wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 03:34 pm BST

PDB ID	:	4D8U
Title	:	Crystal structure of D-Cysteine desulfhydrase from Salmonella typhimurium at 3.3 A in monoclinic space group with 8 subunits in the asymmetric unit
Authors	:	Bharath, S.R.; Shveta, B.; Rajesh, K.H.; Savithri, H.S.; Murthy, M.R.N.
Deposited on	:	2012-01-11
Resolution	:	3.30 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

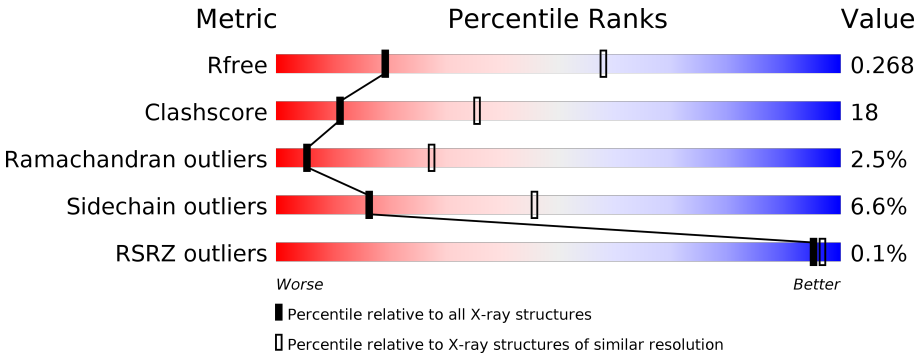
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 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 <=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	342	67% 26% • •
1	B	342	65% 26% • 6%
1	C	342	66% 24% • 6%
1	D	342	68% 24% • •
1	E	342	70% 20% 5% • •
1	F	342	69% 21% • 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	342	
1	H	342	

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	PO4	A	401	-	-	X	-
2	PO4	C	401	-	-	X	-
2	PO4	E	401	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 19156 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 D-cysteine desulfhydrase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	330	Total	C	N	O	P	S	0	0	0
			2466	1567	420	468	1	10			
1	B	323	Total	C	N	O	P	S	1	0	0
			2335	1485	388	452	1	9			
1	C	323	Total	C	N	O	P	S	0	0	0
			2361	1501	392	459	1	8			
1	D	329	Total	C	N	O	P	S	0	0	0
			2464	1567	420	466	1	10			
1	E	330	Total	C	N	O	P	S	0	0	0
			2450	1558	412	470	1	9			
1	F	318	Total	C	N	O	P	S	0	0	0
			2227	1405	375	438	1	8			
1	G	324	Total	C	N	O	P	S	0	0	0
			2341	1481	394	456	1	9			
1	H	331	Total	C	N	O	P	S	0	0	0
			2454	1559	418	466	1	10			

There are 112 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	EXPRESSION TAG	UNP Q8ZNT7
A	-12	ARG	-	EXPRESSION TAG	UNP Q8ZNT7
A	-11	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
A	-10	SER	-	EXPRESSION TAG	UNP Q8ZNT7
A	-9	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
A	-8	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
A	-7	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
A	-6	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
A	-5	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
A	-4	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
A	-3	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
A	-2	MET	-	EXPRESSION TAG	UNP Q8ZNT7
A	-1	ALA	-	EXPRESSION TAG	UNP Q8ZNT7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	EXPRESSION TAG	UNP Q8ZNT7
B	-13	MET	-	EXPRESSION TAG	UNP Q8ZNT7
B	-12	ARG	-	EXPRESSION TAG	UNP Q8ZNT7
B	-11	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
B	-10	SER	-	EXPRESSION TAG	UNP Q8ZNT7
B	-9	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
B	-8	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
B	-7	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
B	-6	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
B	-5	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
B	-4	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
B	-3	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
B	-2	MET	-	EXPRESSION TAG	UNP Q8ZNT7
B	-1	ALA	-	EXPRESSION TAG	UNP Q8ZNT7
B	0	SER	-	EXPRESSION TAG	UNP Q8ZNT7
C	-13	MET	-	EXPRESSION TAG	UNP Q8ZNT7
C	-12	ARG	-	EXPRESSION TAG	UNP Q8ZNT7
C	-11	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
C	-10	SER	-	EXPRESSION TAG	UNP Q8ZNT7
C	-9	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
C	-8	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
C	-7	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
C	-6	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
C	-5	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
C	-4	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
C	-3	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
C	-2	MET	-	EXPRESSION TAG	UNP Q8ZNT7
C	-1	ALA	-	EXPRESSION TAG	UNP Q8ZNT7
C	0	SER	-	EXPRESSION TAG	UNP Q8ZNT7
D	-13	MET	-	EXPRESSION TAG	UNP Q8ZNT7
D	-12	ARG	-	EXPRESSION TAG	UNP Q8ZNT7
D	-11	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
D	-10	SER	-	EXPRESSION TAG	UNP Q8ZNT7
D	-9	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
D	-8	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
D	-7	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
D	-6	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
D	-5	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
D	-4	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
D	-3	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
D	-2	MET	-	EXPRESSION TAG	UNP Q8ZNT7
D	-1	ALA	-	EXPRESSION TAG	UNP Q8ZNT7

Continued on next page...

Continued from previous page...

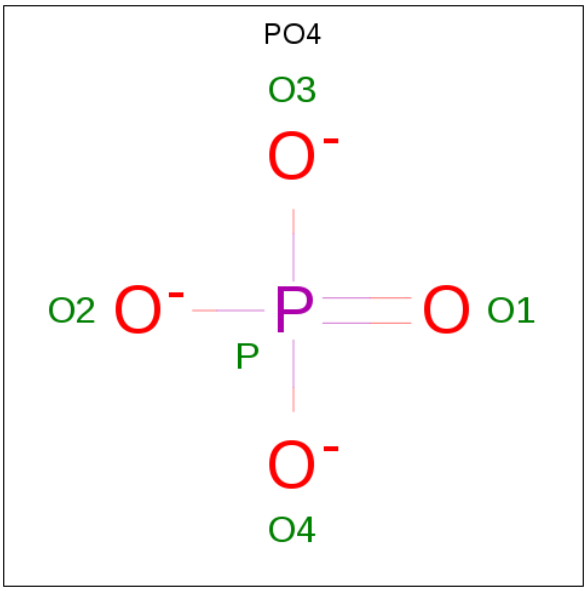
Chain	Residue	Modelled	Actual	Comment	Reference
D	0	SER	-	EXPRESSION TAG	UNP Q8ZNT7
E	-13	MET	-	EXPRESSION TAG	UNP Q8ZNT7
E	-12	ARG	-	EXPRESSION TAG	UNP Q8ZNT7
E	-11	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
E	-10	SER	-	EXPRESSION TAG	UNP Q8ZNT7
E	-9	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
E	-8	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
E	-7	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
E	-6	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
E	-5	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
E	-4	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
E	-3	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
E	-2	MET	-	EXPRESSION TAG	UNP Q8ZNT7
E	-1	ALA	-	EXPRESSION TAG	UNP Q8ZNT7
E	0	SER	-	EXPRESSION TAG	UNP Q8ZNT7
F	-13	MET	-	EXPRESSION TAG	UNP Q8ZNT7
F	-12	ARG	-	EXPRESSION TAG	UNP Q8ZNT7
F	-11	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
F	-10	SER	-	EXPRESSION TAG	UNP Q8ZNT7
F	-9	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
F	-8	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
F	-7	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
F	-6	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
F	-5	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
F	-4	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
F	-3	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
F	-2	MET	-	EXPRESSION TAG	UNP Q8ZNT7
F	-1	ALA	-	EXPRESSION TAG	UNP Q8ZNT7
F	0	SER	-	EXPRESSION TAG	UNP Q8ZNT7
G	-13	MET	-	EXPRESSION TAG	UNP Q8ZNT7
G	-12	ARG	-	EXPRESSION TAG	UNP Q8ZNT7
G	-11	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
G	-10	SER	-	EXPRESSION TAG	UNP Q8ZNT7
G	-9	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
G	-8	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
G	-7	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
G	-6	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
G	-5	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
G	-4	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
G	-3	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
G	-2	MET	-	EXPRESSION TAG	UNP Q8ZNT7
G	-1	ALA	-	EXPRESSION TAG	UNP Q8ZNT7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	0	SER	-	EXPRESSION TAG	UNP Q8ZNT7
H	-13	MET	-	EXPRESSION TAG	UNP Q8ZNT7
H	-12	ARG	-	EXPRESSION TAG	UNP Q8ZNT7
H	-11	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
H	-10	SER	-	EXPRESSION TAG	UNP Q8ZNT7
H	-9	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
H	-8	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
H	-7	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
H	-6	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
H	-5	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
H	-4	HIS	-	EXPRESSION TAG	UNP Q8ZNT7
H	-3	GLY	-	EXPRESSION TAG	UNP Q8ZNT7
H	-2	MET	-	EXPRESSION TAG	UNP Q8ZNT7
H	-1	ALA	-	EXPRESSION TAG	UNP Q8ZNT7
H	0	SER	-	EXPRESSION TAG	UNP Q8ZNT7

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	C	1	Total O P 5 4 1	0	0
2	D	1	Total O P 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	E	1	Total O P 5 4 1	0	0
2	F	1	Total O P 5 4 1	0	0
2	G	1	Total O P 5 4 1	0	0
2	H	1	Total O P 5 4 1	0	0

- Molecule 3 is water.

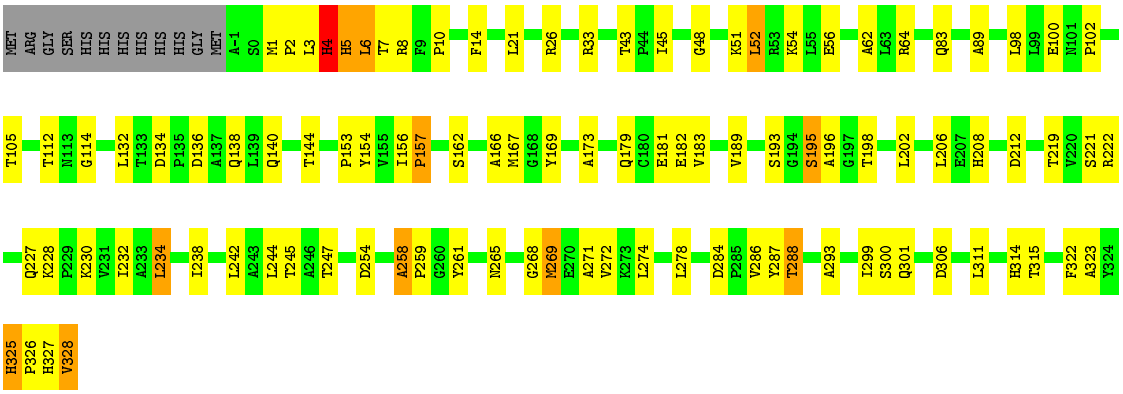
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	6	Total O 6 6	0	0
3	B	1	Total O 1 1	0	0
3	C	1	Total O 1 1	0	0
3	D	2	Total O 2 2	0	0
3	E	1	Total O 1 1	0	0
3	F	1	Total O 1 1	0	0
3	H	6	Total O 6 6	0	0

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.

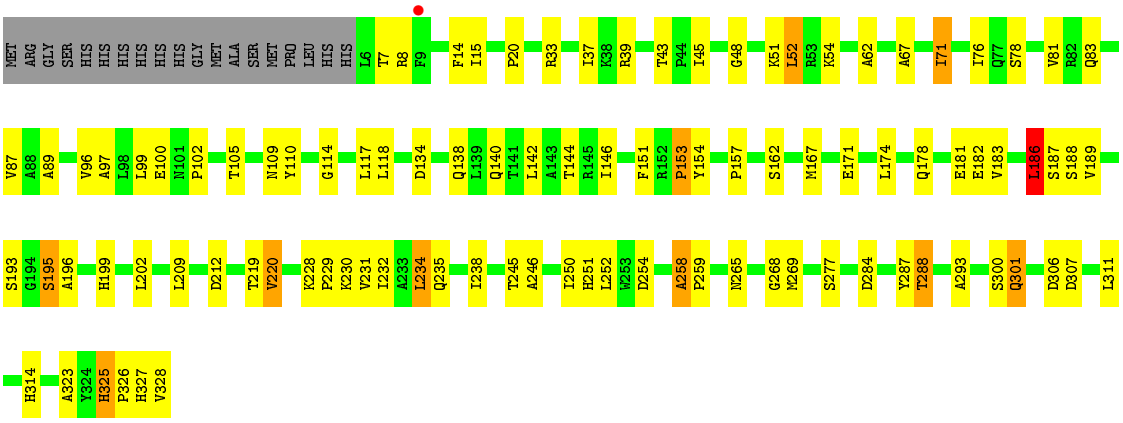
- Molecule 1: D-cysteine desulfhydrase

Chain A:



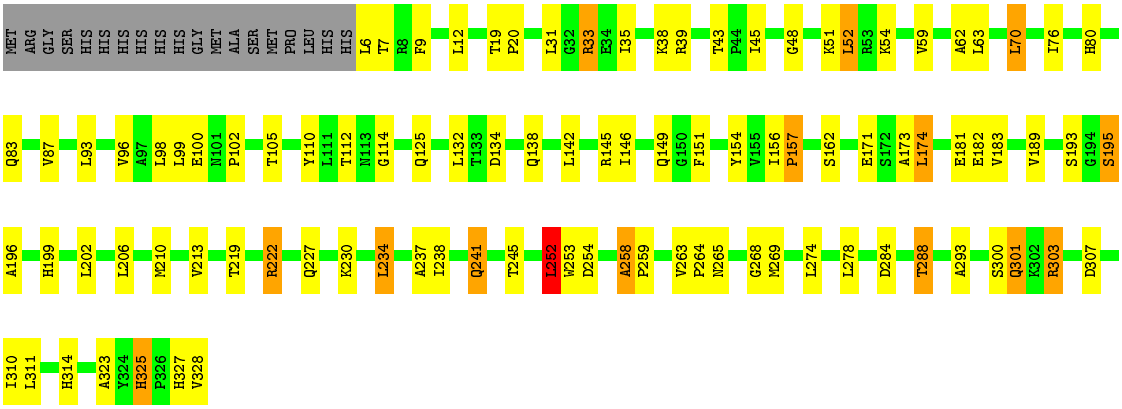
- Molecule 1: D-cysteine desulfhydrase

Chain B:



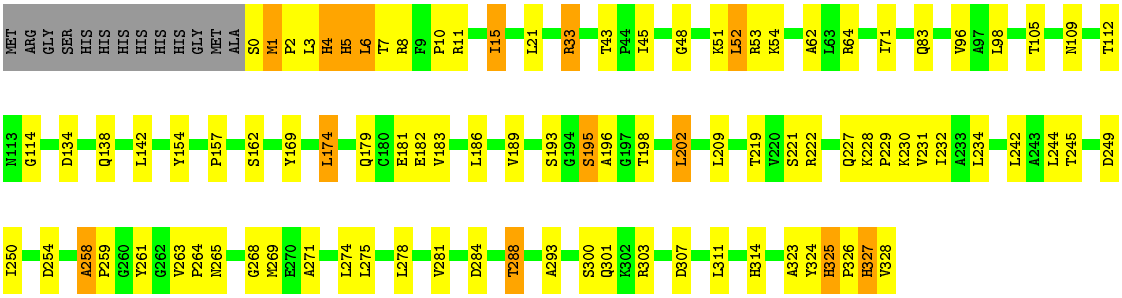
- Molecule 1: D-cysteine desulfhydrase

Chain C:



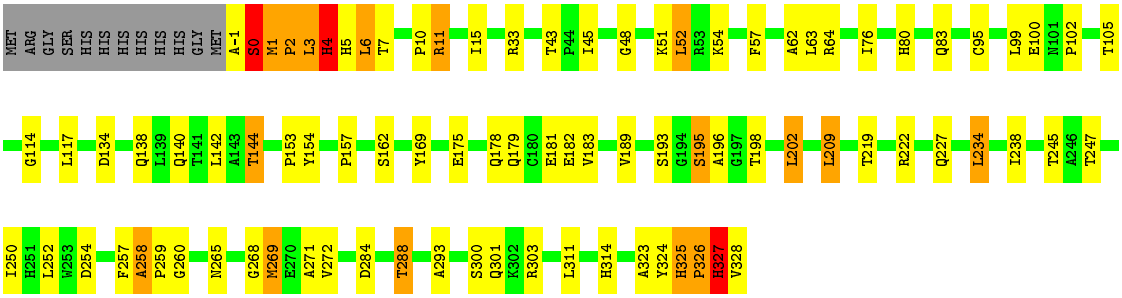
• Molecule 1: D-cysteine desulfhydrase

Chain D:



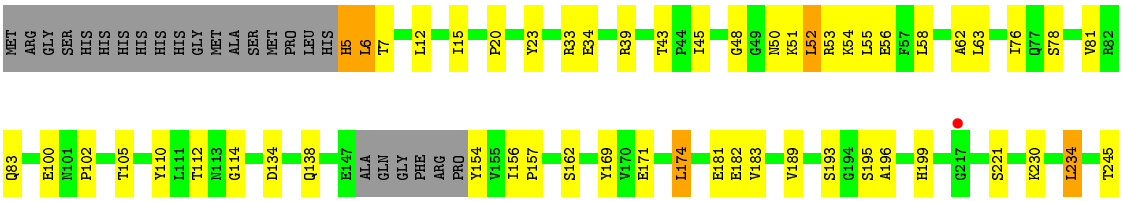
• Molecule 1: D-cysteine desulfhydrase

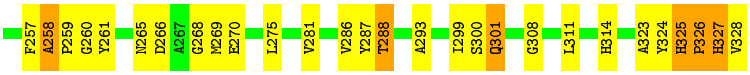
Chain E:



• Molecule 1: D-cysteine desulfhydrase

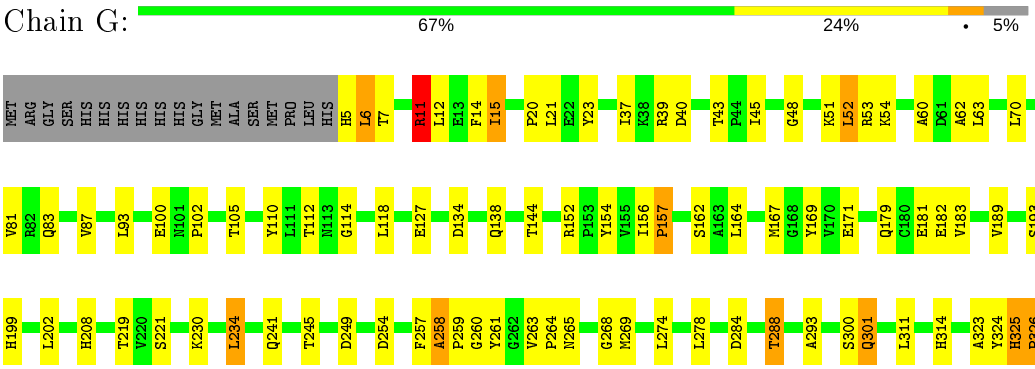
Chain F:





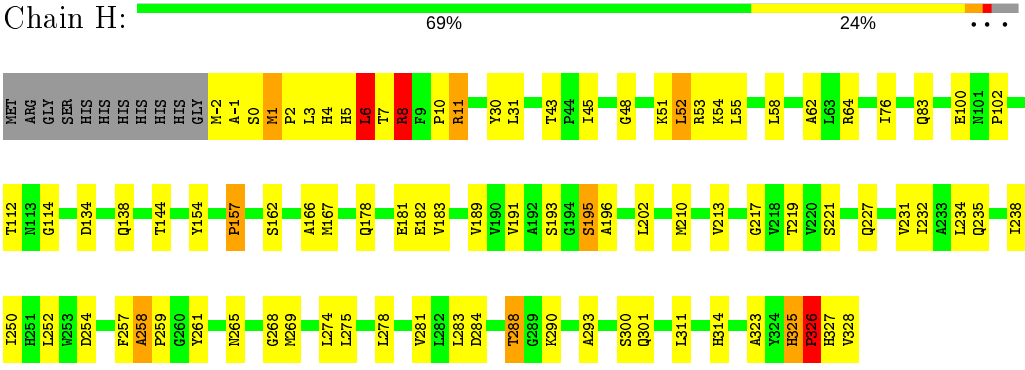
- Molecule 1: D-cysteine desulfhydrase

Chain G:



- Molecule 1: D-cysteine desulfhydrase

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	181.46Å 158.14Å 181.93Å 90.00° 94.10° 90.00°	Depositor
Resolution (Å)	52.04 – 3.30 52.04 – 3.30	Depositor EDS
% Data completeness (in resolution range)	90.4 (52.04-3.30) 90.4 (52.04-3.30)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.89 (at 3.33Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.246 , 0.275 0.244 , 0.268	Depositor DCC
R_{free} test set	3499 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	74.8	Xtriage
Anisotropy	0.150	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 30.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	19156	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: PO4, LLP

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	0.67	1/2489 (0.0%)	0.79	5/3394 (0.1%)
1	B	0.67	0/2354	0.72	1/3220 (0.0%)
1	C	0.65	0/2380	0.74	3/3252 (0.1%)
1	D	0.66	0/2487	0.76	5/3391 (0.1%)
1	E	0.66	1/2473 (0.0%)	0.73	2/3376 (0.1%)
1	F	0.74	0/2241	0.73	0/3070
1	G	0.72	0/2361	0.75	2/3231 (0.1%)
1	H	0.67	2/2477 (0.1%)	0.74	3/3381 (0.1%)
All	All	0.68	4/19262 (0.0%)	0.75	21/26315 (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	E	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	326	PRO	N-CD	-5.48	1.40	1.47
1	E	95	CYS	CB-SG	-5.37	1.73	1.81
1	H	8	ARG	N-CA	-5.07	1.36	1.46
1	A	6	LEU	N-CA	-5.01	1.36	1.46

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	6	LEU	CB-CA-C	-7.56	95.83	110.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	269	MET	CB-CG-SD	7.19	133.98	112.40
1	D	6	LEU	CB-CG-CD1	7.13	123.11	111.00
1	D	269	MET	CB-CG-SD	7.09	133.67	112.40
1	H	8	ARG	N-CA-C	-6.96	92.20	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	0	SER	Peptide

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	2466	0	2456	93	0
1	B	2335	0	2253	79	0
1	C	2361	0	2296	80	0
1	D	2464	0	2462	90	0
1	E	2450	0	2422	91	0
1	F	2227	0	2064	73	0
1	G	2341	0	2237	84	0
1	H	2454	0	2425	109	0
2	A	5	0	0	2	0
2	B	5	0	0	0	0
2	C	5	0	0	2	0
2	D	5	0	0	0	0
2	E	5	0	0	2	0
2	F	5	0	0	0	0
2	G	5	0	0	0	0
2	H	5	0	0	0	0
3	A	6	0	0	1	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	2	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	6	0	0	1	0
All	All	19156	0	18615	671	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 18.

The worst 5 of 671 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:H:-2:MET:CB	1:H:-1:ALA:HA	1.53	1.36
1:H:6:LEU:O	1:H:6:LEU:CD1	1.75	1.34
1:G:325:HIS:HB3	1:G:326:PRO:CD	1.66	1.24
1:F:325:HIS:CD2	1:F:326:PRO:HD3	1.71	1.23
1:G:5:HIS:C	1:G:6:LEU:HD12	1.60	1.18

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	327/342 (96%)	297 (91%)	22 (7%)	8 (2%)	6	28
1	B	320/342 (94%)	288 (90%)	25 (8%)	7 (2%)	6	30
1	C	320/342 (94%)	292 (91%)	22 (7%)	6 (2%)	8	34
1	D	326/342 (95%)	293 (90%)	26 (8%)	7 (2%)	7	31
1	E	327/342 (96%)	292 (89%)	22 (7%)	13 (4%)	3	18
1	F	313/342 (92%)	277 (88%)	28 (9%)	8 (3%)	5	27
1	G	321/342 (94%)	285 (89%)	29 (9%)	7 (2%)	6	30
1	H	328/342 (96%)	297 (90%)	22 (7%)	9 (3%)	5	26
All	All	2582/2736 (94%)	2321 (90%)	196 (8%)	65 (2%)	5	27

5 of 65 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	HIS
1	A	258	ALA
1	B	258	ALA
1	C	258	ALA
1	D	1	MET

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	251/266 (94%)	239 (95%)	12 (5%)	25	56
1	B	227/266 (85%)	209 (92%)	18 (8%)	12	37
1	C	233/266 (88%)	216 (93%)	17 (7%)	14	41
1	D	252/266 (95%)	237 (94%)	15 (6%)	19	49
1	E	248/266 (93%)	230 (93%)	18 (7%)	14	41
1	F	203/266 (76%)	192 (95%)	11 (5%)	22	53
1	G	227/266 (85%)	209 (92%)	18 (8%)	12	37
1	H	247/266 (93%)	231 (94%)	16 (6%)	17	46
All	All	1888/2128 (89%)	1763 (93%)	125 (7%)	16	46

5 of 125 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	245	THR
1	E	178	GLN
1	H	178	GLN
1	D	284	ASP
1	E	11	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	83	GLN
1	E	80	HIS
1	H	208	HIS
1	D	138	GLN
1	D	301	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	LLP	A	51	1	23,24,25	1.79	6 (26%)	25,32,34	1.29	2 (8%)
1	LLP	G	51	1	23,24,25	1.73	4 (17%)	25,32,34	1.44	3 (12%)
1	LLP	D	51	1	23,24,25	1.77	5 (21%)	25,32,34	1.38	4 (16%)
1	LLP	E	51	1	23,24,25	1.72	5 (21%)	25,32,34	1.42	3 (12%)
1	LLP	B	51	1	23,24,25	1.72	5 (21%)	25,32,34	1.40	5 (20%)
1	LLP	H	51	1	23,24,25	1.75	4 (17%)	25,32,34	1.56	6 (24%)
1	LLP	F	51	1	23,24,25	1.76	6 (26%)	25,32,34	1.29	3 (12%)
1	LLP	C	51	1	23,24,25	1.79	4 (17%)	25,32,34	1.36	4 (16%)

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
1	LLP	A	51	1	-	6/16/17/19	0/1/1/1
1	LLP	G	51	1	-	5/16/17/19	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	LLP	D	51	1	-	5/16/17/19	0/1/1/1
1	LLP	E	51	1	-	6/16/17/19	0/1/1/1
1	LLP	B	51	1	-	8/16/17/19	0/1/1/1
1	LLP	H	51	1	-	6/16/17/19	0/1/1/1
1	LLP	F	51	1	-	7/16/17/19	0/1/1/1
1	LLP	C	51	1	-	5/16/17/19	0/1/1/1

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	51	LLP	O3-C3	-5.80	1.23	1.37
1	A	51	LLP	O3-C3	-5.76	1.23	1.37
1	B	51	LLP	O3-C3	-5.50	1.24	1.37
1	G	51	LLP	O3-C3	-5.48	1.24	1.37
1	D	51	LLP	O3-C3	-5.37	1.24	1.37

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	51	LLP	C5-C6-N1	-2.86	119.05	123.82
1	G	51	LLP	C5-C6-N1	-2.70	119.32	123.82
1	F	51	LLP	OP3-P-OP2	2.69	117.92	107.64
1	H	51	LLP	C5-C6-N1	-2.68	119.36	123.82
1	G	51	LLP	OP3-P-OP2	2.64	117.73	107.64

There are no chirality outliers.

5 of 48 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	51	LLP	C4-C4'-NZ-CE
1	A	51	LLP	C5'-OP4-P-OP2
1	A	51	LLP	C5'-OP4-P-OP3
1	G	51	LLP	C4-C4'-NZ-CE
1	G	51	LLP	C5'-OP4-P-OP2

There are no ring outliers.

8 monomers are involved in 34 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	51	LLP	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	G	51	LLP	4	0
1	D	51	LLP	4	0
1	E	51	LLP	4	0
1	B	51	LLP	4	0
1	H	51	LLP	4	0
1	F	51	LLP	3	0
1	C	51	LLP	4	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

8 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	PO4	D	401	-	4,4,4	0.63	0	6,6,6	0.97	0
2	PO4	B	401	-	4,4,4	0.82	0	6,6,6	0.68	0
2	PO4	A	401	-	4,4,4	1.20	0	6,6,6	0.96	0
2	PO4	C	401	-	4,4,4	0.80	0	6,6,6	0.79	0
2	PO4	F	401	-	4,4,4	0.92	0	6,6,6	0.48	0
2	PO4	E	401	-	4,4,4	1.02	0	6,6,6	1.21	1 (16%)
2	PO4	G	401	-	4,4,4	0.83	0	6,6,6	0.79	0
2	PO4	H	401	-	4,4,4	0.84	0	6,6,6	0.71	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	401	PO4	O4-P-O1	-2.66	101.17	110.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	PO4	2	0
2	C	401	PO4	2	0
2	E	401	PO4	2	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^{th} 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	329/342 (96%)	0.10	0 100 100	53, 68, 80, 87	0
1	B	322/342 (94%)	-0.11	1 (0%) 94 94	55, 68, 81, 86	1 (0%)
1	C	322/342 (94%)	-0.06	0 100 100	55, 68, 81, 87	0
1	D	328/342 (95%)	0.06	0 100 100	53, 68, 80, 86	0
1	E	329/342 (96%)	0.03	0 100 100	53, 68, 80, 86	0
1	F	317/342 (92%)	-0.15	1 (0%) 94 94	55, 68, 80, 88	0
1	G	323/342 (94%)	-0.15	0 100 100	55, 68, 81, 86	0
1	H	330/342 (96%)	0.02	0 100 100	54, 68, 81, 87	0
All	All	2600/2736 (95%)	-0.03	2 (0%) 95 97	53, 68, 80, 88	1 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	9	PHE	2.5
1	F	217	GLY	2.2

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

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^{th} 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^2)	Q<0.9
1	LLP	G	51	24/25	0.97	0.21	55,56,57,58	0
1	LLP	B	51	24/25	0.97	0.21	55,56,57,58	0
1	LLP	D	51	24/25	0.98	0.22	55,56,57,58	0
1	LLP	E	51	24/25	0.98	0.20	55,56,57,58	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	LLP	A	51	24/25	0.98	0.23	55,56,57,58	0
1	LLP	H	51	24/25	0.98	0.20	55,56,57,58	0
1	LLP	F	51	24/25	0.98	0.20	55,57,57,58	0
1	LLP	C	51	24/25	0.98	0.22	55,56,57,58	0

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^{th} 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^2)	Q<0.9
2	PO4	G	401	5/5	0.92	0.19	108,108,109,109	0
2	PO4	F	401	5/5	0.94	0.18	154,154,155,155	0
2	PO4	B	401	5/5	0.94	0.18	123,123,124,124	0
2	PO4	A	401	5/5	0.95	0.19	81,81,82,83	0
2	PO4	E	401	5/5	0.95	0.17	75,76,76,77	0
2	PO4	C	401	5/5	0.95	0.15	97,98,98,98	0
2	PO4	H	401	5/5	0.95	0.18	112,113,113,113	0
2	PO4	D	401	5/5	0.96	0.17	83,85,85,85	0

6.5 Other polymers (i)

There are no such residues in this entry.