



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 03:06 pm GMT

PDB ID : 3GYB  
Title : Crystal structure of a LacI-family transcriptional regulatory protein from *Corynebacterium glutamicum*  
Authors : Palani, K.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)  
Deposited on : 2009-04-03  
Resolution : 1.60 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](mailto: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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

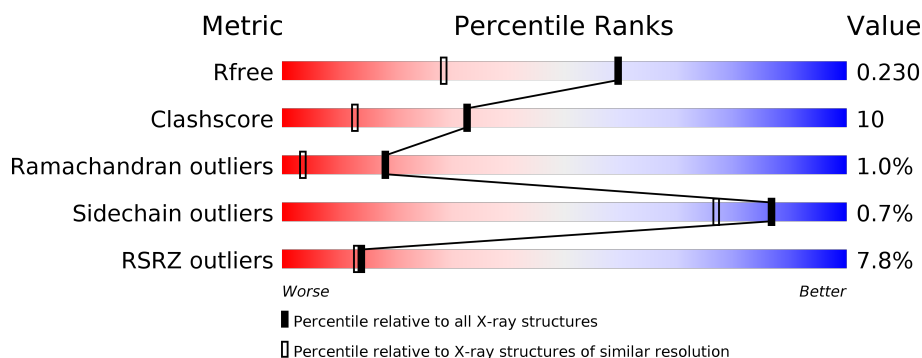
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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))
$R_{free}$	100719	2696 (1.60-1.60)
Clashscore	112137	2967 (1.60-1.60)
Ramachandran outliers	110173	2887 (1.60-1.60)
Sidechain outliers	110143	2886 (1.60-1.60)
RSRZ outliers	101464	2714 (1.60-1.60)

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	280	<div> <div>5%</div> <div> <div></div> <div>80%</div> <div>13%</div> <div>7%</div> </div> </div>
1	B	280	<div> <div>9%</div> <div> <div></div> <div>70%</div> <div>17%</div> <div>• 12%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4144 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 Transcriptional regulators (LACI-FAMILY TRANSCRIPTIONAL REGULATORY PROTEIN).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	261	Total	C	N	O	S	Se	0	0	0
			1980	1247	342	386	1	4			
1	B	246	Total	C	N	O	S	Se	0	0	0
			1865	1176	316	368	1	4			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	59	MSE	-	expression tag	UNP Q8NTY1
A	60	SER	-	expression tag	UNP Q8NTY1
A	61	LEU	-	expression tag	UNP Q8NTY1
A	331	GLU	-	expression tag	UNP Q8NTY1
A	332	GLY	-	expression tag	UNP Q8NTY1
A	333	HIS	-	expression tag	UNP Q8NTY1
A	334	HIS	-	expression tag	UNP Q8NTY1
A	335	HIS	-	expression tag	UNP Q8NTY1
A	336	HIS	-	expression tag	UNP Q8NTY1
A	337	HIS	-	expression tag	UNP Q8NTY1
A	338	HIS	-	expression tag	UNP Q8NTY1
B	59	MSE	-	expression tag	UNP Q8NTY1
B	60	SER	-	expression tag	UNP Q8NTY1
B	61	LEU	-	expression tag	UNP Q8NTY1
B	331	GLU	-	expression tag	UNP Q8NTY1
B	332	GLY	-	expression tag	UNP Q8NTY1
B	333	HIS	-	expression tag	UNP Q8NTY1
B	334	HIS	-	expression tag	UNP Q8NTY1
B	335	HIS	-	expression tag	UNP Q8NTY1
B	336	HIS	-	expression tag	UNP Q8NTY1
B	337	HIS	-	expression tag	UNP Q8NTY1
B	338	HIS	-	expression tag	UNP Q8NTY1

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

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

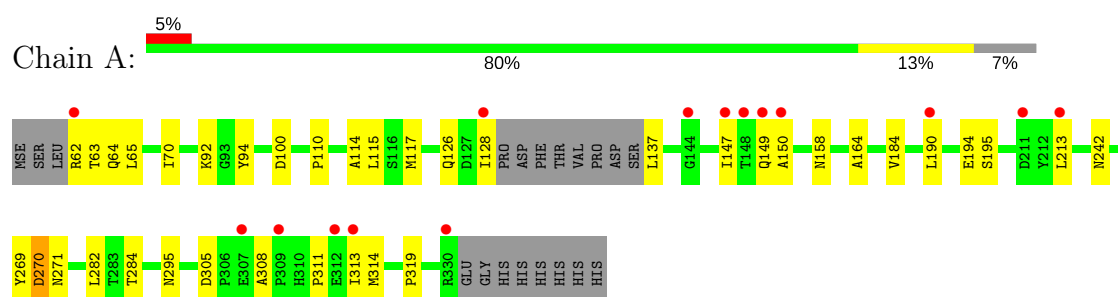
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	152	Total 152	O 152	0	0
3	B	145	Total 145	O 145	0	0

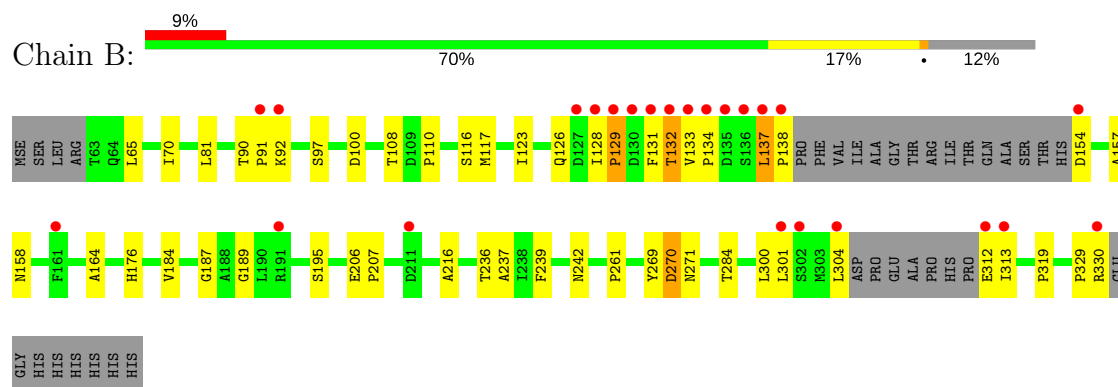
### 3 Residue-property plots [i](#)

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: Transcriptional regulators (LACI-FAMILY TRANSCRIPTIONAL REGULATORY PROTEIN)



- Molecule 1: Transcriptional regulators (LACI-FAMILY TRANSCRIPTIONAL REGULATORY PROTEIN)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	38.07Å 61.88Å 65.60Å 115.83° 106.37° 90.50°	Depositor
Resolution (Å)	32.55 – 1.60 32.55 – 1.60	Depositor EDS
% Data completeness (in resolution range)	90.4 (32.55-1.60) 76.7 (32.55-1.60)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 1.60Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.215 , 0.231 0.214 , 0.230	Depositor DCC
$R_{free}$ test set	3088 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	16.7	Xtriage
Anisotropy	0.328	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 56.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-h-l 0.016 for -h,k,-k-l 0.007 for -h,-k,h+k+l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4144	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.01% 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 [i](#)

### 5.1 Standard geometry [i](#)

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	0.27	0/2014	0.61	0/2742
1	B	0.27	0/1894	0.60	0/2576
All	All	0.27	0/3908	0.61	0/5318

There are no bond length outliers.

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	A	1980	0	1974	33	0
1	B	1865	0	1855	45	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	152	0	0	1	0
3	B	145	0	0	3	0
All	All	4144	0	3829	76	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 10.

All (76) 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:65:LEU:HB3	1:A:117:MSE:HE2	1.22	1.12
1:B:312:GLU:HG2	1:B:313:ILE:H	1.30	0.97
1:A:65:LEU:HB3	1:A:117:MSE:CE	2.04	0.87
1:A:65:LEU:HD13	1:A:117:MSE:HE3	1.58	0.85
1:A:65:LEU:CB	1:A:117:MSE:HE2	2.10	0.77
1:B:128:ILE:HG23	1:B:129:PRO:HD3	1.73	0.70
1:A:311:PRO:HG2	1:A:313:ILE:HD11	1.74	0.69
1:A:65:LEU:HD22	1:A:117:MSE:CE	2.24	0.67
1:A:295:ASN:ND2	1:A:314:MSE:H	1.92	0.67
1:B:128:ILE:HG13	1:B:187:GLY:N	2.10	0.67
1:B:312:GLU:HG2	1:B:313:ILE:N	2.09	0.66
1:B:128:ILE:O	1:B:187:GLY:HA3	1.97	0.65
1:B:133:VAL:C	3:B:451:HOH:O	2.34	0.65
1:B:216:ALA:HB1	1:B:242:ASN:HD21	1.62	0.64
1:A:65:LEU:HD22	1:A:117:MSE:HE1	1.78	0.64
1:A:65:LEU:HD13	1:A:117:MSE:CE	2.27	0.64
1:B:300:LEU:O	1:B:304:LEU:HD13	1.97	0.63
1:A:65:LEU:HD21	1:B:116:SER:OG	1.99	0.62
1:B:128:ILE:CG2	1:B:129:PRO:HD3	2.30	0.61
1:A:184:VAL:HG21	1:A:242:ASN:ND2	2.17	0.59
1:B:97:SER:OG	1:B:117:MSE:HE1	2.02	0.59
1:A:92:LYS:HE2	3:A:364:HOH:O	2.02	0.59
1:B:65:LEU:CD2	1:B:117:MSE:SE	3.01	0.58
1:A:158:ASN:HD22	1:A:319:PRO:HB3	1.69	0.57
1:A:115:LEU:HD21	1:A:137:LEU:HD12	1.84	0.57
1:B:123:ILE:HA	3:B:451:HOH:O	2.04	0.57
1:A:184:VAL:HG21	1:A:242:ASN:HD21	1.71	0.55
1:B:184:VAL:HG23	1:B:189:GLY:HA3	1.89	0.55
1:B:176:HIS:HD2	1:B:236:THR:OG1	1.89	0.55
1:A:213:LEU:HD23	1:A:213:LEU:H	1.70	0.54
1:B:92:LYS:HE3	1:B:301:LEU:HD21	1.90	0.53
1:A:269:TYR:O	1:A:270:ASP:CB	2.56	0.53
1:B:70:ILE:O	1:B:100:ASP:HA	2.08	0.53
1:B:206:GLU:OE1	1:B:207:PRO:HD2	2.09	0.53
1:B:133:VAL:HG13	1:B:154:ASP:O	2.09	0.52
1:A:271:ASN:HA	1:A:284:THR:HG21	1.92	0.51
1:B:330:ARG:HG3	1:B:330:ARG:HH11	1.74	0.51
1:A:149:GLN:HG3	1:A:150:ALA:N	2.25	0.51
1:B:216:ALA:CB	1:B:242:ASN:HD21	2.23	0.51
1:B:65:LEU:HD21	1:B:117:MSE:SE	2.60	0.51
1:B:269:TYR:O	1:B:270:ASP:CB	2.59	0.50

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:149:GLN:HG3	1:A:150:ALA:H	1.77	0.50
1:B:312:GLU:CG	1:B:313:ILE:H	2.09	0.49
1:A:147:ILE:HG22	1:A:149:GLN:HG2	1.94	0.49
1:A:70:ILE:O	1:A:100:ASP:HA	2.13	0.49
1:A:62:ARG:NH2	1:A:94:TYR:OH	2.46	0.48
1:B:137:LEU:HB3	1:B:138:PRO:HD3	1.94	0.48
1:B:70:ILE:C	1:B:126:GLN:HG2	2.35	0.47
1:B:90:THR:OG1	1:B:91:PRO:HD3	2.13	0.47
1:B:158:ASN:HD22	1:B:319:PRO:HB3	1.80	0.46
1:B:184:VAL:HG21	3:B:384:HOH:O	2.16	0.46
1:B:108:THR:C	1:B:110:PRO:HD3	2.37	0.45
1:A:117:MSE:HB2	1:B:117:MSE:HB2	1.98	0.45
1:B:97:SER:OG	1:B:117:MSE:SE	2.84	0.45
1:A:190:LEU:O	1:A:194:GLU:HG3	2.16	0.45
1:B:164:ALA:HB3	1:B:195:SER:HB3	1.99	0.45
1:B:271:ASN:HA	1:B:284:THR:HG21	1.99	0.44
1:B:329:PRO:O	1:B:330:ARG:HB2	2.16	0.44
1:A:126:GLN:HE21	1:A:128:ILE:HD11	1.83	0.44
1:A:269:TYR:O	1:A:270:ASP:HB2	2.18	0.44
1:B:131:PHE:O	1:B:132:THR:CB	2.65	0.43
1:B:128:ILE:HG13	1:B:187:GLY:CA	2.48	0.43
1:B:128:ILE:N	1:B:129:PRO:CD	2.82	0.43
1:A:269:TYR:O	1:A:270:ASP:CG	2.57	0.43
1:B:269:TYR:O	1:B:270:ASP:CG	2.57	0.43
1:B:137:LEU:H	1:B:138:PRO:CD	2.32	0.42
1:A:305:ASP:OD2	1:A:308:ALA:HB2	2.19	0.42
1:A:114:ALA:O	1:A:117:MSE:HG2	2.20	0.42
1:A:164:ALA:HB3	1:A:195:SER:HB3	2.01	0.42
1:B:176:HIS:CD2	1:B:237:ALA:HB2	2.55	0.42
1:A:110:PRO:HG3	1:A:126:GLN:HE22	1.85	0.41
1:B:81:LEU:C	1:B:81:LEU:HD23	2.41	0.41
1:B:134:PRO:HD2	1:B:154:ASP:O	2.21	0.41
1:B:330:ARG:NH1	1:B:330:ARG:HG3	2.36	0.41
1:A:63:THR:O	1:A:64:GLN:HB2	2.20	0.41
1:B:131:PHE:CZ	1:B:157:ALA:HB1	2.56	0.40

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	257/280 (92%)	251 (98%)	5 (2%)	1 (0%)	38	16
1	B	240/280 (86%)	230 (96%)	6 (2%)	4 (2%)	11	1
All	All	497/560 (89%)	481 (97%)	11 (2%)	5 (1%)	18	4

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	137	LEU
1	A	270	ASP
1	B	270	ASP
1	B	132	THR
1	B	129	PRO

### 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	216/229 (94%)	215 (100%)	1 (0%)	91	84
1	B	205/229 (90%)	203 (99%)	2 (1%)	80	65
All	All	421/458 (92%)	418 (99%)	3 (1%)	87	77

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	282	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	239	PHE
1	B	261	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	126	GLN
1	A	149	GLN
1	A	158	ASN
1	A	295	ASN
1	B	158	ASN
1	B	176	HIS

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 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<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	257/280 (91%)	0.38	15 (5%) 24 22	10, 17, 34, 50	0
1	B	242/280 (86%)	0.61	24 (9%) 8 7	10, 17, 40, 56	0
All	All	499/560 (89%)	0.49	39 (7%) 14 13	10, 17, 35, 56	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	128	ILE	18.1
1	B	129	PRO	11.9
1	B	131	PHE	10.9
1	B	133	VAL	9.0
1	B	137	LEU	8.5
1	A	149	GLN	7.2
1	B	130	ASP	6.8
1	B	330	ARG	6.5
1	B	136	SER	5.9
1	A	213	LEU	5.5
1	A	62	ARG	4.6
1	A	128	ILE	4.3
1	A	150	ALA	4.2
1	B	132	THR	4.2
1	B	211	ASP	4.2
1	A	211	ASP	4.1
1	A	147	ILE	4.0
1	A	330	ARG	3.9
1	A	312	GLU	3.8
1	B	304	LEU	3.7
1	A	309	PRO	3.5
1	B	312	GLU	3.3
1	B	127	ASP	3.3
1	B	138	PRO	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	307	GLU	3.1
1	B	301	LEU	3.0
1	A	144	GLY	2.9
1	B	91	PRO	2.9
1	A	313	ILE	2.7
1	A	190	LEU	2.7
1	A	148	THR	2.7
1	B	154	ASP	2.5
1	B	302	SER	2.5
1	B	191	ARG	2.3
1	B	92	LYS	2.3
1	B	313	ILE	2.3
1	B	134	PRO	2.1
1	B	161	PHE	2.1
1	B	135	ASP	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 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	MG	A	1	1/1	0.99	0.10	0.42	12,12,12,12	0
2	MG	B	2	1/1	0.99	0.05	-3.19	12,12,12,12	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.