wwPDB X-ray Structure Validation Summary Report (i)

Aug 14, 2023 – 02:29 PM EDT

PDB ID	:	1S2V
Title	:	Crystal structure of phosphoenolpyruvate mutase complexed with Mg(II)
Authors	:	Liu, S.; Lu, Z.; Han, Y.; Jia, Y.; Howard, A.; Dunaway-Mariano, D.; Herzberg, O.
Deposited on	:	2004-01-11
Resolution	:	2.10 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

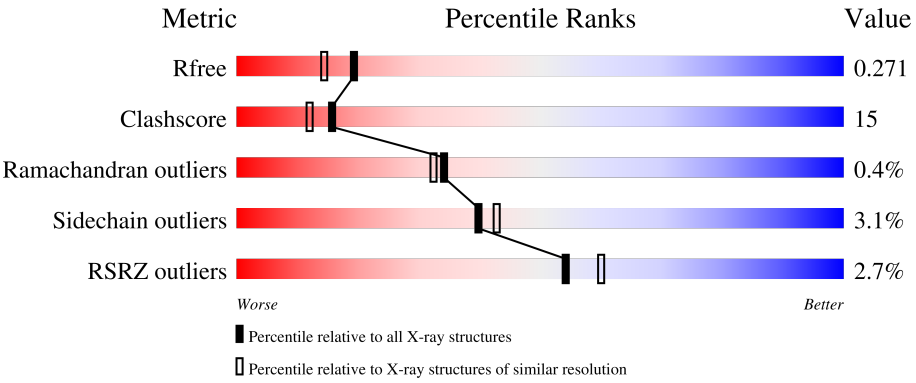
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 <=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	295	3% 62% 32%
1	B	295	2% 74% 21%
1	C	295	3% 67% 26%
1	D	295	2% 73% 21% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9824 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 Phosphoenolpyruvate phosphomutase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	282	Total	C	N	O	S	23	0	0
			2213	1394	382	427	10			
1	B	287	Total	C	N	O	S	36	0	0
			2248	1414	391	433	10			
1	C	282	Total	C	N	O	S	25	0	0
			2216	1398	383	425	10			
1	D	281	Total	C	N	O	S	32	0	0
			2207	1392	381	424	10			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	14	MET	ALA	SEE REMARK 999	UNP P56839
A	24	MET	ALA	SEE REMARK 999	UNP P56839
A	74	MET	ALA	SEE REMARK 999	UNP P56839
A	189	MET	ALA	SEE REMARK 999	UNP P56839
A	203	MET	ALA	SEE REMARK 999	UNP P56839
A	230	MET	ALA	SEE REMARK 999	UNP P56839
A	234	MET	ALA	SEE REMARK 999	UNP P56839
B	14	MET	ALA	SEE REMARK 999	UNP P56839
B	24	MET	ALA	SEE REMARK 999	UNP P56839
B	74	MET	ALA	SEE REMARK 999	UNP P56839
B	189	MET	ALA	SEE REMARK 999	UNP P56839
B	203	MET	ALA	SEE REMARK 999	UNP P56839
B	230	MET	ALA	SEE REMARK 999	UNP P56839
B	234	MET	ALA	SEE REMARK 999	UNP P56839
C	14	MET	ALA	SEE REMARK 999	UNP P56839
C	24	MET	ALA	SEE REMARK 999	UNP P56839
C	74	MET	ALA	SEE REMARK 999	UNP P56839
C	189	MET	ALA	SEE REMARK 999	UNP P56839
C	203	MET	ALA	SEE REMARK 999	UNP P56839
C	230	MET	ALA	SEE REMARK 999	UNP P56839
C	234	MET	ALA	SEE REMARK 999	UNP P56839

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	14	MET	ALA	SEE REMARK 999	UNP P56839
D	24	MET	ALA	SEE REMARK 999	UNP P56839
D	74	MET	ALA	SEE REMARK 999	UNP P56839
D	189	MET	ALA	SEE REMARK 999	UNP P56839
D	203	MET	ALA	SEE REMARK 999	UNP P56839
D	230	MET	ALA	SEE REMARK 999	UNP P56839
D	234	MET	ALA	SEE REMARK 999	UNP P56839

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0
2	B	1	Total Mg 1 1	0	0
2	C	1	Total Mg 1 1	0	0
2	D	1	Total Mg 1 1	0	0

- Molecule 3 is water.

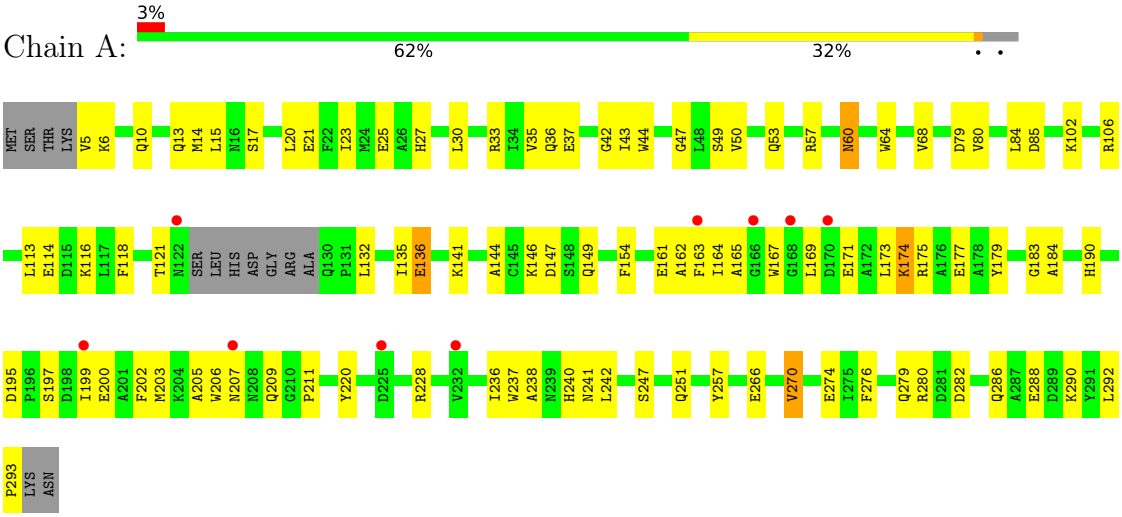
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	224	Total O 224 224	0	0
3	B	236	Total O 236 236	0	0
3	C	240	Total O 240 240	0	0
3	D	236	Total O 236 236	0	0

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.

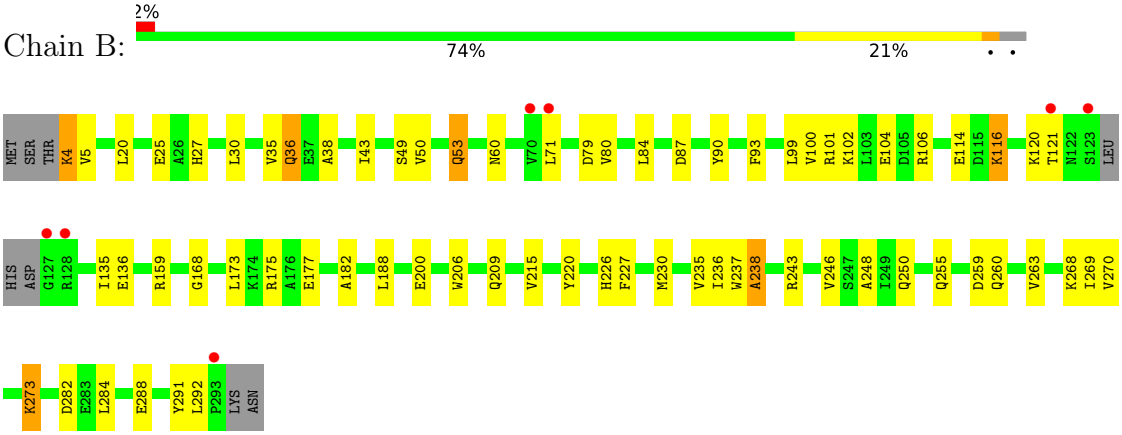
• Molecule 1: Phosphoenolpyruvate phosphomutase

Chain A:



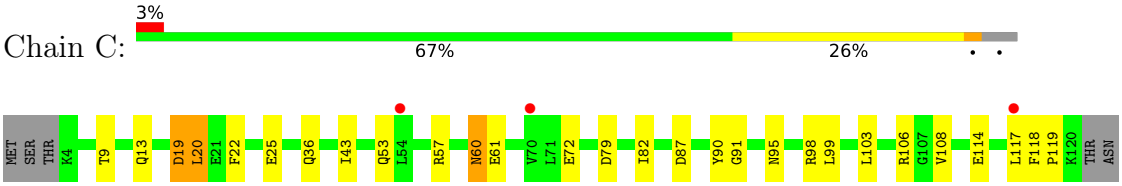
• Molecule 1: Phosphoenolpyruvate phosphomutase

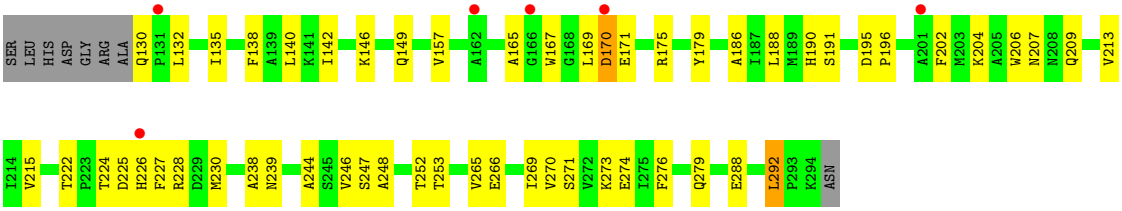
Chain B:



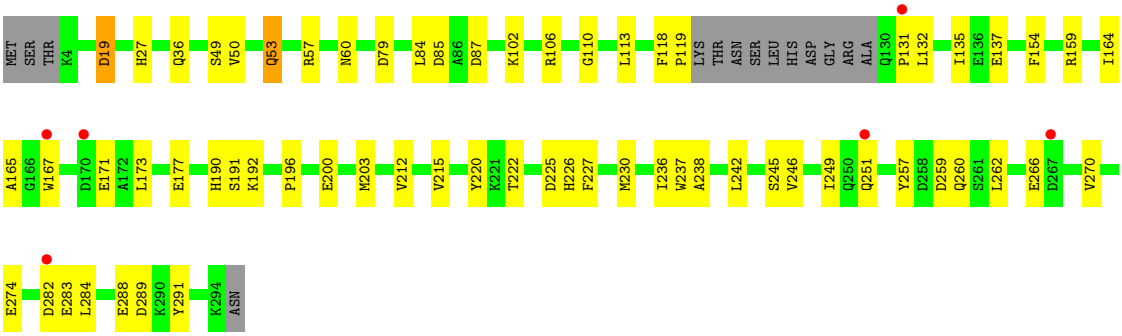
• Molecule 1: Phosphoenolpyruvate phosphomutase

Chain C:





• Molecule 1: Phosphoenolpyruvate phosphomutase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	122.52Å 86.47Å 104.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.00 – 2.10 37.96 – 2.10	Depositor EDS
% Data completeness (in resolution range)	81.3 (38.00-2.10) 96.2 (37.96-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.31 (at 2.10Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.185 , 0.257 0.204 , 0.271	Depositor DCC
R_{free} test set	5106 reflections (8.15%)	wwPDB-VP
Wilson B-factor (Å^2)	29.2	Xtriage
Anisotropy	0.330	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 63.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9824	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.47 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1398e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: MG

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	0.48	0/2251	0.68	2/3042 (0.1%)
1	B	0.49	0/2286	0.70	1/3087 (0.0%)
1	C	0.49	0/2254	0.68	1/3043 (0.0%)
1	D	0.51	0/2245	0.68	2/3032 (0.1%)
All	All	0.49	0/9036	0.69	6/12204 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	238	ALA	N-CA-C	6.70	129.08	111.00
1	D	238	ALA	N-CA-C	6.22	127.81	111.00
1	B	238	ALA	N-CA-C	6.05	127.34	111.00
1	D	237	TRP	N-CA-C	-5.67	95.70	111.00
1	A	238	ALA	N-CA-C	5.58	126.07	111.00

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	2213	0	2194	91	0
1	B	2248	0	2233	67	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2216	0	2207	65	0
1	D	2207	0	2194	54	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	224	0	0	12	0
3	B	236	0	0	8	0
3	C	240	0	0	9	0
3	D	236	0	0	8	0
All	All	9824	0	8828	255	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 15.

The worst 5 of 255 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:B:71:LEU:HD21	1:B:84:LEU:HD13	1.32	1.05
1:A:36:GLN:HE22	1:A:79:ASP:H	1.02	1.00
1:C:225:ASP:HA	1:C:228:ARG:HD2	1.44	1.00
1:C:36:GLN:NE2	1:C:79:ASP:H	1.66	0.93
1:D:36:GLN:HE21	1:D:79:ASP:H	1.07	0.93

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	278/295 (94%)	262 (94%)	16 (6%)	0	100	100
1	B	283/295 (96%)	269 (95%)	12 (4%)	2 (1%)	22	18

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	278/295 (94%)	264 (95%)	12 (4%)	2 (1%)	22 18
1	D	277/295 (94%)	264 (95%)	13 (5%)	0	100 100
All	All	1116/1180 (95%)	1059 (95%)	53 (5%)	4 (0%)	34 32

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	119	PRO
1	B	168	GLY
1	C	239	ASN
1	B	263	VAL

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	238/249 (96%)	232 (98%)	6 (2%)	47 52
1	B	241/249 (97%)	232 (96%)	9 (4%)	34 35
1	C	238/249 (96%)	230 (97%)	8 (3%)	37 39
1	D	237/249 (95%)	230 (97%)	7 (3%)	41 44
All	All	954/996 (96%)	924 (97%)	30 (3%)	40 43

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	282	ASP
1	D	171	GLU
1	C	25	GLU
1	D	289	ASP
1	D	53	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	226	HIS
1	D	13	GLN
1	D	190	HIS
1	D	36	GLN
1	D	10	GLN

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 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 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, 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	282/295 (95%)	0.19	9 (3%)	47	54	20, 36, 55, 63	6 (2%)
1	B	287/295 (97%)	-0.02	7 (2%)	59	64	15, 29, 46, 59	9 (3%)
1	C	282/295 (95%)	0.16	9 (3%)	47	54	18, 32, 54, 62	6 (2%)
1	D	281/295 (95%)	-0.05	6 (2%)	63	68	17, 30, 48, 56	10 (3%)
All	All	1132/1180 (95%)	0.07	31 (2%)	54	60	15, 32, 53, 63	31 (2%)

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	131	PRO	5.3
1	B	123	SER	4.4
1	A	166	GLY	4.3
1	D	167	TRP	3.7
1	B	127	GLY	3.6

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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	MG	B	1002	1/1	0.87	0.21	33,33,33,33	1
2	MG	C	1003	1/1	0.88	0.07	35,35,35,35	0
2	MG	D	1004	1/1	0.95	0.06	37,37,37,37	0
2	MG	A	1001	1/1	0.97	0.14	48,48,48,48	0

6.5 Other polymers (i)

There are no such residues in this entry.