



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 17, 2017 – 06:32 PM EST

PDB ID : 5TFT  
Title : Structure of cytochrome P450 2D6 (CYP2D6) BACE1 inhibitor complex  
Authors : Hsu, M.H.; Johnson, E.F.  
Deposited on : 2016-09-26  
Resolution : 2.71 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

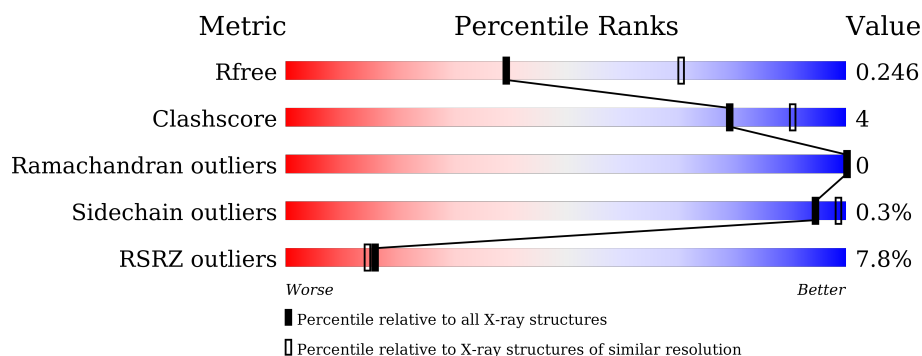
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.71 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	2439 (2.74-2.70)
Clashscore	102246	2771 (2.74-2.70)
Ramachandran outliers	100387	2726 (2.74-2.70)
Sidechain outliers	100360	2727 (2.74-2.70)
RSRZ outliers	91569	2443 (2.74-2.70)

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	479	<div> <div>9%</div> <div>89%</div> <div>8%</div> <div>.</div> </div>
1	B	479	<div> <div>5%</div> <div>88%</div> <div>8%</div> <div>.</div> </div>
1	C	479	<div> <div>8%</div> <div>83%</div> <div>10%</div> <div>7%</div> </div>
1	D	479	<div> <div>8%</div> <div>88%</div> <div>7%</div> <div>5%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 14779 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 Cytochrome P450 2D6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	467	Total	C	N	O	S	0	0	0
			3690	2367	654	655	14			
1	B	460	Total	C	N	O	S	0	0	0
			3642	2334	646	648	14			
1	C	447	Total	C	N	O	S	0	0	0
			3542	2269	628	631	14			
1	D	453	Total	C	N	O	S	0	0	0
			3586	2298	636	638	14			

There are 60 discrepancies between the modelled and reference sequences:

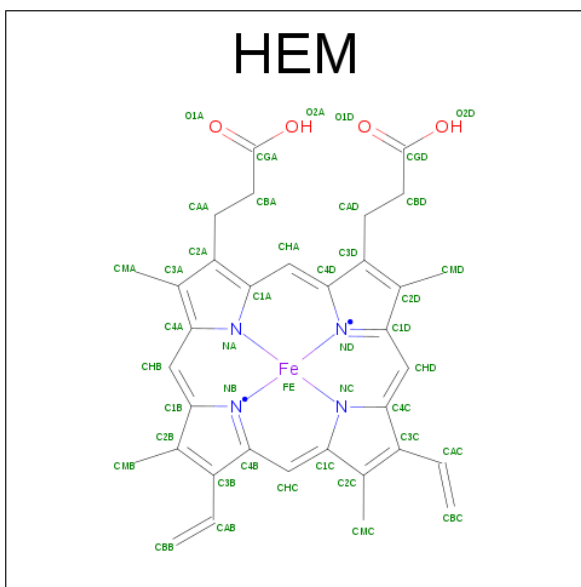
Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MET	-	initiating methionine	UNP P10635
A	24	ALA	-	expression tag	UNP P10635
A	25	LYS	-	expression tag	UNP P10635
A	26	LYS	-	expression tag	UNP P10635
A	27	THR	-	expression tag	UNP P10635
A	28	SER	-	expression tag	UNP P10635
A	29	SER	-	expression tag	UNP P10635
A	30	LYS	-	expression tag	UNP P10635
A	31	GLY	-	expression tag	UNP P10635
A	32	LYS	-	expression tag	UNP P10635
A	33	LEU	-	expression tag	UNP P10635
A	498	HIS	-	expression tag	UNP P10635
A	499	HIS	-	expression tag	UNP P10635
A	500	HIS	-	expression tag	UNP P10635
A	501	HIS	-	expression tag	UNP P10635
B	23	MET	-	initiating methionine	UNP P10635
B	24	ALA	-	expression tag	UNP P10635
B	25	LYS	-	expression tag	UNP P10635
B	26	LYS	-	expression tag	UNP P10635
B	27	THR	-	expression tag	UNP P10635
B	28	SER	-	expression tag	UNP P10635

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	29	SER	-	expression tag	UNP P10635
B	30	LYS	-	expression tag	UNP P10635
B	31	GLY	-	expression tag	UNP P10635
B	32	LYS	-	expression tag	UNP P10635
B	33	LEU	-	expression tag	UNP P10635
B	498	HIS	-	expression tag	UNP P10635
B	499	HIS	-	expression tag	UNP P10635
B	500	HIS	-	expression tag	UNP P10635
B	501	HIS	-	expression tag	UNP P10635
C	23	MET	-	initiating methionine	UNP P10635
C	24	ALA	-	expression tag	UNP P10635
C	25	LYS	-	expression tag	UNP P10635
C	26	LYS	-	expression tag	UNP P10635
C	27	THR	-	expression tag	UNP P10635
C	28	SER	-	expression tag	UNP P10635
C	29	SER	-	expression tag	UNP P10635
C	30	LYS	-	expression tag	UNP P10635
C	31	GLY	-	expression tag	UNP P10635
C	32	LYS	-	expression tag	UNP P10635
C	33	LEU	-	expression tag	UNP P10635
C	498	HIS	-	expression tag	UNP P10635
C	499	HIS	-	expression tag	UNP P10635
C	500	HIS	-	expression tag	UNP P10635
C	501	HIS	-	expression tag	UNP P10635
D	23	MET	-	initiating methionine	UNP P10635
D	24	ALA	-	expression tag	UNP P10635
D	25	LYS	-	expression tag	UNP P10635
D	26	LYS	-	expression tag	UNP P10635
D	27	THR	-	expression tag	UNP P10635
D	28	SER	-	expression tag	UNP P10635
D	29	SER	-	expression tag	UNP P10635
D	30	LYS	-	expression tag	UNP P10635
D	31	GLY	-	expression tag	UNP P10635
D	32	LYS	-	expression tag	UNP P10635
D	33	LEU	-	expression tag	UNP P10635
D	498	HIS	-	expression tag	UNP P10635
D	499	HIS	-	expression tag	UNP P10635
D	500	HIS	-	expression tag	UNP P10635
D	501	HIS	-	expression tag	UNP P10635

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).

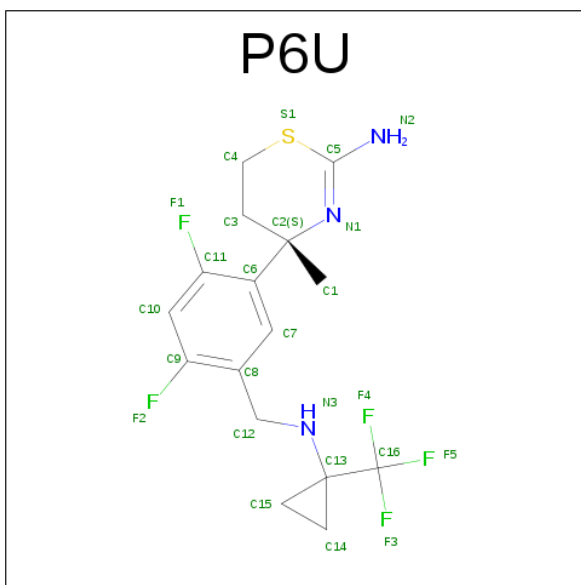


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	B	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	D	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

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

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

- Molecule 4 is (4S)-4-[2,4-difluoro-5-({[1-(trifluoromethyl)cyclopropyl]amino}methyl)phenyl]-4-methyl-5,6-dihydro-4H-1,3-thiazin-2-amine (three-letter code: P6U) (formula: C<sub>16</sub>H<sub>18</sub>F<sub>5</sub>N<sub>3</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	F	N	S	0	0
			25	16	5	3	1		
4	B	1	Total	C	F	N	S	0	0
			25	16	5	3	1		
4	C	1	Total	C	F	N	S	0	0
			25	16	5	3	1		
4	D	1	Total	C	F	N	S	0	0
			25	16	5	3	1		

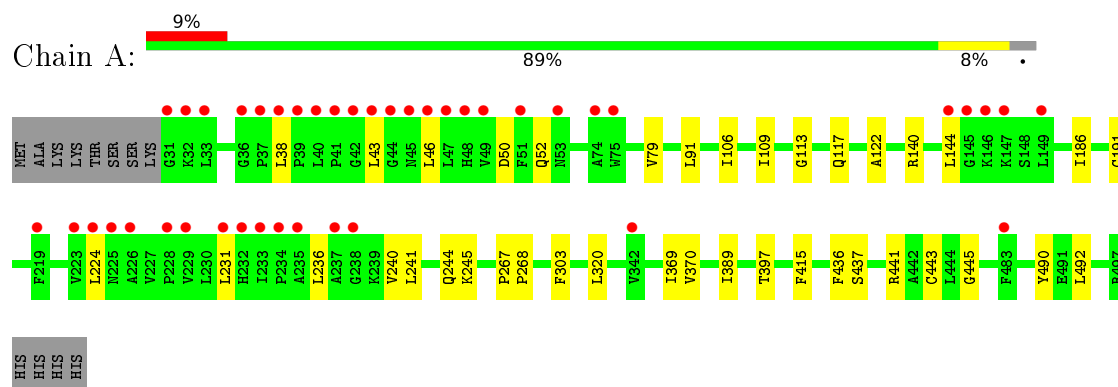
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	13	Total	O	0	0
			13	13		
5	B	12	Total	O	0	0
			12	12		
5	C	10	Total	O	0	0
			10	10		
5	D	7	Total	O	0	0
			7	7		

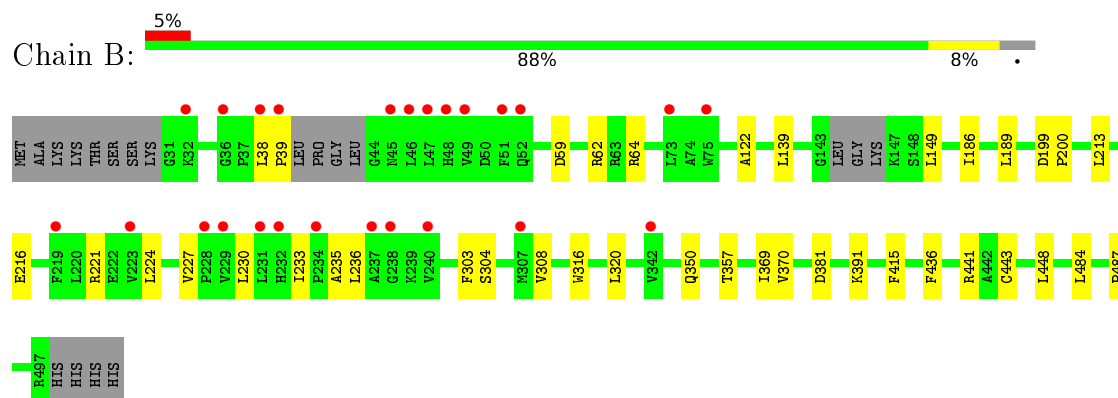
### 3 Residue-property plots

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.

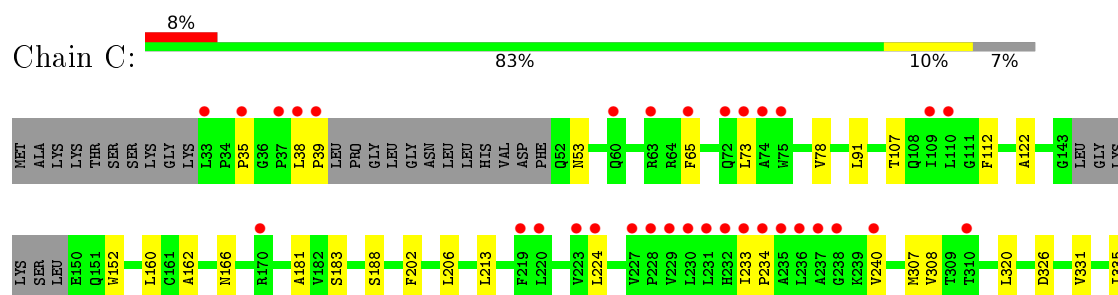
#### • Molecule 1: Cytochrome P450 2D6

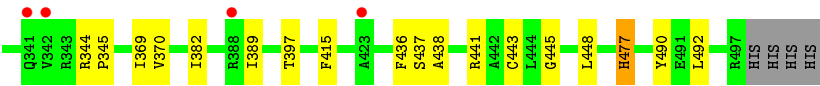


#### • Molecule 1: Cytochrome P450 2D6

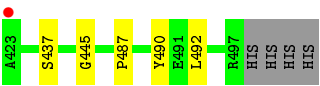
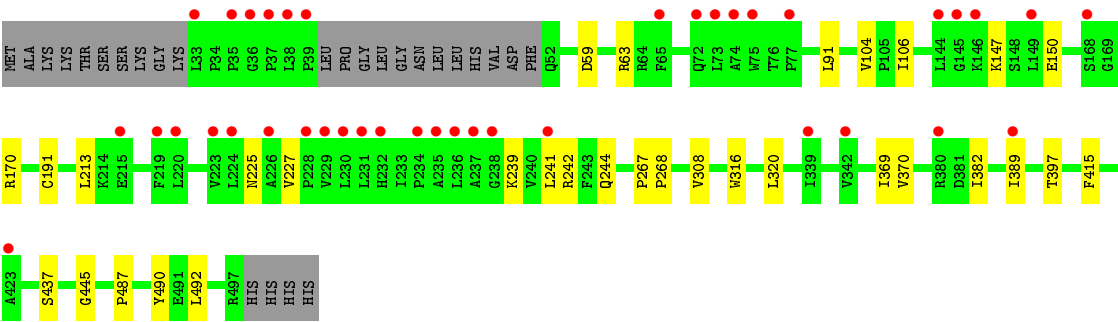
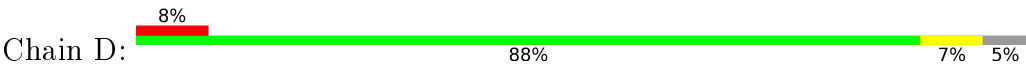


#### • Molecule 1: Cytochrome P450 2D6





● Molecule 1: Cytochrome P450 2D6





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.47Å 192.56Å 247.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.24 – 2.71 39.24 – 2.71	Depositor EDS
% Data completeness (in resolution range)	99.4 (39.24-2.71) 99.5 (39.24-2.71)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.95 (at 2.73Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.211 , 0.245 0.212 , 0.246	Depositor DCC
$R_{free}$ test set	3760 reflections (4.98%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.1	Xtriage
Anisotropy	0.568	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 46.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	14779	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	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 23.09 % 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 4.9974e-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, 0.333 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: HEM, ZN, P6U

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	0.22	0/3790	0.38	0/5155
1	B	0.22	0/3739	0.37	0/5083
1	C	0.22	0/3637	0.37	0/4947
1	D	0.22	0/3682	0.37	0/5007
All	All	0.22	0/14848	0.37	0/20192

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	3690	0	3679	25	0
1	B	3642	0	3618	24	0
1	C	3542	0	3513	29	0
1	D	3586	0	3570	18	0
2	A	43	0	30	5	0
2	B	43	0	30	4	0
2	C	43	0	30	6	0
2	D	43	0	30	5	0
3	A	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	25	0	0	0	0
4	B	25	0	0	1	0
4	C	25	0	0	0	0
4	D	25	0	0	0	0
5	A	13	0	0	1	0
5	B	12	0	0	0	0
5	C	10	0	0	0	0
5	D	7	0	0	1	0
All	All	14779	0	14500	103	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 4.

All (103) 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:186:ILE:HG21	1:A:303:PHE:HA	1.60	0.80
1:B:186:ILE:HG21	1:B:303:PHE:HA	1.67	0.75
1:C:122:ALA:O	1:C:441:ARG:NH2	2.22	0.72
1:D:191:CYS:SG	5:D:707:HOH:O	2.51	0.68
1:B:122:ALA:O	1:B:441:ARG:NH2	2.28	0.67
1:A:122:ALA:O	1:A:441:ARG:NH2	2.23	0.66
1:B:441:ARG:NH1	2:B:601:HEM:O2D	2.32	0.63
1:D:369:ILE:HG13	1:D:370:VAL:HG23	1.81	0.61
1:B:213:LEU:HD22	1:B:308:VAL:HG21	1.81	0.60
1:C:369:ILE:HG13	1:C:370:VAL:HG23	1.84	0.59
1:A:369:ILE:HG13	1:A:370:VAL:HG23	1.85	0.59
1:C:477:HIS:CD2	1:C:477:HIS:H	2.21	0.59
1:B:64:ARG:HH22	1:C:53:ASN:HD22	1.50	0.58
1:D:490:TYR:HE1	1:D:492:LEU:HD23	1.69	0.58
2:A:601:HEM:HMB2	2:A:601:HEM:HBB2	1.86	0.57
1:A:240:VAL:HG12	1:A:241:LEU:HG	1.87	0.57
1:C:490:TYR:HE1	1:C:492:LEU:HD23	1.72	0.55
1:D:147:LYS:HG2	1:D:150:GLU:HB2	1.88	0.55
1:A:43:LEU:HD23	1:A:46:LEU:HD22	1.87	0.55
1:C:441:ARG:NH1	2:C:601:HEM:O2D	2.40	0.55
2:C:601:HEM:HMC2	2:C:601:HEM:HBC2	1.89	0.55
1:B:216:GLU:HA	1:B:221:ARG:HD2	1.88	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:601:HEM:HBC2	2:B:601:HEM:HMC2	1.89	0.55
2:D:601:HEM:HMC2	2:D:601:HEM:HBC2	1.90	0.54
1:B:369:ILE:HG13	1:B:370:VAL:HG23	1.90	0.52
1:C:213:LEU:HD22	1:C:308:VAL:HG21	1.92	0.52
1:A:109:ILE:HD11	1:A:231:LEU:HD22	1.92	0.51
1:A:224:LEU:HD21	1:A:236:LEU:HD21	1.93	0.51
1:C:35:PRO:HG2	1:C:65:PHE:HB3	1.91	0.51
1:B:381:ASP:OD1	1:B:391:LYS:N	2.43	0.50
1:C:38:LEU:HD12	1:C:39:PRO:HD2	1.92	0.50
1:C:152:TRP:CD1	1:C:188:SER:HB3	2.46	0.50
1:C:91:LEU:HD11	1:C:397:THR:HG21	1.93	0.49
1:A:140:ARG:HG2	1:A:144:LEU:HD23	1.94	0.49
1:D:213:LEU:HD22	1:D:308:VAL:HG21	1.94	0.49
1:C:224:LEU:HD11	1:C:240:VAL:HG11	1.95	0.48
1:D:91:LEU:HD11	1:D:397:THR:HG21	1.95	0.48
1:A:79:VAL:HG21	1:A:389:ILE:HD12	1.94	0.48
1:B:149:LEU:HD11	1:B:448:LEU:HD13	1.94	0.48
1:A:490:TYR:HE1	1:A:492:LEU:HD23	1.78	0.48
1:C:326:ASP:OD1	1:C:326:ASP:N	2.46	0.48
2:D:601:HEM:HBB2	2:D:601:HEM:HMB2	1.94	0.48
1:C:437:SER:OG	1:C:438:ALA:N	2.42	0.48
1:B:59:ASP:HA	1:B:62:ARG:HG2	1.97	0.47
1:A:241:LEU:HB3	1:A:244:GLN:HB2	1.95	0.47
2:B:601:HEM:HMB2	2:B:601:HEM:HBB2	1.95	0.47
2:A:601:HEM:HMC2	2:A:601:HEM:HBC2	1.96	0.46
1:D:59:ASP:O	1:D:63:ARG:HG2	2.15	0.46
1:C:445:GLY:HA3	2:C:601:HEM:C3C	2.49	0.46
1:B:227:VAL:HG22	1:B:230:LEU:H	1.81	0.46
1:C:183:SER:OG	1:C:307:MET:SD	2.66	0.46
1:B:149:LEU:HD13	1:B:189:LEU:HD21	1.97	0.46
1:D:445:GLY:HA3	2:D:601:HEM:C3C	2.50	0.46
1:A:106:ILE:HG22	1:A:231:LEU:HD11	1.97	0.46
1:C:73:LEU:HD12	1:C:78:VAL:HG21	1.98	0.46
1:A:91:LEU:HD11	1:A:397:THR:HG21	1.97	0.45
1:D:239:LYS:HD2	1:D:242:ARG:HD3	1.97	0.45
1:A:50:ASP:OD1	1:A:50:ASP:N	2.44	0.45
1:A:236:LEU:O	1:A:240:VAL:HG23	2.17	0.44
1:B:320:LEU:HD13	1:B:415:PHE:CD1	2.51	0.44
1:D:104:VAL:HG12	1:D:106:ILE:HG22	1.99	0.44
1:B:38:LEU:HD12	1:B:39:PRO:HD2	1.98	0.44
1:B:448:LEU:HD23	2:B:601:HEM:HBC2	2.00	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:308:VAL:HG22	1:B:484:LEU:HD22	1.99	0.44
1:D:267:PRO:HA	1:D:268:PRO:HD3	1.91	0.44
1:C:160:LEU:HD13	1:C:181:ALA:HB2	1.98	0.44
1:B:139:LEU:HD22	1:B:189:LEU:HD13	2.00	0.43
1:C:162:ALA:O	1:C:166:ASN:ND2	2.48	0.43
1:C:320:LEU:HD13	1:C:415:PHE:CD1	2.53	0.43
1:C:448:LEU:HD23	2:C:601:HEM:HBC2	2.00	0.43
1:A:436:PHE:HB3	1:A:443:CYS:HB3	1.99	0.43
1:D:382:ILE:HG13	1:D:389:ILE:HB	1.99	0.43
1:C:107:THR:HB	1:C:112:PHE:CD2	2.54	0.43
1:D:320:LEU:HD13	1:D:415:PHE:CD1	2.54	0.43
1:C:344:ARG:HA	1:C:345:PRO:HD3	1.91	0.43
1:D:316:TRP:CD2	1:D:487:PRO:HG3	2.53	0.42
2:C:601:HEM:HMB2	2:C:601:HEM:HBB2	2.01	0.42
1:A:267:PRO:HA	1:A:268:PRO:HD3	1.93	0.42
1:C:233:ILE:HA	1:C:234:PRO:HD3	1.92	0.42
1:C:437:SER:HB3	2:C:601:HEM:HBA1	2.01	0.42
1:A:437:SER:HB3	2:A:601:HEM:HBA1	2.02	0.42
1:A:191:CYS:SG	5:A:713:HOH:O	2.53	0.42
1:A:320:LEU:HD13	1:A:415:PHE:CE1	2.55	0.42
1:B:304:SER:HB2	4:B:603:P6U:C10	2.50	0.42
1:A:113:GLY:O	1:A:117:GLN:HG2	2.19	0.42
1:D:241:LEU:HB3	1:D:244:GLN:HB2	2.01	0.42
1:B:436:PHE:HB3	1:B:443:CYS:HB3	2.02	0.41
1:D:104:VAL:HG13	1:D:225:ASN:HD21	1.85	0.41
1:A:443:CYS:HB2	2:A:601:HEM:NA	2.34	0.41
1:D:437:SER:HB3	2:D:601:HEM:HBA1	2.02	0.41
1:C:436:PHE:HB3	1:C:443:CYS:HB3	2.02	0.41
1:A:50:ASP:C	1:A:52:GLN:H	2.24	0.41
1:D:369:ILE:HD11	2:D:601:HEM:HMB2	2.02	0.41
1:B:224:LEU:HD11	1:B:236:LEU:HD23	2.03	0.41
1:B:233:ILE:HG22	1:B:235:ALA:H	1.85	0.41
1:B:199:ASP:HA	1:B:200:PRO:HD3	1.88	0.41
1:B:350:GLN:HE21	1:B:357:THR:HG23	1.86	0.41
1:C:382:ILE:HG13	1:C:389:ILE:HB	2.02	0.41
1:A:109:ILE:HG22	1:A:245:LYS:HE3	2.02	0.40
1:C:202:PHE:O	1:C:206:LEU:HG	2.21	0.40
1:A:445:GLY:HA3	2:A:601:HEM:C3C	2.57	0.40
1:B:316:TRP:CD2	1:B:487:PRO:HG3	2.55	0.40
1:C:331:VAL:O	1:C:335:ILE:HG13	2.22	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	465/479 (97%)	452 (97%)	13 (3%)	0	100	100
1	B	454/479 (95%)	442 (97%)	12 (3%)	0	100	100
1	C	441/479 (92%)	432 (98%)	9 (2%)	0	100	100
1	D	449/479 (94%)	435 (97%)	14 (3%)	0	100	100
All	All	1809/1916 (94%)	1761 (97%)	48 (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	398/409 (97%)	397 (100%)	1 (0%)	94	98
1	B	393/409 (96%)	393 (100%)	0	100	100
1	C	382/409 (93%)	381 (100%)	1 (0%)	94	98
1	D	387/409 (95%)	385 (100%)	2 (0%)	92	97
All	All	1560/1636 (95%)	1556 (100%)	4 (0%)	94	98

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	LEU
1	C	477	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	170	ARG
1	D	227	VAL

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

Mol	Chain	Res	Type
1	C	341	GLN
1	C	477	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 5 are monoatomic - leaving 8 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	$\# Z  > 2$
2	HEM	A	601	1	24,50,50	1.83	5 (20%)	16,82,82	1.22	0
4	P6U	A	604	-	24,27,27	0.26	0	24,43,43	1.39	2 (8%)
2	HEM	B	601	1	24,50,50	1.83	5 (20%)	16,82,82	1.29	0
4	P6U	B	603	-	24,27,27	0.26	0	24,43,43	1.37	2 (8%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	HEM	C	601	1	24,50,50	1.82	5 (20%)	16,82,82	1.29	2 (12%)
4	P6U	C	603	-	24,27,27	0.24	0	24,43,43	1.33	2 (8%)
2	HEM	D	601	1	24,50,50	1.80	5 (20%)	16,82,82	1.35	3 (18%)
4	P6U	D	603	-	24,27,27	0.24	0	24,43,43	1.34	2 (8%)

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	HEM	A	601	1	-	0/6/54/54	0/0/8/8
4	P6U	A	604	-	-	0/14/37/37	0/2/3/3
2	HEM	B	601	1	-	0/6/54/54	0/0/8/8
4	P6U	B	603	-	-	0/14/37/37	0/2/3/3
2	HEM	C	601	1	-	0/6/54/54	0/0/8/8
4	P6U	C	603	-	-	0/14/37/37	0/2/3/3
2	HEM	D	601	1	-	0/6/54/54	0/0/8/8
4	P6U	D	603	-	-	0/14/37/37	0/2/3/3

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	601	HEM	C3C-C2C	-4.06	1.35	1.40
2	B	601	HEM	C3B-C2B	-4.06	1.35	1.40
2	A	601	HEM	C3C-C2C	-4.03	1.35	1.40
2	C	601	HEM	C3B-C2B	-4.03	1.35	1.40
2	B	601	HEM	C3C-C2C	-3.95	1.35	1.40
2	A	601	HEM	C3B-C2B	-3.93	1.35	1.40
2	D	601	HEM	C3C-C2C	-3.92	1.35	1.40
2	D	601	HEM	C3B-C2B	-3.85	1.35	1.40
2	B	601	HEM	CAD-C3D	2.06	1.54	1.52
2	C	601	HEM	CAD-C3D	2.07	1.54	1.52
2	D	601	HEM	CAD-C3D	2.08	1.54	1.52
2	A	601	HEM	CAD-C3D	2.28	1.55	1.52
2	A	601	HEM	C3B-CAB	3.62	1.55	1.47
2	C	601	HEM	C3B-CAB	3.66	1.55	1.47
2	B	601	HEM	C3B-CAB	3.67	1.55	1.47
2	D	601	HEM	C3B-CAB	3.72	1.55	1.47
2	C	601	HEM	C3C-CAC	3.79	1.55	1.47
2	D	601	HEM	C3C-CAC	3.84	1.55	1.47

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	601	HEM	C3C-CAC	3.85	1.55	1.47
2	A	601	HEM	C3C-CAC	3.87	1.55	1.47

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	601	HEM	CBD-CAD-C3D	-2.22	108.57	112.47
2	D	601	HEM	CAA-CBA-CGA	-2.20	108.51	112.78
2	C	601	HEM	CAA-CBA-CGA	-2.14	108.62	112.78
2	C	601	HEM	CBD-CAD-C3D	-2.07	108.84	112.47
2	D	601	HEM	CMB-C2B-C3B	2.12	129.24	125.09
4	D	603	P6U	C14-C13-N3	3.30	120.44	115.47
4	C	603	P6U	C14-C13-N3	3.32	120.48	115.47
4	B	603	P6U	C14-C13-N3	3.55	120.83	115.47
4	A	604	P6U	C14-C13-N3	3.80	121.20	115.47
4	A	604	P6U	C15-C13-N3	4.95	122.93	115.47
4	C	603	P6U	C15-C13-N3	5.00	123.01	115.47
4	D	603	P6U	C15-C13-N3	5.07	123.11	115.47
4	B	603	P6U	C15-C13-N3	5.10	123.16	115.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	HEM	5	0
2	B	601	HEM	4	0
4	B	603	P6U	1	0
2	C	601	HEM	6	0
2	D	601	HEM	5	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	467/479 (97%)	0.42	42 (8%) 12 10	21, 42, 101, 115	0
1	B	460/479 (96%)	0.25	25 (5%) 29 28	21, 42, 92, 118	0
1	C	447/479 (93%)	0.42	37 (8%) 14 12	25, 45, 95, 119	0
1	D	453/479 (94%)	0.44	39 (8%) 13 11	25, 49, 95, 119	0
All	All	1827/1916 (95%)	0.38	143 (7%) 16 14	21, 45, 97, 119	0

All (143) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	75	TRP	9.1
1	C	231	LEU	9.1
1	C	228	PRO	8.5
1	C	219	PHE	8.0
1	A	43	LEU	7.4
1	A	228	PRO	7.4
1	A	38	LEU	7.1
1	B	237	ALA	6.9
1	A	42	GLY	6.6
1	A	223	VAL	6.3
1	C	237	ALA	6.2
1	A	231	LEU	6.1
1	D	235	ALA	6.0
1	A	47	LEU	5.8
1	A	144	LEU	5.8
1	A	40	LEU	5.7
1	A	219	PHE	5.7
1	B	229	VAL	5.6
1	D	219	PHE	5.5
1	B	47	LEU	5.5
1	C	229	VAL	5.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	236	LEU	5.3
1	C	75	TRP	5.3
1	C	38	LEU	5.3
1	A	75	TRP	5.1
1	C	234	PRO	5.1
1	C	236	LEU	5.0
1	A	48	HIS	4.8
1	D	238	GLY	4.8
1	C	238	GLY	4.8
1	C	33	LEU	4.7
1	D	229	VAL	4.6
1	D	231	LEU	4.6
1	B	46	LEU	4.6
1	A	41	PRO	4.5
1	D	234	PRO	4.5
1	D	38	LEU	4.5
1	C	224	LEU	4.3
1	B	48	HIS	4.3
1	C	223	VAL	4.3
1	A	234	PRO	4.2
1	B	45	ASN	4.2
1	A	32	LYS	4.1
1	B	52	GLN	4.1
1	B	219	PHE	4.0
1	A	53	ASN	4.0
1	A	229	VAL	4.0
1	B	342	VAL	4.0
1	C	37	PRO	4.0
1	D	220	LEU	3.9
1	A	237	ALA	3.9
1	A	235	ALA	3.8
1	B	49	VAL	3.8
1	C	35	PRO	3.8
1	B	75	TRP	3.8
1	D	228	PRO	3.6
1	A	226	ALA	3.6
1	D	72	GLN	3.5
1	A	342	VAL	3.4
1	A	232	HIS	3.4
1	B	38	LEU	3.4
1	D	237	ALA	3.4
1	D	380	ARG	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	32	LYS	3.3
1	D	33	LEU	3.3
1	A	145	GLY	3.3
1	D	37	PRO	3.2
1	D	149	LEU	3.2
1	B	228	PRO	3.2
1	C	220	LEU	3.2
1	D	145	GLY	3.2
1	A	33	LEU	3.1
1	D	223	VAL	3.1
1	C	232	HIS	3.1
1	D	389	ILE	3.1
1	A	74	ALA	3.1
1	A	39	PRO	3.1
1	A	225	ASN	3.0
1	B	234	PRO	3.0
1	A	44	GLY	3.0
1	C	342	VAL	3.0
1	C	230	LEU	3.0
1	C	233	ILE	3.0
1	A	36	GLY	3.0
1	D	144	LEU	2.9
1	C	240	VAL	2.9
1	B	39	PRO	2.9
1	D	65	PHE	2.8
1	A	233	ILE	2.8
1	A	37	PRO	2.8
1	D	168	SER	2.8
1	A	224	LEU	2.8
1	D	77	PRO	2.8
1	A	146	LYS	2.8
1	D	226	ALA	2.7
1	D	232	HIS	2.7
1	B	51	PHE	2.7
1	C	235	ALA	2.7
1	A	238	GLY	2.7
1	A	51	PHE	2.6
1	A	147	LYS	2.6
1	C	73	LEU	2.5
1	D	342	VAL	2.5
1	C	39	PRO	2.5
1	A	483	PHE	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	232	HIS	2.5
1	D	146	LYS	2.5
1	A	149	LEU	2.4
1	B	73	LEU	2.4
1	C	109	ILE	2.4
1	C	72	GLN	2.3
1	D	73	LEU	2.3
1	A	45	ASN	2.3
1	A	46	LEU	2.3
1	C	423	ALA	2.3
1	B	238	GLY	2.2
1	C	170	ARG	2.2
1	B	240	VAL	2.2
1	C	388	ARG	2.2
1	C	60	GLN	2.2
1	C	310	THR	2.2
1	C	341	GLN	2.2
1	D	74	ALA	2.2
1	D	36	GLY	2.2
1	B	231	LEU	2.2
1	C	74	ALA	2.2
1	A	31	GLY	2.1
1	C	227	VAL	2.1
1	C	63	ARG	2.1
1	D	423	ALA	2.1
1	D	339	ILE	2.1
1	D	215	GLU	2.1
1	B	307	MET	2.1
1	C	65	PHE	2.1
1	D	35	PRO	2.1
1	D	230	LEU	2.1
1	D	224	LEU	2.0
1	D	39	PRO	2.0
1	C	110	LEU	2.0
1	D	241	LEU	2.0
1	A	49	VAL	2.0
1	B	223	VAL	2.0
1	B	36	GLY	2.0

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [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
4	P6U	A	604	25/25	0.87	0.24	1.11	56,74,86,96	0
2	HEM	D	601	43/43	0.97	0.22	0.63	22,33,43,44	0
2	HEM	B	601	43/43	0.98	0.21	0.50	17,25,31,36	0
4	P6U	B	603	25/25	0.89	0.23	0.44	51,71,84,93	0
2	HEM	A	601	43/43	0.97	0.20	0.19	18,25,29,31	0
2	HEM	C	601	43/43	0.98	0.21	0.17	24,31,40,42	0
4	P6U	D	603	25/25	0.91	0.21	-0.02	50,75,86,92	0
4	P6U	C	603	25/25	0.90	0.19	-0.18	58,76,84,93	0
3	ZN	A	603	1/1	0.98	0.15	-0.96	57,57,57,57	0
3	ZN	A	602	1/1	0.99	0.14	-1.14	36,36,36,36	0
3	ZN	B	602	1/1	0.99	0.10	-1.60	26,26,26,26	0
3	ZN	C	602	1/1	1.00	0.13	-1.65	35,35,35,35	0
3	ZN	D	602	1/1	0.99	0.09	-4.07	34,34,34,34	0

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

There are no such residues in this entry.