



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:47 PM GMT

PDB ID : 1LYW
Title : CATHEPSIN D AT PH 7.5
Authors : Lee, A.Y.; Gulnik, S.V.; Erickson, J.W.
Deposited on : 1998-06-30
Resolution : 2.50 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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

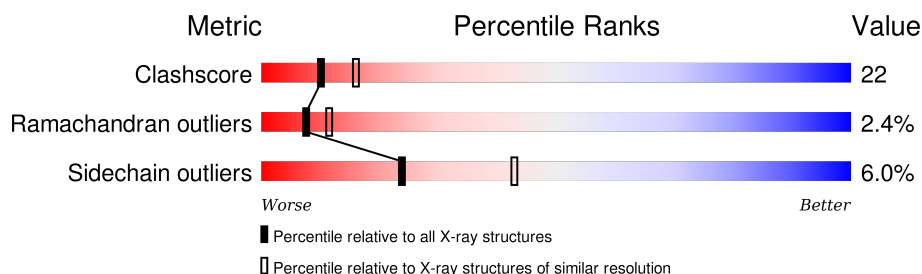
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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))
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)


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 $\leq 5\%$. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	97	
1	C	97	
1	E	97	
1	G	97	
2	B	241	
2	D	241	
2	F	241	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	241	 A horizontal bar chart showing the quality of chain H. The bar is divided into two segments: a green segment on the left labeled '52%' and a yellow segment on the right labeled '44%'. A small orange dot is visible at the far right end of the bar.

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
3	EPE	C	98	-	-	X	-
3	EPE	G	98	-	-	X	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14242 atoms, of which 3296 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 CATHEPSIN D.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	95	Total	C	H	N	O	S	0	0	0
			890	471	153	117	144	5			
1	C	95	Total	C	H	N	O	S	0	0	0
			890	471	153	117	144	5			
1	E	95	Total	C	H	N	O	S	0	0	0
			890	471	153	117	144	5			
1	G	95	Total	C	H	N	O	S	0	0	0
			890	471	153	117	144	5			

- Molecule 2 is a protein called CATHEPSIN D.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	241	Total	C	H	N	O	S	0	0	0
			2235	1184	390	302	348	11			
2	D	241	Total	C	H	N	O	S	0	0	0
			2235	1184	390	302	348	11			
2	F	241	Total	C	H	N	O	S	0	0	0
			2235	1184	390	302	348	11			
2	H	241	Total	C	H	N	O	S	0	0	0
			2235	1184	390	302	348	11			

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C₈H₁₈N₂O₄S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total 17	C 8	H 2	N 2	O 4	S 1	0	0
3	C	1	Total 17	C 8	H 2	N 2	O 4	S 1	0	0
3	G	1	Total 17	C 8	H 2	N 2	O 4	S 1	0	0
3	E	1	Total 17	C 8	H 2	N 2	O 4	S 1	0	0

- Molecule 4 is water.

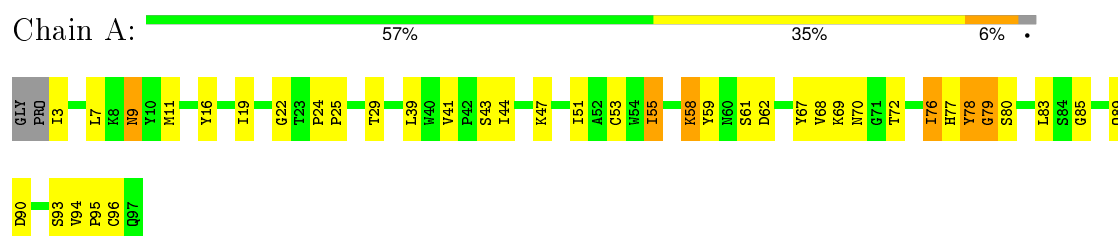
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	55	Total	H	O	0	0
			165	110	55		
4	B	110	Total	H	O	0	0
			330	220	110		
4	C	55	Total	H	O	0	0
			165	110	55		
4	D	78	Total	H	O	0	0
			234	156	78		
4	E	46	Total	H	O	0	0
			138	92	46		
4	F	90	Total	H	O	0	0
			270	180	90		
4	G	39	Total	H	O	0	0
			117	78	39		
4	H	85	Total	H	O	0	0
			255	170	85		

3 Residue-property plots [i](#)

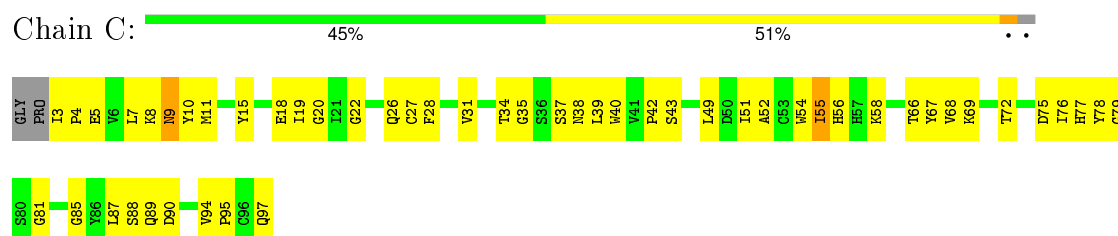
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

Note EDS was not executed.

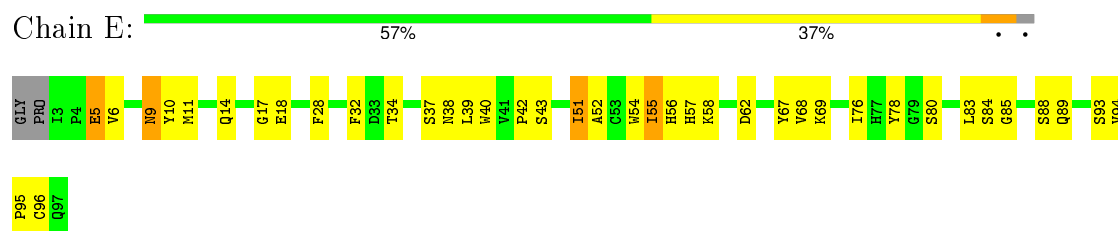
• Molecule 1: CATHEPSIN D



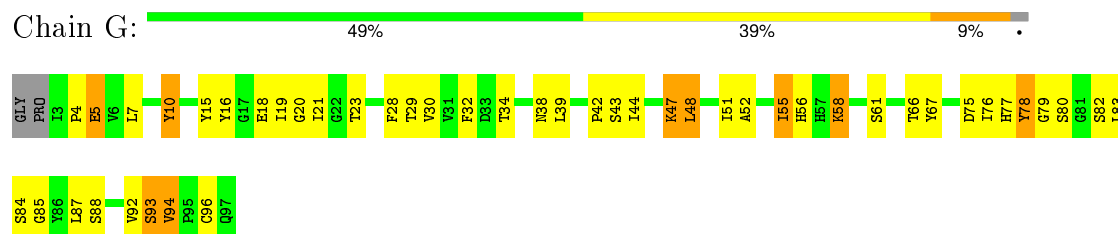
• Molecule 1: CATHEPSIN D



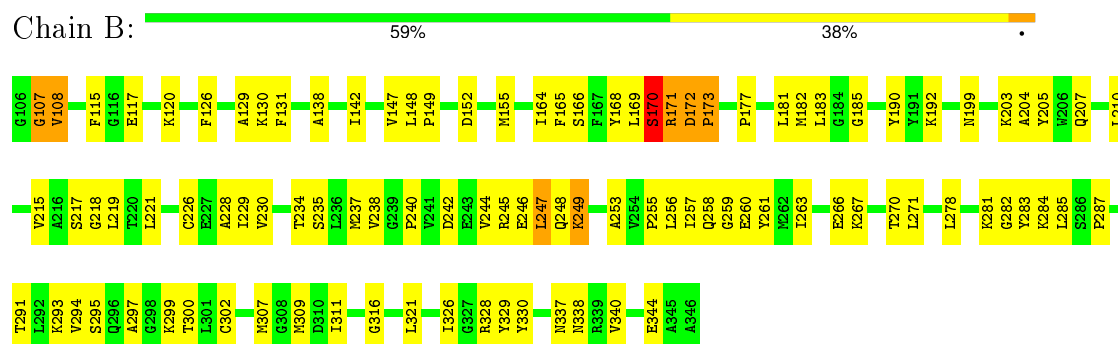
• Molecule 1: CATHEPSIN D



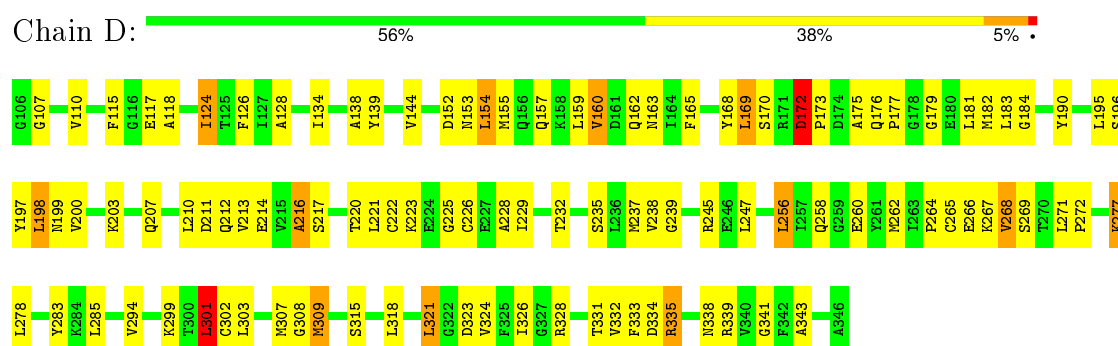
• Molecule 1: CATHEPSIN D



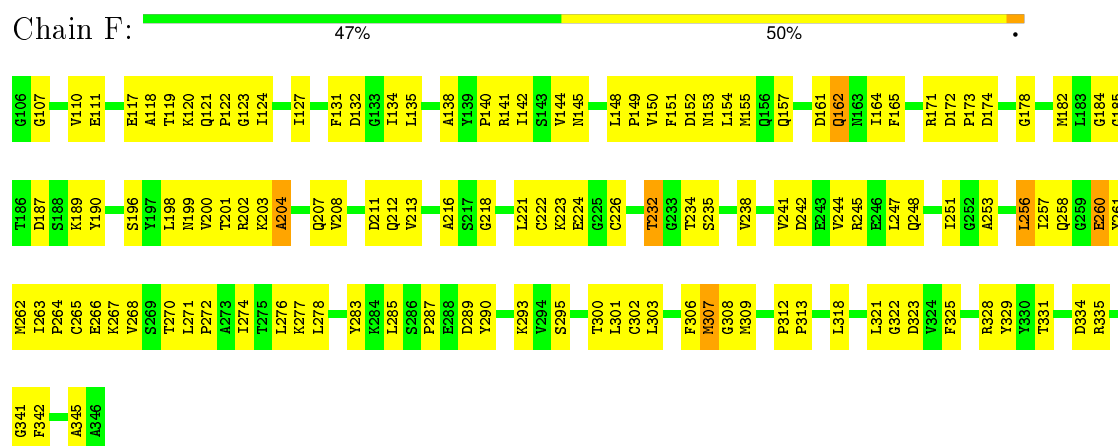
- Molecule 2: CATHEPSIN D



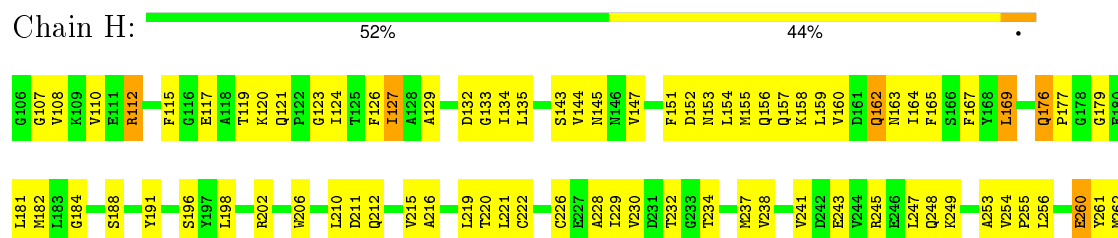
- Molecule 2: CATHEPSIN D



- Molecule 2: CATHEPSIN D



- Molecule 2: CATHEPSIN D



I263	P264	C265	E266	K267	V268		L271	P272		L276	K277	L278	G279	G280		Y283	K284	L285		Y290		L301	C302	L303		M307	G308	M309	D310	I311	P312	P313	P314	S315		L318	G319	I320	L321	G322	D323	V324	P325	I326		Y330		R335		M338	R339	V340	G341	F342	L343		G346
------	------	------	------	------	------	--	------	------	--	------	------	------	------	------	--	------	------	------	--	------	--	------	------	------	--	------	------	------	------	------	------	------	------	------	--	------	------	------	------	------	------	------	------	------	--	------	--	------	--	------	------	------	------	------	------	--	------

4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	140.26Å 136.80Å 140.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.50	Depositor
% Data completeness (in resolution range)	73.0 (8.00-2.50)	Depositor
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.195 , 0.257	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	14242	wwPDB-VP
Average B, all atoms (Å ²)	16.0	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: EPE

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	0.46	0/759	0.78	1/1034 (0.1%)
1	C	0.47	0/759	0.78	0/1034
1	E	0.51	0/759	0.79	2/1034 (0.2%)
1	G	0.50	0/759	0.84	0/1034
2	B	0.56	0/1884	0.83	4/2551 (0.2%)
2	D	0.58	1/1884 (0.1%)	0.91	10/2551 (0.4%)
2	F	0.58	2/1884 (0.1%)	0.85	6/2551 (0.2%)
2	H	0.55	0/1884	0.88	9/2551 (0.4%)
All	All	0.55	3/10572 (0.0%)	0.85	32/14340 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	335	ARG	CZ-NH1	6.71	1.41	1.33
2	F	335	ARG	CZ-NH1	5.72	1.40	1.33
2	F	202	ARG	CZ-NH2	5.22	1.39	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	301	LEU	CB-CG-CD2	-8.75	96.12	111.00
2	F	276	LEU	CB-CG-CD2	-8.02	97.36	111.00
2	D	321	LEU	CB-CG-CD1	7.39	123.56	111.00
2	H	135	LEU	CB-CG-CD1	-7.14	98.87	111.00
2	D	321	LEU	CB-CG-CD2	-7.04	99.03	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	737	153	690	37	0
1	C	737	153	690	56	0
1	E	737	153	690	40	0
1	G	737	153	690	49	0
2	B	1845	390	1850	88	0
2	D	1845	390	1850	83	0
2	F	1845	390	1850	88	0
2	H	1845	390	1850	92	0
3	A	15	2	18	4	0
3	C	15	2	18	8	0
3	E	15	2	18	2	0
3	G	15	2	18	11	0
4	A	55	110	0	0	0
4	B	110	220	0	0	0
4	C	55	110	0	2	0
4	D	78	156	0	0	0
4	E	46	92	0	0	0
4	F	90	180	0	2	0
4	G	39	78	0	0	0
4	H	85	170	0	2	0
All	All	10946	3296	10232	451	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 22.

The worst 5 of 451 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:G:7:LEU:HD11	2:H:309:MET:HG2	1.29	1.05
2:H:253:ALA:HB1	2:H:261:TYR:HB3	1.39	1.04
1:E:43:SER:HB2	2:F:117:GLU:HG2	1.39	1.03
1:G:83:LEU:HD23	3:G:98:EPE:H92	1.42	1.02
2:B:172:ASP:HB2	2:B:173:PRO:HD3	1.44	1.00

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	93/97 (96%)	80 (86%)	10 (11%)	3 (3%)	5	6
1	C	93/97 (96%)	83 (89%)	9 (10%)	1 (1%)	17	31
1	E	93/97 (96%)	83 (89%)	9 (10%)	1 (1%)	17	31
1	G	93/97 (96%)	79 (85%)	12 (13%)	2 (2%)	8	13
2	B	239/241 (99%)	209 (87%)	23 (10%)	7 (3%)	6	8
2	D	239/241 (99%)	215 (90%)	18 (8%)	6 (2%)	7	10
2	F	239/241 (99%)	218 (91%)	13 (5%)	8 (3%)	5	6
2	H	239/241 (99%)	211 (88%)	24 (10%)	4 (2%)	11	19
All	All	1328/1352 (98%)	1178 (89%)	118 (9%)	32 (2%)	7	11

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	78	TYR
2	B	171	ARG
2	B	173	PRO
1	C	9	ASN
2	D	216	ALA

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	84/86 (98%)	76 (90%)	8 (10%)	11	20

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	84/86 (98%)	82 (98%)	2 (2%)	57	82
1	E	84/86 (98%)	78 (93%)	6 (7%)	18	34
1	G	84/86 (98%)	70 (83%)	14 (17%)	3	5
2	B	199/199 (100%)	190 (96%)	9 (4%)	34	59
2	D	199/199 (100%)	189 (95%)	10 (5%)	30	53
2	F	199/199 (100%)	189 (95%)	10 (5%)	30	53
2	H	199/199 (100%)	190 (96%)	9 (4%)	34	59
All	All	1132/1140 (99%)	1064 (94%)	68 (6%)	24	43

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	11	MET
2	F	232	THR
2	H	176	GLN
1	E	55	ILE
1	E	96	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	145	ASN
2	F	258	GLN
2	H	199	ASN
1	C	77	HIS
2	H	176	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 ⓘ

4 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$
3	EPE	A	98	-	14,15,15	1.34	1 (7%)	18,20,20	2.51	8 (44%)
3	EPE	C	98	-	14,15,15	1.52	1 (7%)	18,20,20	2.49	6 (33%)
3	EPE	E	98	-	14,15,15	1.46	1 (7%)	18,20,20	2.37	6 (33%)
3	EPE	G	98	-	14,15,15	1.02	1 (7%)	18,20,20	2.46	8 (44%)

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
3	EPE	A	98	-	-	0/9/19/19	0/1/1/1
3	EPE	C	98	-	-	0/9/19/19	0/1/1/1
3	EPE	E	98	-	-	0/9/19/19	0/1/1/1
3	EPE	G	98	-	-	0/9/19/19	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	98	EPE	O3S-S	3.45	1.55	1.46
3	A	98	EPE	O3S-S	4.20	1.57	1.46
3	E	98	EPE	O3S-S	5.22	1.59	1.46
3	C	98	EPE	O3S-S	5.53	1.60	1.46

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	98	EPE	C5-N4-C3	-5.50	96.99	108.90
3	C	98	EPE	C5-N4-C3	-5.10	97.84	108.90
3	C	98	EPE	O3S-S-O2S	-4.94	100.11	111.61
3	E	98	EPE	C9-N1-C2	-4.06	100.86	111.27
3	G	98	EPE	C7-N4-C3	-4.03	100.92	111.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 25 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	98	EPE	4	0
3	C	98	EPE	8	0
3	E	98	EPE	2	0
3	G	98	EPE	11	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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.