



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:34 PM GMT

PDB ID : 4CMI  
Title : Crystal structure of pteridine reductase 1 (PTR1) from Trypanosoma brucei  
in ternary complex with cofactor and inhibitor  
Authors : Barrack, K.L.; Hunter, W.N.  
Deposited on : 2014-01-16  
Resolution : 1.90 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

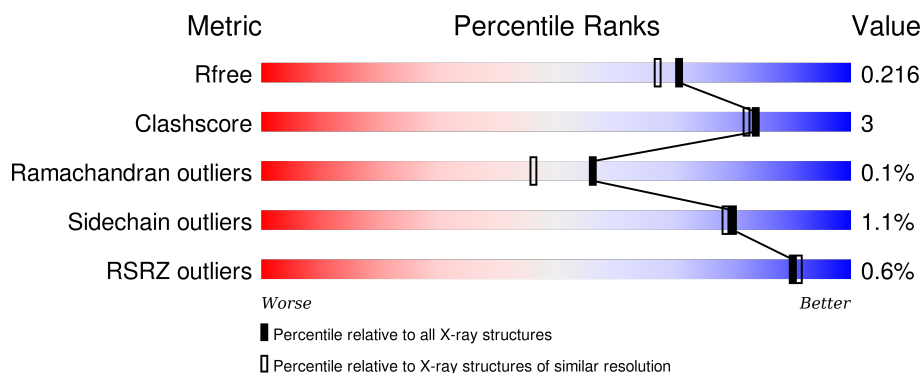
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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	288	<div> <div></div> <div>80%6%13%</div> </div>
1	D	288	<div> <div>%</div> <div>81%6%13%</div> </div>
2	B	288	<div> <div>%</div> <div>84%••13%</div> </div>
2	C	288	<div> <div></div> <div>80%7%13%</div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8524 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 PTERIDINE REDUCTASE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	250	Total	C	N	O	S	0	1	0
			1867	1171	329	356	11			
1	D	250	Total	C	N	O	S	0	1	0
			1865	1172	328	354	11			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP O76290
A	-18	GLY	-	EXPRESSION TAG	UNP O76290
A	-17	SER	-	EXPRESSION TAG	UNP O76290
A	-16	SER	-	EXPRESSION TAG	UNP O76290
A	-15	HIS	-	EXPRESSION TAG	UNP O76290
A	-14	HIS	-	EXPRESSION TAG	UNP O76290
A	-13	HIS	-	EXPRESSION TAG	UNP O76290
A	-12	HIS	-	EXPRESSION TAG	UNP O76290
A	-11	HIS	-	EXPRESSION TAG	UNP O76290
A	-10	HIS	-	EXPRESSION TAG	UNP O76290
A	-9	SER	-	EXPRESSION TAG	UNP O76290
A	-8	SER	-	EXPRESSION TAG	UNP O76290
A	-7	GLY	-	EXPRESSION TAG	UNP O76290
A	-6	LEU	-	EXPRESSION TAG	UNP O76290
A	-5	VAL	-	EXPRESSION TAG	UNP O76290
A	-4	PRO	-	EXPRESSION TAG	UNP O76290
A	-3	ARG	-	EXPRESSION TAG	UNP O76290
A	-2	GLY	-	EXPRESSION TAG	UNP O76290
A	-1	SER	-	EXPRESSION TAG	UNP O76290
A	0	HIS	-	EXPRESSION TAG	UNP O76290
D	-19	MET	-	EXPRESSION TAG	UNP O76290
D	-18	GLY	-	EXPRESSION TAG	UNP O76290
D	-17	SER	-	EXPRESSION TAG	UNP O76290
D	-16	SER	-	EXPRESSION TAG	UNP O76290
D	-15	HIS	-	EXPRESSION TAG	UNP O76290

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-14	HIS	-	EXPRESSION TAG	UNP O76290
D	-13	HIS	-	EXPRESSION TAG	UNP O76290
D	-12	HIS	-	EXPRESSION TAG	UNP O76290
D	-11	HIS	-	EXPRESSION TAG	UNP O76290
D	-10	HIS	-	EXPRESSION TAG	UNP O76290
D	-9	SER	-	EXPRESSION TAG	UNP O76290
D	-8	SER	-	EXPRESSION TAG	UNP O76290
D	-7	GLY	-	EXPRESSION TAG	UNP O76290
D	-6	LEU	-	EXPRESSION TAG	UNP O76290
D	-5	VAL	-	EXPRESSION TAG	UNP O76290
D	-4	PRO	-	EXPRESSION TAG	UNP O76290
D	-3	ARG	-	EXPRESSION TAG	UNP O76290
D	-2	GLY	-	EXPRESSION TAG	UNP O76290
D	-1	SER	-	EXPRESSION TAG	UNP O76290
D	0	HIS	-	EXPRESSION TAG	UNP O76290

- Molecule 2 is a protein called PTERIDINE REDUCTASE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	251	Total	C	N	O	S	0	7	0
			1914	1198	339	366	11			
2	C	250	Total	C	N	O	S	0	4	0
			1884	1182	331	360	11			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	EXPRESSION TAG	UNP O76290
B	-18	GLY	-	EXPRESSION TAG	UNP O76290
B	-17	SER	-	EXPRESSION TAG	UNP O76290
B	-16	SER	-	EXPRESSION TAG	UNP O76290
B	-15	HIS	-	EXPRESSION TAG	UNP O76290
B	-14	HIS	-	EXPRESSION TAG	UNP O76290
B	-13	HIS	-	EXPRESSION TAG	UNP O76290
B	-12	HIS	-	EXPRESSION TAG	UNP O76290
B	-11	HIS	-	EXPRESSION TAG	UNP O76290
B	-10	HIS	-	EXPRESSION TAG	UNP O76290
B	-9	SER	-	EXPRESSION TAG	UNP O76290
B	-8	SER	-	EXPRESSION TAG	UNP O76290
B	-7	GLY	-	EXPRESSION TAG	UNP O76290
B	-6	LEU	-	EXPRESSION TAG	UNP O76290
B	-5	VAL	-	EXPRESSION TAG	UNP O76290

*Continued on next page...*

*Continued from previous page...*

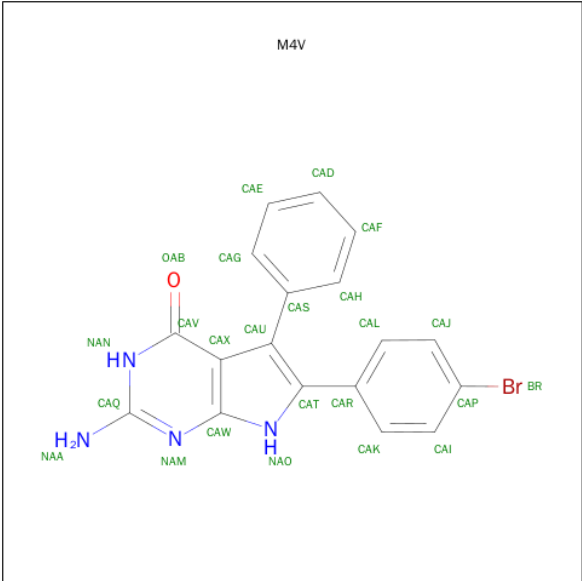
Chain	Residue	Modelled	Actual	Comment	Reference
B	-4	PRO	-	EXPRESSION TAG	UNP O76290
B	-3	ARG	-	EXPRESSION TAG	UNP O76290
B	-2	GLY	-	EXPRESSION TAG	UNP O76290
B	-1	SER	-	EXPRESSION TAG	UNP O76290
B	0	HIS	-	EXPRESSION TAG	UNP O76290
C	-19	MET	-	EXPRESSION TAG	UNP O76290
C	-18	GLY	-	EXPRESSION TAG	UNP O76290
C	-17	SER	-	EXPRESSION TAG	UNP O76290
C	-16	SER	-	EXPRESSION TAG	UNP O76290
C	-15	HIS	-	EXPRESSION TAG	UNP O76290
C	-14	HIS	-	EXPRESSION TAG	UNP O76290
C	-13	HIS	-	EXPRESSION TAG	UNP O76290
C	-12	HIS	-	EXPRESSION TAG	UNP O76290
C	-11	HIS	-	EXPRESSION TAG	UNP O76290
C	-10	HIS	-	EXPRESSION TAG	UNP O76290
C	-9	SER	-	EXPRESSION TAG	UNP O76290
C	-8	SER	-	EXPRESSION TAG	UNP O76290
C	-7	GLY	-	EXPRESSION TAG	UNP O76290
C	-6	LEU	-	EXPRESSION TAG	UNP O76290
C	-5	VAL	-	EXPRESSION TAG	UNP O76290
C	-4	PRO	-	EXPRESSION TAG	UNP O76290
C	-3	ARG	-	EXPRESSION TAG	UNP O76290
C	-2	GLY	-	EXPRESSION TAG	UNP O76290
C	-1	SER	-	EXPRESSION TAG	UNP O76290
C	0	HIS	-	EXPRESSION TAG	UNP O76290

- Molecule 3 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: C<sub>21</sub>H<sub>28</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



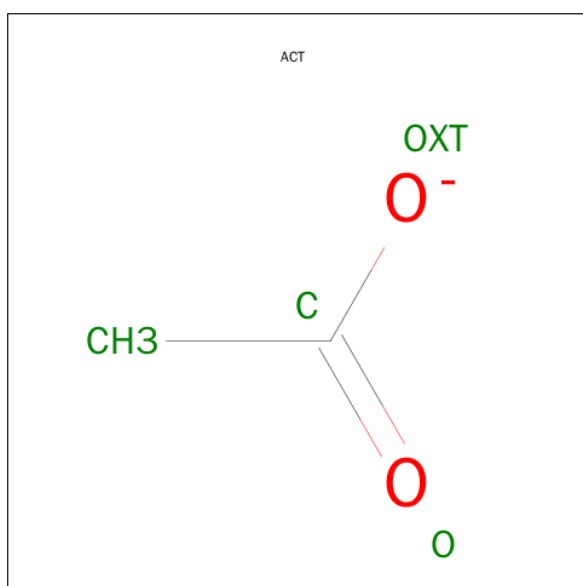
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	C	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	D	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 4 is 2-AMINO-6-(4-BROMOPHENYL)-5-PHENYL-3H-PYRROLO[2,3-D]PYRIMIDIN-4(7H)-ONE (three-letter code: M4V) (formula: C<sub>18</sub>H<sub>13</sub>BrN<sub>4</sub>O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	Br	C	N	O	0	0
			24	1	18	4	1		
4	B	1	Total	Br	C	N	O	0	0
			24	1	18	4	1		
4	C	1	Total	Br	C	N	O	0	0
			24	1	18	4	1		
4	D	1	Total	Br	C	N	O	0	0
			24	1	18	4	1		

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		

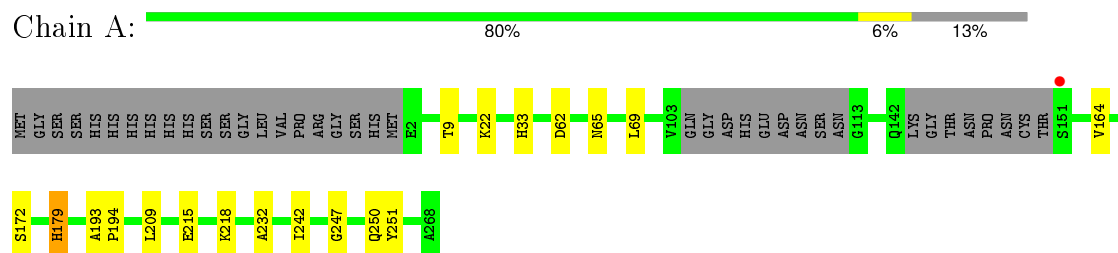
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	197	Total	O	0	0
			197	197		
6	B	198	Total	O	0	0
			198	198		
6	C	166	Total	O	0	0
			166	166		
6	D	137	Total	O	0	0
			137	137		

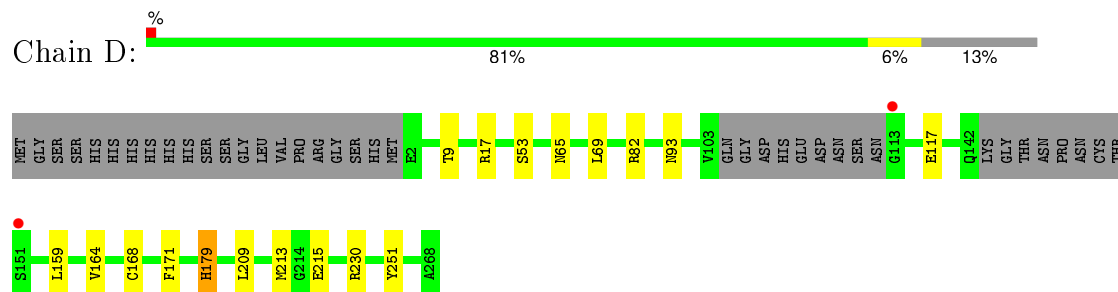
### 3 Residue-property plots

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.

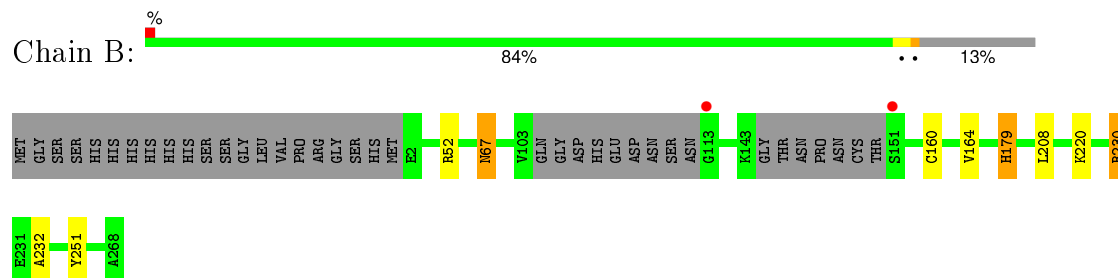
#### • Molecule 1: PTERIDINE REDUCTASE 1



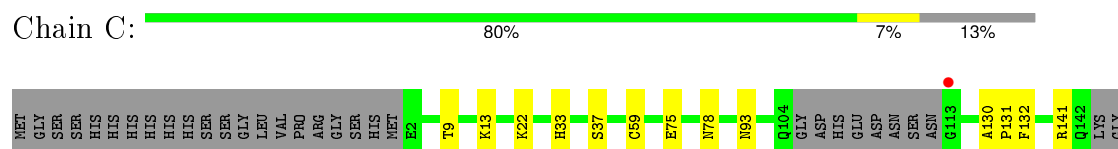
#### • Molecule 1: PTERIDINE REDUCTASE 1



#### • Molecule 2: PTERIDINE REDUCTASE 1



#### • Molecule 2: PTERIDINE REDUCTASE 1





THR	ASN	PRO	ASN	CYS	THR	SER	S162	D161	V164	H179	E216	R223	A232	E235	A268
-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.46 Å 90.26 Å 82.09 Å 90.00° 115.59° 90.00°	Depositor
Resolution (Å)	57.24 – 1.90 57.24 – 1.90	Depositor EDS
% Data completeness (in resolution range)	94.7 (57.24-1.90) 94.8 (57.24-1.90)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.58 (at 1.90 Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.160 , 0.210 0.171 , 0.216	Depositor DCC
$R_{free}$ test set	3712 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	16.7	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.3	EDS
Estimated twinning fraction	0.009 for h,-k,-h-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 73124 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8524	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.96 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 5.1256e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: NAP, M4V, CSX, ACT

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	0.71	0/1886	0.80	1/2557 (0.0%)
1	D	0.69	0/1887	0.83	5/2559 (0.2%)
2	B	0.73	0/1941	0.80	0/2632
2	C	0.69	0/1917	0.77	1/2600 (0.0%)
All	All	0.71	0/7631	0.80	7/10348 (0.1%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	159	LEU	CA-CB-CG	-7.15	98.86	115.30
1	D	17	ARG	NE-CZ-NH1	5.87	123.23	120.30
1	D	17	ARG	NE-CZ-NH2	-5.75	117.42	120.30
1	A	62	ASP	CB-CG-OD2	-5.36	113.47	118.30
1	D	82	ARG	NE-CZ-NH1	5.30	122.95	120.30
2	C	161	ASP	CB-CG-OD1	5.05	122.84	118.30
1	D	82	ARG	NE-CZ-NH2	-5.00	117.80	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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	1867	0	1886	12	0
1	D	1865	0	1891	9	0
2	B	1914	0	1929	12	0
2	C	1884	0	1904	13	0
3	A	48	0	25	1	0
3	B	48	0	25	0	0
3	C	48	0	25	1	0
3	D	48	0	25	1	0
4	A	24	0	13	1	0
4	B	24	0	13	1	0
4	C	24	0	13	2	0
4	D	24	0	13	4	0
5	B	4	0	3	0	0
5	C	4	0	3	0	0
6	A	197	0	0	1	0
6	B	198	0	0	5	0
6	C	166	0	0	2	0
6	D	137	0	0	1	0
All	All	8524	0	7768	46	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 (46) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:52:ARG:NH1	6:B:2024:HOH:O	1.90	1.03
2:C:78:ASN:OD1	2:C:141:ARG:NH1	2.11	0.83
1:D:213:MET:O	6:D:2119:HOH:O	2.00	0.78
2:C:75:GLU:OE1	6:C:2074:HOH:O	2.01	0.76
2:B:67[B]:ASN:ND2	1:D:117:GLU:HG3	2.12	0.64
1:A:247:GLY:O	1:A:250:GLN:HG3	1.98	0.64
2:C:9:THR:HA	2:C:33:HIS:HB3	1.82	0.61
1:A:215:GLU:OE1	6:A:2152:HOH:O	2.17	0.58
4:C:1270:M4V:CAL	4:C:1270:M4V:CAH	2.84	0.56
1:D:168:CSX:HB2	1:D:171:PHE:CD1	2.42	0.54
2:B:67[A]:ASN:HD22	2:B:67[A]:ASN:H	1.55	0.53
2:B:67[A]:ASN:HD22	2:B:67[A]:ASN:N	2.06	0.52
1:A:172[A]:SER:HG	2:C:132:PHE:HE1	1.57	0.51
1:A:65:ASN:HA	1:A:69:LEU:HD22	1.93	0.50
2:C:130:ALA:HB3	2:C:131:PRO:HD3	1.93	0.50
2:C:13:LYS:HE2	2:C:37:SER:OG	2.11	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:1270:M4V:CAH	4:D:1270:M4V:CAL	2.91	0.49
2:B:230:ARG:NH1	6:B:2170:HOH:O	2.46	0.48
1:A:251:TYR:CE2	2:B:232:ALA:HB2	2.49	0.47
3:A:1269:NAP:C7N	4:A:1270:M4V:HAG	2.45	0.47
2:C:232:ALA:HB2	1:D:251:TYR:CD2	2.50	0.47
2:C:223:ARG:NH1	6:C:2126:HOH:O	2.19	0.47
2:C:164:VAL:HG22	2:C:179:HIS:CD2	2.50	0.47
2:B:164:VAL:HG22	2:B:179:HIS:CD2	2.51	0.46
1:D:65:ASN:HA	1:D:69:LEU:HD22	1.97	0.45
2:B:220:LYS:CE	6:B:2162:HOH:O	2.64	0.45
2:B:160:CYS:HB3	6:B:2124:HOH:O	2.17	0.44
1:A:193:ALA:N	1:A:194:PRO:CD	2.81	0.43
2:C:33:HIS:HA	2:C:59:CYS:O	2.17	0.43
1:A:9:THR:HA	1:A:33:HIS:HB3	2.00	0.43
2:C:232:ALA:HB2	1:D:251:TYR:CE2	2.53	0.43
1:D:164:VAL:HG22	1:D:179:HIS:CD2	2.53	0.43
3:D:1269:NAP:C7N	4:D:1270:M4V:HAG	2.48	0.43
1:A:164:VAL:HG22	1:A:179:HIS:CD2	2.53	0.43
1:D:9:THR:O	1:D:93:ASN:HB3	2.19	0.42
1:A:22:LYS:HG2	1:A:242:ILE:HG13	2.01	0.42
4:B:1270:M4V:HAK	4:B:1270:M4V:CAS	2.49	0.42
4:D:1270:M4V:CAS	4:D:1270:M4V:HAL	2.50	0.42
2:C:9:THR:O	2:C:93:ASN:HB3	2.20	0.41
3:C:1269:NAP:C7N	4:C:1270:M4V:HAG	2.50	0.41
2:B:220:LYS:HE3	6:B:2162:HOH:O	2.20	0.41
2:C:22:LYS:CE	2:C:235[B]:GLU:HG3	2.50	0.41
1:A:232:ALA:HB2	2:B:251:TYR:CD2	2.55	0.41
1:A:232:ALA:HB2	2:B:251:TYR:CE2	2.55	0.41
1:A:209:LEU:HD13	1:A:218:LYS:HB3	2.03	0.41
1:D:209:LEU:HD23	4:D:1270:M4V:HAD	2.02	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	244/288 (85%)	235 (96%)	9 (4%)	0	100	100
1	D	244/288 (85%)	237 (97%)	7 (3%)	0	100	100
2	B	252/288 (88%)	245 (97%)	6 (2%)	1 (0%)	39	27
2	C	248/288 (86%)	240 (97%)	8 (3%)	0	100	100
All	All	988/1152 (86%)	957 (97%)	30 (3%)	1 (0%)	56	46

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	208	LEU

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	198/230 (86%)	197 (100%)	1 (0%)	92	92
1	D	198/230 (86%)	194 (98%)	4 (2%)	63	57
2	B	205/231 (89%)	201 (98%)	4 (2%)	63	57
2	C	202/231 (87%)	201 (100%)	1 (0%)	92	92
All	All	803/922 (87%)	793 (99%)	10 (1%)	80	76

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

Mol	Chain	Res	Type
1	A	179	HIS
2	B	67[A]	ASN
2	B	67[B]	ASN
2	B	179	HIS
2	B	230	ARG
2	C	216	GLU
1	D	53	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	179	HIS
1	D	215	GLU
1	D	230	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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	CSX	A	168	1	3,6,7	0.42	0	3,6,8	2.58	2 (66%)
1	CSX	D	168	1	3,6,7	0.49	0	3,6,8	2.34	2 (66%)

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	CSX	A	168	1	-	0/1/5/7	0/0/0/0
1	CSX	D	168	1	-	0/1/5/7	0/0/0/0

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	168	CSX	CB-CA-C	-3.16	102.80	111.46
1	D	168	CSX	CB-CA-C	-2.81	103.77	111.46
1	A	168	CSX	CA-CB-SG	-2.51	105.17	113.01
1	D	168	CSX	CA-CB-SG	-2.17	106.26	113.01

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	168	CSX	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

10 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	$\# Z  > 2$
3	NAP	A	1269	-	42,52,52	1.00	4 (9%)	54,80,80	2.21	12 (22%)
4	M4V	A	1270	-	26,27,27	1.53	6 (23%)	25,39,39	2.12	8 (32%)
3	NAP	B	1269	-	42,52,52	1.00	2 (4%)	54,80,80	2.10	11 (20%)
4	M4V	B	1270	-	26,27,27	1.37	4 (15%)	25,39,39	1.94	4 (16%)
5	ACT	B	1271	-	1,3,3	1.04	0	0,3,3	0.00	-
3	NAP	C	1269	-	42,52,52	0.84	2 (4%)	54,80,80	1.89	10 (18%)
4	M4V	C	1270	-	26,27,27	1.45	5 (19%)	25,39,39	2.10	7 (28%)
5	ACT	C	1271	-	1,3,3	1.97	0	0,3,3	0.00	-
3	NAP	D	1269	-	42,52,52	0.89	0	54,80,80	2.17	14 (25%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	M4V	D	1270	-	26,27,27	1.54	5 (19%)	25,39,39	2.55	7 (28%)

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	NAP	A	1269	-	-	0/27/67/67	0/5/5/5
4	M4V	A	1270	-	-	0/8/8/8	0/4/4/4
3	NAP	B	1269	-	-	0/27/67/67	0/5/5/5
4	M4V	B	1270	-	-	0/8/8/8	0/4/4/4
5	ACT	B	1271	-	-	0/0/0/0	0/0/0/0
3	NAP	C	1269	-	-	0/27/67/67	0/5/5/5
4	M4V	C	1270	-	-	0/8/8/8	0/4/4/4
5	ACT	C	1271	-	-	0/0/0/0	0/0/0/0
3	NAP	D	1269	-	-	0/27/67/67	0/5/5/5
4	M4V	D	1270	-	-	0/8/8/8	0/4/4/4

All (28) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	1270	M4V	CAU-CAX	-3.86	1.37	1.42
4	A	1270	M4V	CAU-CAX	-3.67	1.37	1.42
4	D	1270	M4V	CAU-CAX	-3.31	1.38	1.42
4	B	1270	M4V	CAU-CAX	-2.98	1.38	1.42
4	C	1270	M4V	CAU-CAT	-2.54	1.38	1.41
3	A	1269	NAP	O4B-C4B	-2.07	1.40	1.45
4	A	1270	M4V	CAU-CAT	-2.06	1.39	1.41
4	C	1270	M4V	CAU-CAS	2.01	1.52	1.49
4	B	1270	M4V	CAV-NAN	2.07	1.36	1.33
3	C	1269	NAP	C2A-N3A	2.15	1.36	1.32
4	A	1270	M4V	CAQ-NAA	2.19	1.38	1.34
3	B	1269	NAP	O3D-C3D	2.28	1.48	1.43
3	A	1269	NAP	C5A-C4A	2.29	1.45	1.40
3	A	1269	NAP	C2A-N3A	2.30	1.36	1.32
4	C	1270	M4V	CAV-NAN	2.37	1.37	1.33
4	B	1270	M4V	CAU-CAS	2.44	1.52	1.49
4	D	1270	M4V	CAV-NAN	2.49	1.37	1.33
3	A	1269	NAP	O4D-C1D	2.52	1.44	1.41
4	D	1270	M4V	BR-CAP	2.57	1.96	1.90
3	C	1269	NAP	C5A-C4A	2.60	1.46	1.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1270	M4V	CAU-CAS	2.71	1.53	1.49
4	A	1270	M4V	BR-CAP	2.72	1.96	1.90
3	B	1269	NAP	C5A-C4A	2.75	1.46	1.40
4	A	1270	M4V	CAV-NAN	2.76	1.38	1.33
4	B	1270	M4V	BR-CAP	2.95	1.96	1.90
4	C	1270	M4V	CAQ-NAA	3.13	1.40	1.34
4	D	1270	M4V	CAW-NAO	3.18	1.40	1.34
4	D	1270	M4V	CAQ-NAA	3.40	1.41	1.34

All (73) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1269	NAP	N3A-C2A-N1A	-10.61	120.77	128.89
3	B	1269	NAP	N3A-C2A-N1A	-8.60	122.31	128.89
3	D	1269	NAP	N3A-C2A-N1A	-8.36	122.49	128.89
3	C	1269	NAP	N3A-C2A-N1A	-8.30	122.54	128.89
4	D	1270	M4V	CAX-CAV-NAN	-7.60	118.43	124.19
4	B	1270	M4V	CAX-CAV-NAN	-7.39	118.59	124.19
4	C	1270	M4V	CAX-CAV-NAN	-5.38	120.12	124.19
4	A	1270	M4V	CAX-CAV-NAN	-5.18	120.26	124.19
4	D	1270	M4V	NAM-CAQ-NAN	-4.52	120.56	127.44
3	D	1269	NAP	C4A-C5A-N7A	-4.35	105.47	109.48
3	D	1269	NAP	O7N-C7N-C3N	-4.32	114.87	119.59
4	A	1270	M4V	CAL-CAR-CAT	-4.10	114.20	120.60
3	B	1269	NAP	C4B-O4B-C1B	-4.06	105.26	109.72
4	A	1270	M4V	NAM-CAQ-NAN	-3.91	121.49	127.44
3	C	1269	NAP	C4B-O4B-C1B	-3.91	105.42	109.72
4	C	1270	M4V	CAS-CAU-CAT	-3.81	119.32	126.88
3	A	1269	NAP	C4B-O4B-C1B	-3.73	105.62	109.72
4	C	1270	M4V	NAM-CAQ-NAN	-3.72	121.77	127.44
4	D	1270	M4V	CAK-CAR-CAT	-3.51	115.12	120.60
4	A	1270	M4V	CAS-CAU-CAT	-3.48	119.98	126.88
4	C	1270	M4V	CAK-CAR-CAT	-3.47	115.19	120.60
3	C	1269	NAP	C1B-N9A-C4A	-3.45	121.74	126.94
4	D	1270	M4V	CAS-CAU-CAT	-3.30	120.33	126.88
4	C	1270	M4V	CAU-CAX-CAW	-3.17	105.01	107.54
3	A	1269	NAP	C1B-N9A-C4A	-3.16	122.18	126.94
3	B	1269	NAP	C1B-N9A-C4A	-2.89	122.58	126.94
4	B	1270	M4V	NAM-CAQ-NAN	-2.76	123.25	127.44
3	A	1269	NAP	C4D-O4D-C1D	-2.73	106.72	109.72
3	B	1269	NAP	O7N-C7N-N7N	-2.68	118.82	122.59
3	A	1269	NAP	C4A-C5A-N7A	-2.52	107.16	109.48

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1269	NAP	C4A-C5A-N7A	-2.52	107.16	109.48
3	D	1269	NAP	O3-PA-O5B	-2.42	96.50	102.94
3	B	1269	NAP	PN-O3-PA	-2.40	125.99	132.73
3	D	1269	NAP	C1B-N9A-C4A	-2.35	123.39	126.94
3	A	1269	NAP	O7N-C7N-C3N	-2.30	117.08	119.59
3	D	1269	NAP	PN-O3-PA	-2.26	126.38	132.73
4	B	1270	M4V	CAS-CAU-CAT	-2.18	122.55	126.88
3	A	1269	NAP	O4B-C4B-C3B	-2.13	100.86	105.15
3	C	1269	NAP	O2B-P2B-O1X	-2.11	101.84	107.11
3	D	1269	NAP	C4B-O4B-C1B	-2.05	107.47	109.72
3	B	1269	NAP	O7N-C7N-C3N	-2.04	117.36	119.59
4	A	1270	M4V	CAF-CAH-CAS	2.02	123.23	120.56
4	A	1270	M4V	NAA-CAQ-NAN	2.14	120.74	117.20
3	A	1269	NAP	O2N-PN-O1N	2.21	124.50	112.53
3	D	1269	NAP	O2A-PA-O3	2.25	115.28	105.09
3	D	1269	NAP	O2X-P2B-O1X	2.30	117.99	110.58
3	D	1269	NAP	O2N-PN-O1N	2.40	125.51	112.53
3	C	1269	NAP	O2A-PA-O1A	2.43	125.71	112.53
3	B	1269	NAP	O3B-C3B-C2B	2.47	118.30	111.16
3	C	1269	NAP	C2A-N1A-C6A	2.49	123.22	118.77
3	B	1269	NAP	C2A-N1A-C6A	2.54	123.31	118.77
3	D	1269	NAP	O4D-C1D-N1N	2.71	111.11	108.13
3	A	1269	NAP	C2A-N1A-C6A	2.73	123.64	118.77
4	C	1270	M4V	CAL-CAR-CAT	3.06	125.37	120.60
4	C	1270	M4V	CAV-NAN-CAQ	3.07	120.20	115.94
4	A	1270	M4V	CAV-NAN-CAQ	3.09	120.23	115.94
4	B	1270	M4V	CAV-NAN-CAQ	3.14	120.30	115.94
3	A	1269	NAP	O3X-P2B-O2X	3.15	119.39	107.38
3	C	1269	NAP	O3X-P2B-O1X	3.37	121.43	110.58
3	D	1269	NAP	P2B-O2B-C2B	3.47	129.87	121.56
3	C	1269	NAP	O4B-C1B-N9A	3.57	115.58	108.10
4	D	1270	M4V	CAL-CAR-CAT	3.60	126.22	120.60
4	D	1270	M4V	CAV-NAN-CAQ	3.92	121.38	115.94
4	D	1270	M4V	NAA-CAQ-NAN	3.99	123.81	117.20
4	A	1270	M4V	CAK-CAR-CAT	4.03	126.89	120.60
3	B	1269	NAP	O4B-C1B-N9A	4.18	116.85	108.10
3	C	1269	NAP	O4D-C1D-N1N	4.30	112.85	108.13
3	A	1269	NAP	O4B-C1B-N9A	4.38	117.27	108.10
3	D	1269	NAP	C3N-C7N-N7N	4.96	123.25	117.82
3	B	1269	NAP	O4D-C1D-N1N	5.06	113.69	108.13
3	A	1269	NAP	O4D-C1D-N1N	5.41	114.08	108.13
3	B	1269	NAP	C3N-C7N-N7N	5.48	123.81	117.82

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	D	1269	NAP	O4B-C1B-N9A	5.57	119.75	108.10

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1269	NAP	1	0
4	A	1270	M4V	1	0
4	B	1270	M4V	1	0
3	C	1269	NAP	1	0
4	C	1270	M4V	2	0
3	D	1269	NAP	1	0
4	D	1270	M4V	4	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 [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	249/288 (86%)	-0.51	1 (0%) 93 93	9, 16, 31, 57	0
1	D	249/288 (86%)	-0.42	2 (0%) 87 88	11, 18, 37, 58	0
2	B	251/288 (87%)	-0.47	2 (0%) 87 88	11, 17, 32, 55	0
2	C	250/288 (86%)	-0.52	1 (0%) 93 93	10, 17, 35, 59	0
All	All	999/1152 (86%)	-0.48	6 (0%) 90 91	9, 17, 34, 59	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	113	GLY	4.1
1	D	113	GLY	3.7
1	D	151	SER	3.3
2	B	151	SER	2.7
1	A	151	SER	2.6
2	B	113	GLY	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
1	CSX	D	168	7/8	0.91	0.10	-	20,22,38,45	0
1	CSX	A	168	7/8	0.90	0.11	-	17,21,32,39	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
5	ACT	C	1271	4/4	0.95	0.09	1.03	24,26,26,27	0
4	M4V	B	1270	24/24	0.98	0.09	0.02	14,17,22,22	0
4	M4V	D	1270	24/24	0.99	0.08	-0.05	18,20,23,24	0
3	NAP	A	1269	48/48	0.98	0.07	-0.40	11,13,15,18	0
3	NAP	D	1269	48/48	0.98	0.07	-0.55	13,16,19,20	0
3	NAP	C	1269	48/48	0.98	0.07	-0.63	11,14,17,19	0
3	NAP	B	1269	48/48	0.98	0.07	-0.75	10,13,15,19	0
4	M4V	C	1270	24/24	0.99	0.07	-0.91	14,16,18,19	0
5	ACT	B	1271	4/4	0.98	0.06	-1.12	15,16,16,16	0
4	M4V	A	1270	24/24	0.99	0.06	-1.22	14,17,20,21	0

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

There are no such residues in this entry.