



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 03:15 pm GMT

PDB ID : 1GWS
Title : HEXADECAHEME HIGH MOLECULAR WEIGHT CYTOCHROME HMC
FROM DESULFOVIBRIO VULGARIS HILDENBOROUGH
Authors : Czjzek, M.; Haser, R.; Bruschi, M.
Deposited on : 2002-03-25
Resolution : 2.40 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

<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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

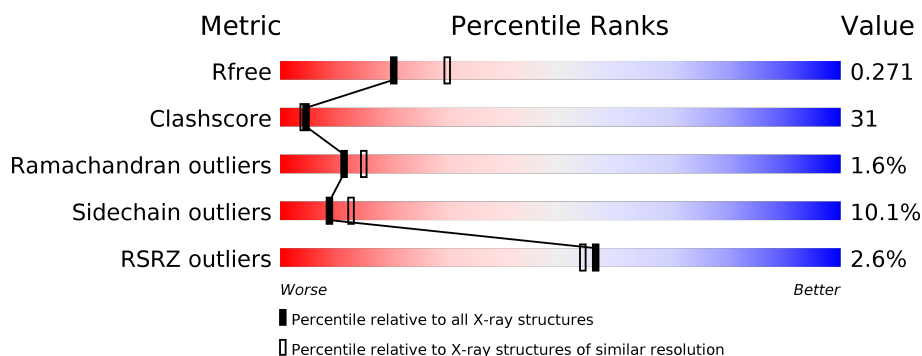
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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}	100719	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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 ≥ 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	545	<div> <div>2%</div> <div>57%</div> <div>28%</div> <div>6%</div> <div>8%</div> </div>

2 Entry composition [i](#)

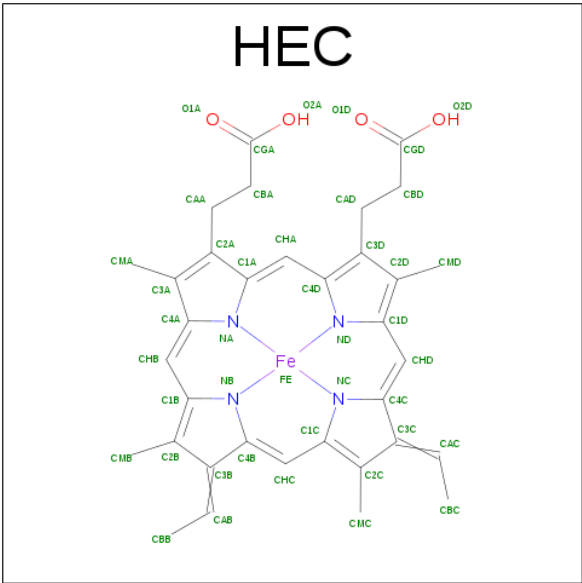
There are 4 unique types of molecules in this entry. The entry contains 4632 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 HIGH-MOLECULAR-WEIGHT CYTOCHROME C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	503	Total	C	N	O	S	0	0	0
			3793	2331	715	703	44			

- Molecule 2 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).



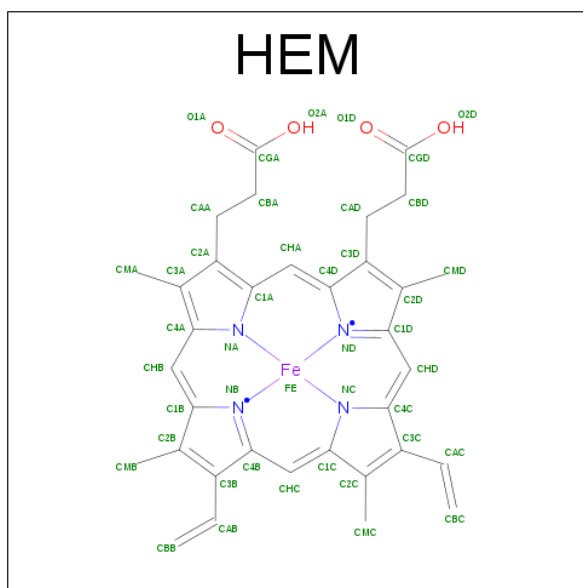
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

Continued on next page...

Continued from previous page...

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

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	Fe	N	O	
			43	34	1	4	4	
3	A	1	Total	C	Fe	N	O	
			43	34	1	4	4	
3	A	1	Total	C	Fe	N	O	
			43	34	1	4	4	
3	A	1	Total	C	Fe	N	O	
			43	34	1	4	4	

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	151	Total 151	O 151	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 62	Depositor
Cell constants a, b, c, α , β , γ	108.39Å 108.39Å 102.81Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.52 – 2.40 33.54 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (33.52-2.40) 99.0 (33.54-2.40)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.51 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.0.36	Depositor
R, R_{free}	0.201 , 0.276 0.199 , 0.271	Depositor DCC
R_{free} test set	1281 reflections (5.06%)	DCC
Wilson B-factor (Å ²)	46.9	Xtriage
Anisotropy	0.280	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 35.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.044 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4632	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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, HEC

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	1.16	12/3876 (0.3%)	1.29	34/5221 (0.7%)

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	31	VAL	CB-CG1	6.94	1.67	1.52
1	A	489	MET	SD-CE	-5.97	1.44	1.77
1	A	14	LEU	CG-CD2	5.66	1.72	1.51
1	A	483	ALA	CA-CB	-5.58	1.40	1.52
1	A	306	GLU	CD-OE1	5.56	1.31	1.25

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	155	ASP	CB-CG-OD2	9.55	126.89	118.30
1	A	240	ASP	CB-CG-OD2	9.47	126.83	118.30
1	A	394	ASP	CB-CG-OD2	9.06	126.45	118.30
1	A	236	ARG	C-N-CA	-8.73	103.97	122.30
1	A	69	ASP	CB-CG-OD2	8.61	126.05	118.30

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	3793	0	3721	220	0
2	A	516	0	367	95	0
3	A	172	0	120	30	0
4	A	151	0	0	9	1
All	All	4632	0	4208	267	1

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 31.

The worst 5 of 267 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:321:CYS:SG	2:A:610:HEC:CAC	2.10	1.39
1:A:508:CYS:SG	2:A:616:HEC:CAC	2.10	1.39
1:A:449:CYS:SG	2:A:613:HEC:CAC	2.10	1.39
1:A:280:CYS:SG	3:A:608:HEM:CAC	2.12	1.37
1:A:107:CYS:SG	2:A:603:HEC:CAC	2.18	1.31

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:2028:HOH:O	4:A:2136:HOH:O[4_675]	0.64	1.56

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	501/545 (92%)	460 (92%)	33 (7%)	8 (2%)	11 15

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	112	PRO
1	A	181	ASP
1	A	257	LYS
1	A	477	ARG
1	A	115	ALA

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	405/435 (93%)	364 (90%)	41 (10%)	9 12

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	263	VAL
1	A	325	ARG
1	A	470	PHE
1	A	284	ARG
1	A	297	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	311	GLN
1	A	323	ASN
1	A	452	ASN
1	A	282	HIS
1	A	447	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

16 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	HEC	A	601	1	28,50,50	2.54	10 (35%)	16,82,82	1.79	3 (18%)
3	HEM	A	602	1	28,50,50	2.56	11 (39%)	17,82,82	2.35	5 (29%)
2	HEC	A	603	1	28,50,50	2.92	9 (32%)	16,82,82	2.38	10 (62%)
3	HEM	A	604	1	28,50,50	2.27	8 (28%)	17,82,82	1.97	2 (11%)
2	HEC	A	605	1	28,50,50	2.68	8 (28%)	16,82,82	1.44	2 (12%)
2	HEC	A	606	1	28,50,50	2.90	7 (25%)	16,82,82	2.15	6 (37%)
3	HEM	A	607	1	28,50,50	2.43	10 (35%)	17,82,82	2.37	7 (41%)
3	HEM	A	608	1	28,50,50	2.67	8 (28%)	17,82,82	2.08	6 (35%)
2	HEC	A	609	1	28,50,50	2.96	9 (32%)	16,82,82	2.92	5 (31%)
2	HEC	A	610	1	28,50,50	2.53	12 (42%)	16,82,82	2.54	5 (31%)
2	HEC	A	611	1	28,50,50	3.25	10 (35%)	16,82,82	3.02	6 (37%)
2	HEC	A	612	1	28,50,50	2.48	10 (35%)	16,82,82	2.82	6 (37%)
2	HEC	A	613	1	28,50,50	2.68	11 (39%)	16,82,82	2.28	5 (31%)
2	HEC	A	614	1	28,50,50	2.46	8 (28%)	16,82,82	1.93	5 (31%)
2	HEC	A	615	1	28,50,50	2.60	9 (32%)	16,82,82	3.19	7 (43%)
2	HEC	A	616	1	28,50,50	2.99	10 (35%)	16,82,82	2.39	4 (25%)

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	HEC	A	601	1	-	0/6/54/54	0/0/8/8
3	HEM	A	602	1	-	0/6/54/54	0/0/8/8
2	HEC	A	603	1	-	0/6/54/54	0/0/8/8
3	HEM	A	604	1	-	0/6/54/54	0/0/8/8
2	HEC	A	605	1	-	0/6/54/54	0/0/8/8
2	HEC	A	606	1	-	0/6/54/54	0/0/8/8
3	HEM	A	607	1	-	0/6/54/54	0/0/8/8
3	HEM	A	608	1	-	0/6/54/54	0/0/8/8
2	HEC	A	609	1	-	0/6/54/54	0/0/8/8
2	HEC	A	610	1	-	0/6/54/54	0/0/8/8
2	HEC	A	611	1	-	0/6/54/54	0/0/8/8
2	HEC	A	612	1	-	0/6/54/54	0/0/8/8
2	HEC	A	613	1	-	0/6/54/54	0/0/8/8
2	HEC	A	614	1	-	0/6/54/54	0/0/8/8
2	HEC	A	615	1	-	0/6/54/54	0/0/8/8
2	HEC	A	616	1	-	0/6/54/54	0/0/8/8

The worst 5 of 150 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	609	HEC	C3C-C2C	-11.53	1.28	1.40
2	A	611	HEC	C3B-C2B	-9.98	1.30	1.40
2	A	606	HEC	C3B-C2B	-9.28	1.30	1.40
2	A	616	HEC	C3C-C2C	-9.19	1.31	1.40
2	A	603	HEC	C3C-C2C	-8.82	1.31	1.40

The worst 5 of 84 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	611	HEC	CAA-CBA-CGA	-8.82	97.59	112.66
2	A	609	HEC	CAD-CBD-CGD	-8.43	98.26	112.66
3	A	604	HEM	CBA-CAA-C2A	-6.83	99.42	112.48
2	A	615	HEC	CMC-C2C-C1C	-6.58	118.35	128.46
3	A	602	HEM	C1D-C2D-C3D	-6.30	102.61	107.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

16 monomers are involved in 125 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	HEC	5	0
3	A	602	HEM	3	0
2	A	603	HEC	10	0
3	A	604	HEM	9	0
2	A	605	HEC	6	0
2	A	606	HEC	9	0
3	A	607	HEM	7	0
3	A	608	HEM	11	0
2	A	609	HEC	7	0
2	A	610	HEC	10	0
2	A	611	HEC	6	0
2	A	612	HEC	7	0
2	A	613	HEC	12	0
2	A	614	HEC	11	0
2	A	615	HEC	6	0
2	A	616	HEC	7	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, 95th 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	503/545 (92%)	-0.15	13 (2%) 56 54	19, 33, 64, 91	11 (2%)

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	469	PRO	10.2
1	A	470	PHE	9.1
1	A	468	LYS	6.0
1	A	473	ASP	5.4
1	A	474	ARG	5.3

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [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, 95th 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 ²)	Q<0.9
2	HEC	A	615	43/43	0.94	0.17	0.68	16,23,42,46	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	HEM	A	602	43/43	0.98	0.13	0.51	15,21,31,42	0
2	HEC	A	606	43/43	0.97	0.15	0.50	16,24,34,35	0
3	HEM	A	607	43/43	0.97	0.15	0.26	20,26,40,43	0
3	HEM	A	608	43/43	0.98	0.12	0.05	9,21,27,30	0
3	HEM	A	604	43/43	0.98	0.13	0.03	22,27,33,37	0
2	HEC	A	610	43/43	0.97	0.12	0.03	13,21,32,38	0
2	HEC	A	605	43/43	0.96	0.14	-0.03	19,27,41,53	0
2	HEC	A	614	43/43	0.95	0.15	-0.09	17,27,40,44	0
2	HEC	A	613	43/43	0.96	0.14	-0.14	11,25,39,43	0
2	HEC	A	609	43/43	0.97	0.13	-0.17	14,27,50,58	0
2	HEC	A	611	43/43	0.98	0.12	-0.19	13,20,25,28	0
2	HEC	A	616	43/43	0.98	0.12	-0.30	12,25,31,34	0
2	HEC	A	612	43/43	0.98	0.11	-0.43	15,22,29,35	0
2	HEC	A	601	43/43	0.98	0.11	-0.61	6,22,30,42	0
2	HEC	A	603	43/43	0.97	0.11	-0.71	16,22,25,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.