOL	A	1001	-	-	0/4/4/4	-
6	A2G	N	101	2	-	0/6/23/26	0/1/1/1
3	GOL	H	702	-	-	0/4/4/4	-
6	A2G	M	103	2	-	0/6/23/26	0/1/1/1
6	A2G	N	102	2	-	0/6/23/26	0/1/1/1
6	A2G	K	102	2	-	0/6/23/26	0/1/1/1
3	GOL	D	701	-	-	4/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	A2G	M	101	2	-	0/6/23/26	0/1/1/1
6	A2G	L	101	2	-	0/6/23/26	0/1/1/1
3	GOL	E	701	-	-	4/4/4/4	-
6	A2G	L	103	2	-	0/6/23/26	0/1/1/1
6	A2G	O	101	2	-	0/6/23/26	0/1/1/1
6	A2G	P	102	2	-	0/6/23/26	0/1/1/1
6	A2G	J	103	2	-	0/6/23/26	0/1/1/1
3	GOL	H	703	-	-	4/4/4/4	-
3	GOL	H	701	-	-	2/4/4/4	-
6	A2G	L	102	2	-	0/6/23/26	0/1/1/1

The worst 5 of 68 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	L	103	A2G	O5-C5	6.64	1.56	1.43
6	L	101	A2G	C1-C2	5.99	1.61	1.52
6	M	103	A2G	C1-C2	5.77	1.61	1.52
6	K	103	A2G	C1-C2	5.71	1.60	1.52
6	J	101	A2G	O5-C5	5.31	1.54	1.43

The worst 5 of 66 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	P	101	A2G	C1-O5-C5	4.74	118.61	112.19
6	M	103	A2G	O5-C5-C6	-4.57	100.04	107.20
6	K	102	A2G	C3-C4-C5	-4.11	102.90	110.24
6	N	102	A2G	O5-C5-C6	4.10	113.63	107.20
6	K	102	A2G	O5-C1-C2	-3.87	105.17	111.29

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	701	GOL	O1-C1-C2-C3
3	F	701	GOL	C1-C2-C3-O3
3	G	701	GOL	C1-C2-C3-O3
3	H	701	GOL	O1-C1-C2-C3
3	H	703	GOL	C1-C2-C3-O3

There are no ring outliers.

14 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	M	102	A2G	1	0
6	O	102	A2G	3	0
3	A	1002	GOL	1	0
6	I	103	A2G	1	0
3	B	1001	GOL	2	0
3	G	701	GOL	5	0
6	I	102	A2G	1	0
5	F	702	MES	4	0
6	N	102	A2G	1	0
3	E	701	GOL	1	0
6	O	101	A2G	1	0
6	P	102	A2G	1	0
3	H	703	GOL	3	0
6	L	102	A2G	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^{th} 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^2)	Q<0.9
1	A	342/627 (54%)	0.12	8 (2%) 60 65	19, 27, 44, 77	0
1	B	343/627 (54%)	0.25	11 (3%) 47 52	17, 27, 54, 96	0
1	C	342/627 (54%)	0.35	24 (7%) 16 19	18, 29, 53, 91	0
1	D	342/627 (54%)	0.39	23 (6%) 17 20	22, 34, 56, 85	0
1	E	344/627 (54%)	0.50	28 (8%) 12 14	21, 35, 64, 91	0
1	F	344/627 (54%)	0.34	21 (6%) 21 23	20, 31, 54, 92	0
1	G	341/627 (54%)	0.53	31 (9%) 9 10	22, 36, 66, 100	0
1	H	342/627 (54%)	0.26	15 (4%) 34 38	21, 31, 53, 83	0
2	I	11/15 (73%)	0.29	2 (18%) 1 1	24, 27, 40, 53	0
2	J	11/15 (73%)	-0.14	1 (9%) 9 10	19, 23, 35, 50	0
2	K	11/15 (73%)	-0.08	1 (9%) 9 10	22, 23, 34, 46	0
2	L	11/15 (73%)	0.27	1 (9%) 9 10	25, 29, 40, 45	0
2	M	11/15 (73%)	0.38	1 (9%) 9 10	30, 34, 54, 63	0
2	N	10/15 (66%)	0.00	1 (10%) 7 8	26, 29, 33, 44	0
2	O	11/15 (73%)	0.15	1 (9%) 9 10	28, 32, 47, 60	0
2	P	11/15 (73%)	0.52	1 (9%) 9 10	26, 31, 40, 54	0
All	All	2827/5136 (55%)	0.34	170 (6%) 21 24	17, 31, 57, 100	0

The worst 5 of 170 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	296	ASN	7.1
1	C	275	ASN	5.8
1	E	109	SER	5.7
2	P	4	ALA	5.6
1	E	351	THR	5.5

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^{th} 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^2)	Q<0.9
3	GOL	H	701	6/6	0.82	0.11	33,43,44,44	0
3	GOL	F	701	6/6	0.85	0.19	40,49,54,58	0
3	GOL	E	701	6/6	0.86	0.15	38,46,50,55	0
3	GOL	C	702	6/6	0.87	0.12	27,32,36,37	0
3	GOL	H	702	6/6	0.87	0.12	35,38,47,54	0
3	GOL	H	703	6/6	0.87	0.24	27,34,59,60	0
3	GOL	B	1001	6/6	0.88	0.11	30,36,43,47	0
3	GOL	A	1002	6/6	0.89	0.14	38,45,47,55	0
6	A2G	P	102	14/15	0.89	0.11	28,34,37,41	0
3	GOL	A	1001	6/6	0.90	0.18	35,38,43,44	0
6	A2G	M	102	14/15	0.90	0.11	31,37,40,40	0
3	GOL	D	701	6/6	0.90	0.12	34,45,49,50	0
6	A2G	N	102	14/15	0.91	0.08	29,33,35,36	0
6	A2G	O	101	14/15	0.91	0.10	30,35,38,40	0
6	A2G	O	102	14/15	0.91	0.09	28,32,37,37	0
3	GOL	G	701	6/6	0.91	0.12	40,48,50,50	0
6	A2G	I	102	14/15	0.92	0.08	26,30,33,33	0
6	A2G	L	101	14/15	0.93	0.08	26,28,32,34	0
5	MES	F	702	12/12	0.93	0.19	38,55,72,73	0
6	A2G	P	103	14/15	0.93	0.08	26,28,32,32	0
6	A2G	J	101	14/15	0.94	0.09	19,21,23,23	0
6	A2G	L	102	14/15	0.94	0.09	27,33,35,36	0
6	A2G	K	102	14/15	0.94	0.07	24,28,31,32	0
6	A2G	O	103	14/15	0.94	0.09	27,29,30,32	0
6	A2G	M	103	14/15	0.94	0.08	32,33,37,39	0
6	A2G	N	101	14/15	0.94	0.08	22,27,30,30	0
6	A2G	N	103	14/15	0.95	0.07	25,29,32,33	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	A2G	K	103	14/15	0.95	0.08	20,22,24,25	0
6	A2G	M	101	14/15	0.95	0.08	27,29,31,31	0
6	A2G	I	103	14/15	0.96	0.06	21,24,25,26	0
3	GOL	C	701	6/6	0.96	0.09	35,41,43,43	0
6	A2G	J	102	14/15	0.96	0.06	21,25,27,28	0
6	A2G	L	103	14/15	0.97	0.07	22,24,26,26	0
6	A2G	J	103	14/15	0.97	0.10	23,23,26,28	0
6	A2G	P	101	14/15	0.97	0.06	23,25,26,27	0
6	A2G	K	101	14/15	0.97	0.09	20,22,25,25	0
4	ZN	F	703	1/1	0.97	0.04	31,31,31,31	0
4	ZN	F	704	1/1	0.98	0.04	26,26,26,26	0
6	A2G	I	101	14/15	0.98	0.05	20,23,26,27	0
4	ZN	G	702	1/1	0.98	0.03	29,29,29,29	0
4	ZN	G	703	1/1	0.98	0.02	31,31,31,31	0
4	ZN	H	704	1/1	0.98	0.03	29,29,29,29	0
4	ZN	B	1002	1/1	0.99	0.04	25,25,25,25	0
4	ZN	E	702	1/1	0.99	0.03	30,30,30,30	0
4	ZN	A	1003	1/1	1.00	0.03	27,27,27,27	0

6.5 Other polymers (i)

There are no such residues in this entry.