Training *s0*

```
Model (+):
ACGTAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA
```

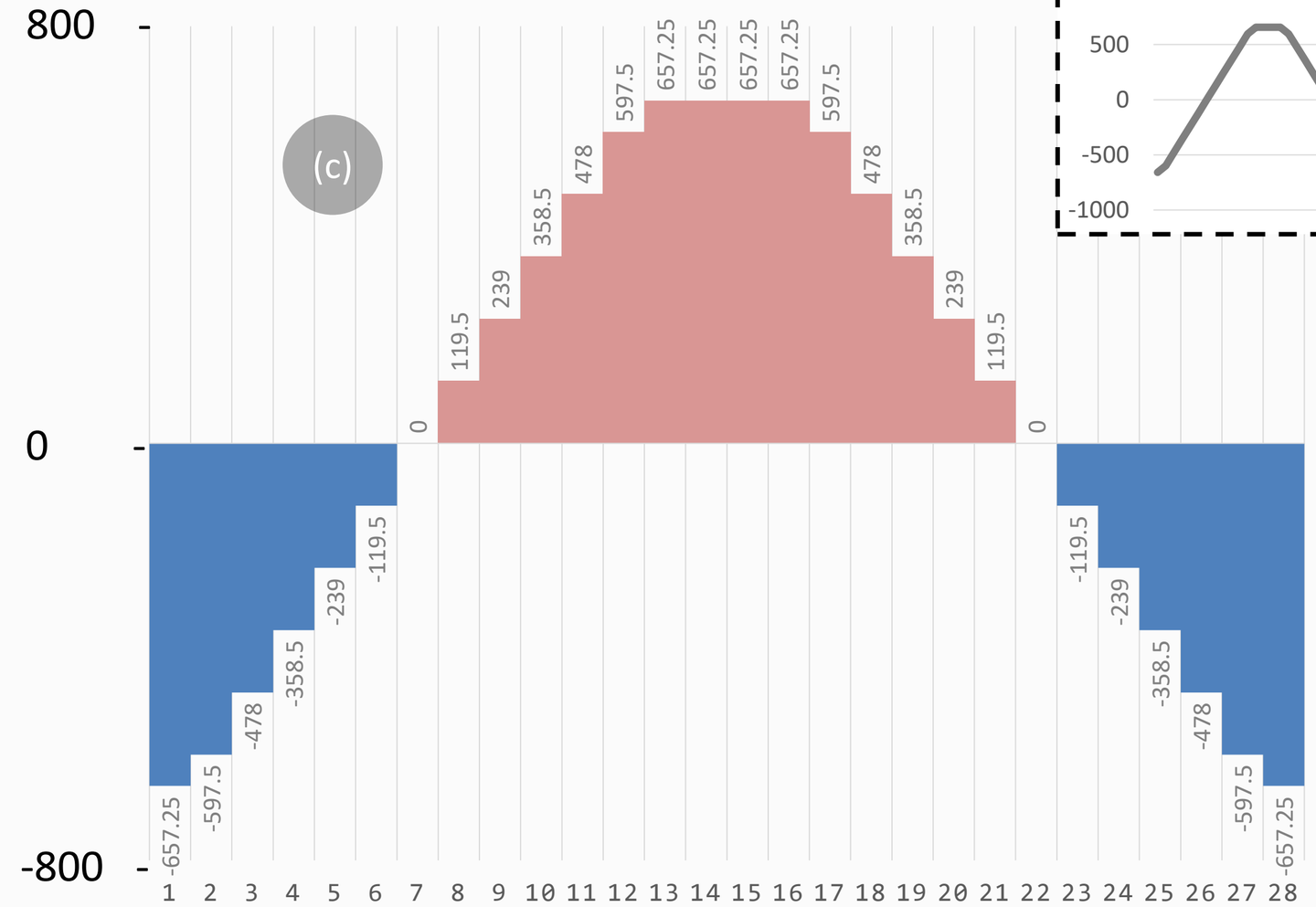
Training *s1*

```
Model (-):
ACGTGGGGGGGGGGGGGGGGG
GGGGGGGGGGGGGGGGGGGGG
```

z-sequence

Signal

Sliding window (sw)



z = GGGGGGGGGGGGAAAAAAAAAAAAAAAAGGGGGGGGGGGGG

(b)

```
GGGGGGGGGGGG = -657.25
GGGGGGGGGGGA = -597.50
GGGGGGGGGGAA = -478.00
GGGGGGGGGAAA = -358.50
GGGGGGGGAAAA = -239.00
GGGGGGGAAAAA = -119.50
GGGGGGAAAAAA = 0.00
GGGGGAAAAAAA = 119.50
GGGGAAAAAAAA = 239.00
GGGAAAAAAAAA = 358.50
GGAAAAAAAAAA = 478.00
GAAAAAAAAAAA = 597.50
AAAAAAAAAAAA = 657.25
AAAAAAAAAAAA = 657.25
AAAAAAAAAAAA = 657.25
AAAAAAAAAAAA = 657.25
AAAAAAAAAAAG = 597.50
AAAAAAAAAAGG = 478.00
AAAAAAAAAGGG = 358.50
AAAAAAAAGGGG = 239.00
AAAAAAAGGGGG = 119.50
AAAAAAGGGGGG = 0.00
AAAAAGGGGGGG = -119.50
AAAAGGGGGGGG = -239.00
AAAGGGGGGGGG = -358.50
AAGGGGGGGGGG = -478.00
AGGGGGGGGGGG = -597.50
GGGGGGGGGGGG = -657.25
GGGGGGGGGGGG = -657.25
```

(a)

An experiment for understanding scores

An experiment for understanding scores. (**a**) shows the contents of the sliding windows above the z-sequence; and (**b**) the score values calculated from the LLM for these sliding windows. (**c**) shows the plot of the score values on a bar chart whose axis contains both positive and negative values. The first half (top) of the chart shows positive values that indicate a resemblance of the region from the z-sequence to the s0 sequence of the "+" model. The second half (bottom) of the chart shows negative values that indicate the regions in the z-sequence that resemble the s1 sequence of the "-" model. Notice the columns of the chart correlate with the positions shown by the a and b region of the panel. For a reference system, the same chart is represented by a line chart in a small window in the upper right corner of the main chart. The left side of the figure indicates the meaning of the elements on the right side of the figure. The z-sequence is placed on a relative position below the chart and there is no direct correlation between the two, because the number of sliding windows is equal to the number of letters in the z-sequence minus the total number of positions in a sliding window.

Cite this work as:

Paul A. Gagniuc. *Algorithms in Bioinformatics: Theory and Implementation*. John Wiley & Sons, Hoboken, NJ, USA, 2021, ISBN: 9781119697961.