



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 12:04 PM GMT

PDB ID : 1EGM
Title : CRYSTAL STRUCTURE OF DIOL DEHYDRATASE-CYANOCOBALAMIN COMPLEX AT 100K.
Authors : Masuda, J.; Shibata, N.; Toraya, T.; Morimoto, Y.; Yasuoka, N.
Deposited on : 2000-02-15
Resolution : 1.85 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

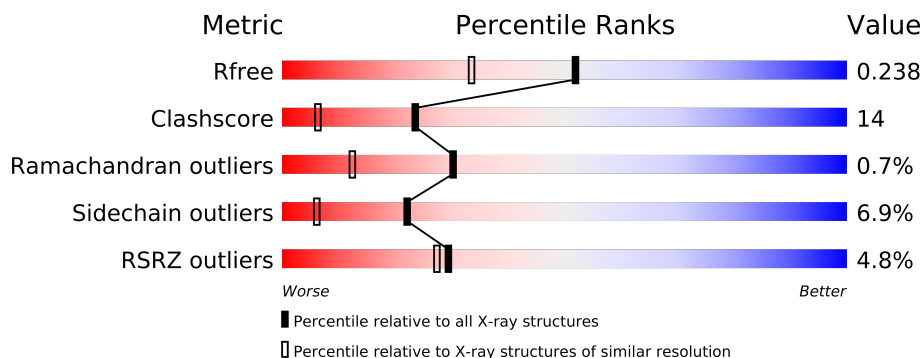
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.85 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}	66092	1269 (1.86-1.86)
Clashscore	79885	1470 (1.86-1.86)
Ramachandran outliers	78287	1451 (1.86-1.86)
Sidechain outliers	78261	1451 (1.86-1.86)
RSRZ outliers	66119	1269 (1.86-1.86)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	554	
1	L	554	
2	B	224	
2	E	224	
3	G	173	
3	M	173	

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 14737 atoms, of which 0 are hydrogen and 0 are deuterium.

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 PROPANEDIOL DEHYDRATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	551	Total	C	N	O	S	0	2	0
			4209	2625	727	828	29			
1	L	551	Total	C	N	O	S	0	1	0
			4206	2623	727	827	29			

- Molecule 2 is a protein called PROPANEDIOL DEHYDRATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	178	Total	C	N	O	S	0	0	0
			1358	859	244	253	2			
2	E	177	Total	C	N	O	S	0	0	0
			1353	856	243	252	2			

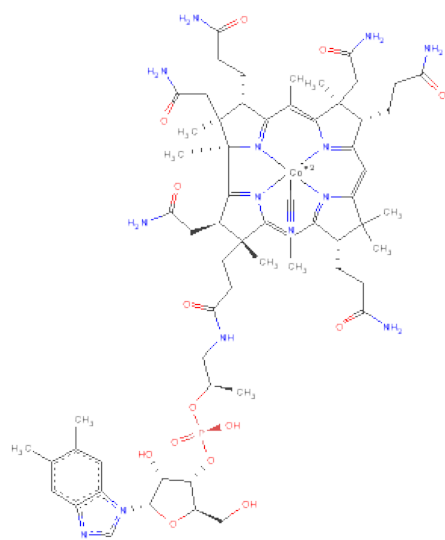
- Molecule 3 is a protein called PROPANEDIOL DEHYDRATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	137	Total	C	N	O	S	0	3	0
			1108	691	196	218	3			
3	M	137	Total	C	N	O	S	0	1	0
			1098	684	195	216	3			

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

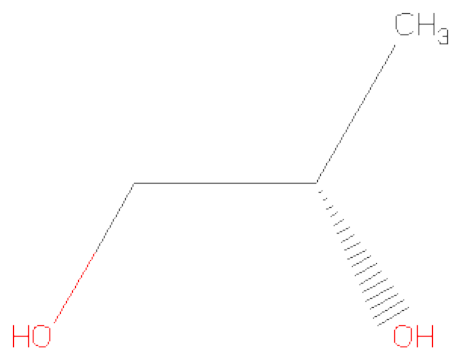
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	K	0	0
			1	1		
4	L	1	Total	K	0	0
			1	1		

- Molecule 5 is CO-CYANOCOBALAMIN (three-letter code: CNC) (formula: C₆₃H₈₈CoN₁₄O₁₄P).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	A	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		
5	L	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		

- Molecule 6 is S-1,2-PROPANEDIOL (three-letter code: PGO) (formula: C₃H₈O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			5	3	2		
6	L	1	Total	C	O	0	0
			5	3	2		

- Molecule 7 is water.

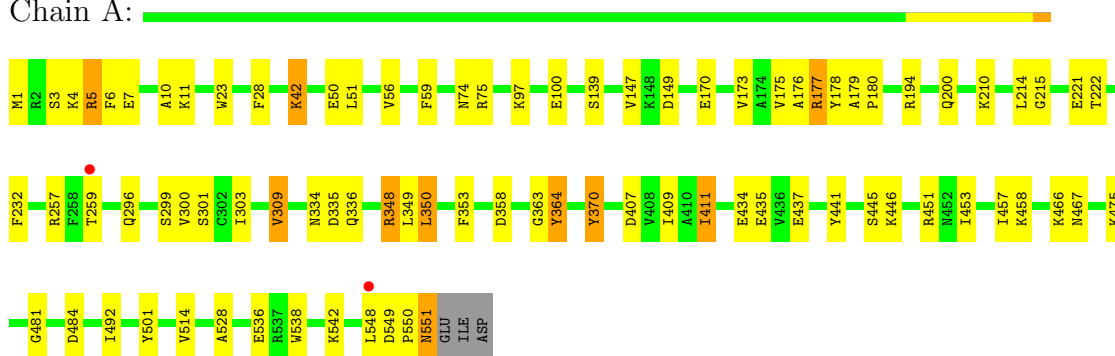
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	490	Total 490	O 490	0	0
7	B	118	Total 118	O 118	0	0
7	E	37	Total 37	O 37	0	0
7	G	152	Total 152	O 152	0	0
7	L	319	Total 319	O 319	0	0
7	M	95	Total 95	O 95	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

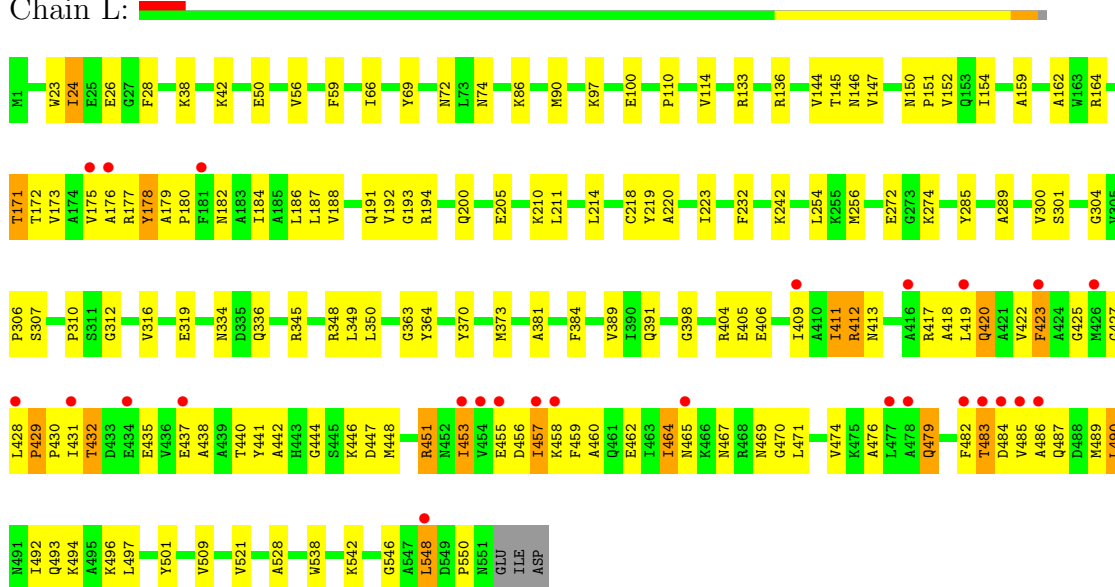
• Molecule 1: PROPANEDIOL DEHYDRATASE

Chain A:



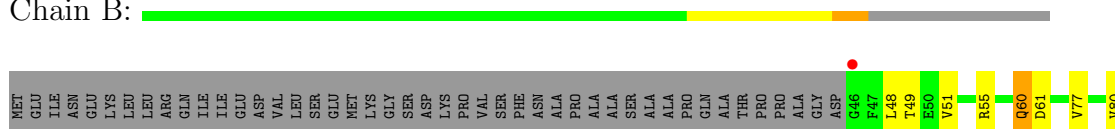
• Molecule 1: PROPANEDIOL DEHYDRATASE

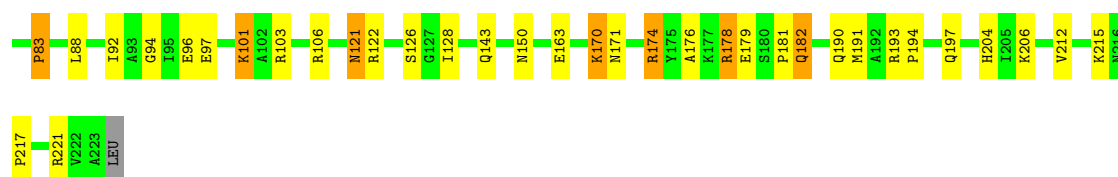
Chain L:



• Molecule 2: PROPANEDIOL DEHYDRATASE

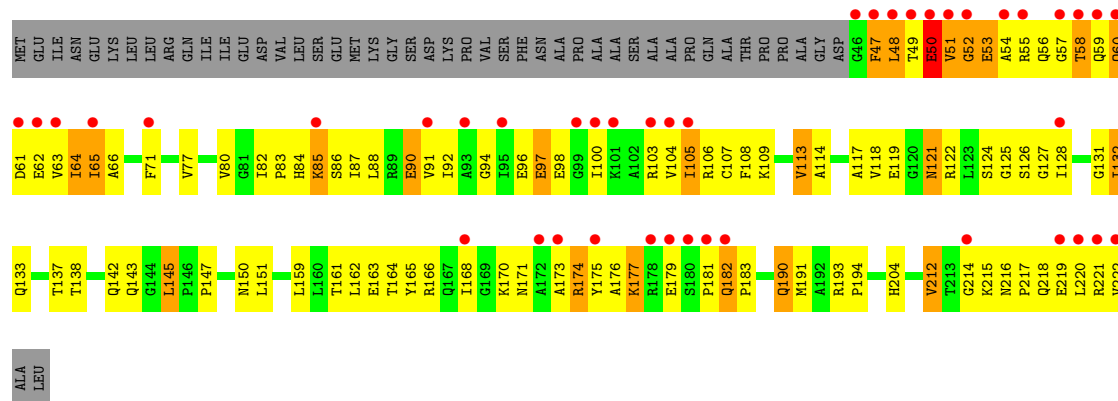
Chain B:





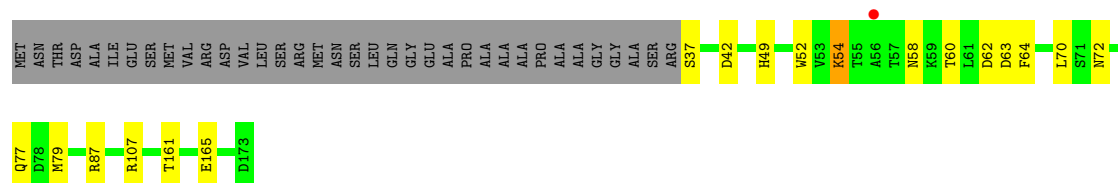
• Molecule 2: PROPANEDIOL DEHYDRATASE

Chain E:



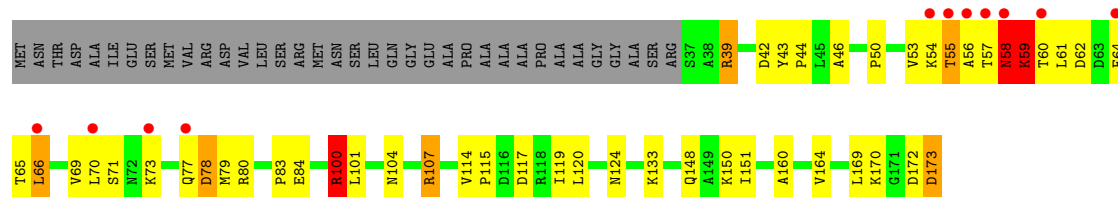
• Molecule 3: PROPANEDIOL DEHYDRATASE

Chain G:



• Molecule 3: PROPANEDIOL DEHYDRATASE

Chain M:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	74.00Å 121.60Å 207.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.85 36.43 – 1.85	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-1.85) 88.7 (36.43-1.85)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.25 (at 1.85Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.182 , 0.249 0.177 , 0.238	Depositor DCC
R_{free} test set	7098 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	18.9	Xtriage
Anisotropy	0.704	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 68.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 142048 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	14737	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.67% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PGO, K, CNC

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	0.36	1/4289 (0.0%)	1.00	9/5809 (0.2%)
1	L	0.33	0/4282	1.00	6/5799 (0.1%)
2	B	0.32	0/1380	0.95	0/1868
2	E	0.31	0/1375	0.86	0/1861
3	G	0.33	0/1135	1.02	4/1532 (0.3%)
3	M	0.31	0/1117	1.04	3/1509 (0.2%)
All	All	0.34	1/13578 (0.0%)	0.99	22/18378 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	309	VAL	CB-CG2	-5.71	1.40	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	364	TYR	CB-CG-CD1	9.35	126.61	121.00
1	A	75	ARG	NE-CZ-NH1	-9.11	115.75	120.30
1	L	164	ARG	NE-CZ-NH1	8.20	124.40	120.30
1	A	364	TYR	CG-CD1-CE1	7.91	127.62	121.30
1	L	164	ARG	CD-NE-CZ	7.25	133.76	123.60

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4209	0	4145	56	0
1	L	4206	0	4142	136	0
2	B	1358	0	1408	35	0
2	E	1353	0	1403	112	0
3	G	1108	0	1114	11	0
3	M	1098	0	1103	48	0
4	A	1	0	0	0	0
4	L	1	0	0	0	0
5	A	91	0	87	6	0
5	L	91	0	87	11	0
6	A	5	0	6	0	0
6	L	5	0	6	0	0
7	A	490	0	0	10	0
7	B	118	0	0	8	0
7	E	37	0	0	3	0
7	G	152	0	0	2	0
7	L	319	0	0	13	0
7	M	95	0	0	3	0
All	All	14737	0	13501	390	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 14.

The worst 5 of 390 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:60:GLN:HG2	2:B:126:SER:HA	1.44	0.99
2:E:190:GLN:H	2:E:190:GLN:HE21	1.10	0.92
2:E:77:VAL:HG12	2:E:83:PRO:HA	1.58	0.84
3:M:56:ALA:HB2	3:M:80:ARG:HB2	1.59	0.82
2:B:174:ARG:HB2	2:B:181:PRO:HG3	1.62	0.81

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	A	551/554 (100%)	531 (96%)	18 (3%)	2 (0%)	43	24
1	L	550/554 (99%)	516 (94%)	30 (6%)	4 (1%)	30	12
2	B	176/224 (79%)	169 (96%)	7 (4%)	0	100	100
2	E	175/224 (78%)	158 (90%)	13 (7%)	4 (2%)	10	1
3	G	138/173 (80%)	135 (98%)	3 (2%)	0	100	100
3	M	136/173 (79%)	131 (96%)	3 (2%)	2 (2%)	15	4
All	All	1726/1902 (91%)	1640 (95%)	74 (4%)	12 (1%)	30	12

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	50	GLU
2	E	52	GLY
2	E	58	THR
3	M	58	ASN
1	L	483	THR

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	452/453 (100%)	434 (96%)	18 (4%)	42	21
1	L	451/453 (100%)	423 (94%)	28 (6%)	26	7
2	B	146/183 (80%)	135 (92%)	11 (8%)	19	4
2	E	146/183 (80%)	119 (82%)	27 (18%)	2	0
3	G	119/141 (84%)	117 (98%)	2 (2%)	73	59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	M	117/141 (83%)	105 (90%)	12 (10%)	10	1
All	All	1431/1554 (92%)	1333 (93%)	98 (7%)	22	6

5 of 98 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	423	PHE
1	L	483	THR
3	M	62	ASP
1	L	432	THR
1	L	457	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	334	ASN
1	L	465	ASN
2	E	190	GLN
1	L	352	GLN
1	L	467	ASN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	$\# Z > 2$
5	CNC	A	601	-	101,101,103	3.22	6 (5%)	149,166,171	1.42	14 (9%)
6	PGO	A	602	4	4,4,4	0.82	0	4,4,4	0.68	0
5	CNC	L	601	-	101,101,103	3.15	7 (6%)	149,166,171	1.30	16 (10%)
6	PGO	L	602	4	4,4,4	0.96	0	4,4,4	0.61	0

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
5	CNC	A	601	-	-	0/56/223/235	0/1/11/11
6	PGO	A	602	4	-	0/2/2/2	0/0/0/0
5	CNC	L	601	-	-	0/56/223/235	0/1/11/11
6	PGO	L	602	4	-	0/2/2/2	0/0/0/0

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	CNC	CO-N21	29.48	1.85	1.79
5	L	601	CNC	CO-N21	29.07	1.85	1.79
5	A	601	CNC	C19-N24	7.10	1.51	1.33
5	L	601	CNC	C19-N24	6.71	1.50	1.33
5	A	601	CNC	C1-C19	5.48	1.57	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	601	CNC	C19-C1-N21	-9.29	100.28	105.78
5	L	601	CNC	C19-C1-N21	-6.27	102.06	105.78
5	L	601	CNC	C55-C17-C16	4.40	123.70	110.84
5	A	601	CNC	C35-C5-C6	4.25	124.79	118.64
5	A	601	CNC	C55-C17-C16	3.85	122.11	110.84

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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	A	551/554 (99%)	-0.37	2 (0%) 90 91	10, 17, 34, 86	0
1	L	551/554 (99%)	0.08	26 (4%) 30 28	10, 24, 68, 100	0
2	B	178/224 (79%)	-0.19	1 (0%) 86 87	15, 28, 51, 78	0
2	E	177/224 (79%)	1.26	43 (24%) 1 1	21, 49, 100, 112	0
3	G	137/173 (79%)	-0.44	1 (0%) 84 85	14, 22, 44, 70	0
3	M	137/173 (79%)	0.18	11 (8%) 12 12	19, 30, 70, 95	0
All	All	1731/1902 (91%)	-0.01	84 (4%) 29 27	10, 23, 65, 112	0

The worst 5 of 84 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	51	VAL	7.1
3	M	66	LEU	6.3
1	L	454	VAL	5.9
2	E	47	PHE	5.5
2	E	100	ILE	5.5

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 carbohydrates 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. 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, 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	PGO	L	602	5/5	0.16	1.91	24,24,26,29	0
6	PGO	A	602	5/5	0.16	1.10	11,11,14,17	0
5	CNC	L	601	91/93	0.12	0.07	16,28,44,67	0
5	CNC	A	601	91/93	0.11	0.05	9,17,25,51	0
4	K	A	603	1/1	0.07	-2.44	13,13,13,13	0
4	K	L	603	1/1	0.07	-4.64	20,20,20,20	0

6.5 Other polymers

There are no such residues in this entry.