



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 11:58 pm GMT

PDB ID : 4C0R  
Title : Molecular and structural basis of glutathione import in Gram-positive bacteria via GshT and the cystine ABC importer TcyBC of Streptococcus mutans.  
Authors : Vergauwen, B.; Verstraete, K.; Senadheera, D.B.; Dansercoer, A.; Cvitkovitch, D.G.; Guedon, E.; Savvides, S.N.  
Deposited on : 2013-08-07  
Resolution : 1.55 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](mailto: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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

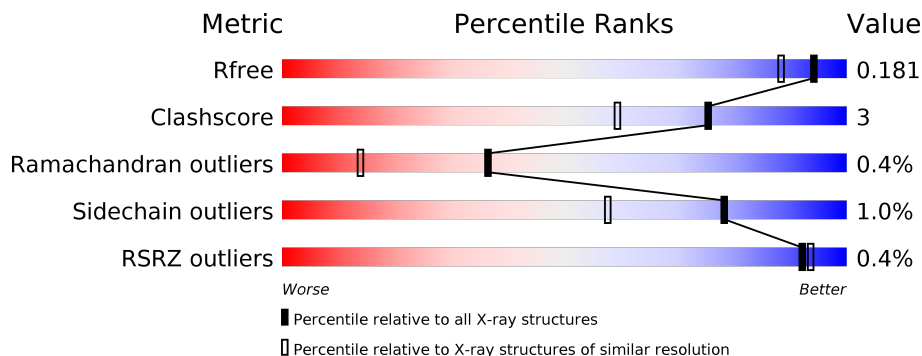
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.55 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))
$R_{free}$	111664	2055 (1.56-1.52)
Clashscore	122126	2124 (1.56-1.52)
Ramachandran outliers	120053	2083 (1.56-1.52)
Sidechain outliers	120020	2081 (1.56-1.52)
RSRZ outliers	108989	2027 (1.56-1.52)

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	248	
1	B	248	

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
2	GDS	A	500	X	-	-	-
2	GDS	B	500	X	-	-	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8081 atoms, of which 3569 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.

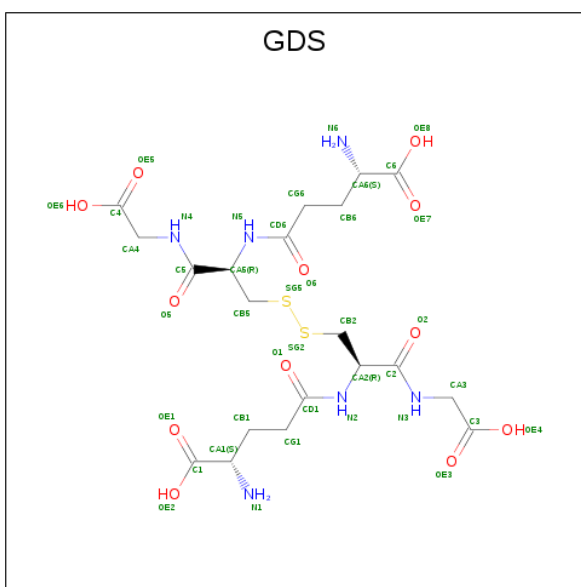
- Molecule 1 is a protein called PUTATIVE AMINO ACID BINDING PROTEIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	239	Total	C	H	N	O	S	0	5	0
			3674	1200	1793	303	375	3			
1	B	238	Total	C	H	N	O	S	0	5	0
			3586	1184	1738	302	360	2			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP M2L4D8
A	2	ASP	-	expression tag	UNP M2L4D8
A	241	LEU	-	expression tag	UNP M2L4D8
A	242	GLU	-	expression tag	UNP M2L4D8
A	243	HIS	-	expression tag	UNP M2L4D8
A	244	HIS	-	expression tag	UNP M2L4D8
A	245	HIS	-	expression tag	UNP M2L4D8
A	246	HIS	-	expression tag	UNP M2L4D8
A	247	HIS	-	expression tag	UNP M2L4D8
A	248	HIS	-	expression tag	UNP M2L4D8
B	1	MET	-	expression tag	UNP M2L4D8
B	2	ASP	-	expression tag	UNP M2L4D8
B	241	LEU	-	expression tag	UNP M2L4D8
B	242	GLU	-	expression tag	UNP M2L4D8
B	243	HIS	-	expression tag	UNP M2L4D8
B	244	HIS	-	expression tag	UNP M2L4D8
B	245	HIS	-	expression tag	UNP M2L4D8
B	246	HIS	-	expression tag	UNP M2L4D8
B	247	HIS	-	expression tag	UNP M2L4D8
B	248	HIS	-	expression tag	UNP M2L4D8

- Molecule 2 is OXIDIZED GLUTATHIONE DISULFIDE (three-letter code: GDS) (formula:  $C_{20}H_{32}N_6O_{12}S_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
2	A	1	Total 50	C 16	H 18	N 5	O 9	S 2	0	0
2	B	1	Total 54	C 17	H 20	N 5	O 10	S 2	0	0

- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	18	Total Cd 18 18	0	0
3	A	22	Total Cd 22 22	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0
4	A	5	Total Cl 5 5	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	359	Total O 359 359	0	0

*Continued on next page...*


*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	312	Total 312	O 312	0	0

### 3 Residue-property plots [i](#)

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.

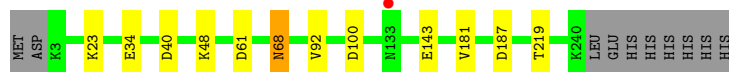
#### ● Molecule 1: PUTATIVE AMINO ACID BINDING PROTEIN

Chain A:  90% 6% .



#### ● Molecule 1: PUTATIVE AMINO ACID BINDING PROTEIN

Chain B:  91% . .



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	37.04Å 103.12Å 64.09Å 90.00° 92.31° 90.00°	Depositor
Resolution (Å)	37.01 – 1.55 37.01 – 1.55	Depositor EDS
% Data completeness (in resolution range)	98.9 (37.01-1.55) 98.9 (37.01-1.55)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.48 (at 1.55Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.155 , 0.183 0.154 , 0.181	Depositor DCC
$R_{free}$ test set	3458 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	13.0	Xtriage
Anisotropy	0.032	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 60.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.079 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8081	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.29% of the height of the origin peak. No significant pseudotranslation is detected.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL, GDS, CD

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	0.81	1/1935 (0.1%)	0.93	5/2626 (0.2%)
1	B	0.77	3/1903 (0.2%)	0.82	2/2589 (0.1%)
All	All	0.79	4/3838 (0.1%)	0.87	7/5215 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	100	ASP	CG-OD2	9.79	1.47	1.25
1	A	100	ASP	CG-OD2	9.15	1.46	1.25
1	B	100	ASP	CA-CB	5.95	1.67	1.53
1	B	143	GLU	CB-CG	5.09	1.61	1.52

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	100	ASP	CB-CG-OD1	-15.91	103.98	118.30
1	A	100	ASP	CB-CG-OD2	13.94	130.85	118.30
1	B	100	ASP	CB-CG-OD1	-11.80	107.68	118.30
1	B	100	ASP	CB-CG-OD2	10.99	128.19	118.30
1	A	40	ASP	CB-CG-OD2	6.00	123.70	118.30

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	1881	1793	1785	15	0
1	B	1848	1738	1724	7	0
2	A	32	18	20	1	0
2	B	34	20	20	1	0
3	A	22	0	0	0	0
3	B	18	0	0	0	0
4	A	5	0	0	2	0
4	B	1	0	0	0	0
5	A	359	0	0	16	1
5	B	312	0	0	4	1
All	All	4512	3569	3549	25	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 3.

The worst 5 of 25 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:150[B]:ARG:HD3	5:A:2250:HOH:O	1.53	1.06
1:A:150[A]:ARG:HD3	5:A:2250:HOH:O	1.60	1.02
4:A:1261:CL:CL	5:A:2351:HOH:O	2.26	0.89
1:A:47:GLN:NE2	5:A:2098:HOH:O	1.84	0.89
1:A:98:ASP:O	5:A:2177:HOH:O	1.95	0.84

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)
5:A:2072:HOH:O	5:B:2150:HOH:O[1_556]	2.18	0.02

## 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	242/248 (98%)	237 (98%)	4 (2%)	1 (0%)	36	14
1	B	241/248 (97%)	234 (97%)	6 (2%)	1 (0%)	36	14
All	All	483/496 (97%)	471 (98%)	10 (2%)	2 (0%)	36	14

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	68	ASN
1	A	68	ASN

### 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	201/226 (89%)	198 (98%)	3 (2%)	67	38
1	B	191/226 (84%)	190 (100%)	1 (0%)	90	79
All	All	392/452 (87%)	388 (99%)	4 (1%)	78	58

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

Mol	Chain	Res	Type
1	A	61	ASP
1	A	68	ASN
1	A	177	ASP
1	B	68	ASN

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 [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 48 ligands modelled in this entry, 46 are monoatomic - leaving 2 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	GDS	A	500	3	20,31,39	2.13	6 (30%)	27,39,50	2.24	12 (44%)
2	GDS	B	500	3	22,33,39	2.16	6 (27%)	30,42,50	2.37	13 (43%)

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	GDS	A	500	3	1/1/12/16	0/32/40/51	0/0/0/0
2	GDS	B	500	3	1/1/13/16	0/34/42/51	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	500	GDS	CA5-N5	-2.97	1.39	1.45
2	B	500	GDS	CA5-C5	-2.64	1.45	1.52
2	A	500	GDS	CA5-N5	-2.33	1.40	1.45
2	A	500	GDS	CA5-C5	-2.32	1.46	1.52
2	A	500	GDS	CA2-N2	-2.00	1.42	1.46

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	A	500	GDS	O6-CD6-N5	-4.06	116.01	122.96
2	A	500	GDS	C3-CA3-N3	-3.50	103.51	110.24
2	B	500	GDS	O1-CD1-N2	-3.21	115.89	121.94
2	B	500	GDS	C3-CA3-N3	-3.17	104.14	110.24
2	A	500	GDS	CA4-N4-C5	-2.92	118.50	122.40

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	500	GDS	CA2
2	A	500	GDS	CA2

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	500	GDS	1	0
2	B	500	GDS	1	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	239/248 (96%)	-0.43	1 (0%) 92 93	7, 14, 23, 32	0
1	B	238/248 (95%)	-0.43	1 (0%) 92 93	8, 16, 25, 32	0
All	All	477/496 (96%)	-0.43	2 (0%) 92 93	7, 15, 24, 32	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	98	ASP	2.4
1	B	133	ASN	2.2

### 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
3	CD	A	1257	1/1	0.78	0.14	104,104,104,104	1
3	CD	A	1249	1/1	0.94	0.10	51,51,51,51	1
2	GDS	B	500	34/40	0.97	0.08	8,12,32,34	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CD	B	1257	1/1	0.97	0.06	50,50,50,50	1
3	CD	A	1255	1/1	0.97	0.07	43,43,43,43	1
2	GDS	A	500	32/40	0.97	0.08	7,11,25,30	0
4	CL	A	1260	1/1	0.97	0.09	19,19,19,19	0
3	CD	B	1252	1/1	0.98	0.11	32,32,32,32	1
3	CD	B	1259	1/1	0.98	0.13	20,20,20,20	1
3	CD	B	1249	1/1	0.98	0.07	31,31,31,31	1
4	CL	A	1261	1/1	0.98	0.06	29,29,29,29	0
3	CD	B	1256	1/1	0.98	0.05	43,43,43,43	1
3	CD	B	1255	1/1	0.98	0.07	39,39,39,39	1
3	CD	B	1251	1/1	0.98	0.05	24,24,24,24	1
4	CL	B	1253	1/1	0.98	0.07	19,19,19,19	0
4	CL	A	1262	1/1	0.98	0.08	15,15,15,15	0
3	CD	A	1254	1/1	0.99	0.07	13,13,13,13	1
3	CD	A	1265	1/1	0.99	0.09	17,17,17,17	1
3	CD	B	1246	1/1	0.99	0.09	14,14,14,14	1
3	CD	B	1258	1/1	0.99	0.11	19,19,19,19	1
4	CL	A	1258	1/1	0.99	0.04	16,16,16,16	0
3	CD	A	1253	1/1	0.99	0.09	17,17,17,17	1
3	CD	A	1256	1/1	0.99	0.10	26,26,26,26	1
3	CD	B	1250	1/1	0.99	0.08	33,33,33,33	1
3	CD	B	1248	1/1	0.99	0.07	29,29,29,29	1
3	CD	A	1266	1/1	0.99	0.10	29,29,29,29	1
3	CD	B	1247	1/1	0.99	0.07	13,13,13,13	1
4	CL	A	1259	1/1	0.99	0.04	13,13,13,13	0
3	CD	A	1267	1/1	0.99	0.11	25,25,25,25	1
3	CD	A	1251	1/1	0.99	0.08	21,21,21,21	1
3	CD	B	1243	1/1	1.00	0.05	15,15,15,15	1
3	CD	A	1250	1/1	1.00	0.06	21,21,21,21	1
3	CD	A	1245	1/1	1.00	0.04	19,19,19,19	1
3	CD	A	1246	1/1	1.00	0.07	12,12,12,12	1
3	CD	A	1240	1/1	1.00	0.06	10,10,10,10	1
3	CD	A	1243	1/1	1.00	0.04	19,19,19,19	1
3	CD	B	1245	1/1	1.00	0.08	15,15,15,15	1
3	CD	A	1247	1/1	1.00	0.10	10,10,10,10	1
3	CD	A	1242	1/1	1.00	0.03	15,15,15,15	1
3	CD	A	1252	1/1	1.00	0.07	12,12,12,12	1
3	CD	B	1242	1/1	1.00	0.05	15,15,15,15	0
3	CD	A	1263	1/1	1.00	0.09	13,13,13,13	1
3	CD	A	1244	1/1	1.00	0.04	15,15,15,15	1
3	CD	B	1241	1/1	1.00	0.07	14,14,14,14	0
3	CD	B	1254	1/1	1.00	0.04	18,18,18,18	1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CD	A	1248	1/1	1.00	0.08	22,22,22,22	1
3	CD	B	1244	1/1	1.00	0.03	25,25,25,25	1
3	CD	A	1241	1/1	1.00	0.05	14,14,14,14	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.