W O R L D W I D E
wwPDB
P R O T E I N D A T A B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 02:21 am GMT

PDB ID	:	1IWB
Title	:	Crystal structure of diol dehydratase
Authors	:	Shibata, N.; Masuda, J.; Morimoto, Y.; Yasuoka, N.; Toraya, T.
Deposited on	:	2002-05-01
Resolution	:	1.85 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
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:

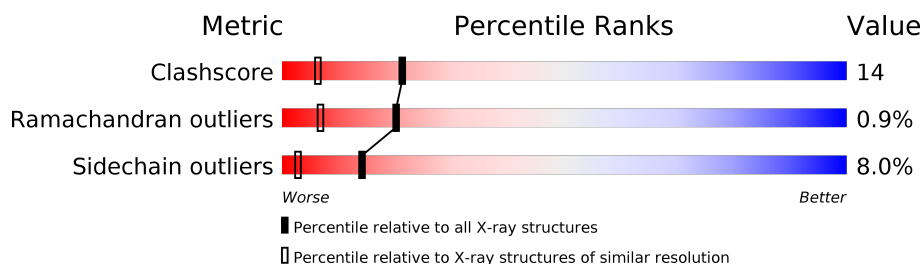|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (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 1.85 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	Percentile Ranks	Value
Clashscore		14
Ramachandran outliers		0.9%
Sidechain outliers		8.0%

Worse — *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	2083 (1.86-1.86)
Ramachandran outliers	110173	2060 (1.86-1.86)
Sidechain outliers	110143	2060 (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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	554	79% 18% • •
1	L	554	76% 21% • •
2	B	224	53% 23% • 21%
2	E	224	38% 33% 8% 21%
3	G	173	53% 21% • • 21%
3	M	173	55% 20% • • 21%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 14446 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 DIOL DEHYDRATASE alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	551	Total	C	N	O	S	0	0	0
			4201	2620	727	825	29			
1	L	551	Total	C	N	O	S	0	0	0
			4200	2620	727	824	29			

- Molecule 2 is a protein called DIOL DEHYDRATASE beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	178	Total	C	N	O	S	0	0	0
			1357	859	244	252	2			
2	E	177	Total	C	N	O	S	0	0	0
			1352	856	243	251	2			

- Molecule 3 is a protein called DIOL DEHYDRATASE gamma chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	137	Total	C	N	O	S	0	0	0
			1093	681	195	214	3			
3	M	137	Total	C	N	O	S	0	0	0
			1093	681	195	214	3			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	K	0	0
			3	3		
4	L	3	Total	K	0	0
			3	3		

- Molecule 5 is COBALAMIN (three-letter code: B12) (formula: $C_{62}H_{89}CoN_{13}O_{14}P$).

B12
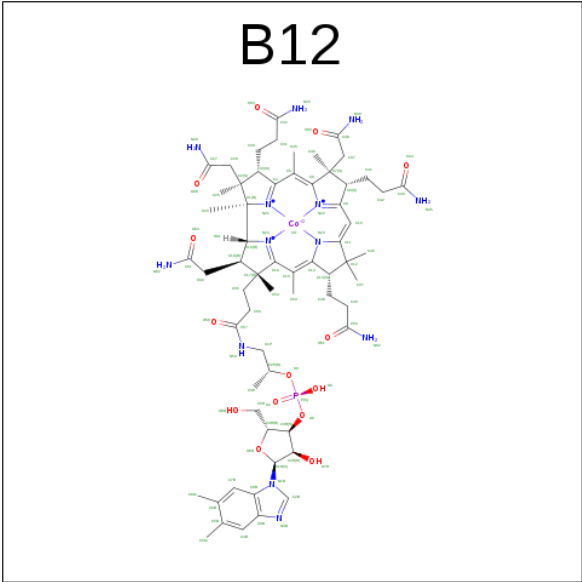
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	B	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		
5	E	1	Total	C	Co	N	O	P	0	0
			91	62	1	13	14	1		

- Molecule 6 is water.

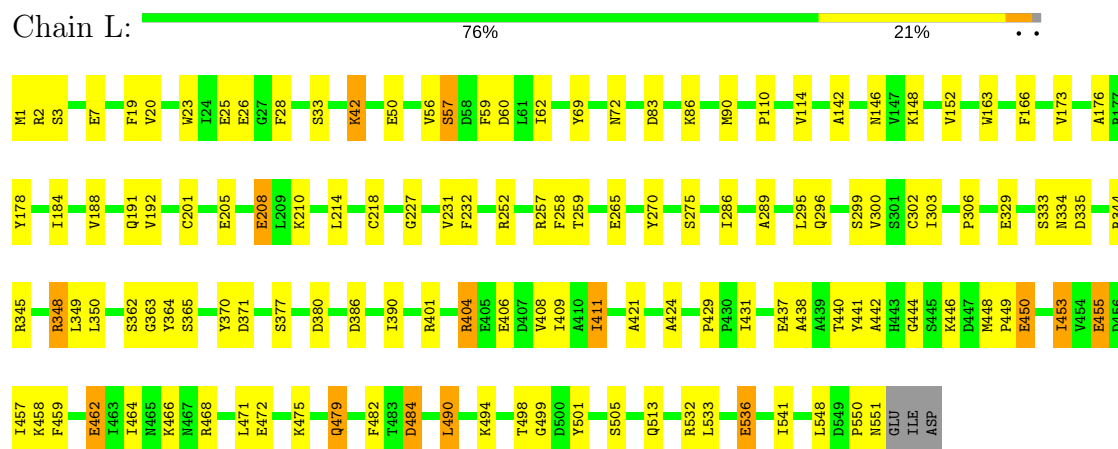| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	335	Total	O	0	0
			335	335		
6	B	89	Total	O	0	0
			89	89		
6	G	70	Total	O	0	0
			70	70		
6	L	330	Total	O	0	0
			330	330		
6	E	54	Total	O	0	0
			54	54		
6	M	84	Total	O	0	0
			84	84		

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 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.
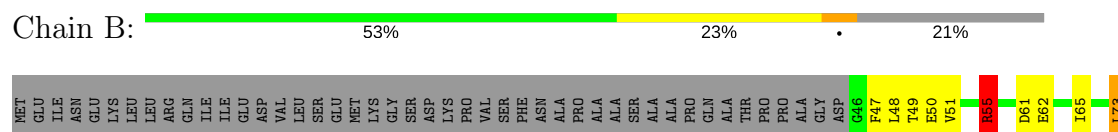
Note EDS was not executed.

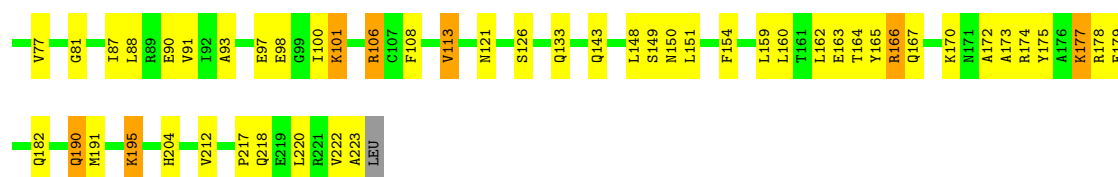• Molecule 1: DIOL DEHYDRATASE alpha chain
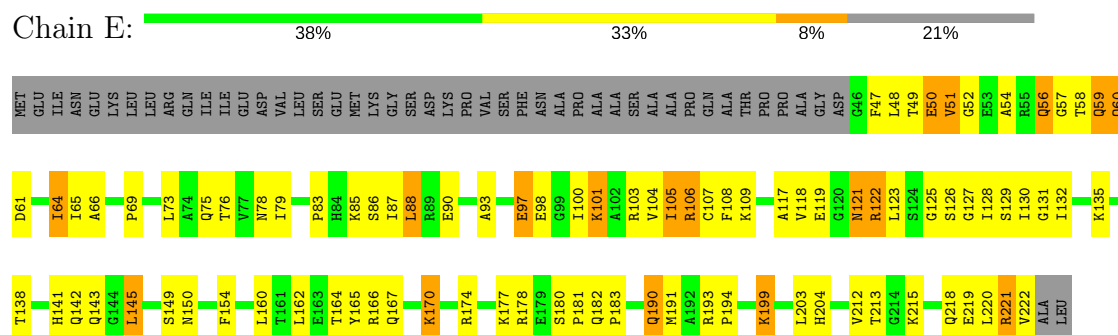
Chain A:

• Molecule 1: DIOL DEHYDRATASE alpha chain
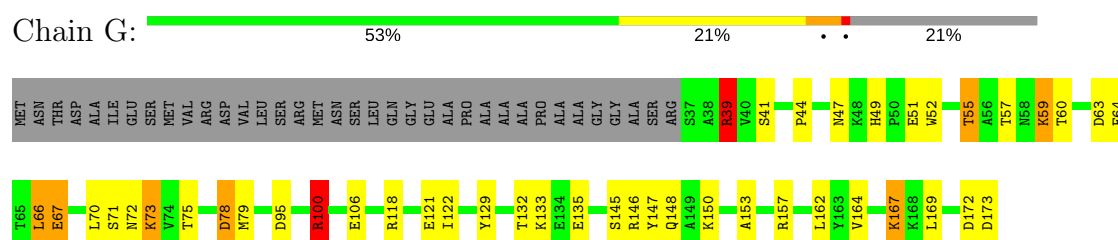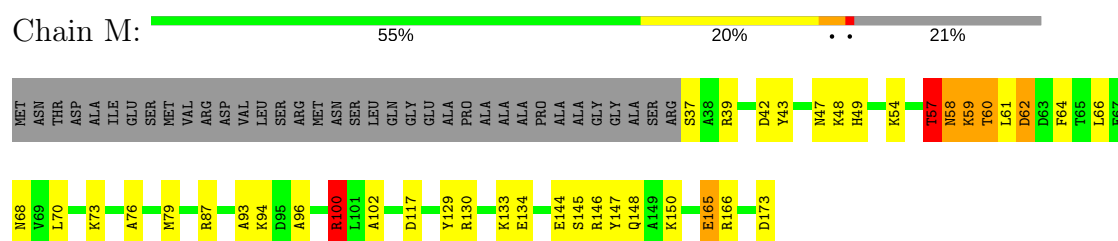
Chain L:

• Molecule 2: DIOL DEHYDRATASE beta chain

Chain B:

V77 G81 I87 L88 R89 E90 V91 I92 A93 E97 E98 G99 I100 K101 R106 C107 F108 V113 N121 S126 Q133 Q143 L148 S149 N150 L151 F154 L159 L160 T161 L162 E163 T164 Y165 R166 Q167 K170 N171 A172 A173 R174 Y175 A176 K177 R178 E179

Q182 Q190 M191 K195 H204 V212 P217 Q218 E219 L220 R221 V222 A223 LEU

● Molecule 2: DIOL DEHYDRATASE beta chain

Chain E:

38%   33%   8%   21%

MET GLU ILE ASN GLU LYS LEU LEU ARG GLN ILE ILE GLU ASP VAL LEU SER GLU MET LYS GLY SER ASP LYS PRO VAL SER PHE ASN ALA PRO ALA ALA SER ALA ALA PRO GLN ALA THR PRO PRO ALA GLY ASP G46 F47 L48 T49 E50 V51 G52 E53 A54 R55 Q56 G57 T58 Q59 Q60

D61 I64 I65 A66 P69 L73 A74 Q75 T76 V77 N78 I79 P83 H84 K85 S86 I87 L88 R89 E90 A93 E97 E98 G99 I100 K101 A102 R103 V104 I105 R106 C107 F108 K109 A117 V118 E119 G120 N121 R122 L123 S124 G125 S126 G127 I128 S129 I130 G131 I132 K135

T138 H141 Q142 Q143 G144 L145 S149 N150 F154 L160 T161 L162 E163 T164 Y165 R166 Q167 K170 R174 K177 R178 E179 S180 P181 Q182 P183 Q190 M191 A192 R193 P194 K199 L203 H204 V212 T213 G214 K215 Q218 E219 L220 R221 V222 ALA LEU

● Molecule 3: DIOL DEHYDRATASE gamma chain

Chain G:

53%   21%   ● ●   21%

MET ASN THR ASP ALA ILE GLU SER MET VAL ARG ASP VAL LEU SER ARG MET ASN SER LEU GLN GLY GLU ALA PRO ALA ALA PRO ALA ALA GLY GLY ALA SER ARG S37 A38 R39 V40 S41 P44 N47 K48 H49 P50 E51 W52 T55 A56 T57 N58 K59 T60 D63 F64

T65 L66 E67 L70 S71 N72 K73 V74 T75 D78 M79 D95 R100 E106 R118 E121 I122 Y129 T132 K133 E134 E135 S145 R146 Y147 Q148 A149 K150 A153 R157 L162 Y163 V164 K167 K168 L169 D172 D173

● Molecule 3: DIOL DEHYDRATASE gamma chain

Chain M:

55%   20%   ● ●   21%

MET ASN THR ASP ALA ILE GLU SER MET VAL ARG ASP VAL LEU SER ARG MET ASN SER LEU GLN GLY GLU ALA PRO ALA ALA PRO ALA ALA GLY GLY ALA SER ARG S37 A38 R39 D42 Y43 N47 K48 H49 K54 T57 N58 K59 T60 L61 D62 D63 F64 T65 L66 E67

N68 V69 L70 K73 A76 M79 R87 A93 K94 D95 A96 R100 L101 A102 D117 Y129 R130 K133 E134 E144 S145 R146 Y147 Q148 A149 K150 E165 R166 D173

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	75.79Å 122.40Å 207.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.85	Depositor
% Data completeness (in resolution range)	(Not available) (10.00-1.85)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.181 , 0.256	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	14446	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: K, B12

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	0.33	0/4273	0.99	9/5787 (0.2%)
1	L	0.34	0/4272	1.01	9/5786 (0.2%)
2	B	0.31	0/1379	0.98	4/1867 (0.2%)
2	E	0.31	0/1374	0.90	1/1860 (0.1%)
3	G	0.31	0/1108	0.96	5/1497 (0.3%)
3	M	0.33	0/1108	1.02	4/1497 (0.3%)
All	All	0.33	0/13514	0.99	32/18294 (0.2%)

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	A	0	1
1	L	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	M	100	ARG	CD-NE-CZ	10.07	137.69	123.60
2	B	55	ARG	NE-CZ-NH1	9.79	125.19	120.30
2	B	55	ARG	CD-NE-CZ	8.70	135.78	123.60
1	A	345	ARG	NE-CZ-NH1	-8.62	115.99	120.30
3	M	87	ARG	CD-NE-CZ	7.50	134.10	123.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	205	GLU	Sidechain
1	L	208	GLU	Sidechain

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	4201	0	4138	73	0
1	L	4200	0	4139	99	0
2	B	1357	0	1408	47	0
2	E	1352	0	1403	87	0
3	G	1093	0	1101	38	0
3	M	1093	0	1101	35	0
4	A	3	0	0	0	0
4	L	3	0	0	0	0
5	B	91	0	88	9	0
5	E	91	0	88	5	0
6	A	335	0	0	7	0
6	B	89	0	0	2	0
6	E	54	0	0	3	0
6	G	70	0	0	1	0
6	L	330	0	0	6	0
6	M	84	0	0	2	0
All	All	14446	0	13466	366	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 14.

The worst 5 of 366 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)
3:M:37:SER:HB3	3:M:39:ARG:HH21	1.32	0.94
2:E:190:GLN:H	2:E:190:GLN:HE21	1.14	0.94
2:B:47:PHE:HA	2:B:223:ALA:HB3	1.52	0.91
1:L:513:GLN:HE22	3:M:47:ASN:HD21	1.13	0.90
2:B:190:GLN:HE21	2:B:190:GLN:H	1.22	0.88

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	549/554 (99%)	526 (96%)	17 (3%)	6 (1%)	17	5
1	L	549/554 (99%)	532 (97%)	14 (3%)	3 (0%)	32	16
2	B	176/224 (79%)	169 (96%)	7 (4%)	0	100	100
2	E	175/224 (78%)	161 (92%)	8 (5%)	6 (3%)	4	0
3	G	135/173 (78%)	131 (97%)	4 (3%)	0	100	100
3	M	135/173 (78%)	132 (98%)	2 (2%)	1 (1%)	25	10
All	All	1719/1902 (90%)	1651 (96%)	52 (3%)	16 (1%)	20	7

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	50	GLU
2	E	59	GLN
3	M	57	THR
2	E	51	VAL
2	E	52	GLY

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	450/453 (99%)	425 (94%)	25 (6%)	25	8

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	L	450/453 (99%)	426 (95%)	24 (5%)	26	10
2	B	146/183 (80%)	130 (89%)	16 (11%)	7	1
2	E	146/183 (80%)	122 (84%)	24 (16%)	2	0
3	G	116/141 (82%)	102 (88%)	14 (12%)	6	1
3	M	116/141 (82%)	105 (90%)	11 (10%)	10	1
All	All	1424/1554 (92%)	1310 (92%)	114 (8%)	14	3

5 of 114 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	121	GLU
1	L	348	ARG
3	M	59	LYS
3	G	150	LYS
1	L	42	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	45	ASN
1	L	513	GLN
2	E	218	GLN
1	L	334	ASN
1	L	352	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 8 ligands modelled in this entry, 6 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
5	B12	B	1601	-	73,101,101	0.98	5 (6%)	111,166,166	1.32	13 (11%)
5	B12	E	1602	-	73,101,101	1.00	3 (4%)	111,166,166	1.33	15 (13%)

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	B12	B	1601	-	-	0/51/223/223	0/3/11/11
5	B12	E	1602	-	-	0/51/223/223	0/3/11/11

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1601	B12	C20-C1	-2.10	1.49	1.53
5	B	1601	B12	C8B-N1B	-2.09	1.36	1.38
5	B	1601	B12	C11-C10	-2.08	1.37	1.41
5	B	1601	B12	C41-C8	2.05	1.57	1.54
5	B	1601	B12	O58-C57	2.19	1.27	1.23

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	1602	B12	C54-C17-C18	-4.02	106.69	112.94
5	E	1602	B12	O58-C57-C56	-3.30	115.80	122.01
5	E	1602	B12	C1-C19-C18	-2.96	116.96	121.90
5	E	1602	B12	C3-C4-C5	-2.89	121.81	131.85
5	B	1601	B12	O58-C57-C56	-2.84	116.68	122.01

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1601	B12	9	0
5	E	1602	B12	5	0

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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.