wwPDB X-ray Structure Validation Summary Report (i)

Sep 24, 2023 – 04:29 AM EDT

PDB ID	:	5THP
Title	:	Rhodocetin in complex with the integrin alpha2-A domain
Authors	:	McDougall, M.; Orriss, G.L.; Stetefeld, J.
Deposited on	:	2016-09-30
Resolution	:	3.01 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 (1)) were used in the production of this report:

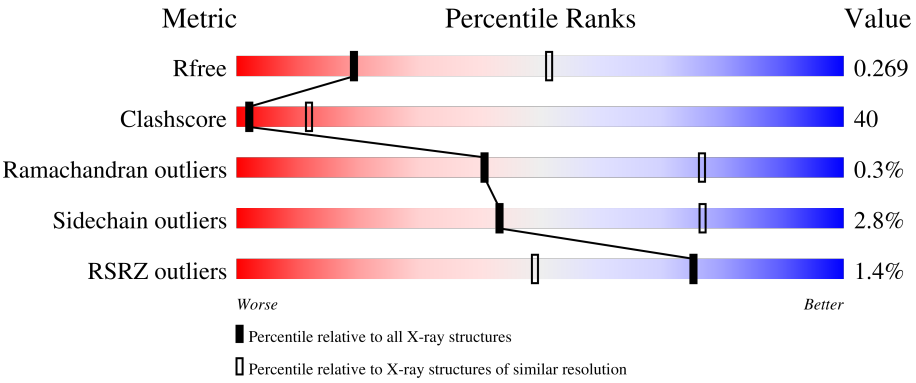
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.01 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.



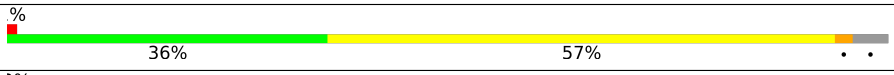
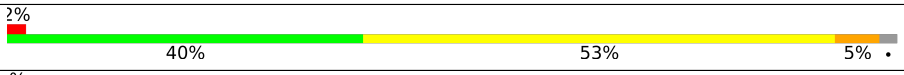
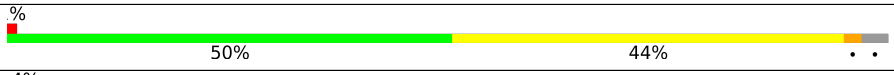
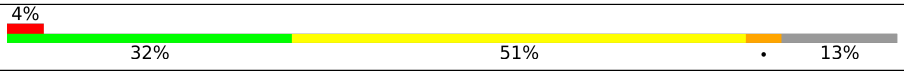
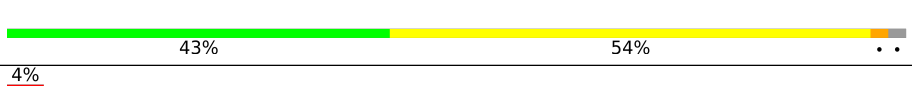
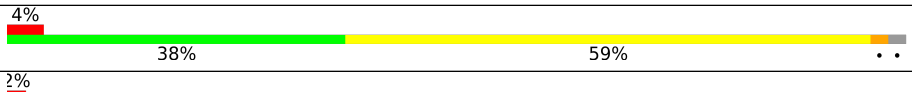
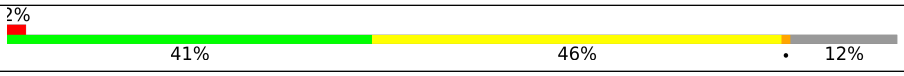
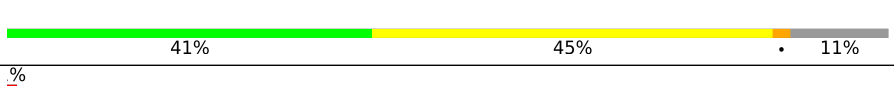
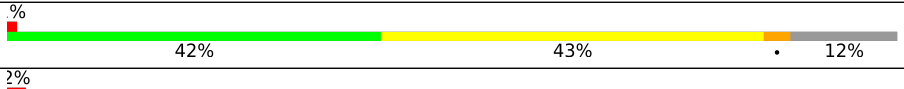
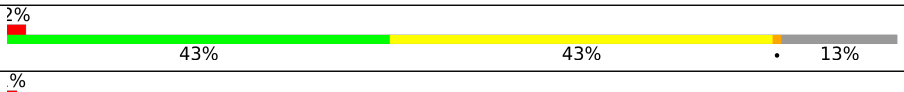
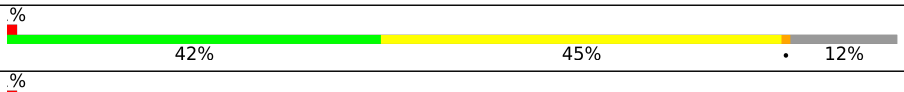
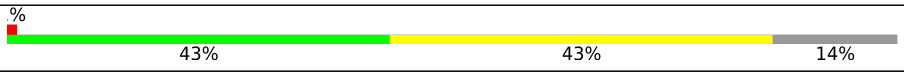
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	135	41% 53%
1	D	135	47% 50%
1	G	135	41% 55%
1	J	135	39% 57%
1	M	135	4% 28% 67%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	P	135	
2	B	124	
2	E	124	
2	H	124	
2	K	124	
2	N	124	
2	Q	124	
3	C	217	
3	F	217	
3	I	217	
3	L	217	
3	O	217	
3	R	217	

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
4	NA	Q	201	-	-	-	X
5	CL	I	404	-	-	-	X
5	CL	O	406	-	-	X	-
6	GOL	J	204	-	-	X	-
7	MG	L	403	-	-	-	X
7	MG	O	405	-	-	-	X
8	SO4	L	406	-	-	X	-
8	SO4	O	407	-	-	X	-

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 21018 atoms, of which 8 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 Snaclec rhodocetin subunit gamma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	129	Total	C	N	O	S	0	0	0
			1006	638	171	188	9			
1	D	130	Total	C	N	O	S	0	0	0
			1022	642	176	196	8			
1	G	132	Total	C	N	O	S	0	0	0
			1064	671	185	199	9			
1	J	131	Total	C	N	O	S	0	0	0
			1046	658	180	199	9			
1	M	132	Total	C	N	O	S	0	0	0
			1041	655	178	199	9			
1	P	132	Total	C	N	O	S	0	0	0
			1058	666	182	201	9			

- Molecule 2 is a protein called Snaclec rhodocetin subunit delta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	119	Total	C	N	O	S	0	0	0
			987	641	167	170	9			
2	E	122	Total	C	N	O	S	0	0	0
			1006	650	169	178	9			
2	H	120	Total	C	N	O	S	0	0	0
			949	609	160	171	9			
2	K	108	Total	C	N	O	S	0	0	0
			885	572	149	157	7			
2	N	122	Total	C	N	O	S	0	0	0
			1029	667	174	179	9			
2	Q	122	Total	C	N	O	S	0	0	0
			1026	664	174	179	9			

- Molecule 3 is a protein called Integrin alpha-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	190	Total	C	N	O	S	0	0	0
			1435	905	244	282	4			
3	F	193	Total	C	N	O	S	0	0	0
			1443	908	246	284	5			
3	I	191	Total	C	N	O	S	0	0	0
			1396	882	232	277	5			
3	L	189	Total	C	N	O	S	0	0	0
			1392	878	236	273	5			
3	O	191	Total	C	N	O	S	0	0	0
			1430	907	235	283	5			
3	R	187	Total	C	N	O	S	0	0	0
			1407	892	236	274	5			

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	150	MET	-	initiating methionine	UNP P17301
C	151	GLY	-	expression tag	UNP P17301
C	152	SER	-	expression tag	UNP P17301
C	153	SER	-	expression tag	UNP P17301
C	154	HIS	-	expression tag	UNP P17301
C	155	HIS	-	expression tag	UNP P17301
C	156	HIS	-	expression tag	UNP P17301
C	157	HIS	-	expression tag	UNP P17301
C	158	HIS	-	expression tag	UNP P17301
C	159	HIS	-	expression tag	UNP P17301
C	160	SER	-	expression tag	UNP P17301
C	161	SER	-	expression tag	UNP P17301
C	162	GLY	-	expression tag	UNP P17301
C	163	LEU	-	expression tag	UNP P17301
C	164	VAL	-	expression tag	UNP P17301
C	165	PRO	-	expression tag	UNP P17301
C	166	ARG	-	expression tag	UNP P17301
C	167	GLY	-	expression tag	UNP P17301
C	168	GLY	-	expression tag	UNP P17301
C	169	SER	-	expression tag	UNP P17301
F	150	MET	-	initiating methionine	UNP P17301
F	151	GLY	-	expression tag	UNP P17301
F	152	SER	-	expression tag	UNP P17301
F	153	SER	-	expression tag	UNP P17301
F	154	HIS	-	expression tag	UNP P17301
F	155	HIS	-	expression tag	UNP P17301
F	156	HIS	-	expression tag	UNP P17301
F	157	HIS	-	expression tag	UNP P17301

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	158	HIS	-	expression tag	UNP P17301
F	159	HIS	-	expression tag	UNP P17301
F	160	SER	-	expression tag	UNP P17301
F	161	SER	-	expression tag	UNP P17301
F	162	GLY	-	expression tag	UNP P17301
F	163	LEU	-	expression tag	UNP P17301
F	164	VAL	-	expression tag	UNP P17301
F	165	PRO	-	expression tag	UNP P17301
F	166	ARG	-	expression tag	UNP P17301
F	167	GLY	-	expression tag	UNP P17301
F	168	GLY	-	expression tag	UNP P17301
F	169	SER	-	expression tag	UNP P17301
I	150	MET	-	initiating methionine	UNP P17301
I	151	GLY	-	expression tag	UNP P17301
I	152	SER	-	expression tag	UNP P17301
I	153	SER	-	expression tag	UNP P17301
I	154	HIS	-	expression tag	UNP P17301
I	155	HIS	-	expression tag	UNP P17301
I	156	HIS	-	expression tag	UNP P17301
I	157	HIS	-	expression tag	UNP P17301
I	158	HIS	-	expression tag	UNP P17301
I	159	HIS	-	expression tag	UNP P17301
I	160	SER	-	expression tag	UNP P17301
I	161	SER	-	expression tag	UNP P17301
I	162	GLY	-	expression tag	UNP P17301
I	163	LEU	-	expression tag	UNP P17301
I	164	VAL	-	expression tag	UNP P17301
I	165	PRO	-	expression tag	UNP P17301
I	166	ARG	-	expression tag	UNP P17301
I	167	GLY	-	expression tag	UNP P17301
I	168	GLY	-	expression tag	UNP P17301
I	169	SER	-	expression tag	UNP P17301
L	150	MET	-	initiating methionine	UNP P17301
L	151	GLY	-	expression tag	UNP P17301
L	152	SER	-	expression tag	UNP P17301
L	153	SER	-	expression tag	UNP P17301
L	154	HIS	-	expression tag	UNP P17301
L	155	HIS	-	expression tag	UNP P17301
L	156	HIS	-	expression tag	UNP P17301
L	157	HIS	-	expression tag	UNP P17301
L	158	HIS	-	expression tag	UNP P17301
L	159	HIS	-	expression tag	UNP P17301

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	160	SER	-	expression tag	UNP P17301
L	161	SER	-	expression tag	UNP P17301
L	162	GLY	-	expression tag	UNP P17301
L	163	LEU	-	expression tag	UNP P17301
L	164	VAL	-	expression tag	UNP P17301
L	165	PRO	-	expression tag	UNP P17301
L	166	ARG	-	expression tag	UNP P17301
L	167	GLY	-	expression tag	UNP P17301
L	168	GLY	-	expression tag	UNP P17301
L	169	SER	-	expression tag	UNP P17301
O	150	MET	-	initiating methionine	UNP P17301
O	151	GLY	-	expression tag	UNP P17301
O	152	SER	-	expression tag	UNP P17301
O	153	SER	-	expression tag	UNP P17301
O	154	HIS	-	expression tag	UNP P17301
O	155	HIS	-	expression tag	UNP P17301
O	156	HIS	-	expression tag	UNP P17301
O	157	HIS	-	expression tag	UNP P17301
O	158	HIS	-	expression tag	UNP P17301
O	159	HIS	-	expression tag	UNP P17301
O	160	SER	-	expression tag	UNP P17301
O	161	SER	-	expression tag	UNP P17301
O	162	GLY	-	expression tag	UNP P17301
O	163	LEU	-	expression tag	UNP P17301
O	164	VAL	-	expression tag	UNP P17301
O	165	PRO	-	expression tag	UNP P17301
O	166	ARG	-	expression tag	UNP P17301
O	167	GLY	-	expression tag	UNP P17301
O	168	GLY	-	expression tag	UNP P17301
O	169	SER	-	expression tag	UNP P17301
R	150	MET	-	initiating methionine	UNP P17301
R	151	GLY	-	expression tag	UNP P17301
R	152	SER	-	expression tag	UNP P17301
R	153	SER	-	expression tag	UNP P17301
R	154	HIS	-	expression tag	UNP P17301
R	155	HIS	-	expression tag	UNP P17301
R	156	HIS	-	expression tag	UNP P17301
R	157	HIS	-	expression tag	UNP P17301
R	158	HIS	-	expression tag	UNP P17301
R	159	HIS	-	expression tag	UNP P17301
R	160	SER	-	expression tag	UNP P17301
R	161	SER	-	expression tag	UNP P17301

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
R	162	GLY	-	expression tag	UNP P17301
R	163	LEU	-	expression tag	UNP P17301
R	164	VAL	-	expression tag	UNP P17301
R	165	PRO	-	expression tag	UNP P17301
R	166	ARG	-	expression tag	UNP P17301
R	167	GLY	-	expression tag	UNP P17301
R	168	GLY	-	expression tag	UNP P17301
R	169	SER	-	expression tag	UNP P17301

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Na 1 1	0	0
4	B	1	Total Na 1 1	0	0
4	C	4	Total Na 4 4	0	0
4	D	3	Total Na 3 3	0	0
4	E	3	Total Na 3 3	0	0
4	F	2	Total Na 2 2	0	0
4	G	1	Total Na 1 1	0	0
4	H	1	Total Na 1 1	0	0
4	I	2	Total Na 2 2	0	0
4	J	2	Total Na 2 2	0	0
4	L	2	Total Na 2 2	0	0
4	M	1	Total Na 1 1	0	0
4	N	1	Total Na 1 1	0	0
4	O	3	Total Na 3 3	0	0
4	P	1	Total Na 1 1	0	0

Continued on next page...

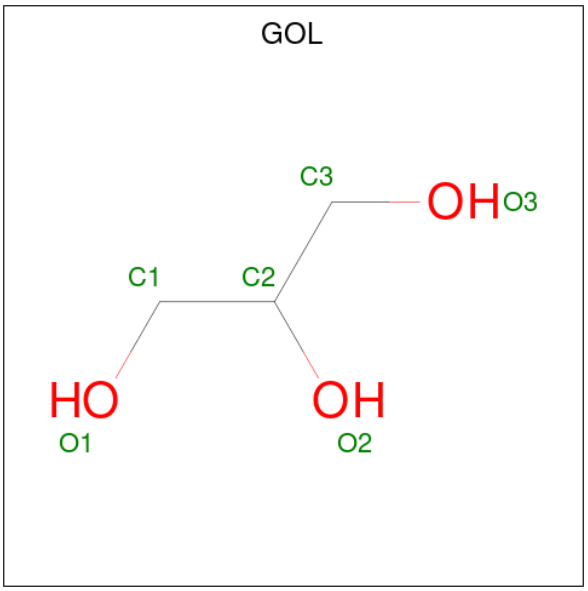
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	Q	2	Total Na 2 2	0	0
4	R	1	Total Na 1 1	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Cl 1 1	0	0
5	C	4	Total Cl 4 4	0	0
5	D	5	Total Cl 5 5	0	0
5	E	3	Total Cl 3 3	0	0
5	F	1	Total Cl 1 1	0	0
5	G	2	Total Cl 2 2	0	0
5	H	2	Total Cl 2 2	0	0
5	I	3	Total Cl 3 3	0	0
5	J	1	Total Cl 1 1	0	0
5	L	2	Total Cl 2 2	0	0
5	M	2	Total Cl 2 2	0	0
5	N	3	Total Cl 3 3	0	0
5	O	1	Total Cl 1 1	0	0
5	P	2	Total Cl 2 2	0	0
5	Q	2	Total Cl 2 2	0	0
5	R	1	Total Cl 1 1	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 6 3 3	0	0
6	C	1	Total C O 6 3 3	0	0
6	G	1	Total C O 6 3 3	0	0
6	G	1	Total C H O 14 3 8 3	0	0
6	H	1	Total C O 6 3 3	0	0
6	J	1	Total C O 6 3 3	0	0
6	J	1	Total C O 6 3 3	0	0
6	P	1	Total C O 6 3 3	0	0

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

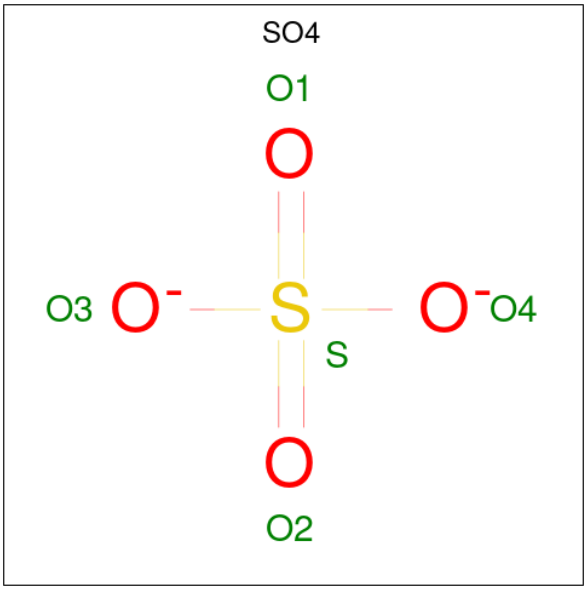
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	1	Total Mg 1 1	0	0
7	F	1	Total Mg 1 1	0	0
7	I	1	Total Mg 1 1	0	0
7	L	1	Total Mg 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	O	1	Total Mg 1 1	0	0
7	R	1	Total Mg 1 1	0	0

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	C	1	Total O S 5 4 1	0	0
8	D	1	Total O S 5 4 1	0	0
8	E	1	Total O S 5 4 1	0	0
8	F	1	Total O S 5 4 1	0	0
8	F	1	Total O S 5 4 1	0	0
8	I	1	Total O S 5 4 1	0	0
8	L	1	Total O S 5 4 1	0	0
8	O	1	Total O S 5 4 1	0	0
8	O	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	R	1	Total O S 5 4 1	0	0

- Molecule 9 is water.

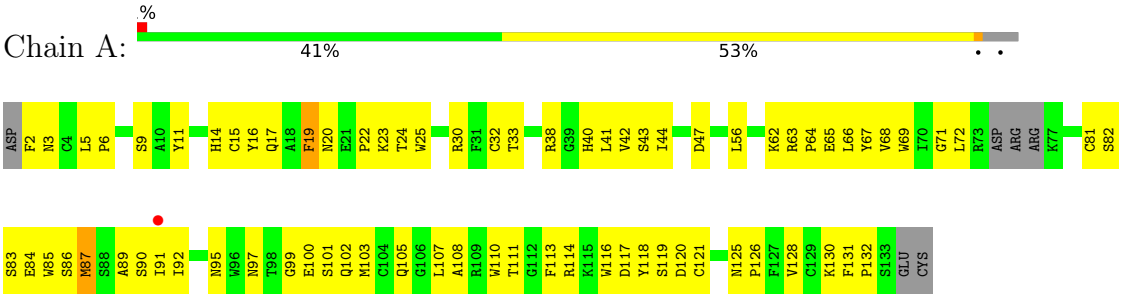
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	16	Total O 16 16	0	0
9	B	7	Total O 7 7	0	0
9	C	15	Total O 15 15	0	0
9	D	19	Total O 19 19	0	0
9	E	9	Total O 9 9	0	0
9	F	11	Total O 11 11	0	0
9	G	16	Total O 16 16	0	0
9	H	7	Total O 7 7	0	0
9	I	16	Total O 16 16	0	0
9	J	10	Total O 10 10	0	0
9	K	8	Total O 8 8	0	0
9	L	17	Total O 17 17	0	0
9	M	6	Total O 6 6	0	0
9	N	7	Total O 7 7	0	0
9	O	20	Total O 20 20	0	0
9	P	6	Total O 6 6	0	0
9	Q	14	Total O 14 14	0	0
9	R	14	Total O 14 14	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.

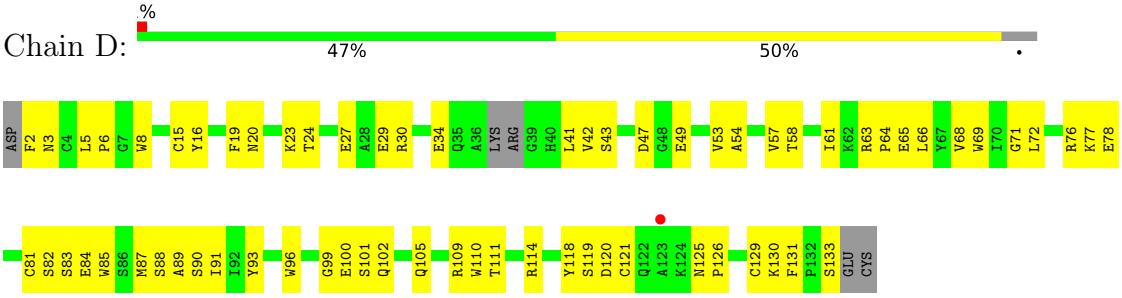
• Molecule 1: Snaclec rhodocetin subunit gamma

Chain A:



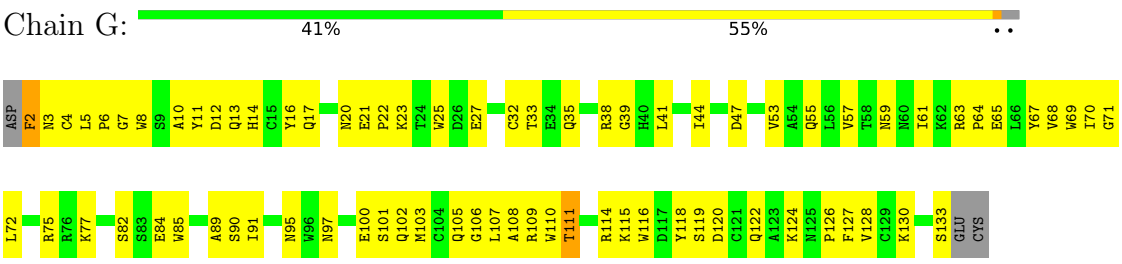
• Molecule 1: Snaclec rhodocetin subunit gamma

Chain D:



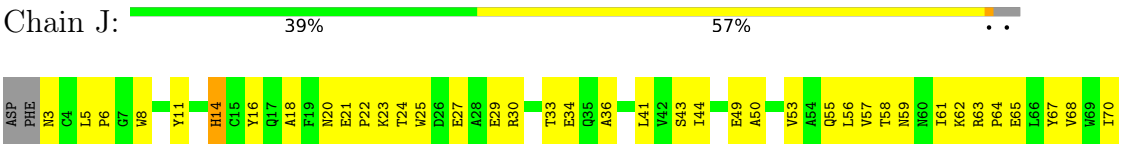
• Molecule 1: Snaclec rhodocetin subunit gamma

Chain G:



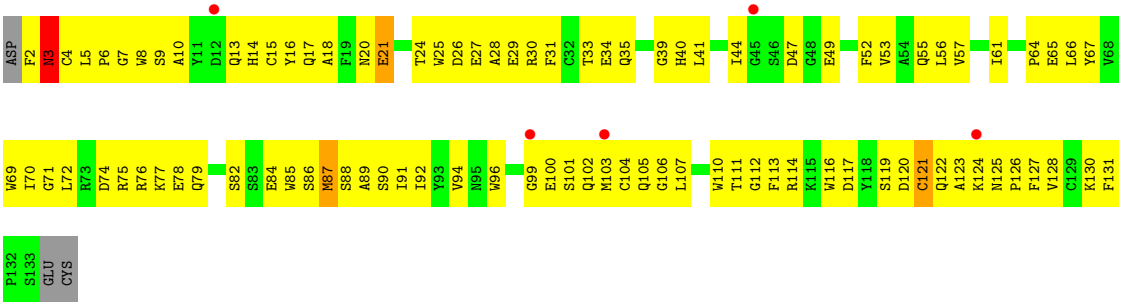
• Molecule 1: Snaclec rhodocetin subunit gamma

Chain J:



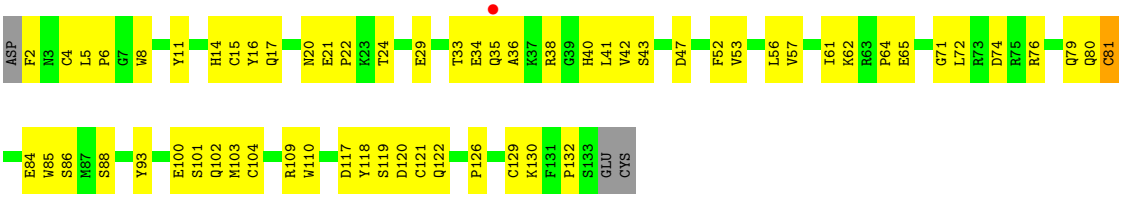
• Molecule 1: Snaclec rhodocetin subunit gamma

Chain M:



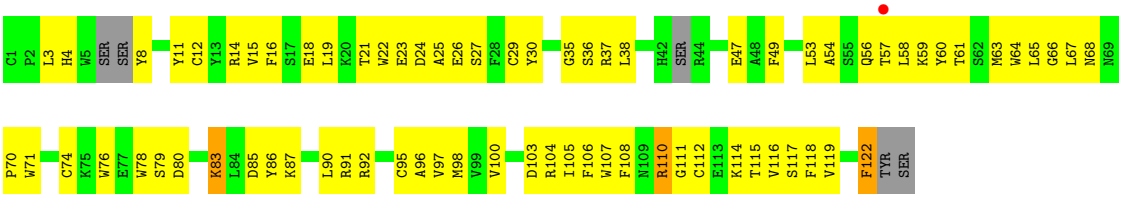
• Molecule 1: Snaclec rhodocetin subunit gamma

Chain P:



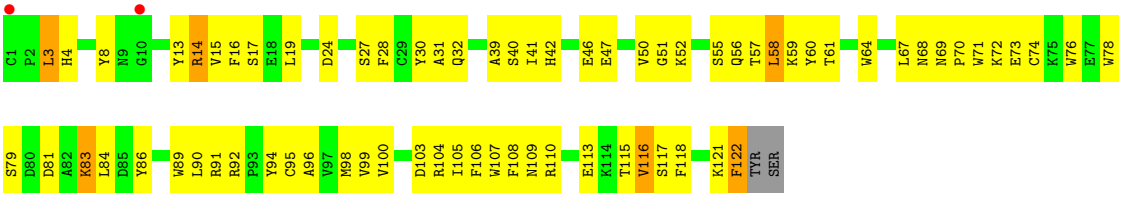
• Molecule 2: Snaclec rhodocetin subunit delta

Chain B:



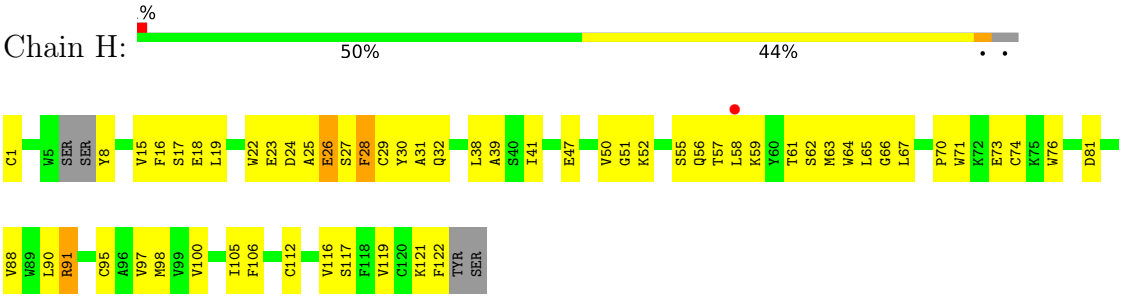
• Molecule 2: Snaclec rhodocetin subunit delta

Chain E:



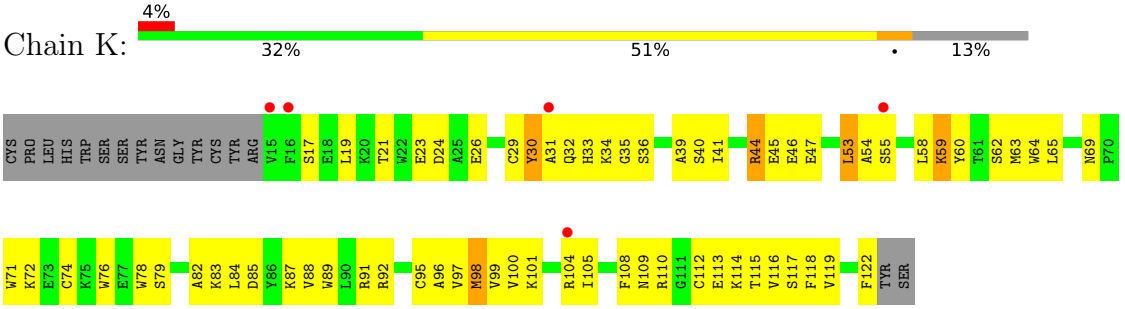
• Molecule 2: Snaclec rhodocetin subunit delta

Chain H:



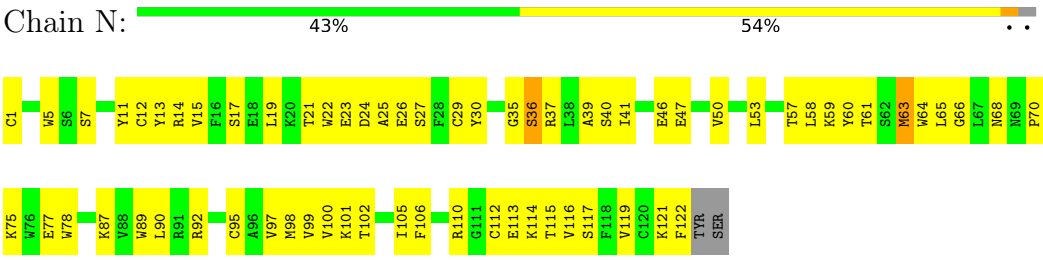
• Molecule 2: Snaclec rhodocetin subunit delta

Chain K:



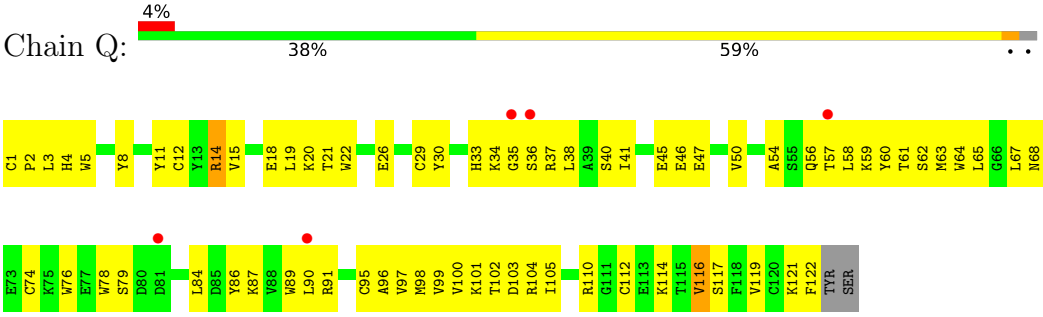
• Molecule 2: Snaclec rhodocetin subunit delta

Chain N:



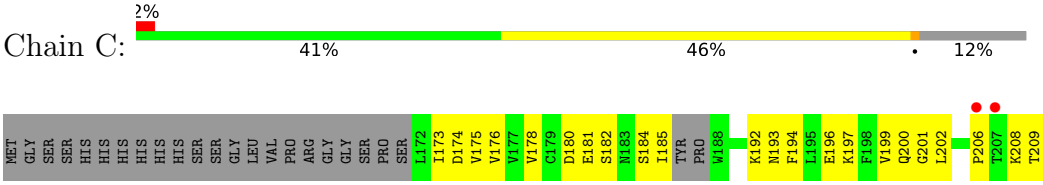
• Molecule 2: Snaclec rhodocetin subunit delta

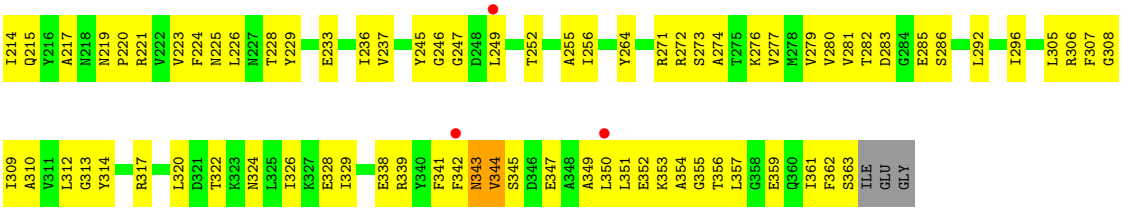
Chain Q:



• Molecule 3: Integrin alpha-2

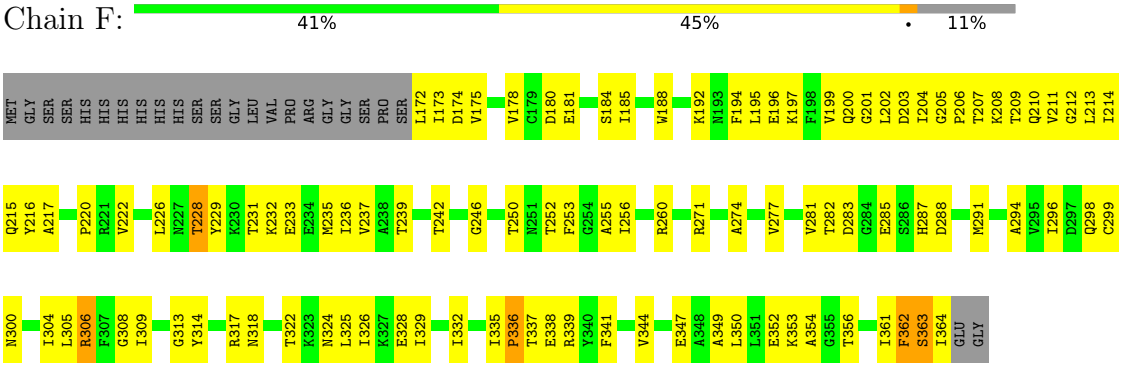
Chain C:





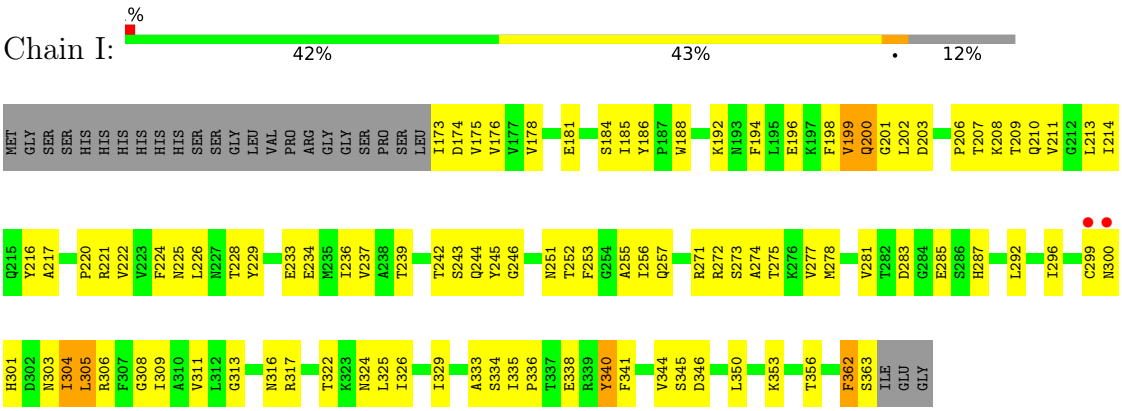
• Molecule 3: Integrin alpha-2

Chain F:



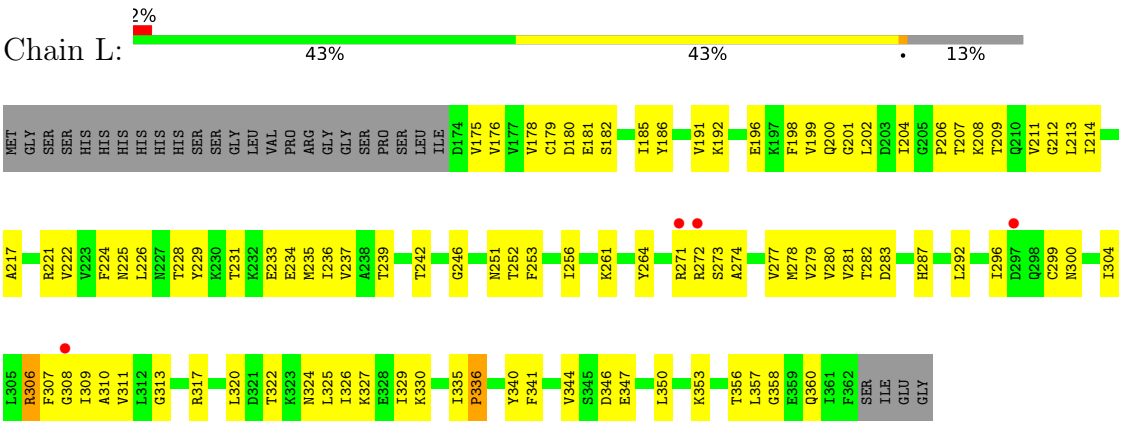
• Molecule 3: Integrin alpha-2

Chain I:



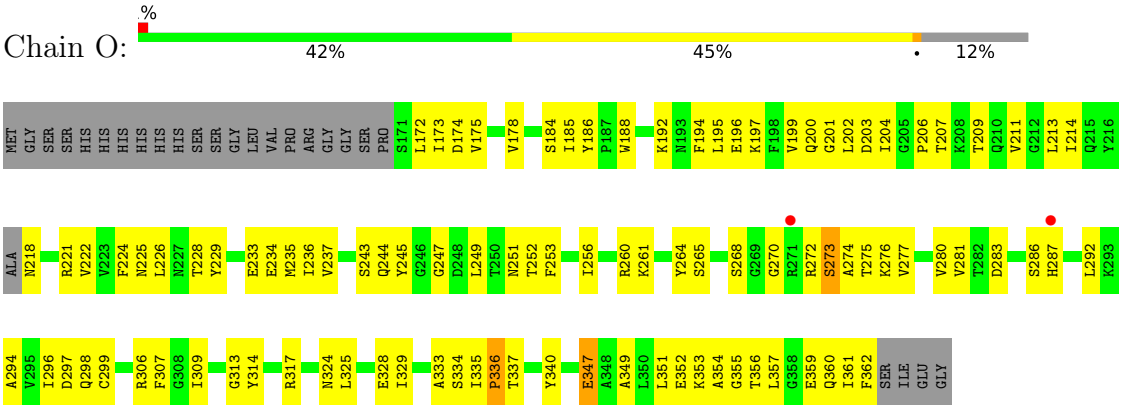
• Molecule 3: Integrin alpha-2

Chain L:



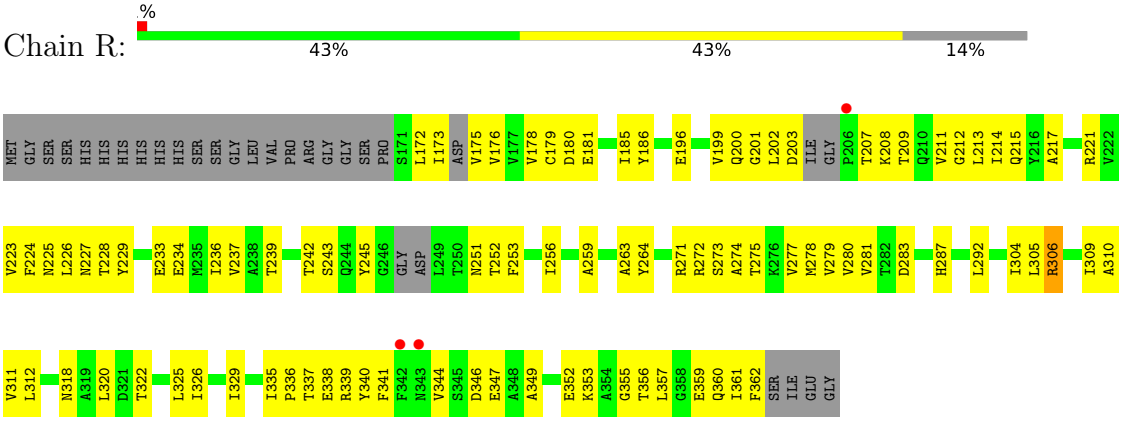
• Molecule 3: Integrin alpha-2

Chain O:



• Molecule 3: Integrin alpha-2

Chain R:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	130.76Å 130.76Å 251.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.87 – 3.01 19.87 – 3.01	Depositor EDS
% Data completeness (in resolution range)	99.6 (19.87-3.01) 99.6 (19.87-3.01)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.27 (at 2.98Å)	Xtriage
Refinement program	PHENIX (1.10.1-2155_2155: ???)	Depositor
R, R_{free}	0.221 , 0.270 0.221 , 0.269	Depositor DCC
R_{free} test set	4300 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	71.1	Xtriage
Anisotropy	0.257	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 59.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.467 for h,-k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	21018	wwPDB-VP
Average B, all atoms (Å^2)	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% of the height of the origin peak. No significant pseudotranslation is detected.*

[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: GOL, SO4, NA, CL, MG

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	0.24	0/1036	0.43	0/1409
1	D	0.24	0/1053	0.46	0/1435
1	G	0.24	0/1096	0.43	0/1488
1	J	0.23	0/1077	0.43	0/1464
1	M	0.23	0/1072	0.45	0/1460
1	P	0.23	0/1090	0.41	0/1481
2	B	0.24	0/1018	0.46	0/1377
2	E	0.23	0/1040	0.41	0/1410
2	H	0.23	0/978	0.45	0/1330
2	K	0.23	0/913	0.46	0/1235
2	N	0.23	0/1064	0.43	0/1439
2	Q	0.23	0/1061	0.43	0/1435
3	C	0.24	0/1456	0.41	0/1976
3	F	0.23	0/1465	0.43	0/1991
3	I	0.25	0/1419	0.44	0/1937
3	L	0.23	0/1414	0.43	0/1927
3	O	0.24	0/1452	0.42	0/1976
3	R	0.23	0/1426	0.40	0/1936
All	All	0.23	0/21130	0.43	0/28706

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	A	0	1
1	M	0	1
2	E	0	1
2	K	0	2
3	F	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
3	I	0	1
All	All	0	8

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	19	PHE	Peptide
2	E	58	LEU	Peptide
3	F	206	PRO	Peptide
3	F	207	THR	Peptide
3	I	202	LEU	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	1006	0	890	91	0
1	D	1022	0	879	73	0
1	G	1064	0	962	98	0
1	J	1046	0	938	96	0
1	M	1041	0	920	114	0
1	P	1058	0	944	74	0
2	B	987	0	922	81	0
2	E	1006	0	929	88	0
2	H	949	0	847	69	0
2	K	885	0	823	82	0
2	N	1029	0	976	77	0
2	Q	1026	0	967	92	0
3	C	1435	0	1384	117	0
3	F	1443	0	1387	109	0
3	I	1396	0	1294	113	0
3	L	1392	0	1323	108	0
3	O	1430	0	1367	129	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	R	1407	0	1358	100	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	4	0	0	0	0
4	D	3	0	0	0	0
4	E	3	0	0	0	0
4	F	2	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
4	I	2	0	0	0	0
4	J	2	0	0	0	0
4	L	2	0	0	0	0
4	M	1	0	0	0	0
4	N	1	0	0	0	0
4	O	3	0	0	0	0
4	P	1	0	0	0	0
4	Q	2	0	0	0	0
4	R	1	0	0	0	0
5	A	1	0	0	0	0
5	C	4	0	0	2	0
5	D	5	0	0	1	0
5	E	3	0	0	1	0
5	F	1	0	0	1	0
5	G	2	0	0	0	0
5	H	2	0	0	1	0
5	I	3	0	0	2	0
5	J	1	0	0	0	0
5	L	2	0	0	0	0
5	M	2	0	0	2	0
5	N	3	0	0	0	0
5	O	1	0	0	3	0
5	P	2	0	0	0	0
5	Q	2	0	0	0	0
5	R	1	0	0	0	0
6	A	6	0	8	0	0
6	C	6	0	8	0	0
6	G	12	8	16	0	0
6	H	6	0	8	2	0
6	J	12	0	16	11	0
6	P	6	0	8	0	0
7	C	1	0	0	0	0
7	F	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	I	1	0	0	0	0
7	L	1	0	0	0	0
7	O	1	0	0	0	0
7	R	1	0	0	0	0
8	C	5	0	0	0	0
8	D	5	0	0	1	0
8	E	5	0	0	1	0
8	F	10	0	0	1	0
8	I	5	0	0	0	0
8	L	5	0	0	2	0
8	O	10	0	0	5	0
8	R	5	0	0	0	0
9	A	16	0	0	10	0
9	B	7	0	0	5	0
9	C	15	0	0	7	0
9	D	19	0	0	6	0
9	E	9	0	0	6	0
9	F	11	0	0	10	0
9	G	16	0	0	10	0
9	H	7	0	0	5	0
9	I	16	0	0	6	0
9	J	10	0	0	5	0
9	K	8	0	0	11	0
9	L	17	0	0	6	0
9	M	6	0	0	3	0
9	N	7	0	0	5	0
9	O	20	0	0	16	0
9	P	6	0	0	2	0
9	Q	14	0	0	9	0
9	R	14	0	0	7	0
All	All	21010	8	19174	1612	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 40.

The worst 5 of 1612 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:A:19:PHE:HA	1:A:20:ASN:HB3	1.23	1.14
2:B:56:GLN:HG3	2:B:57:THR:HG23	1.35	1.08
1:A:87:MET:HB3	2:B:37:ARG:HD2	1.29	1.07

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:K:92:ARG:NH1	2:K:113:GLU:OE2	1.88	1.06
3:I:199:VAL:O	3:I:201:GLY:N	1.90	1.03

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	125/135 (93%)	106 (85%)	19 (15%)	0	100	100
1	D	126/135 (93%)	108 (86%)	18 (14%)	0	100	100
1	G	130/135 (96%)	118 (91%)	12 (9%)	0	100	100
1	J	129/135 (96%)	114 (88%)	15 (12%)	0	100	100
1	M	130/135 (96%)	113 (87%)	17 (13%)	0	100	100
1	P	130/135 (96%)	122 (94%)	8 (6%)	0	100	100
2	B	113/124 (91%)	103 (91%)	10 (9%)	0	100	100
2	E	120/124 (97%)	111 (92%)	8 (7%)	1 (1%)	19	57
2	H	116/124 (94%)	101 (87%)	15 (13%)	0	100	100
2	K	106/124 (86%)	94 (89%)	12 (11%)	0	100	100
2	N	120/124 (97%)	111 (92%)	9 (8%)	0	100	100
2	Q	120/124 (97%)	109 (91%)	10 (8%)	1 (1%)	19	57
3	C	186/217 (86%)	174 (94%)	11 (6%)	1 (0%)	29	68
3	F	191/217 (88%)	168 (88%)	22 (12%)	1 (0%)	29	68
3	I	189/217 (87%)	167 (88%)	19 (10%)	3 (2%)	9	40
3	L	187/217 (86%)	168 (90%)	18 (10%)	1 (0%)	29	68
3	O	187/217 (86%)	169 (90%)	17 (9%)	1 (0%)	29	68
3	R	179/217 (82%)	166 (93%)	13 (7%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	2584/2856 (90%)	2322 (90%)	253 (10%)	9 (0%)	41 76

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	I	199	VAL
3	I	200	GLN
3	I	304	ILE
2	E	116	VAL
2	Q	116	VAL

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	102/118 (86%)	100 (98%)	2 (2%)	55 83
1	D	103/118 (87%)	102 (99%)	1 (1%)	76 91
1	G	111/118 (94%)	109 (98%)	2 (2%)	59 85
1	J	109/118 (92%)	106 (97%)	3 (3%)	43 77
1	M	107/118 (91%)	99 (92%)	8 (8%)	13 43
1	P	110/118 (93%)	105 (96%)	5 (4%)	27 64
2	B	102/112 (91%)	98 (96%)	4 (4%)	32 69
2	E	105/112 (94%)	98 (93%)	7 (7%)	16 49
2	H	94/112 (84%)	90 (96%)	4 (4%)	29 66
2	K	91/112 (81%)	87 (96%)	4 (4%)	28 65
2	N	110/112 (98%)	108 (98%)	2 (2%)	59 85
2	Q	109/112 (97%)	107 (98%)	2 (2%)	59 85
3	C	149/181 (82%)	146 (98%)	3 (2%)	55 83
3	F	149/181 (82%)	144 (97%)	5 (3%)	37 72
3	I	138/181 (76%)	135 (98%)	3 (2%)	52 81
3	L	140/181 (77%)	139 (99%)	1 (1%)	84 94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	O	148/181 (82%)	145 (98%)	3 (2%)	55 83
3	R	146/181 (81%)	145 (99%)	1 (1%)	84 94
All	All	2123/2466 (86%)	2063 (97%)	60 (3%)	43 77

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	305	LEU
1	P	47	ASP
2	K	44	ARG
1	P	15	CYS
3	R	306	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	Q	33	HIS
3	O	200	GLN
3	I	244	GLN
1	M	13	GLN
3	F	257	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 90 ligands modelled in this entry, 72 are monoatomic - leaving 18 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	GOL	H	204	-	5,5,5	0.37	0	5,5,5	0.29	0
6	GOL	G	205	-	5,5,5	0.36	0	5,5,5	0.21	0
8	SO4	F	406	-	4,4,4	0.14	0	6,6,6	0.05	0
8	SO4	D	208	-	4,4,4	0.14	0	6,6,6	0.05	0
8	SO4	F	401	7	4,4,4	0.14	0	6,6,6	0.04	0
6	GOL	J	204	-	5,5,5	0.38	0	5,5,5	0.35	0
8	SO4	O	407	-	4,4,4	0.14	0	6,6,6	0.05	0
8	SO4	I	407	7	4,4,4	0.14	0	6,6,6	0.04	0
8	SO4	L	406	-	4,4,4	0.14	0	6,6,6	0.05	0
8	SO4	E	207	4	4,4,4	0.14	0	6,6,6	0.05	0
8	SO4	C	409	-	4,4,4	0.14	0	6,6,6	0.05	0
6	GOL	P	204	-	5,5,5	0.36	0	5,5,5	0.27	0
8	SO4	R	404	7	4,4,4	0.14	0	6,6,6	0.05	0
6	GOL	J	205	-	5,5,5	0.38	0	5,5,5	0.23	0
8	SO4	O	402	4	4,4,4	0.15	0	6,6,6	0.05	0
6	GOL	C	410	-	5,5,5	0.37	0	5,5,5	0.27	0
6	GOL	G	204	-	5,5,5	0.36	0	5,5,5	0.28	0
6	GOL	A	203	-	5,5,5	0.37	0	5,5,5	0.27	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	GOL	H	204	-	-	2/4/4/4	-
6	GOL	G	205	-	-	1/4/4/4	-
6	GOL	J	204	-	-	2/4/4/4	-
6	GOL	P	204	-	-	2/4/4/4	-
6	GOL	J	205	-	-	2/4/4/4	-
6	GOL	C	410	-	-	2/4/4/4	-
6	GOL	G	204	-	-	2/4/4/4	-
6	GOL	A	203	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	203	GOL	O1-C1-C2-C3
6	C	410	GOL	O1-C1-C2-C3
6	H	204	GOL	O1-C1-C2-C3
6	J	204	GOL	O1-C1-C2-C3
6	G	204	GOL	O1-C1-C2-C3

There are no ring outliers.

9 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	H	204	GOL	2	0
8	D	208	SO4	1	0
8	F	401	SO4	1	0
6	J	204	GOL	9	0
8	O	407	SO4	4	0
8	L	406	SO4	2	0
8	E	207	SO4	1	0
6	J	205	GOL	2	0
8	O	402	SO4	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	129/135 (95%)	-0.12	1 (0%) 86 65	66, 99, 147, 169	0
1	D	130/135 (96%)	-0.26	1 (0%) 86 65	55, 84, 129, 169	0
1	G	132/135 (97%)	-0.32	0 100 100	49, 78, 124, 149	0
1	J	131/135 (97%)	-0.31	0 100 100	59, 88, 114, 128	0
1	M	132/135 (97%)	0.07	5 (3%) 40 16	66, 113, 179, 221	0
1	P	132/135 (97%)	-0.13	1 (0%) 86 65	60, 96, 157, 213	0
2	B	119/124 (95%)	-0.28	1 (0%) 86 65	61, 89, 131, 204	0
2	E	122/124 (98%)	-0.16	2 (1%) 72 44	64, 88, 131, 215	0
2	H	120/124 (96%)	-0.08	1 (0%) 86 65	63, 93, 155, 173	0
2	K	108/124 (87%)	0.08	5 (4%) 32 12	73, 106, 185, 255	0
2	N	122/124 (98%)	-0.32	0 100 100	53, 82, 118, 154	0
2	Q	122/124 (98%)	-0.01	5 (4%) 37 14	66, 99, 163, 248	0
3	C	190/217 (87%)	-0.09	5 (2%) 56 27	57, 94, 147, 199	0
3	F	193/217 (88%)	-0.18	0 100 100	60, 92, 127, 151	0
3	I	191/217 (88%)	-0.26	2 (1%) 82 59	60, 87, 139, 187	0
3	L	189/217 (87%)	-0.14	4 (2%) 63 34	49, 87, 149, 228	0
3	O	191/217 (88%)	-0.15	2 (1%) 82 59	55, 84, 130, 196	0
3	R	187/217 (86%)	-0.18	3 (1%) 72 44	52, 80, 122, 156	0
All	All	2640/2856 (92%)	-0.16	38 (1%) 75 49	49, 90, 149, 255	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	K	55	SER	10.7
2	Q	57	THR	5.4
3	L	308	GLY	5.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	57	THR	4.6
1	A	91	ILE	4.5

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
4	NA	F	403	1/1	0.62	0.25	102,102,102,102	0
7	MG	O	405	1/1	0.62	0.45	226,226,226,226	0
4	NA	E	201	1/1	0.64	0.14	103,103,103,103	0
4	NA	Q	201	1/1	0.67	0.62	95,95,95,95	0
4	NA	A	201	1/1	0.73	0.29	87,87,87,87	0
4	NA	P	201	1/1	0.74	0.18	55,55,55,55	0
5	CL	I	404	1/1	0.79	0.47	137,137,137,137	0
7	MG	L	403	1/1	0.79	0.43	169,169,169,169	0
4	NA	J	201	1/1	0.79	0.16	72,72,72,72	0
5	CL	H	203	1/1	0.82	0.70	165,165,165,165	0
4	NA	D	202	1/1	0.83	0.10	52,52,52,52	0
4	NA	D	209	1/1	0.84	0.37	158,158,158,158	0
4	NA	C	403	1/1	0.84	0.18	67,67,67,67	0
7	MG	F	404	1/1	0.85	0.31	103,103,103,103	0
5	CL	N	202	1/1	0.85	0.15	95,95,95,95	0
6	GOL	G	205	6/6	0.85	0.14	73,91,109,109	0
5	CL	G	203	1/1	0.88	0.18	123,123,123,123	0
4	NA	N	201	1/1	0.88	0.12	81,81,81,81	0
5	CL	L	404	1/1	0.89	0.10	100,100,100,100	0
4	NA	Q	202	1/1	0.90	0.13	40,40,40,40	0
6	GOL	J	205	6/6	0.90	0.14	67,72,82,83	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	CL	C	408	1/1	0.90	0.12	76,76,76,76	0
7	MG	I	403	1/1	0.90	0.13	51,51,51,51	0
5	CL	G	202	1/1	0.90	0.12	91,91,91,91	0
4	NA	B	201	1/1	0.90	0.10	62,62,62,62	0
8	SO4	I	407	5/5	0.90	0.18	101,107,112,115	0
5	CL	N	204	1/1	0.91	0.07	99,99,99,99	0
4	NA	O	404	1/1	0.91	0.16	74,74,74,74	0
4	NA	I	401	1/1	0.91	0.33	115,115,115,115	0
4	NA	E	203	1/1	0.91	0.12	62,62,62,62	0
4	NA	M	201	1/1	0.91	0.08	66,66,66,66	0
4	NA	C	411	1/1	0.91	0.24	117,117,117,117	0
5	CL	E	204	1/1	0.91	0.10	82,82,82,82	0
5	CL	N	203	1/1	0.91	0.17	98,98,98,98	0
5	CL	Q	204	1/1	0.92	0.14	89,89,89,89	0
5	CL	C	405	1/1	0.92	0.26	116,116,116,116	0
5	CL	I	405	1/1	0.92	0.12	72,72,72,72	0
4	NA	G	201	1/1	0.92	0.13	54,54,54,54	0
4	NA	E	202	1/1	0.93	0.10	68,68,68,68	0
4	NA	J	202	1/1	0.93	0.25	60,60,60,60	0
5	CL	D	203	1/1	0.93	0.16	83,83,83,83	0
8	SO4	R	404	5/5	0.93	0.15	95,104,116,118	0
4	NA	O	401	1/1	0.94	0.10	76,76,76,76	0
5	CL	R	403	1/1	0.94	0.09	78,78,78,78	0
5	CL	H	202	1/1	0.94	0.15	77,77,77,77	0
8	SO4	C	409	5/5	0.94	0.15	105,106,113,120	0
8	SO4	E	207	5/5	0.94	0.10	119,123,130,133	0
8	SO4	F	401	5/5	0.94	0.15	91,91,97,121	0
5	CL	L	405	1/1	0.94	0.11	86,86,86,86	0
8	SO4	O	402	5/5	0.94	0.18	108,118,125,130	0
8	SO4	O	407	5/5	0.94	0.11	93,95,103,107	0
5	CL	O	406	1/1	0.94	0.08	69,69,69,69	0
5	CL	Q	203	1/1	0.95	0.07	88,88,88,88	0
5	CL	C	406	1/1	0.95	0.11	82,82,82,82	0
5	CL	C	407	1/1	0.95	0.16	84,84,84,84	0
6	GOL	A	203	6/6	0.95	0.15	55,71,77,78	0
4	NA	H	201	1/1	0.95	0.23	62,62,62,62	0
6	GOL	H	204	6/6	0.95	0.14	47,61,78,89	0
8	SO4	L	406	5/5	0.95	0.13	103,108,109,118	0
5	CL	J	203	1/1	0.95	0.10	78,78,78,78	0
4	NA	F	402	1/1	0.95	0.14	37,37,37,37	0
5	CL	P	202	1/1	0.95	0.09	77,77,77,77	0
4	NA	D	201	1/1	0.96	0.09	90,90,90,90	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NA	I	402	1/1	0.96	0.11	65,65,65,65	0
6	GOL	P	204	6/6	0.96	0.18	69,78,86,87	0
5	CL	I	406	1/1	0.96	0.09	81,81,81,81	0
4	NA	C	401	1/1	0.96	0.13	78,78,78,78	0
5	CL	D	205	1/1	0.97	0.11	80,80,80,80	0
5	CL	M	203	1/1	0.97	0.04	92,92,92,92	0
5	CL	D	207	1/1	0.97	0.11	71,71,71,71	0
6	GOL	C	410	6/6	0.97	0.13	43,66,73,81	0
6	GOL	G	204	6/6	0.97	0.15	65,70,77,87	0
5	CL	A	202	1/1	0.97	0.06	89,89,89,89	0
5	CL	E	205	1/1	0.97	0.09	70,70,70,70	0
6	GOL	J	204	6/6	0.97	0.30	52,66,83,88	0
5	CL	E	206	1/1	0.97	0.08	90,90,90,90	0
4	NA	L	402	1/1	0.97	0.18	54,54,54,54	0
4	NA	L	401	1/1	0.97	0.24	74,74,74,74	0
5	CL	D	204	1/1	0.98	0.10	82,82,82,82	0
7	MG	R	402	1/1	0.98	0.20	47,47,47,47	0
5	CL	P	203	1/1	0.98	0.13	70,70,70,70	0
5	CL	M	202	1/1	0.98	0.11	77,77,77,77	0
4	NA	C	402	1/1	0.98	0.09	61,61,61,61	0
5	CL	F	405	1/1	0.98	0.17	83,83,83,83	0
7	MG	C	404	1/1	0.98	0.19	84,84,84,84	0
5	CL	D	206	1/1	0.98	0.15	68,68,68,68	0
4	NA	O	403	1/1	0.98	0.12	60,60,60,60	0
4	NA	R	401	1/1	0.98	0.10	46,46,46,46	0
8	SO4	F	406	5/5	0.99	0.14	77,79,83,100	0
8	SO4	D	208	5/5	0.99	0.19	55,55,62,85	0

6.5 Other polymers (i)

There are no such residues in this entry.