



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 18, 2016 – 11:30 AM EDT

PDB ID : 5ETX  
Title : Crystal structure of HCV NS3/4A protease A156T variant in complex with 5172-Linear (MK-5172 linear analogue)  
Authors : Soumana, D.; Yilmaz, N.K.; Ali, A.; Prachanronarong, K.L.; Aydin, C.; Schiffer, C.A.  
Deposited on : 2015-11-18  
Resolution : 2.35 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027257  
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) : rb-20027257

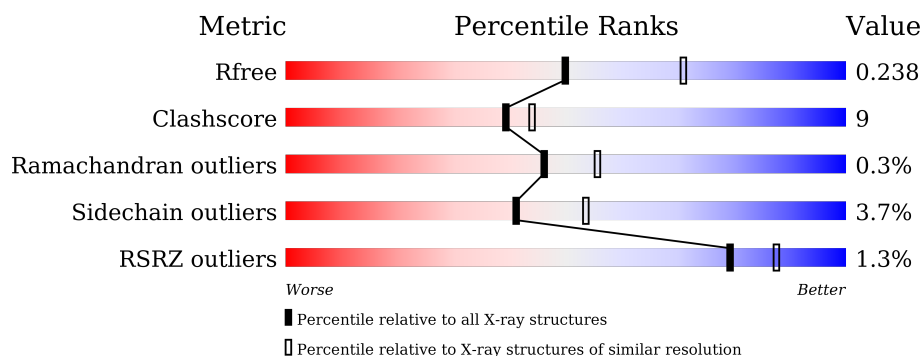
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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))
$R_{free}$	91344	1352 (2.38-2.34)
Clashscore	102246	1456 (2.38-2.34)
Ramachandran outliers	100387	1435 (2.38-2.34)
Sidechain outliers	100360	1436 (2.38-2.34)
RSRZ outliers	91569	1358 (2.38-2.34)

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	194	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 87%, yellow 87%, yellow 96%, grey 96%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>87%</span> <span>9%</span> <span>..</span> </div> </div>
1	B	194	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 2%, green 2%, green 79%, yellow 79%, yellow 98%, grey 98%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>2%</span> <span>79%</span> <span>19%</span> <span>..</span> </div> </div>
1	C	194	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 2%, green 2%, green 84%, yellow 84%, yellow 97%, grey 97%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>2%</span> <span>84%</span> <span>13%</span> <span>..</span> </div> </div>
1	D	194	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, green 1%, green 75%, yellow 75%, yellow 96%, grey 96%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>75%</span> <span>21%</span> <span>..</span> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11129 atoms, of which 5256 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 NS3 protease.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	187	Total	C	H	N	O	S	0	1	0
			2632	834	1289	240	263	6			
1	B	193	Total	C	H	N	O	S	0	2	0
			2688	854	1315	247	268	4			
1	C	193	Total	C	H	N	O	S	0	2	0
			2739	863	1351	247	273	5			
1	D	187	Total	C	H	N	O	S	11	4	0
			2656	844	1301	241	265	5			

There are 116 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	987	LYS	-	expression tag	UNP C1KIK8
A	988	LYS	-	expression tag	UNP C1KIK8
A	989	LYS	-	expression tag	UNP C1KIK8
A	990	GLY	-	expression tag	UNP C1KIK8
A	991	SER	-	expression tag	UNP C1KIK8
A	992	VAL	-	expression tag	UNP C1KIK8
A	993	VAL	-	expression tag	UNP C1KIK8
A	994	ILE	-	expression tag	UNP C1KIK8
A	995	VAL	-	expression tag	UNP C1KIK8
A	996	GLY	-	expression tag	UNP C1KIK8
A	997	ARG	-	expression tag	UNP C1KIK8
A	998	ILE	-	expression tag	UNP C1KIK8
A	999	ASN	-	expression tag	UNP C1KIK8
A	1000	LEU	-	expression tag	UNP C1KIK8
A	1001	SER	-	expression tag	UNP C1KIK8
A	1002	GLY	-	expression tag	UNP C1KIK8
A	1003	ASP	-	expression tag	UNP C1KIK8
A	1013	GLU	LEU	conflict	UNP C1KIK8
A	1014	GLU	LEU	conflict	UNP C1KIK8
A	1017	GLN	ILE	conflict	UNP C1KIK8
A	1018	GLU	ILE	conflict	UNP C1KIK8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1021	GLN	LEU	conflict	UNP C1KIK8
A	1040	THR	ALA	conflict	UNP C1KIK8
A	1047	SER	CYS	conflict	UNP C1KIK8
A	1052	LEU	CYS	conflict	UNP C1KIK8
A	1072	THR	ILE	conflict	UNP C1KIK8
A	1086	GLN	PRO	conflict	UNP C1KIK8
A	1156	THR	ALA	engineered mutation	UNP C1KIK8
A	1159	SER	CYS	engineered mutation	UNP C1KIK8
B	987	LYS	-	expression tag	UNP C1KIK8
B	988	LYS	-	expression tag	UNP C1KIK8
B	989	LYS	-	expression tag	UNP C1KIK8
B	990	GLY	-	expression tag	UNP C1KIK8
B	991	SER	-	expression tag	UNP C1KIK8
B	992	VAL	-	expression tag	UNP C1KIK8
B	993	VAL	-	expression tag	UNP C1KIK8
B	994	ILE	-	expression tag	UNP C1KIK8
B	995	VAL	-	expression tag	UNP C1KIK8
B	996	GLY	-	expression tag	UNP C1KIK8
B	997	ARG	-	expression tag	UNP C1KIK8
B	998	ILE	-	expression tag	UNP C1KIK8
B	999	ASN	-	expression tag	UNP C1KIK8
B	1000	LEU	-	expression tag	UNP C1KIK8
B	1001	SER	-	expression tag	UNP C1KIK8
B	1002	GLY	-	expression tag	UNP C1KIK8
B	1003	ASP	-	expression tag	UNP C1KIK8
B	1013	GLU	LEU	conflict	UNP C1KIK8
B	1014	GLU	LEU	conflict	UNP C1KIK8
B	1017	GLN	ILE	conflict	UNP C1KIK8
B	1018	GLU	ILE	conflict	UNP C1KIK8
B	1021	GLN	LEU	conflict	UNP C1KIK8
B	1040	THR	ALA	conflict	UNP C1KIK8
B	1047	SER	CYS	conflict	UNP C1KIK8
B	1052	LEU	CYS	conflict	UNP C1KIK8
B	1072	THR	ILE	conflict	UNP C1KIK8
B	1086	GLN	PRO	conflict	UNP C1KIK8
B	1156	THR	ALA	engineered mutation	UNP C1KIK8
B	1159	SER	CYS	engineered mutation	UNP C1KIK8
C	987	LYS	-	expression tag	UNP C1KIK8
C	988	LYS	-	expression tag	UNP C1KIK8
C	989	LYS	-	expression tag	UNP C1KIK8
C	990	GLY	-	expression tag	UNP C1KIK8
C	991	SER	-	expression tag	UNP C1KIK8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	992	VAL	-	expression tag	UNP C1KIK8
C	993	VAL	-	expression tag	UNP C1KIK8
C	994	ILE	-	expression tag	UNP C1KIK8
C	995	VAL	-	expression tag	UNP C1KIK8
C	996	GLY	-	expression tag	UNP C1KIK8
C	997	ARG	-	expression tag	UNP C1KIK8
C	998	ILE	-	expression tag	UNP C1KIK8
C	999	ASN	-	expression tag	UNP C1KIK8
C	1000	LEU	-	expression tag	UNP C1KIK8
C	1001	SER	-	expression tag	UNP C1KIK8
C	1002	GLY	-	expression tag	UNP C1KIK8
C	1003	ASP	-	expression tag	UNP C1KIK8
C	1013	GLU	LEU	conflict	UNP C1KIK8
C	1014	GLU	LEU	conflict	UNP C1KIK8
C	1017	GLN	ILE	conflict	UNP C1KIK8
C	1018	GLU	ILE	conflict	UNP C1KIK8
C	1021	GLN	LEU	conflict	UNP C1KIK8
C	1040	THR	ALA	conflict	UNP C1KIK8
C	1047	SER	CYS	conflict	UNP C1KIK8
C	1052	LEU	CYS	conflict	UNP C1KIK8
C	1072	THR	ILE	conflict	UNP C1KIK8
C	1086	GLN	PRO	conflict	UNP C1KIK8
C	1156	THR	ALA	engineered mutation	UNP C1KIK8
C	1159	SER	CYS	engineered mutation	UNP C1KIK8
D	987	LYS	-	expression tag	UNP C1KIK8
D	988	LYS	-	expression tag	UNP C1KIK8
D	989	LYS	-	expression tag	UNP C1KIK8
D	990	GLY	-	expression tag	UNP C1KIK8
D	991	SER	-	expression tag	UNP C1KIK8
D	992	VAL	-	expression tag	UNP C1KIK8
D	993	VAL	-	expression tag	UNP C1KIK8
D	994	ILE	-	expression tag	UNP C1KIK8
D	995	VAL	-	expression tag	UNP C1KIK8
D	996	GLY	-	expression tag	UNP C1KIK8
D	997	ARG	-	expression tag	UNP C1KIK8
D	998	ILE	-	expression tag	UNP C1KIK8
D	999	ASN	-	expression tag	UNP C1KIK8
D	1000	LEU	-	expression tag	UNP C1KIK8
D	1001	SER	-	expression tag	UNP C1KIK8
D	1002	GLY	-	expression tag	UNP C1KIK8
D	1003	ASP	-	expression tag	UNP C1KIK8
D	1013	GLU	LEU	conflict	UNP C1KIK8

*Continued on next page...*

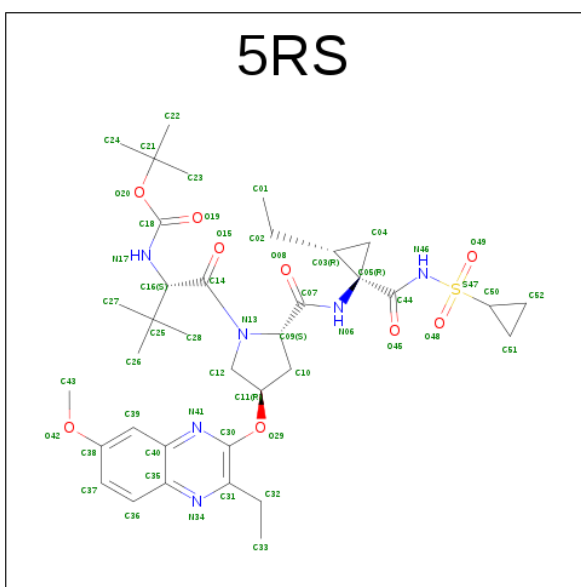
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	1014	GLU	LEU	conflict	UNP C1KIK8
D	1017	GLN	ILE	conflict	UNP C1KIK8
D	1018	GLU	ILE	conflict	UNP C1KIK8
D	1021	GLN	LEU	conflict	UNP C1KIK8
D	1040	THR	ALA	conflict	UNP C1KIK8
D	1047	SER	CYS	conflict	UNP C1KIK8
D	1052	LEU	CYS	conflict	UNP C1KIK8
D	1072	THR	ILE	conflict	UNP C1KIK8
D	1086	GLN	PRO	conflict	UNP C1KIK8
D	1156	THR	ALA	engineered mutation	UNP C1KIK8
D	1159	SER	CYS	engineered mutation	UNP C1KIK8

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Zn 1 1	0	0
2	A	1	Total Zn 1 1	0	0
2	D	1	Total Zn 1 1	0	0
2	C	1	Total Zn 1 1	0	0

- Molecule 3 is {tert}-butyl {N}-[(2 {S})-1-[(2 {S},4 {R})-2-[(1 {R},2 {R})-1-(cyclopropyl sulfonylcarbamoyl)-2-ethyl-cyclopropyl]carbamoyl]-4-(3-ethyl-7-methoxy-quinoxalin-2-yl) oxy-pyrrolidin-1-yl]-3,3-dimethyl-1-oxidanylidene-butan-2-yl]carbamate (three-letter code: 5RS) (formula: C<sub>36</sub>H<sub>52</sub>N<sub>6</sub>O<sub>9</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			52	36	6	9	1		
3	B	1	Total	C	N	O	S	0	0
			52	36	6	9	1		
3	C	1	Total	C	N	O	S	0	0
			52	36	6	9	1		
3	D	1	Total	C	N	O	S	0	0
			52	36	6	9	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Cl	0	0
			1	1		

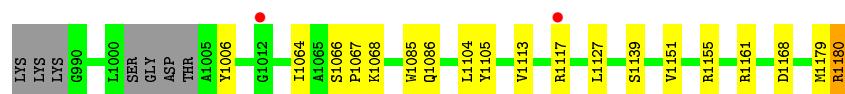
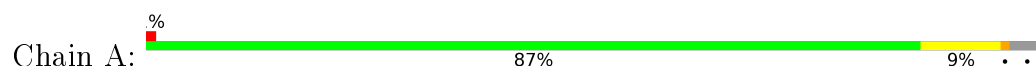
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	58	Total	O	0	0
			58	58		
5	B	52	Total	O	0	0
			52	52		
5	C	42	Total	O	0	0
			42	42		
5	D	49	Total	O	0	0
			49	49		

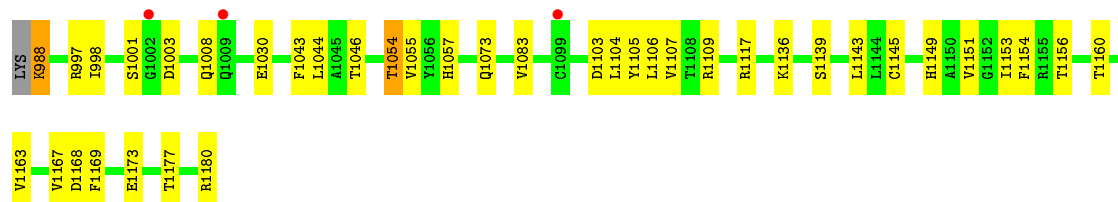
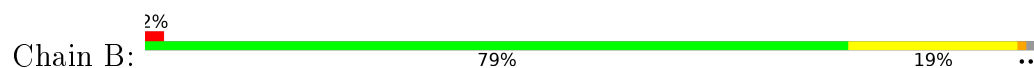
### 3 Residue-property plots [i](#)

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.

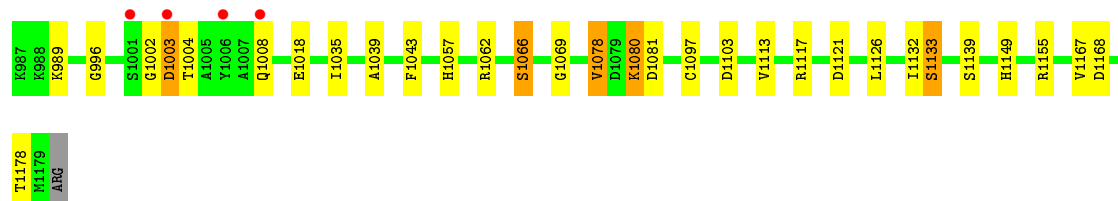
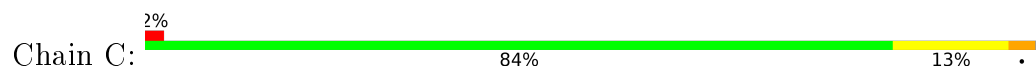
- Molecule 1: NS3 protease



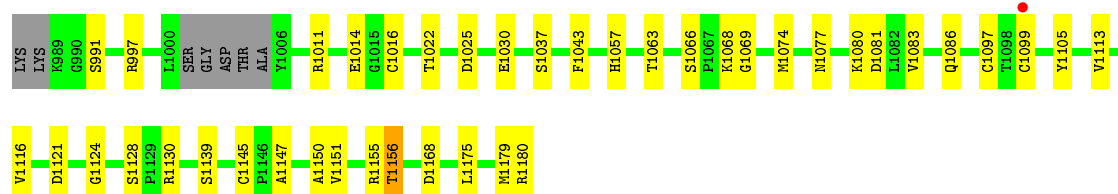
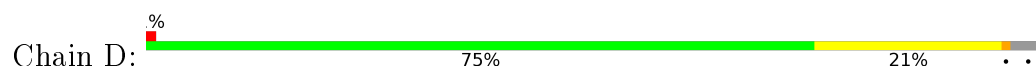
- Molecule 1: NS3 protease



- Molecule 1: NS3 protease



- Molecule 1: NS3 protease





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.44Å 63.23Å 92.30Å 90.00° 91.65° 90.00°	Depositor
Resolution (Å)	27.53 – 2.35 27.66 – 2.33	Depositor EDS
% Data completeness (in resolution range)	98.3 (27.53-2.35) 91.5 (27.66-2.33)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	10.51 (at 2.34Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.200 , 0.254 0.217 , 0.238	Depositor DCC
$R_{free}$ test set	1618 reflections (5.37%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.0	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 14.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.048 for k,h,-l 0.049 for -k,-h,-l 0.298 for h,-k,-l	Xtriage
Reported twinning fraction	0.080 for -k,-h,-l	Depositor
Outliers	1 of 31784 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11129	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.20% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 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: ZN, CL, 5RS

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	0.61	0/1369	0.78	0/1870
1	B	0.59	0/1403	0.79	0/1919
1	C	0.63	0/1418	0.82	0/1934
1	D	0.64	0/1390	0.80	0/1898
All	All	0.62	0/5580	0.80	0/7621

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	1343	1289	1304	12	2
1	B	1373	1315	1331	33	0
1	C	1388	1351	1358	25	1
1	D	1355	1301	1324	25	1
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	52	0	52	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	52	0	52	10	0
3	C	52	0	52	6	0
3	D	52	0	52	5	0
4	B	1	0	0	0	0
5	A	58	0	0	0	0
5	B	52	0	0	1	0
5	C	42	0	0	0	0
5	D	49	0	0	0	0
All	All	5873	5256	5525	100	2

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 (100) 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:1168:ASP:OD2	5:B:1301:HOH:O	1.63	1.14
1:B:1177:THR:HA	1:B:1180:ARG:HD2	1.50	0.91
1:A:1179:MET:O	1:A:1180:ARG:HB2	1.79	0.80
1:C:1066:SER:HB2	1:C:1069:GLY:O	1.84	0.78
1:C:989:LYS:HE2	1:C:1062:ARG:HE	1.54	0.71
1:D:1156:THR:HG21	3:D:1202:5RS:H222	1.72	0.70
1:D:1074:MET:SD	1:D:1086:GLN:HG2	2.30	0.70
1:B:1044:LEU:O	1:B:1054:THR:HG23	1.91	0.69
3:C:1202:5RS:O45	3:C:1202:5RS:H3	1.92	0.69
1:B:1054:THR:HG22	1:B:1055:VAL:H	1.59	0.65
1:B:1154:PHE:CE1	1:B:1167:VAL:HG23	2.30	0.65
1:D:1077:ASN:ND2	1:D:1080:LYS:HD2	2.14	0.63
1:C:1139:SER:HB3	3:C:1202:5RS:C44	2.29	0.61
1:B:1139:SER:HB3	3:B:1202:5RS:C44	2.32	0.60
1:C:1080:LYS:HE3	1:C:1155:ARG:NH2	2.17	0.59
3:A:1202:5RS:O45	3:A:1202:5RS:H3	2.03	0.58
1:C:1078:VAL:HG23	3:C:1202:5RS:H371	1.85	0.58
1:D:1057:HIS:HB2	3:D:1202:5RS:C40	2.34	0.58
1:C:1080:LYS:HE3	1:C:1155:ARG:HH22	1.70	0.57
1:B:1057:HIS:CE1	3:B:1202:5RS:H102	2.39	0.57
3:D:1202:5RS:H3	3:D:1202:5RS:O45	2.04	0.57
1:A:1066:SER:OG	1:A:1067:PRO:CD	2.53	0.56
1:A:1161:ARG:H	1:C:1121:ASP:HB2	1.70	0.55
1:D:1066:SER:OG	1:D:1069:GLY:O	2.20	0.55
1:A:1066:SER:OG	1:A:1067:PRO:HD2	2.08	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:1103:ASP:OD1	1:C:1117:ARG:NH2	2.33	0.54
3:B:1202:5RS:H3	3:B:1202:5RS:O45	2.07	0.54
1:B:1154:PHE:HE1	1:B:1167:VAL:HG23	1.72	0.54
1:C:1132:ILE:HG23	1:C:1133:SER:N	2.24	0.53
1:B:1167:VAL:HG22	1:B:1168:ASP:N	2.24	0.53
1:B:1173:GLU:OE2	1:B:1180:ARG:NH1	2.42	0.52
1:B:998:ILE:HD11	1:B:1109:ARG:HA	1.90	0.52
1:C:1113:VAL:O	1:C:1113:VAL:HG23	2.09	0.51
1:B:988:LYS:N	1:B:988:LYS:CD	2.72	0.51
1:D:1077:ASN:HD22	1:D:1080:LYS:HD2	1.75	0.51
1:B:1106:LEU:HD12	1:B:1107:VAL:N	2.27	0.50
3:B:1202:5RS:O19	3:B:1202:5RS:H223	2.10	0.50
1:B:1154:PHE:CE1	1:B:1167:VAL:CG2	2.95	0.49
1:D:1139:SER:HB3	3:D:1202:5RS:C44	2.42	0.49
3:C:1202:5RS:H122	3:C:1202:5RS:C27	2.43	0.49
1:D:1128:SER:O	1:D:1130:ARG:HD3	2.13	0.48
1:B:1145:CYS:SG	1:B:1149:HIS:HB2	2.52	0.48
1:B:997:ARG:NH2	1:B:1030:GLU:HG2	2.29	0.48
1:D:1156:THR:HG22	1:D:1168:ASP:HB3	1.95	0.48
1:D:1097:CYS:CB	1:D:1151:VAL:HG22	2.44	0.48
1:A:1117:ARG:NH1	1:D:1014:GLU:CB	2.77	0.47
1:B:997:ARG:CZ	1:B:1030:GLU:HG2	2.44	0.47
1:B:988:LYS:N	1:B:988:LYS:HD2	2.29	0.47
1:A:1155:ARG:NE	1:A:1168:ASP:OD2	2.41	0.46
1:A:1155:ARG:NH1	1:A:1168:ASP:OD1	2.37	0.46
1:C:1043:PHE:CD1	1:C:1043:PHE:N	2.84	0.46
1:B:1103:ASP:OD1	1:B:1117:ARG:NH2	2.41	0.46
1:B:1043:PHE:CE2	1:B:1054:THR:HG21	2.51	0.46
1:C:1057:HIS:CE1	3:C:1202:5RS:H102	2.51	0.46
1:C:1080:LYS:HD2	1:C:1080:LYS:HA	1.71	0.46
1:D:997:ARG:CD	1:D:1030:GLU:HG2	2.46	0.46
1:D:1145:CYS:SG	1:D:1147:ALA:HB3	2.57	0.45
1:D:1105:TYR:HB3	1:D:1113:VAL:HG13	1.98	0.45
1:D:1097:CYS:N	1:D:1150:ALA:O	2.46	0.45
1:C:1126:LEU:HD11	1:C:1167:VAL:HG13	1.99	0.45
1:D:1155:ARG:NE	1:D:1168:ASP:OD2	2.39	0.44
1:A:1139:SER:HB3	3:A:1202:5RS:C44	2.47	0.44
1:C:1097:CYS:HA	1:C:1149:HIS:CB	2.48	0.44
1:D:997:ARG:NE	1:D:1030:GLU:HG2	2.33	0.43
1:C:996:GLY:HA3	1:C:1035:ILE:HD12	2.00	0.43
1:C:989:LYS:NZ	1:C:1039:ALA:CB	2.81	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:1057:HIS:N	1:D:1081:ASP:OD1	2.47	0.43
3:B:1202:5RS:H122	3:B:1202:5RS:C27	2.49	0.43
1:D:1016:CYS:SG	1:D:1037:SER:HB2	2.58	0.43
1:C:1003:ASP:OD1	1:C:1003:ASP:N	2.52	0.43
1:D:991:SER:OG	1:D:1063:THR:OG1	2.23	0.43
1:C:989:LYS:CE	1:C:1062:ARG:HE	2.29	0.43
1:A:1105:TYR:HB3	1:A:1113:VAL:HG13	2.01	0.43
1:D:1179:MET:O	1:D:1180:ARG:HB2	2.19	0.42
1:B:1046:THR:HB	1:B:1153:ILE:CD1	2.49	0.42
1:A:1104:LEU:HD22	1:A:1151:VAL:HG21	1.99	0.42
1:D:1011:ARG:NH2	1:D:1025:ASP:OD1	2.52	0.42
1:C:1057:HIS:ND1	1:C:1081:ASP:OD2	2.45	0.42
1:C:989:LYS:HZ3	1:C:1039:ALA:HB3	1.85	0.42
1:A:1064:ILE:HB	1:A:1085:TRP:CE2	2.54	0.41
1:B:1043:PHE:HE2	1:B:1054:THR:HG21	1.85	0.41
1:B:1104:LEU:C	1:B:1105:TYR:CD1	2.94	0.41
1:B:1156:THR:HB	1:B:1168:ASP:HB3	2.01	0.41
1:B:1057:HIS:ND1	3:B:1202:5RS:H102	2.35	0.41
1:C:1080:LYS:CE	1:C:1155:ARG:HH22	2.32	0.41
1:C:1132:ILE:HG23	1:C:1133:SER:H	1.85	0.41
3:C:1202:5RS:H161	3:C:1202:5RS:H121	1.91	0.41
1:B:1143:LEU:HD13	1:B:1169:PHE:HB3	2.02	0.41
1:B:1043:PHE:N	1:B:1043:PHE:CD1	2.89	0.41
1:B:1104:LEU:HD22	1:B:1151:VAL:HG21	2.03	0.41
1:C:1097:CYS:HA	1:C:1149:HIS:HB3	2.03	0.41
1:A:1127:LEU:HB3	1:D:1022:THR:HG22	2.03	0.41
1:B:1154:PHE:CZ	3:B:1202:5RS:C04	3.04	0.41
1:B:1154:PHE:CZ	3:B:1202:5RS:H042	2.56	0.40
1:B:1136:LYS:HA	3:B:1202:5RS:H3	2.03	0.40
1:C:1002:GLY:O	1:C:1004:THR:N	2.54	0.40
1:D:1057:HIS:CD2	3:D:1202:5RS:H102	2.56	0.40
1:D:1116:VAL:CG1	1:D:1124:GLY:HA3	2.50	0.40
1:B:1073:GLN:CD	1:B:1083:VAL:HG13	2.42	0.40
1:B:1057:HIS:CG	3:B:1202:5RS:H102	2.56	0.40

All (2) 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:A:1068:LYS:HZ3	1:D:1068:LYS:HG2[2_455]	1.17	0.43
1:A:1006:TYR:HE2	1:C:1008:GLN:HE21[1_455]	1.24	0.36

## 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	184/194 (95%)	181 (98%)	3 (2%)	0	100	100
1	B	193/194 (100%)	188 (97%)	4 (2%)	1 (0%)	34	39
1	C	193/194 (100%)	188 (97%)	4 (2%)	1 (0%)	34	39
1	D	187/194 (96%)	183 (98%)	4 (2%)	0	100	100
All	All	757/776 (98%)	740 (98%)	15 (2%)	2 (0%)	46	55

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1003	ASP
1	C	1003	ASP

### 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	142/159 (89%)	140 (99%)	2 (1%)	74	86
1	B	142/159 (89%)	136 (96%)	6 (4%)	36	46
1	C	147/159 (92%)	140 (95%)	7 (5%)	31	40
1	D	144/159 (91%)	137 (95%)	7 (5%)	31	39
All	All	575/636 (90%)	553 (96%)	22 (4%)	41	52

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1086	GLN
1	A	1180	ARG
1	B	988	LYS
1	B	1001	SER
1	B	1008	GLN
1	B	1054	THR
1	B	1160	THR
1	B	1163	VAL
1	C	1018	GLU
1	C	1066	SER
1	C	1078	VAL
1	C	1080	LYS
1	C	1133	SER
1	C	1168	ASP
1	C	1178	THR
1	D	1043	PHE
1	D	1083	VAL
1	D	1099	CYS
1	D	1121[A]	ASP
1	D	1121[B]	ASP
1	D	1156	THR
1	D	1175	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1009	GLN
1	C	1149	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 carbohydrates in this entry.

## 5.6 Ligand geometry

Of 9 ligands modelled in this entry, 5 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$
3	5RS	A	1202	-	52,56,56	1.91	10 (19%)	67,87,87	1.58	7 (10%)
3	5RS	B	1202	-	52,56,56	1.95	12 (23%)	67,87,87	1.74	10 (14%)
3	5RS	C	1202	-	52,56,56	1.75	7 (13%)	67,87,87	1.99	13 (19%)
3	5RS	D	1202	-	52,56,56	1.87	9 (17%)	67,87,87	1.56	8 (11%)

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	5RS	A	1202	-	-	0/57/81/81	0/3/5/5
3	5RS	B	1202	-	-	0/57/81/81	0/3/5/5
3	5RS	C	1202	-	-	0/57/81/81	0/3/5/5
3	5RS	D	1202	-	-	0/57/81/81	0/3/5/5

All (38) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1202	5RS	C32-C31	-5.85	1.39	1.51
3	A	1202	5RS	C32-C31	-5.84	1.39	1.51
3	C	1202	5RS	C09-C07	-5.71	1.40	1.52
3	D	1202	5RS	C09-C07	-5.65	1.40	1.52
3	B	1202	5RS	C32-C31	-5.60	1.40	1.51
3	D	1202	5RS	C32-C31	-5.46	1.40	1.51
3	A	1202	5RS	C09-C07	-5.45	1.41	1.52
3	B	1202	5RS	C09-C07	-5.45	1.41	1.52
3	A	1202	5RS	S47-N46	-3.88	1.53	1.62
3	B	1202	5RS	O20-C21	-3.04	1.42	1.48
3	D	1202	5RS	C39-C40	-2.97	1.37	1.41

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1202	5RS	C35-N34	-2.95	1.32	1.37
3	B	1202	5RS	C35-N34	-2.80	1.32	1.37
3	D	1202	5RS	S47-N46	-2.71	1.55	1.62
3	D	1202	5RS	C40-N41	-2.69	1.32	1.37
3	A	1202	5RS	O20-C21	-2.64	1.43	1.48
3	A	1202	5RS	C40-N41	-2.59	1.33	1.37
3	A	1202	5RS	C35-N34	-2.58	1.33	1.37
3	C	1202	5RS	O20-C21	-2.46	1.43	1.48
3	A	1202	5RS	C39-C40	-2.44	1.38	1.41
3	C	1202	5RS	C36-C35	-2.33	1.37	1.41
3	A	1202	5RS	C44-N46	-2.27	1.33	1.38
3	D	1202	5RS	C35-N34	-2.27	1.33	1.37
3	B	1202	5RS	C39-C40	-2.24	1.38	1.41
3	D	1202	5RS	C44-N46	-2.13	1.33	1.38
3	B	1202	5RS	C36-C35	-2.13	1.38	1.41
3	B	1202	5RS	C40-N41	-2.06	1.34	1.37
3	B	1202	5RS	C44-N46	-2.05	1.34	1.38
3	B	1202	5RS	C05-C03	2.10	1.54	1.52
3	B	1202	5RS	C31-N34	2.14	1.34	1.32
3	C	1202	5RS	O49-S47	4.34	1.48	1.43
3	C	1202	5RS	O48-S47	4.64	1.49	1.43
3	D	1202	5RS	O49-S47	5.09	1.49	1.43
3	A	1202	5RS	O49-S47	5.13	1.49	1.43
3	B	1202	5RS	O48-S47	5.79	1.50	1.43
3	D	1202	5RS	O48-S47	5.79	1.50	1.43
3	A	1202	5RS	O48-S47	6.11	1.50	1.43
3	B	1202	5RS	O49-S47	6.11	1.50	1.43

All (38) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1202	5RS	O49-S47-O48	-7.53	106.87	119.31
3	A	1202	5RS	O49-S47-O48	-7.39	107.10	119.31
3	D	1202	5RS	O49-S47-O48	-6.82	108.05	119.31
3	B	1202	5RS	O49-S47-O48	-6.10	109.23	119.31
3	B	1202	5RS	C10-C11-C12	-4.79	99.40	103.56
3	C	1202	5RS	C10-C11-C12	-3.71	100.33	103.56
3	D	1202	5RS	O19-C18-N17	-3.45	118.89	124.89
3	A	1202	5RS	O19-C18-N17	-3.38	119.03	124.89
3	B	1202	5RS	O19-C18-N17	-3.37	119.04	124.89
3	B	1202	5RS	C25-C16-N17	-3.10	108.73	112.28
3	C	1202	5RS	O19-C18-N17	-2.99	119.70	124.89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1202	5RS	C07-C09-N13	-2.85	105.36	112.57
3	D	1202	5RS	C12-N13-C09	-2.79	107.33	111.62
3	C	1202	5RS	C25-C16-N17	-2.78	109.10	112.28
3	D	1202	5RS	C10-C11-C12	-2.77	101.15	103.56
3	C	1202	5RS	C04-C05-N06	-2.77	114.05	117.87
3	C	1202	5RS	O20-C18-O19	-2.64	120.28	125.55
3	C	1202	5RS	C07-C09-N13	-2.29	106.78	112.57
3	C	1202	5RS	C12-N13-C09	-2.22	108.21	111.62
3	B	1202	5RS	C12-N13-C09	-2.15	108.31	111.62
3	C	1202	5RS	C16-C14-N13	-2.03	114.98	118.50
3	D	1202	5RS	C10-C09-N13	2.02	106.13	103.15
3	A	1202	5RS	O29-C11-C10	2.05	114.21	108.63
3	A	1202	5RS	C11-C12-N13	2.20	105.62	102.61
3	A	1202	5RS	O49-S47-N46	2.29	114.28	107.49
3	D	1202	5RS	C31-N34-C35	2.36	120.34	117.67
3	C	1202	5RS	C31-N34-C35	2.56	120.57	117.67
3	A	1202	5RS	C31-N34-C35	2.61	120.62	117.67
3	B	1202	5RS	C38-C39-C40	2.76	120.80	119.19
3	D	1202	5RS	C38-C39-C40	2.78	120.81	119.19
3	B	1202	5RS	C44-C05-N06	3.22	119.79	116.08
3	B	1202	5RS	C14-C16-N17	4.26	111.97	107.45
3	D	1202	5RS	O20-C18-N17	4.61	118.44	110.07
3	C	1202	5RS	C44-C05-N06	5.02	121.88	116.08
3	B	1202	5RS	O20-C18-N17	5.04	119.24	110.07
3	A	1202	5RS	O20-C18-N17	5.14	119.41	110.07
3	C	1202	5RS	O20-C18-N17	5.48	120.02	110.07
3	C	1202	5RS	C14-C16-N17	6.49	114.34	107.45

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1202	5RS	2	0
3	B	1202	5RS	10	0
3	C	1202	5RS	6	0
3	D	1202	5RS	5	0

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	187/194 (96%)	0.09	2 (1%) 82 90	10, 23, 39, 53	0
1	B	193/194 (99%)	0.09	3 (1%) 74 84	11, 24, 43, 73	0
1	C	193/194 (99%)	0.19	4 (2%) 67 79	10, 24, 46, 87	0
1	D	187/194 (96%)	0.02	1 (0%) 91 96	11, 22, 38, 54	0
All	All	760/776 (97%)	0.10	10 (1%) 79 87	10, 23, 42, 87	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1012	GLY	3.0
1	C	1006	TYR	2.9
1	D	1099	CYS	2.7
1	C	1001	SER	2.5
1	C	1003	ASP	2.5
1	B	1099	CYS	2.5
1	A	1117	ARG	2.3
1	C	1008	GLN	2.3
1	B	1009	GLN	2.2
1	B	1002	GLY	2.1

### 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 carbohydrates in this entry.

## 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
3	5RS	D	1202	52/52	0.93	0.15	0.95	20,20,20,20	0
3	5RS	A	1202	52/52	0.94	0.15	0.57	20,20,20,20	0
3	5RS	C	1202	52/52	0.92	0.14	0.53	20,20,20,20	0
3	5RS	B	1202	52/52	0.93	0.14	0.29	20,20,20,20	0
2	ZN	B	1201	1/1	0.53	0.16	0.17	42,42,42,42	0
2	ZN	D	1201	1/1	0.86	0.10	-1.20	30,30,30,30	0
2	ZN	C	1201	1/1	0.98	0.08	-2.58	37,37,37,37	0
2	ZN	A	1201	1/1	0.92	0.07	-3.92	30,30,30,30	0
4	CL	B	1203	1/1	0.92	0.08	-	36,36,36,36	0

## 6.5 Other polymers

There are no such residues in this entry.