Caffeinated StickWRLD

Interpreting the Diagrams - Getting Started
StickWRLD diagrams

StickWRLD diagrams provide 3 types of information

Sequence Family
Composition by Position
StickWRLD diagrams

StickWRLD diagrams provide 3 types of information:

- Sequence Family
- Composition by Position
- Positional Propensity for Physical Properties
StickWRLD diagrams

StickWRLD diagrams provide 3 types of information:

- Sequence Family Composition by Position
- Positional Propensity for Physical Properties
- Interpositional Correlation of Identities
StickWRLD diagrams

StickWRLD diagrams are live 3D objects and can be:

Rotated
StickWRLD diagrams

StickWRLD diagrams are live 3D objects and can be:

Rotated
StickWRLD diagrams

StickWRLD diagrams are live 3D objects and can be:

- Rotated
StickWRLD diagrams

StickWRLD diagrams are live 3D objects and can be:

- Rotated
- Scaled
StickWRLD diagrams

StickWRLD diagrams are live 3D objects and can be:

- Rotated
- and Scaled
StickWRLD diagrams

StickWRLD diagrams are live 3D objects and can be:

- Rotated
- and Scaled

Thresholds can be changed to restrict correlations displayed:

Tr = 0.100
StickWRLD diagrams

StickWRLD diagrams are live 3D objects and can be:

- Rotated
- and Scaled

Thresholds can be changed to restrict correlations displayed:

$Tr = 0.075$
StickWRLD diagrams

StickWRLD diagrams are live 3D objects and can be:

- Rotated
- and Scaled

Thresholds can be changed to restrict correlations displayed:

$Tr = 0.050$
StickWRLD

diagrams

StickWRLD diagrams are live 3D objects and can be:

- Rotated
- and Scaled

Thresholds can be changed to restrict correlations displayed:

$Tr = 0.025$
StickWRLD Balls

StickWRLD Balls are the visual analog of a Position Specific Scoring Matrix (PSSM).

- Colored by Hydrophobicity
- Default Ordering (Hydropathy)
- Ungrouped (individual amino acids)
StickWRLD Balls

StickWRLD Balls are the visual analog of a Position Specific Scoring Matrix (PSSM)

- Colored by Hydrophobicity
- Default Ordering (Hydropathy)
- Ungrouped (individual amino acids)

Positions such as these have strong propensities for a single residue
StickWRLD Balls

StickWRLD Balls are the visual analog of a Position Specific Scoring Matrix (PSSM).

- Colored by Hydrophobicity
- Default Ordering (Hydropathy)
- Ungrouped (individual amino acids)

While the content of these, are more varied, though not uniformly distributed.
StickWRLD Balls

StickWRLD Balls are the visual analog of a Position Specific Scoring Matrix (PSSM).

- Colored by Hydrophobicity
- Ordered by Polarity
- Ungrouped
StickWRLD Balls are the visual analog of a Position Specific Scoring Matrix (PSSM)

- Colored by Charge
- Ordered by Polarity
- Ungrouped
StickWRLD Balls are the visual analog of a Position Specific Scoring Matrix (PSSM).

Grouped by Hydrophobicity

When grouping, the identities at a position are collapsed based on their physicochemical properties. Here, grouping by Hydrophobicity, it’s clear that the sequence requirements at positions 12 and 13 are not as non-specific as they previously appeared.
**StickWRLD Balls**

StickWRLD Balls are the visual analog of a Position Specific Scoring Matrix (PSSM)

- **Grouped by Hydrophobicity**

---

**Hydrophobic grouping and colors:**

<table>
<thead>
<tr>
<th>Group</th>
<th>Hydrophobicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group1</td>
<td>C</td>
</tr>
<tr>
<td>Group2</td>
<td>IVLFMAWG</td>
</tr>
<tr>
<td>Group3</td>
<td>HSTPYNQDE</td>
</tr>
<tr>
<td>Group4</td>
<td>R</td>
</tr>
<tr>
<td>Group5</td>
<td>K</td>
</tr>
</tbody>
</table>
StickWRLD Balls are the visual analog of a Position Specific Scoring Matrix (PSSM)

By-Charge grouping and colors:

- Basic(+): RHK
- Acidic(-): DE
- Polar: NSTY
- Slightly-polar: CWV
- Non-polar: AGILMFV
StickWRLD Edges (links)

StickWRLD edges link residues and positions that are over or under correlated based on the PSSM.
StickWRLD Edges (links)

StickWRLD edges link residues and positions that are over or under correlated based on the PSSM.

Solid edges indicate positive correlations.

Positive links such as these indicate specific pairs of residues at specific positions, whose mutual occurrence in the family of sequences is strongly linked.
StickWRLD Edges (links)

StickWRLD edges link residues and positions that are over or under correlated based on the PSSM.

Solid edges indicate positive correlations.

Here, the Arg at position 9, shows a strong positive correlation with the Glu at 13, the Gln at 12, and pair of Pros at 10 and 11.
StickWRLD Edges (links)

StickWRLD edges link residues and positions that are over or under correlated based on the PSSM.

Solid edges indicate positive correlations.

Likewise the Glu at 13 is strongly correlated with the Arg at 9, and the Pro at 10 and 11. There is enough sequence variability in this family that the correlation between the Glu at 11 and Gln at 12 is not visible with these threshold settings.
StickWRLD edges link residues and positions that are over or under correlated based on the PSSM.

Solid edges indicate positive correlations.

Of note, the positive correlation between Tyr at 9, and Met at 12, has no positive relationships to any of the residues related to Arg at 9.
StickWRLD Edges (links)

StickWRLD edges link residues and positions that are over or under correlated based on the PSSM.

Dashed edges indicate negative correlations.
StickWRLD Edges (links)

StickWRLD edges link residues and positions that are over or under correlated based on the PSSM.

Dashed edges indicate negative correlations.

Here, the presence of Lys 10 shows a negative correlation to Arg 9, and Pro 11. Pro 10 shows a negative correlation to Tyr 9.
StickWRLD Edges (links)

StickWRLD edges link residues and positions that are over or under correlated based on the PSSM.

Negatively linked collections of subfamilies suggest possible alternative requirements on sequence characteristics.

One subfamily probably tends to contain Arg 9, Pro 10, Pro 11, Gln 12, and Glu 13.
The other tends to contain Tyr 9, Lys 10, and Met 12.
Additional relationships become visible at other Tr/Pr/Nr threshold values.
StickWRLD Edges (links)

When grouped, the PSSM is recalculated based on the selected property type.

Because multiple residues collapse into property archetypes in a grouped view, interpositional correlations that were not visible at the level of the individual residue often appear.
When grouped, the PSSM is recalculated based on the selected property type.

Here, the charged (Basic) Arg at 9, can be seen to be positively correlated with the same collection of residues previously noted, as well as some others, such as the “slightly polar” residue at 2.

<table>
<thead>
<tr>
<th>Property Type</th>
<th>Colours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic(+)</td>
<td>Red</td>
</tr>
<tr>
<td>Acidic(−)</td>
<td>Blue</td>
</tr>
<tr>
<td>Polar</td>
<td>Cyan</td>
</tr>
<tr>
<td>Slightly-polar</td>
<td>Green</td>
</tr>
<tr>
<td>Non-polar</td>
<td>Yellow</td>
</tr>
</tbody>
</table>
When grouped, the PSSM is recalculated based on the selected property type.

**Arg** 9 is positively correlated with the negative (Acidic) **Glu** at 13, despite the family consensus being considerably higher for a non-polar residue at that position.
When grouped, the PSSM is recalculated based on the selected property type.

Likewise, the positive correlation between positions 13 and 12 now appears, demonstrating that the subfamily preference is for a polar residue at 12, rather than specifically the Gln observed previously.
When grouped, the PSSM is recalculated based on the selected property type.

The negative relationship between the subfamily containing a positive (predominantly Arg) at 9, and that containing a Polar (predominantly Trp) at 9 also becomes more obvious.
StickWRLD Edges (links)

When grouped, the PSSM is recalculated based on the selected property type. The positive relationships between the subfamily containing (predominantly) Trp at 9, and Lys at 10 also extend and become more apparent. The Polar at 9/Positive at 10 subfamily may even be two subfamilies, each distinct from the Arg at 9 group, as none of these identities share a positive edge.