



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 30, 2021 – 01:11 pm GMT

PDB ID : 7PRP
Title : Crystal Structure of the B subunit of heat labile enterotoxin LT-IIc from Escherichia coli in apo form
Authors : Varrot, A.
Deposited on : 2021-09-22
Resolution : 2.30 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/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	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

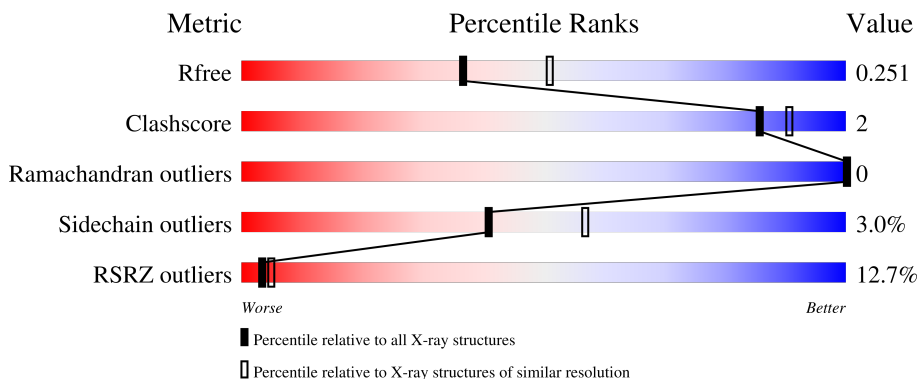
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	DDD	104	<div> <div>12%</div> <div>91%</div> <div>9%</div> </div>
1	EEE	104	<div> <div>12%</div> <div>91%</div> <div>9%</div> </div>
1	FFF	104	<div> <div>16%</div> <div>93%</div> <div>7%</div> </div>
1	GGG	104	<div> <div>12%</div> <div>94%</div> <div>6%</div> </div>
1	HHH	104	<div> <div>10%</div> <div>96%</div> <div>.</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4261 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 Heat-labile enterotoxin IIA, B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	DDD	104	Total	C	N	O	S	0	0	0
			799	503	141	151	4			
1	EEE	104	Total	C	N	O	S	0	0	0
			805	507	143	151	4			
1	FFF	104	Total	C	N	O	S	0	0	0
			798	502	141	151	4			
1	GGG	104	Total	C	N	O	S	0	0	0
			798	502	141	151	4			
1	HHH	104	Total	C	N	O	S	0	0	0
			801	504	142	151	4			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
DDD	99	HIS	-	expression tag	UNP H6W8F2
DDD	100	HIS	-	expression tag	UNP H6W8F2
DDD	101	HIS	-	expression tag	UNP H6W8F2
DDD	102	HIS	-	expression tag	UNP H6W8F2
DDD	103	HIS	-	expression tag	UNP H6W8F2
DDD	104	HIS	-	expression tag	UNP H6W8F2
EEE	99	HIS	-	expression tag	UNP H6W8F2
EEE	100	HIS	-	expression tag	UNP H6W8F2
EEE	101	HIS	-	expression tag	UNP H6W8F2
EEE	102	HIS	-	expression tag	UNP H6W8F2
EEE	103	HIS	-	expression tag	UNP H6W8F2
EEE	104	HIS	-	expression tag	UNP H6W8F2
FFF	99	HIS	-	expression tag	UNP H6W8F2
FFF	100	HIS	-	expression tag	UNP H6W8F2
FFF	101	HIS	-	expression tag	UNP H6W8F2
FFF	102	HIS	-	expression tag	UNP H6W8F2
FFF	103	HIS	-	expression tag	UNP H6W8F2
FFF	104	HIS	-	expression tag	UNP H6W8F2
GGG	99	HIS	-	expression tag	UNP H6W8F2

Continued on next page...

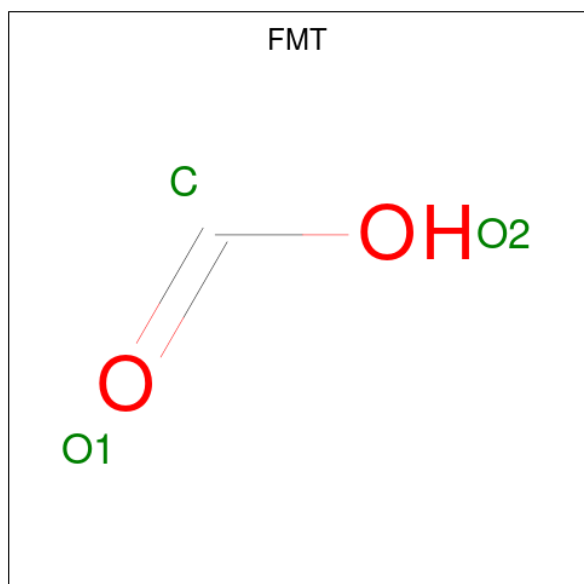
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
GGG	100	HIS	-	expression tag	UNP H6W8F2
GGG	101	HIS	-	expression tag	UNP H6W8F2
GGG	102	HIS	-	expression tag	UNP H6W8F2
GGG	103	HIS	-	expression tag	UNP H6W8F2
GGG	104	HIS	-	expression tag	UNP H6W8F2
HHH	99	HIS	-	expression tag	UNP H6W8F2
HHH	100	HIS	-	expression tag	UNP H6W8F2
HHH	101	HIS	-	expression tag	UNP H6W8F2
HHH	102	HIS	-	expression tag	UNP H6W8F2
HHH	103	HIS	-	expression tag	UNP H6W8F2
HHH	104	HIS	-	expression tag	UNP H6W8F2

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	DDD	1	Total Na 1 1	0	0
2	EEE	1	Total Na 1 1	0	0
2	FFF	1	Total Na 1 1	0	0
2	GGG	2	Total Na 2 2	0	0
2	HHH	2	Total Na 2 2	0	0

- Molecule 3 is FORMIC ACID (three-letter code: FMT) (formula: CH₂O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	EEE	1	Total	C	O	0	0
			3	1	2		
3	EEE	1	Total	C	O	0	0
			3	1	2		

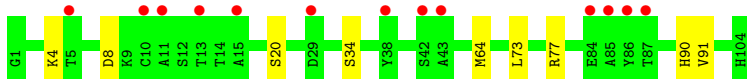
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	DDD	46	Total	O	0	3
			49	49		
4	EEE	47	Total	O	0	1
			48	48		
4	FFF	36	Total	O	0	3
			39	39		
4	GGG	51	Total	O	0	5
			56	56		
4	HHH	51	Total	O	0	4
			55	55		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

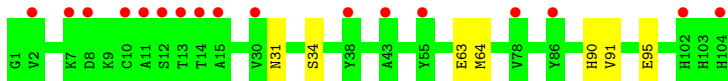
- Molecule 1: Heat-labile enterotoxin IIA, B chain



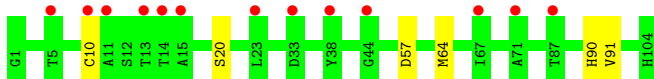
- Molecule 1: Heat-labile enterotoxin IIA, B chain



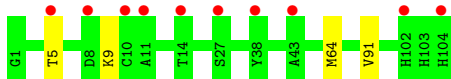
- Molecule 1: Heat-labile enterotoxin IIA, B chain



- Molecule 1: Heat-labile enterotoxin IIA, B chain



- Molecule 1: Heat-labile enterotoxin IIA, B chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	73.40Å 73.69Å 102.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.98 – 2.30 41.98 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.9 (41.98-2.30) 97.9 (41.98-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.72 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.196 , 0.245 0.202 , 0.251	Depositor DCC
R_{free} test set	1266 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å ²)	28.5	Xtriage
Anisotropy	0.159	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4261	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.71% 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: NA, FMT

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	DDD	0.77	0/820	0.84	0/1116
1	EEE	0.74	0/826	0.88	0/1122
1	FFF	0.81	2/819 (0.2%)	0.85	0/1115
1	GGG	0.78	0/819	0.82	0/1115
1	HHH	0.78	0/822	0.82	0/1118
All	All	0.77	2/4106 (0.0%)	0.84	0/5586

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	FFF	95	GLU	CD-OE2	7.10	1.33	1.25
1	FFF	63	GLU	CD-OE1	5.41	1.31	1.25

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	DDD	799	0	762	5	0
1	EEE	805	0	780	5	0
1	FFF	798	0	760	3	0
1	GGG	798	0	760	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	HHH	801	0	769	4	0
2	DDD	1	0	0	0	0
2	EEE	1	0	0	0	0
2	FFF	1	0	0	0	0
2	GGG	2	0	0	0	0
2	HHH	2	0	0	0	0
3	EEE	6	0	2	0	0
4	DDD	49	0	0	0	0
4	EEE	48	0	0	0	0
4	FFF	39	0	0	0	0
4	GGG	56	0	0	0	0
4	HHH	55	0	0	0	0
All	All	4261	0	3833	17	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 2.

All (17) 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:DDD:64:MET:HE1	1:DDD:91:VAL:HG12	1.75	0.69
1:DDD:64:MET:CE	1:DDD:91:VAL:HG12	2.25	0.67
1:EEE:64:MET:CE	1:EEE:91:VAL:HG12	2.25	0.66
1:HHH:64:MET:CE	1:HHH:91:VAL:HG12	2.25	0.66
1:GGG:64:MET:CE	1:GGG:91:VAL:HG12	2.27	0.64
1:FFF:64:MET:CE	1:FFF:91:VAL:HG12	2.32	0.59
1:EEE:64:MET:HE3	1:EEE:91:VAL:HG12	1.85	0.58
1:FFF:64:MET:HE3	1:FFF:91:VAL:HG12	1.89	0.54
1:GGG:64:MET:HE3	1:GGG:91:VAL:HG12	1.91	0.53
1:DDD:4:LYS:HE3	1:DDD:8:ASP:OD2	2.09	0.52
1:HHH:64:MET:HE3	1:HHH:91:VAL:HG12	1.91	0.51
1:DDD:64:MET:CE	1:DDD:91:VAL:CG1	2.95	0.45
1:DDD:34:SER:HB2	1:HHH:9:LYS:HD3	1.99	0.44
1:EEE:45:LYS:HG2	1:EEE:47:TRP:CZ2	2.54	0.42
1:EEE:13:THR:HG21	1:FFF:31:ASN:CG	2.41	0.41
1:HHH:5:THR:O	1:HHH:9:LYS:HG3	2.20	0.41
1:EEE:73:LEU:HA	1:EEE:73:LEU:HD23	1.90	0.40

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	DDD	102/104 (98%)	101 (99%)	1 (1%)	0	100	100
1	EEE	102/104 (98%)	102 (100%)	0	0	100	100
1	FFF	102/104 (98%)	102 (100%)	0	0	100	100
1	GGG	102/104 (98%)	102 (100%)	0	0	100	100
1	HHH	102/104 (98%)	102 (100%)	0	0	100	100
All	All	510/520 (98%)	509 (100%)	1 (0%)	0	100	100

There are no Ramachandran outliers to report.

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	DDD	86/89 (97%)	82 (95%)	4 (5%)	26	37
1	EEE	88/89 (99%)	85 (97%)	3 (3%)	37	51
1	FFF	86/89 (97%)	84 (98%)	2 (2%)	50	67
1	GGG	86/89 (97%)	82 (95%)	4 (5%)	26	37
1	HHH	87/89 (98%)	87 (100%)	0	100	100
All	All	433/445 (97%)	420 (97%)	13 (3%)	41	57

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	DDD	20	SER
1	DDD	73	LEU
1	DDD	77	ARG
1	DDD	90	HIS
1	EEE	12	SER
1	EEE	57	ASP
1	EEE	77	ARG
1	FFF	34	SER
1	FFF	90	HIS
1	GGG	10	CYS
1	GGG	20	SER
1	GGG	57	ASP
1	GGG	90	HIS

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

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 7 are monoatomic - leaving 2 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	DDD	104/104 (100%)	1.19	13 (12%) 3 5	18, 26, 44, 63	0
1	EEE	104/104 (100%)	1.12	13 (12%) 3 5	17, 28, 47, 58	0
1	FFF	104/104 (100%)	1.29	17 (16%) 1 2	17, 27, 49, 69	1 (0%)
1	GGG	104/104 (100%)	1.15	13 (12%) 3 5	17, 26, 50, 75	0
1	HHH	104/104 (100%)	1.11	10 (9%) 8 10	17, 28, 48, 69	0
All	All	520/520 (100%)	1.17	66 (12%) 3 5	17, 27, 51, 75	1 (0%)

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	DDD	85	ALA	5.7
1	DDD	11	ALA	5.5
1	GGG	38	TYR	4.9
1	FFF	10	CYS	4.5
1	GGG	11	ALA	4.5
1	FFF	13	THR	4.3
1	EEE	2	VAL	4.1
1	GGG	33	ASP	3.9
1	FFF	14	THR	3.9
1	FFF	30	VAL	3.9
1	HHH	5	THR	3.8
1	EEE	38	TYR	3.8
1	FFF	38	TYR	3.8
1	DDD	13	THR	3.7
1	GGG	10	CYS	3.6
1	HHH	104	HIS	3.6
1	FFF	86	TYR	3.6
1	EEE	13	THR	3.5
1	FFF	11	ALA	3.5
1	GGG	14	THR	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	HHH	11	ALA	3.4
1	EEE	39	ILE	3.3
1	DDD	10	CYS	3.3
1	GGG	13	THR	3.3
1	FFF	8	ASP	3.2
1	DDD	43	ALA	3.2
1	HHH	14	THR	3.0
1	EEE	10	CYS	3.0
1	DDD	38	TYR	2.9
1	GGG	5	THR	2.9
1	EEE	11	ALA	2.9
1	GGG	23	LEU	2.8
1	FFF	15	ALA	2.8
1	GGG	15	ALA	2.7
1	HHH	102	HIS	2.7
1	EEE	33	ASP	2.7
1	DDD	86	TYR	2.7
1	FFF	2	VAL	2.7
1	HHH	10	CYS	2.6
1	EEE	86	TYR	2.6
1	EEE	101	HIS	2.6
1	EEE	43	ALA	2.6
1	FFF	12	SER	2.5
1	DDD	87	THR	2.5
1	FFF	43	ALA	2.5
1	FFF	104	HIS	2.5
1	HHH	38	TYR	2.5
1	DDD	84	GLU	2.4
1	DDD	42	SER	2.4
1	EEE	14	THR	2.3
1	HHH	8	ASP	2.3
1	HHH	43	ALA	2.3
1	FFF	102	HIS	2.3
1	GGG	87	THR	2.3
1	DDD	15	ALA	2.2
1	DDD	29	ASP	2.2
1	DDD	5	THR	2.2
1	FFF	55	TYR	2.1
1	EEE	85	ALA	2.1
1	GGG	44	GLY	2.1
1	FFF	7	LYS	2.1
1	HHH	27	SER	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	FFF	78	VAL	2.1
1	EEE	67	ILE	2.0
1	GGG	67	ILE	2.0
1	GGG	71	ALA	2.0

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 monosaccharides 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. 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	B-factors(\AA^2)	Q<0.9
2	NA	HHH	201	1/1	0.75	0.12	41,41,41,41	0
3	FMT	EEE	201	3/3	0.81	0.27	45,45,51,53	0
2	NA	EEE	203	1/1	0.89	0.10	39,39,39,39	0
2	NA	HHH	202	1/1	0.90	0.11	40,40,40,40	0
2	NA	GGG	202	1/1	0.90	0.26	37,37,37,37	0
3	FMT	EEE	202	3/3	0.90	0.15	40,40,42,43	0
2	NA	FFF	201	1/1	0.94	0.14	45,45,45,45	0
2	NA	DDD	201	1/1	0.94	0.17	33,33,33,33	0
2	NA	GGG	201	1/1	0.98	0.14	37,37,37,37	0

6.5 Other polymers [i](#)

There are no such residues in this entry.