



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 02:21 am GMT

PDB ID : 1GU6
Title : STRUCTURE OF THE PERIPLASMIC CYTOCHROME C NITRITE REDUCTASE FROM ESCHERICHIA COLI
Authors : Bamford, V.A.; Angove, H.C.; Seward, H.E.; Thomson, A.J.; Cole, J.A.; Butt, J.N.; Hemmings, A.M.; Richardson, D.J.
Deposited on : 2002-01-24
Resolution : 2.50 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)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

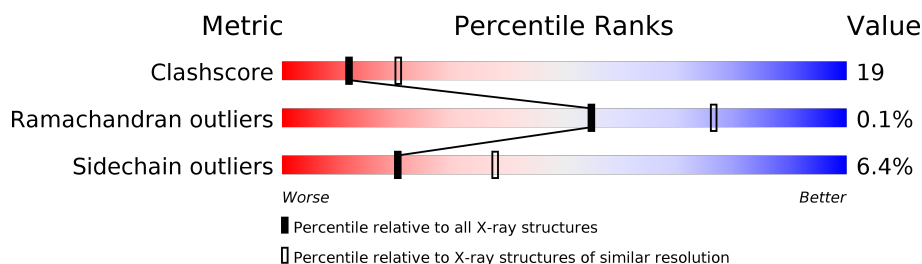
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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))
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	452	 71% 24% . .
1	C	452	 68% 25% . .
1	E	452	 71% 23% . .
1	G	452	 69% 25% . .

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	A	1481	-	X	-	-
3	GOL	C	1481	-	X	-	-
3	GOL	E	1481	-	X	-	-
3	GOL	G	1481	-	X	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15486 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.

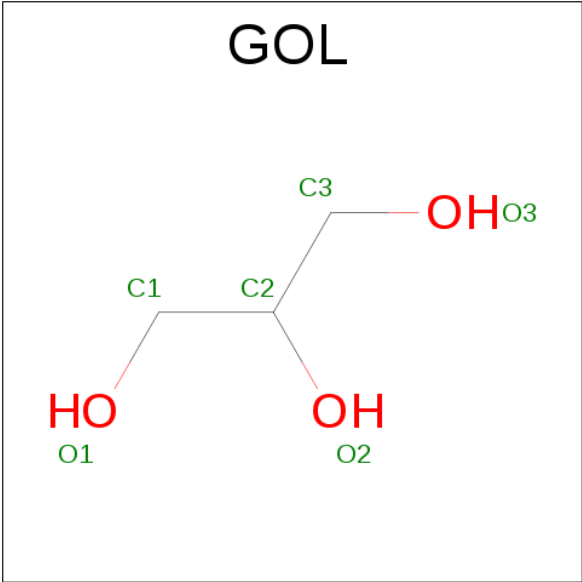
- Molecule 1 is a protein called CYTOCHROME C552.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	441	Total	C	N	O	S	0	0	0
			3479	2179	620	658	22			
1	C	441	Total	C	N	O	S	0	0	0
			3479	2179	620	658	22			
1	E	441	Total	C	N	O	S	0	0	0
			3479	2179	620	658	22			
1	G	441	Total	C	N	O	S	0	0	0
			3479	2179	620	658	22			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

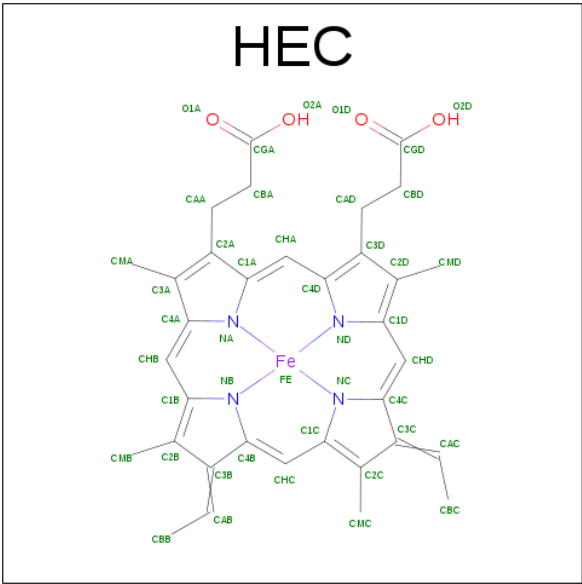
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	2	Total	Ca	0	0
			2	2		
2	A	2	Total	Ca	0	0
			2	2		
2	C	2	Total	Ca	0	0
			2	2		
2	E	2	Total	Ca	0	0
			2	2		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	G	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is HEME C (three-letter code: HEC) (formula: C₃₄H₃₄FeN₄O₄).



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

- Molecule 5 is water.

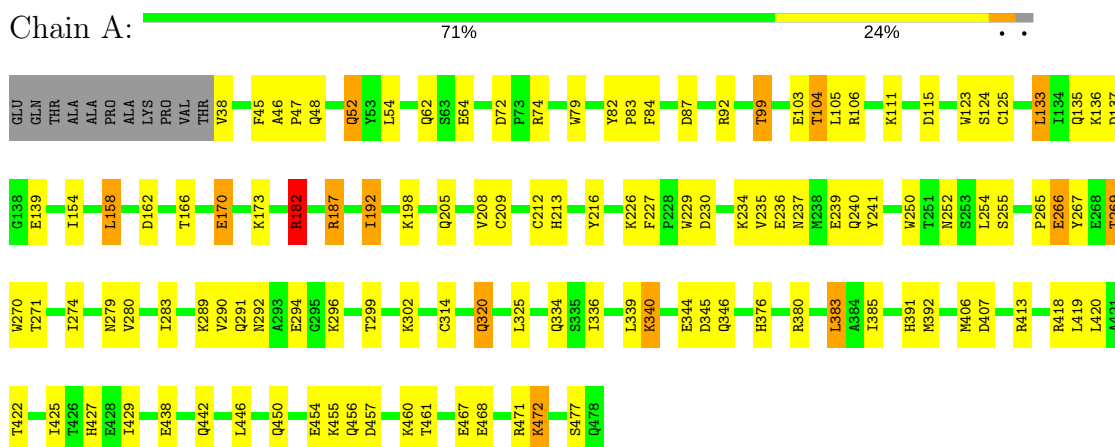
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	163	Total 163	O 163	0	0
5	C	121	Total 121	O 121	0	0
5	E	191	Total 191	O 191	0	0
5	G	203	Total 203	O 203	0	0

3 Residue-property plots

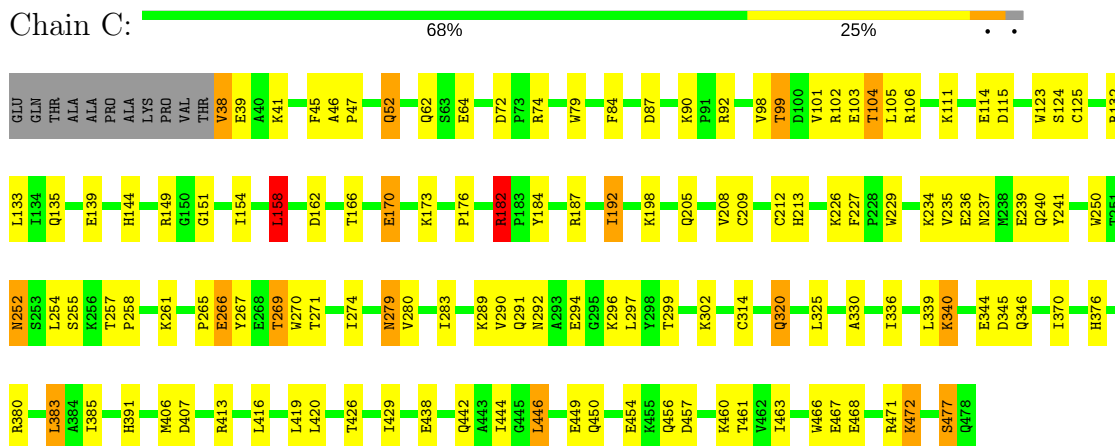
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

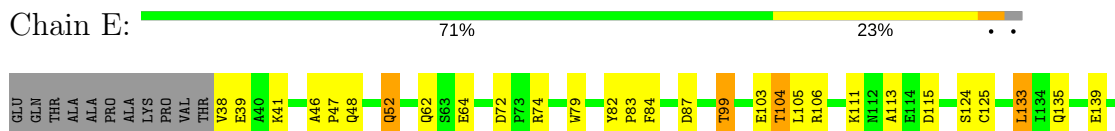
• Molecule 1: CYTOCHROME C552



• Molecule 1: CYTOCHROME C552



• Molecule 1: CYTOCHROME C552



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	81.47Å 90.84Å 293.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50	Depositor
% Data completeness (in resolution range)	91.5 (20.00-2.50)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.203 , 0.243	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	15486	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

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, CA, 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	0.75	0/3563	0.73	4/4817 (0.1%)
1	C	0.82	3/3563 (0.1%)	0.73	3/4817 (0.1%)
1	E	0.76	0/3563	0.75	3/4817 (0.1%)
1	G	0.76	2/3563 (0.1%)	0.74	3/4817 (0.1%)
All	All	0.77	5/14252 (0.0%)	0.74	13/19268 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	G	0	1
All	All	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	449	GLU	CD-OE1	6.10	1.32	1.25
1	G	449	GLU	CD-OE2	5.87	1.32	1.25
1	C	438	GLU	CG-CD	5.79	1.60	1.51
1	C	449	GLU	CD-OE2	5.29	1.31	1.25
1	G	468	GLU	CG-CD	5.08	1.59	1.51

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	158	LEU	CA-CB-CG	7.84	133.34	115.30
1	E	158	LEU	CA-CB-CG	7.75	133.12	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	158	LEU	CA-CB-CG	7.44	132.42	115.30
1	A	158	LEU	CA-CB-CG	7.35	132.21	115.30
1	E	182	ARG	NE-CZ-NH2	-6.67	116.97	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	216	TYR	Sidechain
1	G	216	TYR	Sidechain

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	3479	0	3365	134	0
1	C	3479	0	3365	135	0
1	E	3479	0	3365	130	0
1	G	3479	0	3365	142	0
2	A	2	0	0	0	0
2	C	2	0	0	0	0
2	E	2	0	0	0	0
2	G	2	0	0	0	0
3	A	6	0	5	3	0
3	C	6	0	5	3	0
3	E	6	0	4	3	0
3	G	6	0	5	3	0
4	A	215	0	150	18	0
4	C	215	0	150	17	0
4	E	215	0	150	15	0
4	G	215	0	150	15	0
5	A	163	0	0	15	0
5	C	121	0	0	20	0
5	E	191	0	0	24	0
5	G	203	0	0	28	0
All	All	15486	0	14079	556	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 19.

The worst 5 of 556 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)
3:A:1481:GOL:C1	3:A:1481:GOL:O1	1.66	1.44
3:G:1481:GOL:C1	3:G:1481:GOL:O1	1.65	1.43
3:E:1481:GOL:O1	3:E:1481:GOL:C1	1.65	1.43
3:C:1481:GOL:O1	3:C:1481:GOL:C1	1.68	1.42
1:E:240:GLN:HB2	5:E:2102:HOH:O	1.54	1.07

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	439/452 (97%)	421 (96%)	18 (4%)	0	100	100
1	C	439/452 (97%)	421 (96%)	18 (4%)	0	100	100
1	E	439/452 (97%)	421 (96%)	17 (4%)	1 (0%)	51	73
1	G	439/452 (97%)	419 (95%)	20 (5%)	0	100	100
All	All	1756/1808 (97%)	1682 (96%)	73 (4%)	1 (0%)	55	76

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	113	ALA

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	362/370 (98%)	339 (94%)	23 (6%)	20	38
1	C	362/370 (98%)	338 (93%)	24 (7%)	19	36
1	E	362/370 (98%)	338 (93%)	24 (7%)	19	36
1	G	362/370 (98%)	340 (94%)	22 (6%)	22	40
All	All	1448/1480 (98%)	1355 (94%)	93 (6%)	20	38

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	383	LEU
1	E	135	GLN
1	G	320	GLN
1	C	419	LEU
1	E	38	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 65 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	388	HIS
1	E	252	ASN
1	G	346	GLN
1	C	391	HIS
1	E	62	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

Of 32 ligands modelled in this entry, 8 are monoatomic - leaving 24 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
3	GOL	A	1481	-	5,5,5	4.72	5 (100%)	5,5,5	5.59	3 (60%)
4	HEC	A	1482	1,5	28,50,50	1.94	5 (17%)	16,82,82	1.78	6 (37%)
4	HEC	A	1483	1	28,50,50	1.67	6 (21%)	16,82,82	1.95	6 (37%)
4	HEC	A	1484	1,2	28,50,50	2.03	2 (7%)	16,82,82	2.10	7 (43%)
4	HEC	A	1485	1,2	28,50,50	1.71	3 (10%)	16,82,82	1.84	5 (31%)
4	HEC	A	1486	1	28,50,50	1.83	3 (10%)	16,82,82	1.78	5 (31%)
3	GOL	C	1481	-	5,5,5	4.78	3 (60%)	5,5,5	5.57	3 (60%)
4	HEC	C	1482	1,5	28,50,50	1.64	2 (7%)	16,82,82	1.82	7 (43%)
4	HEC	C	1483	1	28,50,50	1.73	5 (17%)	16,82,82	1.93	6 (37%)
4	HEC	C	1484	1,2	28,50,50	1.76	2 (7%)	16,82,82	1.96	8 (50%)
4	HEC	C	1485	1,2	28,50,50	2.06	4 (14%)	16,82,82	1.90	7 (43%)
4	HEC	C	1486	1	28,50,50	1.72	3 (10%)	16,82,82	1.87	6 (37%)
3	GOL	E	1481	-	5,5,5	4.69	5 (100%)	5,5,5	5.56	3 (60%)
4	HEC	E	1482	1,5	28,50,50	1.76	5 (17%)	16,82,82	1.80	7 (43%)
4	HEC	E	1483	1	28,50,50	1.71	5 (17%)	16,82,82	1.97	7 (43%)
4	HEC	E	1484	1,2	28,50,50	1.81	3 (10%)	16,82,82	2.15	7 (43%)
4	HEC	E	1485	1,2	28,50,50	1.77	3 (10%)	16,82,82	1.88	6 (37%)
4	HEC	E	1486	1	28,50,50	1.86	3 (10%)	16,82,82	1.82	5 (31%)
3	GOL	G	1481	-	5,5,5	4.74	5 (100%)	5,5,5	5.55	3 (60%)
4	HEC	G	1482	1,5	28,50,50	1.49	4 (14%)	16,82,82	1.83	7 (43%)
4	HEC	G	1483	1	28,50,50	1.85	5 (17%)	16,82,82	1.93	7 (43%)
4	HEC	G	1484	1,2	28,50,50	1.99	2 (7%)	16,82,82	2.14	6 (37%)
4	HEC	G	1485	1,2	28,50,50	1.96	3 (10%)	16,82,82	1.83	5 (31%)
4	HEC	G	1486	1	28,50,50	1.94	3 (10%)	16,82,82	1.83	5 (31%)

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	GOL	A	1481	-	-	0/4/4/4	0/0/0/0
4	HEC	A	1482	1,5	-	0/6/54/54	0/0/8/8
4	HEC	A	1483	1	-	0/6/54/54	0/0/8/8
4	HEC	A	1484	1,2	-	0/6/54/54	0/0/8/8
4	HEC	A	1485	1,2	-	0/6/54/54	0/0/8/8
4	HEC	A	1486	1	-	0/6/54/54	0/0/8/8
3	GOL	C	1481	-	-	0/4/4/4	0/0/0/0
4	HEC	C	1482	1,5	-	0/6/54/54	0/0/8/8
4	HEC	C	1483	1	-	0/6/54/54	0/0/8/8
4	HEC	C	1484	1,2	-	0/6/54/54	0/0/8/8
4	HEC	C	1485	1,2	-	0/6/54/54	0/0/8/8
4	HEC	C	1486	1	-	0/6/54/54	0/0/8/8
3	GOL	E	1481	-	-	0/4/4/4	0/0/0/0
4	HEC	E	1482	1,5	-	0/6/54/54	0/0/8/8
4	HEC	E	1483	1	-	0/6/54/54	0/0/8/8
4	HEC	E	1484	1,2	-	0/6/54/54	0/0/8/8
4	HEC	E	1485	1,2	-	0/6/54/54	0/0/8/8
4	HEC	E	1486	1	-	0/6/54/54	0/0/8/8
3	GOL	G	1481	-	-	0/4/4/4	0/0/0/0
4	HEC	G	1482	1,5	-	0/6/54/54	0/0/8/8
4	HEC	G	1483	1	-	0/6/54/54	0/0/8/8
4	HEC	G	1484	1,2	-	0/6/54/54	0/0/8/8
4	HEC	G	1485	1,2	-	0/6/54/54	0/0/8/8
4	HEC	G	1486	1	-	0/6/54/54	0/0/8/8

The worst 5 of 89 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1484	HEC	C3C-C2C	-8.23	1.32	1.40
3	C	1481	GOL	C3-C2	-7.78	1.23	1.52
3	G	1481	GOL	C3-C2	-7.78	1.23	1.52
3	A	1481	GOL	C3-C2	-7.66	1.23	1.52
3	E	1481	GOL	C3-C2	-7.51	1.24	1.52

The worst 5 of 137 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	1486	HEC	CMD-C2D-C1D	-3.80	122.62	128.46
4	E	1486	HEC	CMD-C2D-C1D	-3.53	123.03	128.46
4	A	1486	HEC	CMD-C2D-C1D	-3.48	123.12	128.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	1486	HEC	CMD-C2D-C1D	-3.48	123.12	128.46
4	G	1485	HEC	CMC-C2C-C1C	-3.46	123.15	128.46

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

23 monomers are involved in 73 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1481	GOL	3	0
4	A	1482	HEC	1	0
4	A	1483	HEC	3	0
4	A	1484	HEC	1	0
4	A	1485	HEC	5	0
4	A	1486	HEC	8	0
3	C	1481	GOL	3	0
4	C	1482	HEC	1	0
4	C	1483	HEC	3	0
4	C	1484	HEC	2	0
4	C	1485	HEC	2	0
4	C	1486	HEC	9	0
3	E	1481	GOL	3	0
4	E	1482	HEC	1	0
4	E	1483	HEC	3	0
4	E	1485	HEC	5	0
4	E	1486	HEC	6	0
3	G	1481	GOL	3	0
4	G	1482	HEC	2	0
4	G	1483	HEC	3	0
4	G	1484	HEC	1	0
4	G	1485	HEC	3	0
4	G	1486	HEC	6	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](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.