



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 03:48 PM BST

PDB ID : 6P9U  
Title : Crystal structure of human thrombin mutant W215A  
Authors : Pelc, L.A.; Koester, S.K.; Chen, Z.; Di Cera, E.  
Deposited on : 2019-06-10  
Resolution : 3.30 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

<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

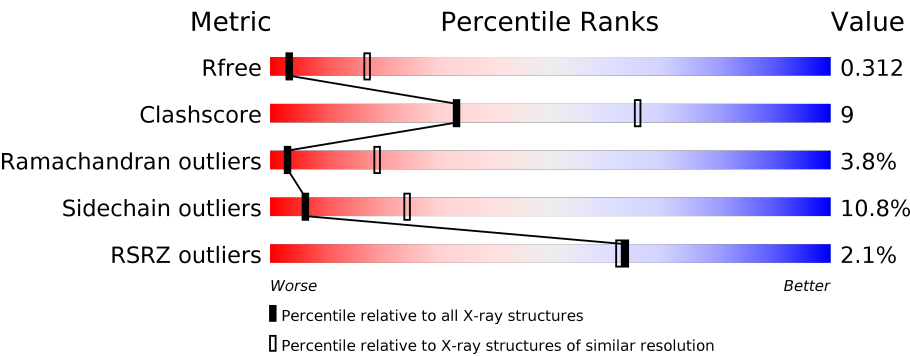
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	A	31	<div><div></div><div><div>77%</div><div>13%</div><div>.</div><div>.</div><div>.</div></div></div>
1	C	31	<div><div></div><div><div>71%</div><div>23%</div><div>6%</div></div></div>
1	E	31	<div><div>3%</div><div></div><div><div>71%</div><div>26%</div><div>.</div></div></div>
1	G	31	<div><div>3%</div><div></div><div><div>97%</div><div>.</div></div></div>
2	B	273	<div><div>4%</div><div></div><div><div>63%</div><div>23%</div><div>.</div><div>10%</div></div></div>
2	D	273	<div><div>2%</div><div></div><div><div>64%</div><div>21%</div><div>.</div><div>11%</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	273	<div><div>%</div><div><div></div><div>63%</div><div>25%</div><div>•</div><div>10%</div></div></div>
2	H	273	<div><div>%</div><div><div></div><div>63%</div><div>24%</div><div>•</div><div>9%</div></div></div>

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 9011 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 Prothrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	30	Total	C	N	O	S	0	0	0
			243	151	39	52	1			
1	C	31	Total	C	N	O	S	0	0	0
			254	157	43	53	1			
1	E	31	Total	C	N	O	S	0	0	0
			254	157	43	53	1			
1	G	31	Total	C	N	O	S	0	0	0
			254	157	43	53	1			

- Molecule 2 is a protein called Prothrombin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	245	Total	C	N	O	S	0	0	0
			1973	1260	348	351	14			
2	D	244	Total	C	N	O	S	0	0	0
			1968	1259	343	352	14			
2	F	247	Total	C	N	O	S	0	0	0
			1998	1278	350	356	14			
2	H	248	Total	C	N	O	S	0	0	0
			1997	1277	349	357	14			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	215	ALA	TRP	engineered mutation	UNP P00734
B	248	TYR	-	expression tag	UNP P00734
B	249	LEU	-	expression tag	UNP P00734
B	250	GLU	-	expression tag	UNP P00734
B	251	ASP	-	expression tag	UNP P00734
B	252	GLN	-	expression tag	UNP P00734
B	253	VAL	-	expression tag	UNP P00734
B	254	ASP	-	expression tag	UNP P00734

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	255	PRO	-	expression tag	UNP P00734
B	256	ARG	-	expression tag	UNP P00734
B	257	LEU	-	expression tag	UNP P00734
B	258	ILE	-	expression tag	UNP P00734
B	259	ASP	-	expression tag	UNP P00734
B	260	GLY	-	expression tag	UNP P00734
B	261	LYS	-	expression tag	UNP P00734
D	215	ALA	TRP	engineered mutation	UNP P00734
D	248	TYR	-	expression tag	UNP P00734
D	249	LEU	-	expression tag	UNP P00734
D	250	GLU	-	expression tag	UNP P00734
D	251	ASP	-	expression tag	UNP P00734
D	252	GLN	-	expression tag	UNP P00734
D	253	VAL	-	expression tag	UNP P00734
D	254	ASP	-	expression tag	UNP P00734
D	255	PRO	-	expression tag	UNP P00734
D	256	ARG	-	expression tag	UNP P00734
D	257	LEU	-	expression tag	UNP P00734
D	258	ILE	-	expression tag	UNP P00734
D	259	ASP	-	expression tag	UNP P00734
D	260	GLY	-	expression tag	UNP P00734
D	261	LYS	-	expression tag	UNP P00734
F	215	ALA	TRP	engineered mutation	UNP P00734
F	248	TYR	-	expression tag	UNP P00734
F	249	LEU	-	expression tag	UNP P00734
F	250	GLU	-	expression tag	UNP P00734
F	251	ASP	-	expression tag	UNP P00734
F	252	GLN	-	expression tag	UNP P00734
F	253	VAL	-	expression tag	UNP P00734
F	254	ASP	-	expression tag	UNP P00734
F	255	PRO	-	expression tag	UNP P00734
F	256	ARG	-	expression tag	UNP P00734
F	257	LEU	-	expression tag	UNP P00734
F	258	ILE	-	expression tag	UNP P00734
F	259	ASP	-	expression tag	UNP P00734
F	260	GLY	-	expression tag	UNP P00734
F	261	LYS	-	expression tag	UNP P00734
H	215	ALA	TRP	engineered mutation	UNP P00734
H	248	TYR	-	expression tag	UNP P00734
H	249	LEU	-	expression tag	UNP P00734
H	250	GLU	-	expression tag	UNP P00734
H	251	ASP	-	expression tag	UNP P00734

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	252	GLN	-	expression tag	UNP P00734
H	253	VAL	-	expression tag	UNP P00734
H	254	ASP	-	expression tag	UNP P00734
H	255	PRO	-	expression tag	UNP P00734
H	256	ARG	-	expression tag	UNP P00734
H	257	LEU	-	expression tag	UNP P00734
H	258	ILE	-	expression tag	UNP P00734
H	259	ASP	-	expression tag	UNP P00734
H	260	GLY	-	expression tag	UNP P00734
H	261	LYS	-	expression tag	UNP P00734

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	1	Total 1	Zn 1	0	0
3	D	2	Total 2	Zn 2	0	0
3	E	1	Total 1	Zn 1	0	0
3	B	1	Total 1	Zn 1	0	0
3	C	1	Total 1	Zn 1	0	0
3	F	1	Total 1	Zn 1	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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


- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	O	0	0
			1	1		
5	D	1	Total	O	0	0
			1	1		
5	E	1	Total	O	0	0
			1	1		
5	F	2	Total	O	0	0
			2	2		
5	H	2	Total	O	0	0
			2	2		

### 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: Prothrombin

Chain A: 



- Molecule 1: Prothrombin

Chain C: 



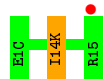
- Molecule 1: Prothrombin

Chain E: 



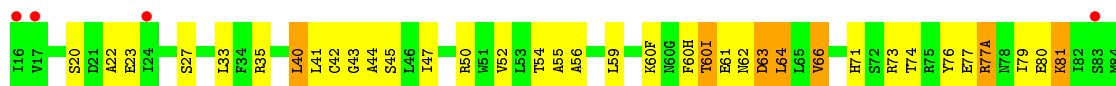
- Molecule 1: Prothrombin

Chain G: 

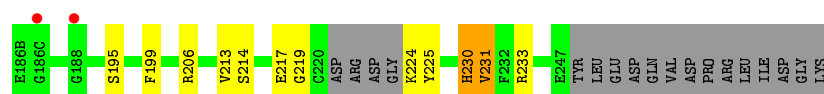


- Molecule 2: Prothrombin

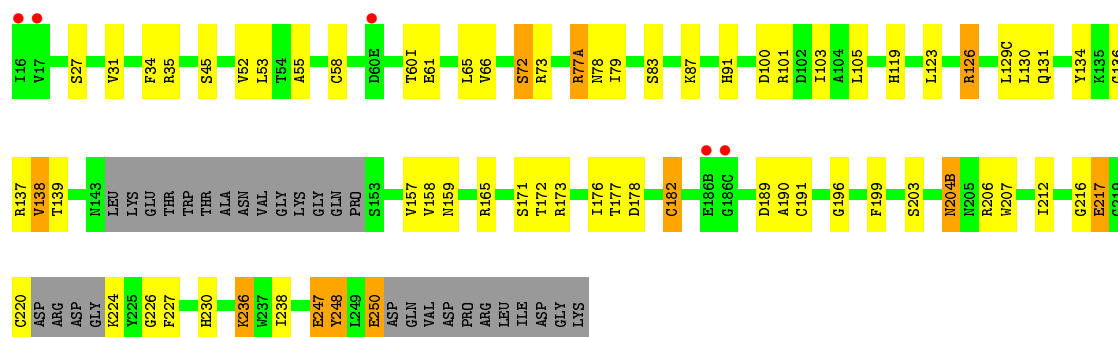
Chain B: 



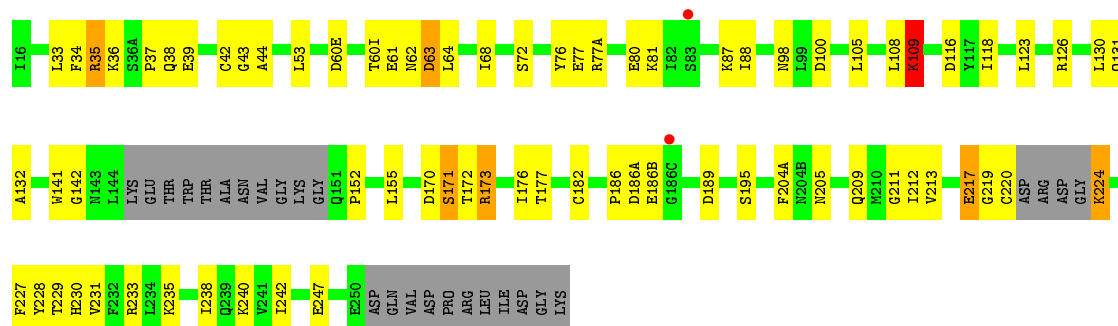




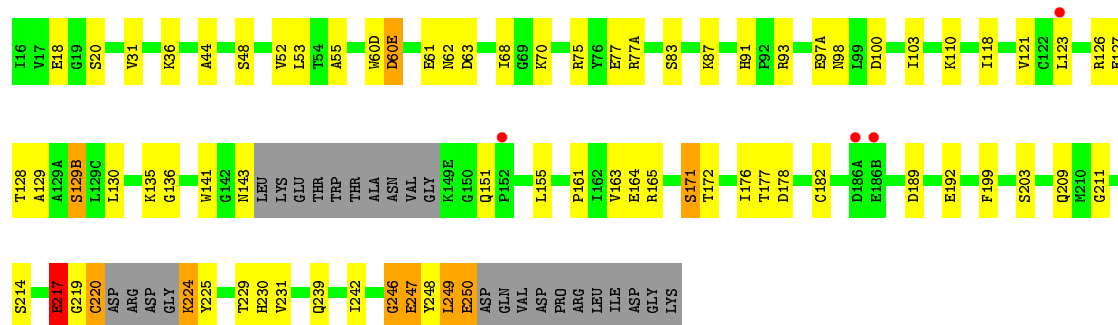
• Molecule 2: Prothrombin



• Molecule 2: Prothrombin



• Molecule 2: Prothrombin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.32Å 44.23Å 136.19Å 90.00° 90.04° 90.00°	Depositor
Resolution (Å)	32.59 – 3.30 32.58 – 3.30	Depositor EDS
% Data completeness (in resolution range)	89.5 (32.59-3.30) 89.6 (32.58-3.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.17	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.12 (at 3.32Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.225 , 0.308 0.228 , 0.312	Depositor DCC
$R_{free}$ test set	1087 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.0	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 11.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.369 for -l,k,h 0.052 for -h,-k,l 0.054 for l,-k,h	Xtriage
Reported twinning fraction	0.534 for H, K, L 0.466 for L, K, -H	Depositor
Outliers	13 of 22632 reflections (0.057%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	9011	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 57.60 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.3112e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: ZN, NAG

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	0.63	0/245	0.89	0/326
1	C	0.71	0/256	1.04	0/340
1	E	0.91	0/256	0.99	0/340
1	G	0.86	0/256	0.97	0/340
2	B	0.77	0/2022	0.92	0/2730
2	D	0.68	0/2017	0.88	0/2724
2	F	0.75	0/2048	0.89	0/2766
2	H	0.73	0/2047	0.90	0/2764
All	All	0.74	0/9147	0.91	0/12330

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	243	0	238	11	0
1	C	254	0	251	5	0
1	E	254	0	251	5	0
1	G	254	0	251	2	0
2	B	1973	0	1953	35	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1968	0	1939	31	0
2	F	1998	0	1976	35	0
2	H	1997	0	1970	39	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	2	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
4	B	14	0	13	0	0
4	D	14	0	13	0	0
4	F	14	0	13	0	0
4	H	14	0	13	0	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
5	F	2	0	0	0	0
5	H	2	0	0	0	0
All	All	9011	0	8881	153	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 9.

All (153) 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)
2:B:42:CYS:HB3	2:B:195:SER:O	1.79	0.81
2:F:100:ASP:OD1	2:F:177:THR:HG21	1.82	0.80
2:H:250:GLU:OE1	2:H:250:GLU:N	2.15	0.80
2:B:33:LEU:O	2:B:41:LEU:HB2	1.86	0.76
1:C:14(D):ARG:NH2	2:H:97(A):GLU:O	2.28	0.67
2:B:61:GLU:HG2	2:B:87:LYS:HA	1.75	0.66
2:F:63:ASP:N	2:F:63:ASP:OD1	2.27	0.66
1:A:14(K):ILE:H	1:A:14(K):ILE:HD13	1.61	0.65
2:F:42:CYS:HB3	2:F:195:SER:O	1.98	0.62
1:A:14(K):ILE:HB	2:F:173:ARG:HB3	1.83	0.61
2:D:100:ASP:OD1	2:D:177:THR:HG21	2.01	0.60
1:E:1(A):ASP:OD1	1:E:9:LYS:NZ	2.27	0.59
2:D:204(B):ASN:OD1	2:D:206:ARG:HD2	2.02	0.59
2:F:230:HIS:CD2	2:F:233:ARG:HG3	2.38	0.59
2:F:76:TYR:CD2	2:F:80:GLU:OE2	2.56	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:126:ARG:O	2:H:127:GLU:C	2.41	0.59
1:E:14(K):ILE:HG22	1:E:14(K):ILE:O	2.03	0.58
2:D:173:ARG:CB	1:E:14(J):TYR:O	2.52	0.58
2:F:77:GLU:OE1	2:F:80:GLU:OE2	2.22	0.58
2:H:53:LEU:HD23	2:H:209:GLN:HE22	1.68	0.57
2:B:43:GLY:O	2:B:44:ALA:HB2	2.04	0.57
2:B:115:SER:O	2:B:118:ILE:O	2.23	0.57
1:A:14(K):ILE:N	1:A:14(K):ILE:HD13	2.19	0.57
2:H:247:GLU:OE1	2:H:249:LEU:HD21	2.05	0.56
2:B:60(I):THR:HG22	2:B:63:ASP:OD1	2.06	0.56
2:B:86:GLU:HB2	2:B:109:LYS:HA	1.88	0.56
2:D:131:GLN:O	2:D:134:TYR:HB2	2.05	0.56
2:H:127:GLU:O	2:H:129(B):SER:HB2	2.06	0.56
1:A:1:CYS:O	2:B:206:ARG:NH1	2.39	0.56
2:D:182:CYS:HB2	2:D:227:PHE:HD2	1.72	0.55
1:G:14(K):ILE:HG22	1:G:14(K):ILE:O	2.07	0.55
2:B:59:LEU:HB2	2:B:90:ILE:HD11	1.89	0.54
1:C:1:CYS:O	2:D:206:ARG:NH1	2.40	0.54
2:H:68:ILE:CG2	2:H:118:ILE:HG23	2.38	0.54
2:D:136:GLY:HA3	2:D:199:PHE:CZ	2.43	0.54
2:F:36:LYS:CD	2:F:62:ASN:O	2.56	0.53
2:H:61:GLU:HG2	2:H:87:LYS:HA	1.90	0.53
2:H:163:VAL:HG21	2:H:225:TYR:CD1	2.42	0.53
1:G:14(K):ILE:CG2	1:G:14(K):ILE:O	2.56	0.53
2:H:224:LYS:HD2	2:H:224:LYS:N	2.23	0.52
1:A:14(J):TYR:O	2:F:173:ARG:HD2	2.10	0.52
2:H:247:GLU:O	2:H:248:TYR:HB3	2.09	0.52
2:F:211:GLY:HA2	2:F:231:VAL:HG23	1.91	0.52
2:B:33:LEU:HA	2:B:66:VAL:HG12	1.92	0.52
2:F:108:LEU:O	2:F:109:LYS:C	2.48	0.52
2:F:35:ARG:HD3	2:F:37:PRO:O	2.09	0.51
2:D:65:LEU:HD12	2:D:66:VAL:N	2.26	0.51
2:F:68:ILE:HG22	2:F:118:ILE:HG23	1.93	0.51
2:F:43:GLY:O	2:F:44:ALA:HB2	2.11	0.51
1:E:5:PRO:HB2	2:F:116:ASP:HA	1.92	0.51
2:H:177:THR:OG1	2:H:178:ASP:N	2.44	0.51
1:A:14(J):TYR:N	1:A:14(J):TYR:HD2	2.09	0.50
1:A:14(J):TYR:N	1:A:14(J):TYR:CD2	2.79	0.50
2:D:177:THR:OG1	2:D:178:ASP:N	2.43	0.50
2:H:171:SER:OG	2:H:172:THR:N	2.45	0.50
2:B:85:LEU:HD23	2:B:108:LEU:HD23	1.93	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:216:GLY:O	2:D:217:GLU:HB2	2.12	0.50
2:H:129:ALA:O	2:H:130:LEU:HB2	2.12	0.49
2:H:163:VAL:HG21	2:H:225:TYR:CE1	2.47	0.49
2:B:89:TYR:N	2:B:105:LEU:O	2.45	0.49
2:B:136:GLY:HA3	2:B:199:PHE:CE1	2.47	0.49
2:D:136:GLY:HA3	2:D:199:PHE:CE1	2.47	0.49
2:F:142:GLY:O	2:F:152:PRO:HG3	2.13	0.49
2:B:87:LYS:O	2:B:107:LYS:N	2.45	0.49
2:F:172:THR:HG21	2:F:227:PHE:CZ	2.47	0.49
2:B:81:LYS:HB2	2:B:118:ILE:HD11	1.95	0.48
2:D:91:HIS:CE1	2:D:101:ARG:HD3	2.49	0.48
2:H:91:HIS:HB2	2:H:103:ILE:HG23	1.95	0.48
2:B:44:ALA:HB1	2:B:52:VAL:CG1	2.43	0.48
1:A:14(J):TYR:O	2:F:173:ARG:CD	2.62	0.48
2:F:36:LYS:HD3	2:F:62:ASN:O	2.14	0.48
2:F:211:GLY:HA2	2:F:229:THR:O	2.13	0.48
2:D:129(C):LEU:HD11	2:D:203:SER:HA	1.96	0.48
2:H:68:ILE:HG21	2:H:118:ILE:HG23	1.96	0.48
2:B:52:VAL:CG2	2:B:108:LEU:HD21	2.44	0.47
2:H:217:GLU:O	2:H:220:CYS:N	2.48	0.47
2:B:50:ARG:HA	2:B:108:LEU:HD12	1.97	0.47
2:B:63:ASP:OD1	2:B:63:ASP:N	2.46	0.47
2:H:211:GLY:HA2	2:H:231:VAL:HG23	1.97	0.47
2:F:61:GLU:HG2	2:F:87:LYS:HA	1.97	0.46
2:H:100:ASP:OD1	2:H:177:THR:HG21	2.15	0.46
2:D:216:GLY:O	2:D:217:GLU:CB	2.64	0.46
2:F:238:ILE:O	2:F:242:ILE:HG12	2.15	0.46
2:B:40:LEU:HD13	2:B:73:ARG:HD2	1.97	0.46
2:F:68:ILE:CG2	2:F:118:ILE:HG23	2.46	0.46
2:H:60(D):TRP:O	2:H:60(E):ASP:C	2.54	0.46
2:H:70:LYS:NZ	2:H:77:GLU:OE1	2.46	0.46
2:F:217:GLU:O	2:F:220:CYS:N	2.49	0.46
2:F:235:LYS:HA	2:F:238:ILE:HD12	1.97	0.46
2:F:224:LYS:N	2:F:224:LYS:HD3	2.31	0.45
1:E:14(K):ILE:O	1:E:14(M):GLY:N	2.49	0.45
2:F:224:LYS:N	2:F:224:LYS:CD	2.80	0.45
2:H:136:GLY:HA3	2:H:199:PHE:CZ	2.51	0.45
2:H:247:GLU:HB3	2:H:249:LEU:CD2	2.46	0.45
2:H:130:LEU:HD13	2:H:230:HIS:CE1	2.52	0.45
2:B:77:GLU:N	2:B:80:GLU:OE2	2.49	0.45
1:A:14(K):ILE:N	1:A:14(K):ILE:CD1	2.77	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:47:ILE:HG21	2:B:123:LEU:HD11	1.98	0.45
2:B:77:GLU:HB3	2:B:79:ILE:HB	1.98	0.44
2:B:85:LEU:CD2	2:B:108:LEU:HD23	2.47	0.44
2:D:53:LEU:HD11	2:D:103:ILE:HG13	2.00	0.44
2:D:189:ASP:OD1	2:D:226:GLY:N	2.49	0.44
2:B:165:ARG:CB	2:B:166:PRO:CD	2.96	0.44
1:C:8:GLU:N	1:C:8:GLU:OE1	2.51	0.44
2:H:135:LYS:HA	2:H:161:PRO:HA	1.98	0.44
2:H:143:ASN:ND2	2:H:192:GLU:O	2.51	0.44
2:H:128:THR:HA	2:H:129(B):SER:HB2	2.00	0.44
1:A:5:PRO:HB2	2:B:116:ASP:HA	2.00	0.44
2:D:55:ALA:O	2:D:58:CYS:HB2	2.18	0.44
2:D:31:VAL:HG11	2:D:52:VAL:HG11	2.00	0.44
2:D:130:LEU:CD1	2:D:230:HIS:CE1	3.00	0.44
2:D:250:GLU:OE1	2:D:250:GLU:N	2.51	0.43
2:B:230:HIS:O	2:B:231:VAL:C	2.56	0.43
2:F:53:LEU:HD13	2:F:105:LEU:HD21	2.00	0.43
2:H:249:LEU:HG	2:H:249:LEU:H	1.74	0.43
2:H:31:VAL:HB	2:H:44:ALA:HB3	2.00	0.43
2:B:163:VAL:HG21	2:B:225:TYR:CD1	2.54	0.43
2:D:105:LEU:HD11	2:D:238:ILE:HG23	2.00	0.43
1:C:8:GLU:OE2	2:D:207:TRP:NE1	2.50	0.43
2:D:61:GLU:HG2	2:D:87:LYS:HA	2.00	0.43
2:B:56:ALA:O	2:B:59:LEU:N	2.49	0.43
2:F:60(I):THR:HA	2:F:88:ILE:CD1	2.49	0.43
2:H:68:ILE:HG22	2:H:118:ILE:HG23	2.00	0.42
2:D:77(A):ARG:O	2:D:79:ILE:N	2.50	0.42
2:F:141:TRP:CZ2	2:F:155:LEU:HB2	2.54	0.42
2:F:171:SER:O	2:F:173:ARG:NH2	2.50	0.42
2:F:33:LEU:HD11	2:F:64:LEU:HD13	2.01	0.42
2:B:60(H):PHE:CD1	2:B:64:LEU:HD21	2.54	0.42
2:D:137:ARG:O	2:D:138:VAL:HB	2.19	0.42
2:B:44:ALA:HB1	2:B:52:VAL:HG12	2.01	0.42
2:H:31:VAL:HG21	2:H:52:VAL:HG13	2.00	0.42
2:H:163:VAL:HG12	2:H:164:GLU:N	2.35	0.42
2:D:247:GLU:O	2:D:248:TYR:HB2	2.19	0.41
2:H:239:GLN:HA	2:H:242:ILE:HB	2.02	0.41
2:B:22:ALA:O	2:B:71:HIS:CE1	2.74	0.41
2:D:196:GLY:HA2	2:D:212:ILE:HG22	2.01	0.41
2:H:36:LYS:CD	2:H:62:ASN:O	2.67	0.41
2:B:54:THR:OG1	2:B:55:ALA:N	2.53	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:14(K):ILE:O	1:A:14(K):ILE:HG22	2.21	0.41
2:B:76:TYR:CE1	2:B:77(A):ARG:HA	2.55	0.41
2:D:190:ALA:O	2:D:216:GLY:HA3	2.21	0.41
2:B:60(F):LYS:HG2	2:B:60(H):PHE:CE2	2.56	0.41
2:D:72:SER:OG	2:D:73:ARG:N	2.52	0.41
2:F:76:TYR:HD2	2:F:80:GLU:OE2	2.01	0.41
2:H:141:TRP:CZ2	2:H:155:LEU:HD13	2.55	0.41
2:H:36:LYS:HD3	2:H:63:ASP:C	2.40	0.41
2:F:209:GLN:NE2	2:F:212:ILE:HG12	2.36	0.41
2:D:158:VAL:HG22	2:D:159:ASN:N	2.36	0.41
2:H:246:GLY:O	2:H:248:TYR:N	2.54	0.41
2:H:98:ASN:OD1	2:H:98:ASN:N	2.54	0.41
2:F:213:VAL:HG22	2:F:228:TYR:HE2	1.85	0.40
2:D:91:HIS:HB2	2:D:103:ILE:HG23	2.03	0.40
1:C:5:PRO:HA	1:C:9:LYS:HB2	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	28/31 (90%)	22 (79%)	4 (14%)	2 (7%)	1	7
1	C	29/31 (94%)	21 (72%)	6 (21%)	2 (7%)	1	8
1	E	29/31 (94%)	24 (83%)	3 (10%)	2 (7%)	1	8
1	G	29/31 (94%)	25 (86%)	4 (14%)	0	100	100
2	B	239/273 (88%)	196 (82%)	35 (15%)	8 (3%)	4	22
2	D	238/273 (87%)	202 (85%)	29 (12%)	7 (3%)	4	24
2	F	241/273 (88%)	197 (82%)	33 (14%)	11 (5%)	2	15
2	H	242/273 (89%)	205 (85%)	28 (12%)	9 (4%)	3	20

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1075/1216 (88%)	892 (83%)	142 (13%)	41 (4%)	3	19

All (41) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	217	GLU
1	E	14(L)	ASP
2	F	247	GLU
2	H	247	GLU
1	A	14(K)	ILE
1	A	14(L)	ASP
2	B	109	LYS
2	B	142	GLY
2	B	219	GLY
1	C	14(M)	GLY
2	D	78	ASN
2	D	126	ARG
2	D	248	TYR
2	F	39	GLU
2	F	171	SER
2	F	219	GLY
2	H	60(E)	ASP
2	H	217	GLU
2	H	219	GLY
2	B	126	ARG
2	B	217	GLU
2	D	138	VAL
2	F	132	ALA
2	D	236	LYS
2	F	126	ARG
2	F	130	LEU
2	F	204(A)	PHE
2	F	217	GLU
2	H	18	GLU
2	H	171	SER
2	F	109	LYS
2	H	55	ALA
2	H	93	ARG
2	D	247	GLU
1	C	14(K)	ILE
1	E	14(M)	GLY
2	B	213	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	231	VAL
2	B	161	PRO
2	H	246	GLY
2	F	186	PRO

### 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	27/28 (96%)	24 (89%)	3 (11%)	6	23
1	C	28/28 (100%)	24 (86%)	4 (14%)	3	15
1	E	28/28 (100%)	26 (93%)	2 (7%)	14	42
1	G	28/28 (100%)	27 (96%)	1 (4%)	35	63
2	B	213/237 (90%)	187 (88%)	26 (12%)	5	20
2	D	212/237 (90%)	188 (89%)	24 (11%)	6	22
2	F	216/237 (91%)	194 (90%)	22 (10%)	7	27
2	H	215/237 (91%)	193 (90%)	22 (10%)	7	27
All	All	967/1060 (91%)	863 (89%)	104 (11%)	6	24

All (104) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	LYS
1	A	14(J)	TYR
1	A	14(K)	ILE
2	B	20	SER
2	B	23	GLU
2	B	27	SER
2	B	35	ARG
2	B	40	LEU
2	B	45	SER
2	B	60(I)	THR
2	B	62	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	63	ASP
2	B	64	LEU
2	B	66	VAL
2	B	74	THR
2	B	77(A)	ARG
2	B	81	LYS
2	B	86	GLU
2	B	119	HIS
2	B	126	ARG
2	B	139	THR
2	B	141	TRP
2	B	176	ILE
2	B	182	CYS
2	B	186(A)	ASP
2	B	214	SER
2	B	224	LYS
2	B	230	HIS
2	B	233	ARG
1	C	14(C)	GLU
1	C	14(D)	ARG
1	C	14(J)	TYR
1	C	14(K)	ILE
2	D	27	SER
2	D	34	PHE
2	D	35	ARG
2	D	45	SER
2	D	60(I)	THR
2	D	72	SER
2	D	77(A)	ARG
2	D	83	SER
2	D	119	HIS
2	D	123	LEU
2	D	126	ARG
2	D	139	THR
2	D	157	VAL
2	D	165	ARG
2	D	171	SER
2	D	172	THR
2	D	176	ILE
2	D	182	CYS
2	D	191	CYS
2	D	204(B)	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	D	220	CYS
2	D	224	LYS
2	D	236	LYS
2	D	250	GLU
1	E	10	LYS
1	E	14(I)	SER
2	F	34	PHE
2	F	35	ARG
2	F	38	GLN
2	F	60(E)	ASP
2	F	63	ASP
2	F	72	SER
2	F	77(A)	ARG
2	F	81	LYS
2	F	98	ASN
2	F	109	LYS
2	F	123	LEU
2	F	131	GLN
2	F	170	ASP
2	F	173	ARG
2	F	176	ILE
2	F	182	CYS
2	F	186(A)	ASP
2	F	186(B)	GLU
2	F	189	ASP
2	F	205	ASN
2	F	224	LYS
2	F	240	LYS
1	G	14(K)	ILE
2	H	20	SER
2	H	48	SER
2	H	75	ARG
2	H	77(A)	ARG
2	H	83	SER
2	H	110	LYS
2	H	121	VAL
2	H	123	LEU
2	H	129(B)	SER
2	H	151	GLN
2	H	165	ARG
2	H	176	ILE
2	H	182	CYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	H	189	ASP
2	H	203	SER
2	H	214	SER
2	H	217	GLU
2	H	220	CYS
2	H	224	LYS
2	H	229	THR
2	H	249	LEU
2	H	250	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	57	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 7 are monoatomic - leaving 4 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
4	NAG	F	302	2	14,14,15	0.72	0	17,19,21	1.78	4 (23%)
4	NAG	H	701	2	14,14,15	0.85	0	17,19,21	1.92	5 (29%)
4	NAG	D	403	2	14,14,15	0.81	0	17,19,21	1.50	3 (17%)
4	NAG	B	402	2	14,14,15	0.89	0	17,19,21	1.53	4 (23%)

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
4	NAG	F	302	2	-	0/6/23/26	0/1/1/1
4	NAG	H	701	2	-	0/6/23/26	0/1/1/1
4	NAG	D	403	2	-	0/6/23/26	0/1/1/1
4	NAG	B	402	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	701	NAG	O5-C1-C2	-4.84	103.64	111.29
4	F	302	NAG	O5-C1-C2	-3.92	105.09	111.29
4	F	302	NAG	C1-C2-N2	3.68	116.78	110.49
4	H	701	NAG	C1-C2-N2	3.23	116.01	110.49
4	F	302	NAG	C3-C4-C5	3.07	115.71	110.24
4	H	701	NAG	C3-C4-C5	3.01	115.61	110.24
4	D	403	NAG	C4-C3-C2	2.84	115.18	111.02
4	D	403	NAG	C2-N2-C7	2.74	126.81	122.90
4	D	403	NAG	O5-C5-C6	2.64	111.34	107.20
4	B	402	NAG	O3-C3-C2	2.60	114.85	109.47
4	H	701	NAG	O5-C5-C4	2.58	117.10	110.83
4	H	701	NAG	C1-O5-C5	2.33	115.34	112.19
4	B	402	NAG	O4-C4-C3	-2.26	105.12	110.35
4	B	402	NAG	O4-C4-C5	2.12	114.56	109.30
4	F	302	NAG	C4-C3-C2	-2.06	107.99	111.02
4	B	402	NAG	C1-O5-C5	2.02	114.93	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	402	NAG	O5-C5-C6-O6
4	B	402	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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	A	30/31 (96%)	-0.22	0 100 100	52, 67, 82, 87	0
1	C	31/31 (100%)	-0.16	0 100 100	55, 67, 81, 90	0
1	E	31/31 (100%)	0.10	1 (3%) 47 46	55, 70, 84, 87	0
1	G	31/31 (100%)	0.16	1 (3%) 47 46	62, 70, 83, 89	0
2	B	245/273 (89%)	0.14	10 (4%) 37 35	45, 68, 90, 109	0
2	D	244/273 (89%)	0.03	5 (2%) 65 64	45, 66, 87, 107	0
2	F	247/273 (90%)	0.07	2 (0%) 86 86	49, 66, 88, 103	0
2	H	248/273 (90%)	0.10	4 (1%) 72 70	49, 67, 88, 111	0
All	All	1107/1216 (91%)	0.07	23 (2%) 63 62	45, 67, 88, 111	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	83	SER	3.9
2	B	152	PRO	3.2
2	B	24	ILE	3.2
2	H	186(B)	GLU	3.2
2	B	186(C)	GLY	3.2
2	F	83	SER	3.0
2	B	17	VAL	2.7
2	B	16	ILE	2.7
2	D	186(B)	GLU	2.7
1	E	14(M)	GLY	2.6
2	H	186(A)	ASP	2.6
2	D	186(C)	GLY	2.3
2	F	186(C)	GLY	2.3
2	B	188	GLY	2.3
2	B	126	ARG	2.2
2	D	17	VAL	2.2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	16	ILE	2.1
2	B	97(A)	GLU	2.1
2	H	152	PRO	2.1
1	G	15	ARG	2.1
2	H	123	LEU	2.1
2	B	98	ASN	2.0
2	D	60(E)	ASP	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	H	701	14/15	0.86	0.27	72,86,92,92	0
4	NAG	B	402	14/15	0.89	0.23	57,72,80,80	0
4	NAG	D	403	14/15	0.90	0.19	64,72,79,88	0
4	NAG	F	302	14/15	0.91	0.19	73,84,86,88	0
3	ZN	D	402	1/1	0.96	0.10	81,81,81,81	1
3	ZN	E	401	1/1	0.96	0.12	70,70,70,70	0
3	ZN	C	101	1/1	0.98	0.07	68,68,68,68	1
3	ZN	F	301	1/1	0.98	0.09	65,65,65,65	1
3	ZN	B	401	1/1	0.98	0.12	62,62,62,62	0
3	ZN	D	401	1/1	0.99	0.12	56,56,56,56	0
3	ZN	G	401	1/1	0.99	0.12	66,66,66,66	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.