



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 30, 2017 – 01:29 PM EST

PDB ID : 4RV8  
Title : Co-Crystal Structure of the Catalytic Domain of the Inosine Monophosphate Dehydrogenase from *Cryptosporidium parvum* and the inhibitor p131  
Authors : Kim, Y.; Makowska-Grzyska, M.; Gu, M.; Kavitha, M.; Hedstrom, L.; Anderson, W.F.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2014-11-25  
Resolution : 2.05 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-20028442
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-20028442

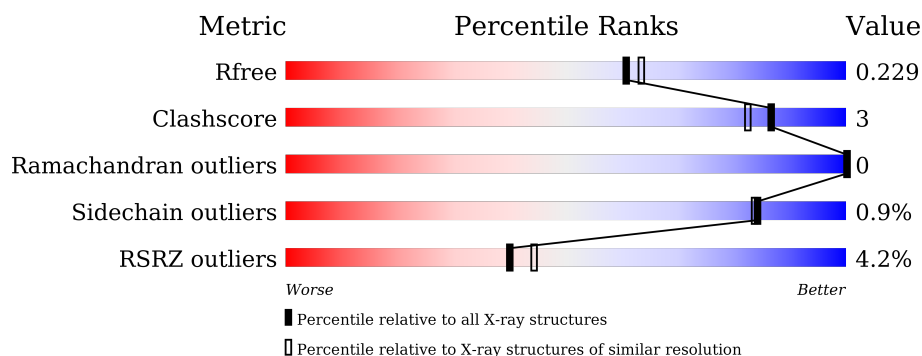
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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	1799 (2.08-2.04)
Clashscore	102246	1910 (2.08-2.04)
Ramachandran outliers	100387	1893 (2.08-2.04)
Sidechain outliers	100360	1893 (2.08-2.04)
RSRZ outliers	91569	1802 (2.08-2.04)

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	361	<div> <div>2%</div> <div>86% 5% 9%</div> </div>
1	B	361	<div> <div>5%</div> <div>85% 6% 9%</div> </div>
1	C	361	<div> <div>4%</div> <div>83% 7% 10%</div> </div>
1	D	361	<div> <div>4%</div> <div>83% 7% 9%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	I13	A	501	-	-	-	X
4	GOL	A	502	-	-	-	X
4	GOL	B	503	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10414 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 Inosine-5'-monophosphate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	327	Total	C	N	O	S	0	5	0
			2459	1546	417	480	16			
1	B	327	Total	C	N	O	S	0	6	0
			2469	1552	418	482	17			
1	C	326	Total	C	N	O	S	0	3	0
			2436	1532	413	475	16			
1	D	327	Total	C	N	O	S	0	5	0
			2461	1546	417	482	16			

There are 24 discrepancies between the modelled and reference sequences:

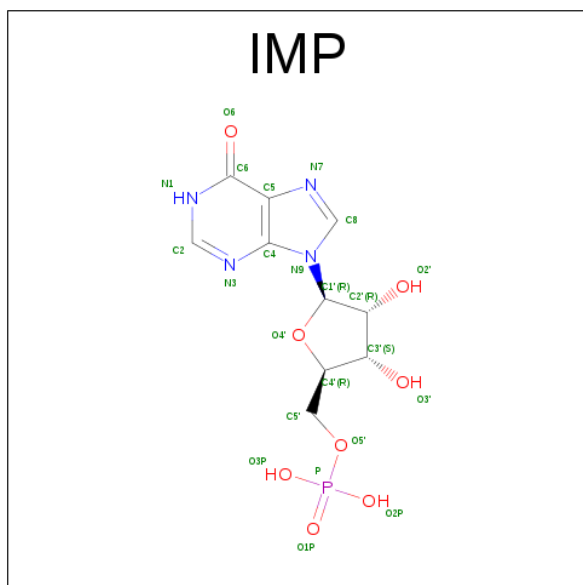
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q8T6T2
A	-1	ASN	-	EXPRESSION TAG	UNP Q8T6T2
A	0	ALA	-	EXPRESSION TAG	UNP Q8T6T2
A	90	SER	-	LINKER	UNP Q8T6T2
A	91	GLY	-	LINKER	UNP Q8T6T2
A	92	GLY	-	LINKER	UNP Q8T6T2
B	-2	SER	-	EXPRESSION TAG	UNP Q8T6T2
B	-1	ASN	-	EXPRESSION TAG	UNP Q8T6T2
B	0	ALA	-	EXPRESSION TAG	UNP Q8T6T2
B	90	SER	-	LINKER	UNP Q8T6T2
B	91	GLY	-	LINKER	UNP Q8T6T2
B	92	GLY	-	LINKER	UNP Q8T6T2
C	-2	SER	-	EXPRESSION TAG	UNP Q8T6T2
C	-1	ASN	-	EXPRESSION TAG	UNP Q8T6T2
C	0	ALA	-	EXPRESSION TAG	UNP Q8T6T2
C	90	SER	-	LINKER	UNP Q8T6T2
C	91	GLY	-	LINKER	UNP Q8T6T2
C	92	GLY	-	LINKER	UNP Q8T6T2
D	-2	SER	-	EXPRESSION TAG	UNP Q8T6T2
D	-1	ASN	-	EXPRESSION TAG	UNP Q8T6T2
D	0	ALA	-	EXPRESSION TAG	UNP Q8T6T2

*Continued on next page...*

Continued from previous page...

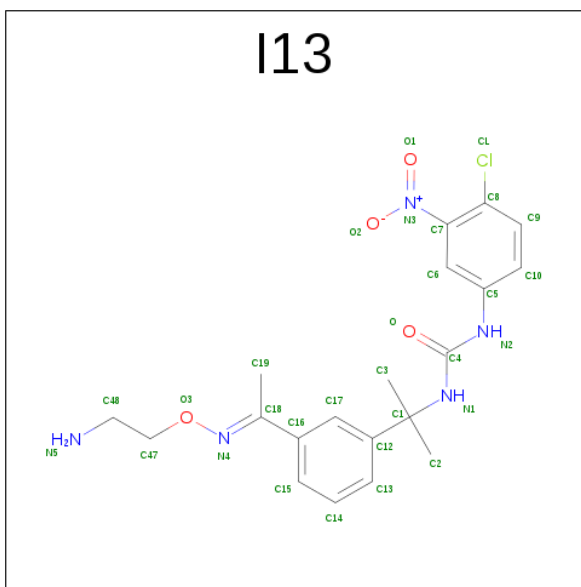
Chain	Residue	Modelled	Actual	Comment	Reference
D	90	SER	-	LINKER	UNP Q8T6T2
D	91	GLY	-	LINKER	UNP Q8T6T2
D	92	GLY	-	LINKER	UNP Q8T6T2

- Molecule 2 is INOSINIC ACID (three-letter code: IMP) (formula:  $C_{10}H_{13}N_4O_8P$ ).



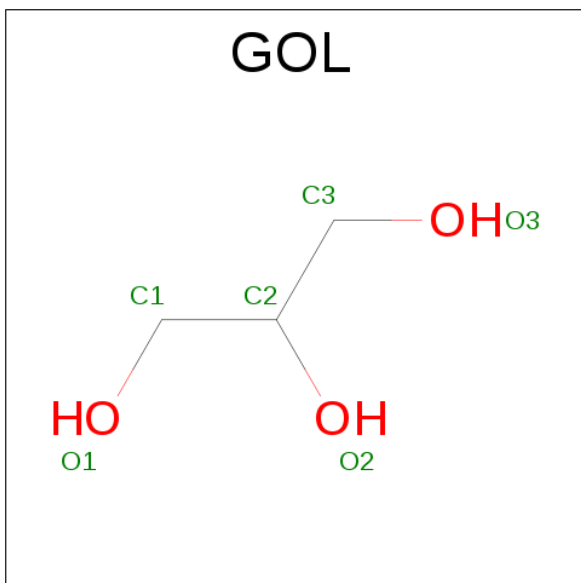
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			23	10	4	8	1		
2	B	1	Total	C	N	O	P	0	0
			23	10	4	8	1		
2	C	1	Total	C	N	O	P	0	0
			23	10	4	8	1		
2	D	1	Total	C	N	O	P	0	0
			23	10	4	8	1		

- Molecule 3 is 1-(2-{3-[(1E)-N-(2-AMINOETHOXY)ETHANIMIDOYL]PHENYL}PROPAN-2-YL)-3-(4-CHLORO-3-NITROPHENYL)UREA (three-letter code: I13) (formula:  $C_{20}H_{24}ClN_5O_4$ ).



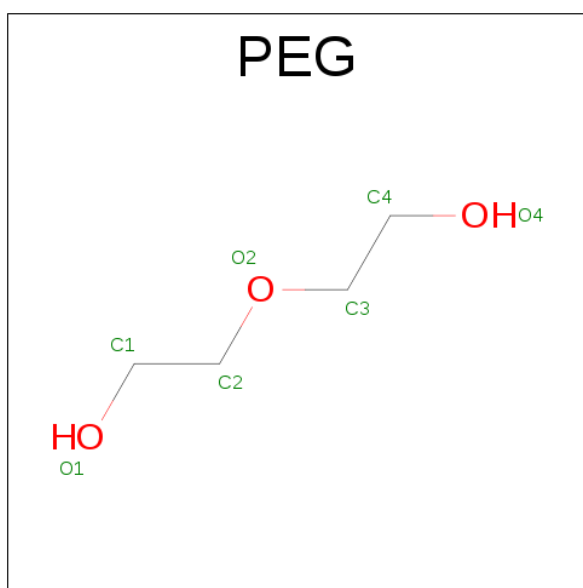
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 30	C 20	Cl 1	N 5	O 4	0	0
3	B	1	Total 30	C 20	Cl 1	N 5	O 4	0	0
3	B	1	Total 30	C 20	Cl 1	N 5	O 4	0	0
3	C	1	Total 30	C 20	Cl 1	N 5	O 4	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			7	4	3		

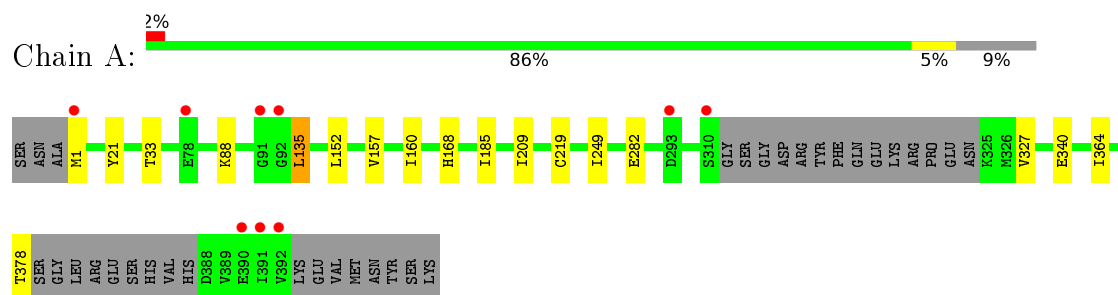
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	95	Total	O	0	0
			95	95		
6	B	81	Total	O	0	0
			81	81		
6	C	84	Total	O	0	0
			84	84		
6	D	79	Total	O	0	0
			79	79		

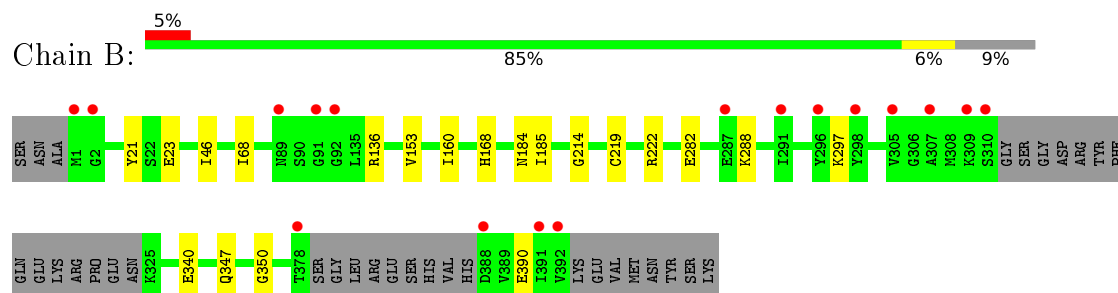
### 3 Residue-property plots [i](#)

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.

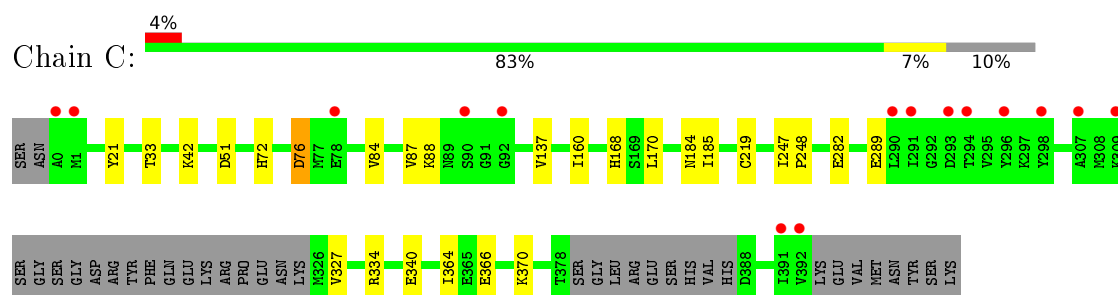
- Molecule 1: Inosine-5'-monophosphate dehydrogenase



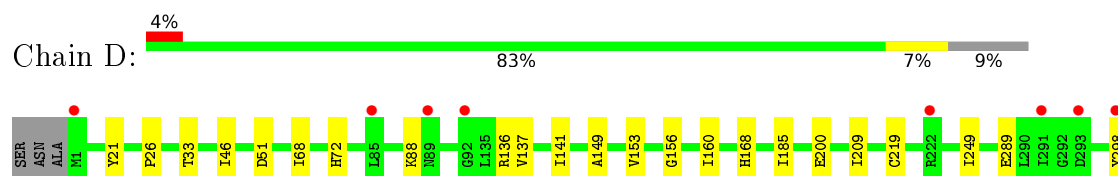
- Molecule 1: Inosine-5'-monophosphate dehydrogenase



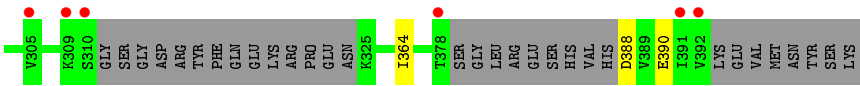
- Molecule 1: Inosine-5'-monophosphate dehydrogenase



- Molecule 1: Inosine-5'-monophosphate dehydrogenase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.19Å 91.83Å 92.02Å 90.00° 103.75° 90.00°	Depositor
Resolution (Å)	40.84 – 2.05 40.84 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.6 (40.84-2.05) 99.6 (40.84-2.05)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.14	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.05Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1745)	Depositor
R, $R_{free}$	0.187 , 0.229 0.187 , 0.229	Depositor DCC
$R_{free}$ test set	4498 reflections (5.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	27.5	Xtriage
Anisotropy	0.524	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 49.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.011 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10414	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.41% 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, 0.333 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: GOL, OCS, PEG, I13, IMP

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	0.25	0/2471	0.45	0/3321
1	B	0.25	0/2481	0.45	0/3332
1	C	0.25	0/2448	0.45	0/3291
1	D	0.25	0/2473	0.45	0/3323
All	All	0.25	0/9873	0.45	0/13267

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	2459	0	2566	11	0
1	B	2469	0	2575	13	0
1	C	2436	0	2541	15	0
1	D	2461	0	2563	17	0
2	A	23	0	11	1	0
2	B	23	0	11	1	0
2	C	23	0	11	1	0
2	D	23	0	11	1	0
3	A	30	0	24	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	60	0	48	3	0
3	C	30	0	24	1	0
4	A	6	0	8	0	0
4	B	6	0	8	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
5	A	7	0	10	2	0
5	B	7	0	10	0	0
6	A	95	0	0	0	0
6	B	81	0	0	1	0
6	C	84	0	0	0	0
6	D	79	0	0	1	0
All	All	10414	0	10437	56	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 3.

All (56) 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:160:ILE:HG12	1:B:185:ILE:HD13	1.72	0.71
1:D:160:ILE:HG12	1:D:185:ILE:HD13	1.74	0.69
3:C:503:I13:CL	3:C:503:I13:O2	2.49	0.68
3:B:505:I13:CL	3:B:505:I13:O2	2.51	0.66
1:B:219:OCS:OD3	2:B:501:IMP:H2	1.99	0.63
1:D:219:OCS:OD1	2:D:501:IMP:H2	2.00	0.61
1:C:289:GLU:OE2	1:C:334:ARG:NH1	2.33	0.61
3:A:501:I13:O2	3:A:501:I13:CL	2.56	0.60
1:B:153:VAL:HG22	1:B:185:ILE:HD12	1.84	0.59
1:C:366:GLU:HG2	1:C:370:LYS:HD2	1.85	0.58
1:D:209:ILE:HB	1:D:249:ILE:HD13	1.87	0.57
5:A:503:PEG:H11	1:D:26:PRO:HG2	1.87	0.56
1:D:289:GLU:HB2	1:D:298:TYR:HE1	1.71	0.56
1:B:23:GLU:HG2	1:C:170:LEU:HB2	1.88	0.56
1:A:219:OCS:OD1	2:A:500:IMP:H2	2.05	0.55
1:D:388:ASP:OD2	1:D:388:ASP:N	2.41	0.54
1:B:282[A]:GLU:HG2	1:B:340:GLU:HA	1.89	0.53
1:B:136:ARG:NH2	6:B:643:HOH:O	2.42	0.52
1:A:152:LEU:O	1:A:157[B]:VAL:HG23	2.12	0.49
1:C:87:VAL:HG11	1:C:137:VAL:HB	1.94	0.49
1:C:282:GLU:HG2	1:C:340:GLU:HA	1.95	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:153:VAL:HG22	1:D:185:ILE:HD12	1.95	0.48
1:C:219:OCS:OD2	2:C:501:IMP:H2	2.14	0.48
1:D:168:HIS:NE2	1:D:200:GLU:OE2	2.46	0.47
1:B:350:GLY:HA3	1:C:327:VAL:HG21	1.96	0.47
1:D:141:ILE:HD13	1:D:149:ALA:HB2	1.96	0.46
1:A:327:VAL:HB	5:A:503:PEG:H31	1.97	0.46
1:C:160:ILE:HG12	1:C:185:ILE:HG12	1.97	0.46
1:B:214:GLY:CA	1:B:222[B]:ARG:HE	2.29	0.46
1:D:88:LYS:HE3	1:D:156:GLY:O	2.16	0.46
1:D:88:LYS:HE2	1:D:137:VAL:HG12	1.97	0.45
1:A:88:LYS:HA	1:A:88:LYS:HD3	1.82	0.45
1:D:51:ASP:HA	1:D:72:HIS:CD2	2.51	0.45
3:B:502:I13:O	3:B:502:I13:H7	2.17	0.45
1:A:21:TYR:HB2	1:B:168:HIS:CE1	2.52	0.44
1:C:84:VAL:O	1:C:88:LYS:HG2	2.17	0.44
1:B:214:GLY:HA3	1:B:222[B]:ARG:HE	1.83	0.44
1:B:288:LYS:HG3	1:B:297:LYS:HG2	2.00	0.44
1:D:136:ARG:NH2	6:D:634:HOH:O	2.50	0.44
1:C:76:ASP:OD2	1:C:76:ASP:N	2.42	0.43
1:A:168:HIS:CE1	1:D:21:TYR:HB2	2.54	0.43
3:A:501:I13:O	3:A:501:I13:H7	2.18	0.43
1:A:282:GLU:HG2	1:A:340:GLU:HA	2.00	0.43
1:C:21:TYR:HB2	1:D:168:HIS:CE1	2.54	0.42
1:C:247:ILE:HA	1:C:248:PRO:HD3	1.91	0.42
1:C:51:ASP:HA	1:C:72:HIS:CD2	2.55	0.41
1:B:46:ILE:HG23	1:B:68:ILE:HG23	2.02	0.41
1:A:160:ILE:HG12	1:A:185:ILE:HG12	2.03	0.41
1:D:46:ILE:HG23	1:D:68:ILE:HG23	2.03	0.41
1:A:209:ILE:HB	1:A:249:ILE:HD13	2.02	0.41
1:A:33:THR:HB	1:A:364:ILE:HD12	2.02	0.41
1:C:33:THR:HB	1:C:364:ILE:HD12	2.02	0.40
1:D:33:THR:HB	1:D:364:ILE:HD12	2.02	0.40
1:A:135:LEU:HA	1:A:135:LEU:HD12	1.94	0.40
3:B:505:I13:O	3:B:505:I13:H7	2.22	0.40
1:B:21:TYR:HB2	1:C:168:HIS:CE1	2.56	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	325/361 (90%)	317 (98%)	8 (2%)	0	100	100
1	B	326/361 (90%)	317 (97%)	9 (3%)	0	100	100
1	C	322/361 (89%)	314 (98%)	8 (2%)	0	100	100
1	D	325/361 (90%)	315 (97%)	10 (3%)	0	100	100
All	All	1298/1444 (90%)	1263 (97%)	35 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	269/294 (92%)	266 (99%)	3 (1%)	80	79
1	B	270/294 (92%)	267 (99%)	3 (1%)	80	79
1	C	265/294 (90%)	262 (99%)	3 (1%)	80	79
1	D	269/294 (92%)	268 (100%)	1 (0%)	93	94
All	All	1073/1176 (91%)	1063 (99%)	10 (1%)	84	84

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	135	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	378	THR
1	B	184	ASN
1	B	347	GLN
1	B	390	GLU
1	C	42	LYS
1	C	76	ASP
1	C	184	ASN
1	D	390	GLU

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

Mol	Chain	Res	Type
1	B	347	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	OCS	A	219	1	6,8,9	1.30	1 (16%)	7,11,13	3.04	3 (42%)
1	OCS	B	219	1	6,8,9	1.37	1 (16%)	7,11,13	3.58	3 (42%)
1	OCS	C	219	1	6,8,9	1.33	1 (16%)	7,11,13	3.33	1 (14%)
1	OCS	D	219	1	6,8,9	1.30	1 (16%)	7,11,13	1.99	3 (42%)

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
1	OCS	A	219	1	-	1/4/7/9	0/0/0/0
1	OCS	B	219	1	-	1/4/7/9	0/0/0/0
1	OCS	C	219	1	-	1/4/7/9	0/0/0/0
1	OCS	D	219	1	-	1/4/7/9	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	219	OCS	CB-SG	2.59	1.80	1.77
1	D	219	OCS	CB-SG	2.63	1.80	1.77
1	A	219	OCS	CB-SG	2.67	1.80	1.77
1	B	219	OCS	CB-SG	2.71	1.80	1.77

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	219	OCS	OD1-SG-CB	-5.68	102.92	106.92
1	B	219	OCS	OD3-SG-CB	-3.26	104.63	106.92
1	B	219	OCS	OD2-SG-OD3	-3.05	104.52	111.26
1	A	219	OCS	OD3-SG-OD1	-2.53	106.81	113.96
1	D	219	OCS	OD1-SG-CB	-2.48	105.18	106.92
1	D	219	OCS	OD3-SG-OD1	-2.42	107.14	113.96
1	D	219	OCS	OD2-SG-CB	3.23	111.71	104.99
1	A	219	OCS	OD2-SG-CB	4.52	114.38	104.99
1	B	219	OCS	OD1-SG-CB	7.80	112.41	106.92
1	C	219	OCS	OD1-SG-CB	8.38	112.82	106.92

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	219	OCS	SG-CB-CA-N
1	A	219	OCS	SG-CB-CA-N
1	B	219	OCS	SG-CB-CA-N
1	D	219	OCS	SG-CB-CA-N

There are no ring outliers.

4 monomers are involved in 4 short contacts:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	219	OCS	1	0
1	B	219	OCS	1	0
1	C	219	OCS	1	0
1	D	219	OCS	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

14 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	IMP	A	500	-	21,25,25	1.20	3 (14%)	22,38,38	2.65	4 (18%)
3	I13	A	501	-	28,31,31	2.44	5 (17%)	36,43,43	2.46	3 (8%)
4	GOL	A	502	-	5,5,5	0.36	0	5,5,5	0.17	0
5	PEG	A	503	-	6,6,6	0.38	0	5,5,5	0.33	0
2	IMP	B	501	-	21,25,25	1.19	3 (14%)	22,38,38	2.45	4 (18%)
3	I13	B	502	-	28,31,31	2.48	5 (17%)	36,43,43	2.33	3 (8%)
4	GOL	B	503	-	5,5,5	0.38	0	5,5,5	0.27	0
5	PEG	B	504	-	6,6,6	0.46	0	5,5,5	0.35	0
3	I13	B	505	-	28,31,31	2.47	6 (21%)	36,43,43	2.34	3 (8%)
2	IMP	C	501	-	21,25,25	1.17	3 (14%)	22,38,38	2.61	3 (13%)
4	GOL	C	502	-	5,5,5	0.38	0	5,5,5	0.21	0
3	I13	C	503	-	28,31,31	2.47	5 (17%)	36,43,43	2.32	3 (8%)
2	IMP	D	501	-	21,25,25	1.14	3 (14%)	22,38,38	2.48	4 (18%)
4	GOL	D	502	-	5,5,5	0.33	0	5,5,5	0.30	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
2	IMP	A	500	-	-	0/6/26/26	0/3/3/3
3	I13	A	501	-	-	0/27/28/28	0/2/2/2
4	GOL	A	502	-	-	0/4/4/4	0/0/0/0
5	PEG	A	503	-	-	0/4/4/4	0/0/0/0
2	IMP	B	501	-	-	0/6/26/26	0/3/3/3
3	I13	B	502	-	-	0/27/28/28	0/2/2/2
4	GOL	B	503	-	-	0/4/4/4	0/0/0/0
5	PEG	B	504	-	-	0/4/4/4	0/0/0/0
3	I13	B	505	-	-	0/27/28/28	0/2/2/2
2	IMP	C	501	-	-	0/6/26/26	0/3/3/3
4	GOL	C	502	-	-	0/4/4/4	0/0/0/0
3	I13	C	503	-	-	0/27/28/28	0/2/2/2
2	IMP	D	501	-	-	0/6/26/26	0/3/3/3
4	GOL	D	502	-	-	0/4/4/4	0/0/0/0

All (33) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	505	I13	O3-N4	-8.20	1.25	1.42
3	B	502	I13	O3-N4	-8.18	1.26	1.42
3	A	501	I13	O3-N4	-8.13	1.26	1.42
3	C	503	I13	O3-N4	-8.09	1.26	1.42
3	B	502	I13	C5-N2	-2.93	1.36	1.41
3	C	503	I13	C5-N2	-2.85	1.36	1.41
3	B	505	I13	C5-N2	-2.80	1.36	1.41
3	A	501	I13	C5-N2	-2.70	1.36	1.41
3	C	503	I13	C18-N4	-2.29	1.26	1.28
3	B	502	I13	C18-N4	-2.15	1.26	1.28
3	B	505	I13	C18-N4	-2.14	1.26	1.28
3	B	505	I13	C1-C12	-2.06	1.50	1.53
3	A	501	I13	C18-N4	-2.00	1.26	1.28
2	C	501	IMP	C2-N1	2.30	1.38	1.33
2	A	500	IMP	C2-N1	2.31	1.38	1.33
2	B	501	IMP	C2-N1	2.31	1.38	1.33
3	B	502	I13	C8-CL	2.32	1.79	1.73
2	D	501	IMP	C2-N1	2.35	1.38	1.33
3	C	503	I13	C8-CL	2.39	1.79	1.73
3	A	501	I13	C8-CL	2.39	1.79	1.73
3	B	505	I13	C8-CL	2.48	1.80	1.73
2	D	501	IMP	C6-N1	2.80	1.38	1.33
2	C	501	IMP	C6-N1	2.88	1.38	1.33
2	B	501	IMP	C6-N1	2.93	1.38	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	500	IMP	C6-N1	3.02	1.38	1.33
2	D	501	IMP	C2-N3	3.21	1.37	1.32
2	C	501	IMP	C2-N3	3.35	1.38	1.32
2	A	500	IMP	C2-N3	3.45	1.38	1.32
2	B	501	IMP	C2-N3	3.47	1.38	1.32
3	A	501	I13	O1-N3	8.58	1.39	1.22
3	C	503	I13	O1-N3	8.58	1.39	1.22
3	B	505	I13	O1-N3	8.60	1.39	1.22
3	B	502	I13	O1-N3	8.68	1.39	1.22

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	500	IMP	N3-C2-N1	-11.17	120.09	128.87
2	C	501	IMP	N3-C2-N1	-10.90	120.31	128.87
2	B	501	IMP	N3-C2-N1	-10.34	120.75	128.87
2	D	501	IMP	N3-C2-N1	-10.19	120.87	128.87
2	C	501	IMP	C1'-N9-C4	-3.49	122.91	126.81
2	D	501	IMP	C1'-N9-C4	-3.43	122.97	126.81
2	A	500	IMP	C1'-N9-C4	-3.05	123.40	126.81
2	B	501	IMP	C1'-N9-C4	-2.93	123.53	126.81
3	A	501	I13	C19-C18-N4	-2.82	118.12	123.71
3	C	503	I13	C19-C18-N4	-2.72	118.32	123.71
3	B	502	I13	C19-C18-N4	-2.60	118.56	123.71
3	B	505	I13	C19-C18-N4	-2.54	118.68	123.71
2	B	501	IMP	C2-N1-C6	2.24	119.98	116.13
2	D	501	IMP	C2-N1-C6	2.37	120.22	116.13
2	B	501	IMP	O2P-P-O1P	2.40	118.46	110.63
2	A	500	IMP	C2-N1-C6	2.44	120.34	116.13
2	C	501	IMP	C2-N1-C6	2.58	120.58	116.13
2	A	500	IMP	O3P-P-O2P	2.60	116.98	107.44
2	D	501	IMP	O3P-P-O2P	2.83	117.82	107.44
3	A	501	I13	O3-N4-C18	4.07	120.53	111.75
3	C	503	I13	O3-N4-C18	4.39	121.22	111.75
3	B	505	I13	O3-N4-C18	4.55	121.56	111.75
3	B	502	I13	O3-N4-C18	4.55	121.57	111.75
3	C	503	I13	C47-O3-N4	11.90	119.27	108.61
3	B	502	I13	C47-O3-N4	12.05	119.40	108.61
3	B	505	I13	C47-O3-N4	12.05	119.40	108.61
3	A	501	I13	C47-O3-N4	12.88	120.15	108.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	500	IMP	1	0
3	A	501	I13	2	0
5	A	503	PEG	2	0
2	B	501	IMP	1	0
3	B	502	I13	1	0
3	B	505	I13	2	0
2	C	501	IMP	1	0
3	C	503	I13	1	0
2	D	501	IMP	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<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	326/361 (90%)	-0.02	9 (2%) 56 62	18, 32, 59, 93	0
1	B	326/361 (90%)	0.05	17 (5%) 31 34	22, 32, 59, 82	0
1	C	325/361 (90%)	0.19	15 (4%) 36 40	20, 33, 60, 84	0
1	D	326/361 (90%)	-0.03	14 (4%) 39 43	18, 31, 58, 79	0
All	All	1303/1444 (90%)	0.05	55 (4%) 40 44	18, 32, 60, 93	0

All (55) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	0	ALA	9.3
1	D	1	MET	5.1
1	D	310	SER	4.7
1	C	391	ILE	4.5
1	C	291	ILE	4.1
1	D	291	ILE	4.0
1	D	92	GLY	3.8
1	B	310	SER	3.6
1	A	1	MET	3.5
1	A	391	ILE	3.4
1	A	92	GLY	3.4
1	C	92	GLY	3.2
1	B	92	GLY	3.2
1	B	91	GLY	3.1
1	D	293	ASP	3.1
1	D	309	LYS	3.1
1	B	89	ASN	3.0
1	B	309	LYS	2.9
1	B	298	TYR	2.9
1	A	310	SER	2.9
1	A	293	ASP	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	293	ASP	2.8
1	C	294	THR	2.7
1	C	296	TYR	2.7
1	D	85	LEU	2.7
1	B	378	THR	2.7
1	D	392	VAL	2.7
1	B	2	GLY	2.7
1	B	388	ASP	2.7
1	D	378	THR	2.6
1	C	1	MET	2.6
1	B	291	ILE	2.5
1	D	305	VAL	2.4
1	B	296	TYR	2.4
1	C	309	LYS	2.4
1	C	298	TYR	2.3
1	A	91	GLY	2.3
1	A	390	GLU	2.3
1	B	1	MET	2.2
1	B	307	ALA	2.2
1	D	89	ASN	2.2
1	A	78	GLU	2.2
1	B	287	GLU	2.2
1	C	78	GLU	2.2
1	B	391	ILE	2.1
1	A	392	VAL	2.1
1	C	307	ALA	2.1
1	B	392	VAL	2.1
1	C	392	VAL	2.1
1	C	90	SER	2.1
1	D	222[A]	ARG	2.1
1	D	298	TYR	2.0
1	C	290	LEU	2.0
1	D	391	ILE	2.0
1	B	305	VAL	2.0

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

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( $\text{\AA}^2$ )	Q<0.9
1	OCS	B	219	9/10	0.93	0.13	-	25,31,45,50	0
1	OCS	A	219	9/10	0.95	0.10	-	30,35,47,53	0
1	OCS	D	219	9/10	0.93	0.15	-	31,34,54,59	0
1	OCS	C	219	9/10	0.96	0.10	-	31,33,50,51	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. 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( $\text{\AA}^2$ )	Q<0.9
4	GOL	A	502	6/6	0.83	0.29	5.18	63,65,69,70	0
4	GOL	B	503	6/6	0.89	0.23	3.44	52,65,70,72	0
3	I13	A	501	30/30	0.89	0.16	2.25	22,37,59,70	0
3	I13	B	502	30/30	0.92	0.13	1.77	26,34,56,61	0
3	I13	C	503	30/30	0.93	0.13	1.67	22,31,60,62	0
4	GOL	C	502	6/6	0.86	0.19	1.67	55,63,66,66	0
3	I13	B	505	30/30	0.88	0.15	1.58	25,34,58,64	0
5	PEG	B	504	7/7	0.88	0.18	1.13	57,63,66,67	0
5	PEG	A	503	7/7	0.77	0.16	0.13	64,66,67,69	0
2	IMP	B	501	23/23	0.96	0.10	-0.30	20,25,29,30	0
2	IMP	A	500	23/23	0.97	0.09	-0.40	19,25,28,36	0
2	IMP	D	501	23/23	0.97	0.09	-0.64	19,25,28,30	0
2	IMP	C	501	23/23	0.98	0.10	-1.07	20,27,30,32	0
4	GOL	D	502	6/6	0.76	0.18	-	49,58,59,62	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.