Oct 17, 2021 – 09:55 AM EDT

PDB ID : 1JAA
Title : Solution structure of lactam analogue (DapE) of HIV gp41 600-612 loop.
Authors : Phan Chan Du, A.; Limal, D.; Semetey, V.; Dali, H.; Jolivet, M.; Desgranges, C.; Cung, M.T.; Briand, J.P.; Petit, M.C.; Muller, S.
Deposited on : 2001-05-30

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
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 : 1.8.5 (274361), CSD as541be (2020)
- Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
- RCI : v_In_11_5_13_A (Berjanski et al., 2005)
- PANAV : Wang et al. (2010)
- ShiftChecker : 2.23.2
- Ideal geometry (proteins) : Engh & Huber (2001)
- Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
- Validation Pipeline (wwPDB-VP) : 2.23.2
1 Overall quality at a glance

The following experimental techniques were used to determine the structure: SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Whole archive (#Entries)</th>
<th>NMR archive (#Entries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clashscore</td>
<td>158937</td>
<td>12864</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td>154571</td>
<td>11451</td>
</tr>
<tr>
<td>Sidechain outliers</td>
<td>154315</td>
<td>11428</td>
</tr>
</tbody>
</table>

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Length</th>
<th>Quality of chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>14</td>
<td>57%</td>
</tr>
</tbody>
</table>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA and RNA chains that are outliers for geometric criteria:

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Compound</th>
<th>Res</th>
<th>Total models with violations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Chirality</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>SET</td>
<td>5</td>
<td>50</td>
</tr>
</tbody>
</table>
2 Ensemble composition and analysis

This entry contains 50 models. Model 27 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: lowest energy.

The following residues are included in the computation of the global validation metrics.

<table>
<thead>
<tr>
<th>Well-defined (core) protein residues</th>
<th>Residue range (total)</th>
<th>Backbone RMSD (Å)</th>
<th>Medoid model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-defined core</td>
<td>A:3-A:4, A:6-A:12 (9)</td>
<td>0.77</td>
<td>27</td>
</tr>
</tbody>
</table>

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 8 clusters. No single-model clusters were found.

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 49, 50</td>
</tr>
<tr>
<td>2</td>
<td>33, 34, 35, 44, 45, 46, 47, 48</td>
</tr>
<tr>
<td>3</td>
<td>12, 13, 15, 24, 25, 26, 27, 28</td>
</tr>
<tr>
<td>4</td>
<td>14, 16, 17, 18, 19, 20, 21</td>
</tr>
<tr>
<td>5</td>
<td>1, 29, 30, 31, 32</td>
</tr>
<tr>
<td>6</td>
<td>36, 37, 38, 39, 40</td>
</tr>
<tr>
<td>7</td>
<td>41, 42, 43</td>
</tr>
<tr>
<td>8</td>
<td>22, 23</td>
</tr>
</tbody>
</table>
3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 199 atoms, of which 101 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Residues</th>
<th>Atoms</th>
<th>Trace</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>14</td>
<td>Total C H N O</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>199 63 101 16 19</td>
<td></td>
</tr>
</tbody>
</table>

There are 3 discrepancies between the modelled and reference sequences:

<table>
<thead>
<tr>
<th>Chain</th>
<th>Residue</th>
<th>Modelled</th>
<th>Actual</th>
<th>Comment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>ACE</td>
<td>-</td>
<td>acetylation</td>
<td>UNP P12488</td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>SET</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P12488</td>
</tr>
<tr>
<td>A</td>
<td>11</td>
<td>GLU</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P12488</td>
</tr>
</tbody>
</table>
4  Residue-property plots

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.2 Score per residue for model 2

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
4.2.3 Score per residue for model 3

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.4 Score per residue for model 4

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.5 Score per residue for model 5

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.6 Score per residue for model 6

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.7 Score per residue for model 7

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
4.2.8 Score per residue for model 8

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.9 Score per residue for model 9

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.10 Score per residue for model 10

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.11 Score per residue for model 11

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.12 Score per residue for model 12

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
4.2.13  Score per residue for model 13

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.14  Score per residue for model 14

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.15  Score per residue for model 15

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.16  Score per residue for model 16

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.17  Score per residue for model 17

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
4.2.18 Score per residue for model 18

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.19 Score per residue for model 19

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.20 Score per residue for model 20

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.21 Score per residue for model 21

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.22 Score per residue for model 22

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
4.2.23 Score per residue for model 23

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

```
ACE1 I2 W3 G4 S5 T13 A14
```

4.2.24 Score per residue for model 24

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

```
ACE1 I2 W3 G4 S5 T13 A14
```

4.2.25 Score per residue for model 25

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

```
ACE1 I2 W3 G4 S5 T13 A14
```

4.2.26 Score per residue for model 26

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

```
ACE1 I2 W3 G4 S5 T13 A14
```

4.2.27 Score per residue for model 27 (medoid)

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

```
ACE1 I2 W3 G4 S5 T13 A14
```
4.2.28 Score per residue for model 28

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.29 Score per residue for model 29

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.30 Score per residue for model 30

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.31 Score per residue for model 31

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.32 Score per residue for model 32

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
4.2.33  Score per residue for model 33

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

50%  7%  7%  36%

4.2.34  Score per residue for model 34

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

50%  14%  36%

4.2.35  Score per residue for model 35

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

50%  14%  36%

4.2.36  Score per residue for model 36

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

43%  21%  36%

4.2.37  Score per residue for model 37

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

36%  29%  36%
4.2.38 Score per residue for model 38

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.39 Score per residue for model 39

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.40 Score per residue for model 40

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.41 Score per residue for model 41

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.42 Score per residue for model 42

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
4.2.43  Score per residue for model 43

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.44  Score per residue for model 44

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.45  Score per residue for model 45

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.46  Score per residue for model 46

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.47  Score per residue for model 47

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
4.2.48 Score per residue for model 48

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.49 Score per residue for model 49

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:

4.2.50 Score per residue for model 50

- Molecule 1: DapE : (Ace)IWG(Dap)SGKLIETTA ANALOGUE OF HIV GP41

Chain A:
5 Refinement protocol and experimental data overview

The models were refined using the following method: *Torsion angle dynamics, molecular dynamics, energy minimization*.

Of the 50 calculated structures, 50 were deposited, based on the following criterion: *all calculated structures submitted*.

The following table shows the software used for structure solution, optimisation and refinement.

<table>
<thead>
<tr>
<th>Software name</th>
<th>Classification</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYANA</td>
<td>refinement</td>
<td>1.5</td>
</tr>
<tr>
<td>Discover</td>
<td>refinement</td>
<td>3</td>
</tr>
</tbody>
</table>

No chemical shift data was provided.
6 Model quality

6.1 Standard geometry

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

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 (average) root-mean-square of all Z scores of the bond lengths (or angles).

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Bond lengths</th>
<th>Bond angles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSZ</td>
<td>#Z&gt;5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.29±0.02</td>
<td>0±0/69 (0.0±0.0%)</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>1.29</td>
<td>0/3450 (0.0%)</td>
</tr>
</tbody>
</table>

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>Atoms</th>
<th>Z</th>
<th>Observed(°)</th>
<th>Ideal(°)</th>
<th>Models</th>
<th>Worst</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>3</td>
<td>TRP</td>
<td>CD1-NE1-CE2</td>
<td>-6.05</td>
<td>103.55</td>
<td>109.00</td>
<td>8</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>3</td>
<td>TRP</td>
<td>CA-CB-CG</td>
<td>5.56</td>
<td>124.26</td>
<td>113.70</td>
<td>7</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

There are no chirality outliers.

There are no planarity outliers.

6.2 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 each 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 averaged over the ensemble.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Non-H</th>
<th>H(model)</th>
<th>H(added)</th>
<th>Clashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>All</td>
<td>3400</td>
<td>3450</td>
<td>3400</td>
<td>-</td>
</tr>
</tbody>
</table>

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 -.

There are no clashes.
6.3 Torsion angles

6.3.1 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 PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Analysed</th>
<th>Favoured</th>
<th>Allowed</th>
<th>Outliers</th>
<th>Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>8/14 (57%)</td>
<td>4±1 (45±12%)</td>
<td>4±1 (50±15%)</td>
<td>0±1 (6±7%)</td>
<td>3 23</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>400/700 (57%)</td>
<td>180 (45%)</td>
<td>198 (50%)</td>
<td>22 (6%)</td>
<td>3 23</td>
</tr>
</tbody>
</table>

All 2 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>Models (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>10</td>
<td>ILE</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>4</td>
<td>GLY</td>
<td>5</td>
</tr>
</tbody>
</table>

6.3.2 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 PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Analysed</th>
<th>Rotameric</th>
<th>Outliers</th>
<th>Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>7/9 (78%)</td>
<td>6±1 (87±13%)</td>
<td>1±1 (13±13%)</td>
<td>8 50</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>350/450 (78%)</td>
<td>306 (87%)</td>
<td>44 (13%)</td>
<td>8 50</td>
</tr>
</tbody>
</table>

All 4 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>Models (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>11</td>
<td>GLU</td>
<td>19</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>8</td>
<td>LYS</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>3</td>
<td>TRP</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>10</td>
<td>ILE</td>
<td>4</td>
</tr>
</tbody>
</table>
6.3.3 RNA

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Type</th>
<th>Chain</th>
<th>Res</th>
<th>Link</th>
<th>Counts</th>
<th>Bond lengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SET</td>
<td>A</td>
<td>5</td>
<td>1</td>
<td>4,5,6</td>
<td>2.55±0.03</td>
</tr>
</tbody>
</table>

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Type</th>
<th>Chain</th>
<th>Res</th>
<th>Link</th>
<th>Counts</th>
<th>Bond angles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SET</td>
<td>A</td>
<td>5</td>
<td>1</td>
<td>2,5,7</td>
<td>1.85±0.10</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Type</th>
<th>Chain</th>
<th>Res</th>
<th>Link</th>
<th>Chirals</th>
<th>Torsions</th>
<th>Rings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SET</td>
<td>A</td>
<td>5</td>
<td>1</td>
<td>1±0,1,1,2</td>
<td>0±0,4,4,6</td>
<td>-</td>
</tr>
</tbody>
</table>

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.
All unique angle outliers are listed below.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>Atoms</th>
<th>Z</th>
<th>Observed(Å)</th>
<th>Ideal(Å)</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>5</td>
<td>SET</td>
<td>OG-CB</td>
<td>4.41</td>
<td>1.23</td>
<td>1.42</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>5</td>
<td>SET</td>
<td>CB-CA</td>
<td>2.76</td>
<td>1.56</td>
<td>1.52</td>
<td>44</td>
</tr>
</tbody>
</table>

All unique chiral outliers are listed below.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>Atoms</th>
<th>Z</th>
<th>Observed(°)</th>
<th>Ideal(°)</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>5</td>
<td>SET</td>
<td>OG-CB-CA</td>
<td>2.77</td>
<td>122.48</td>
<td>111.52</td>
<td>13</td>
</tr>
</tbody>
</table>

There are no torsion outliers.

There are no ring outliers.

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.
7 Chemical shift validation

No chemical shift data were provided