A Context Free Grammar for RNA Secondary Structures

Danielle Dees, Laura Slaybaugh

July 23, 2001

Note, this grammar assumes that the energy for each stacked pair is -1, and that the energy of each unbonded nucleotide in a bulge or internal loop is 1. The base case is a harpin loop.

$$
\begin{aligned}
W &\rightarrow tW | V_0 \\
V_0 &\rightarrow tV_1\bar{t} | tI_0\bar{t} | tM_0\bar{t} \\
V_1 &\rightarrow tV_2\bar{t} | H_1 | tI_1\bar{t} | tM_1\bar{t} \\
V_2 &\rightarrow tV_3\bar{t} | H_2 | tI_2\bar{t} | tM_2\bar{t} \\
V_3 &\rightarrow H_3 | tI_3\bar{t} | tM_3\bar{t} \\
H_3 &\rightarrow tt_1t_2t_3\bar{t} \\
H_2 &\rightarrow tt_1t_2\bar{t} \\
H_1 &\rightarrow tt_1\bar{t} \\
I_1 &\rightarrow tV_0 | V_0t \\
I_2 &\rightarrow tV_1 | V_1t | tI_1 | I_1t \\
I_3 &\rightarrow tV_2 | V_2t | tI_2 | I_2t \\
M_0 &\rightarrow WM_0WM_0 \\
M_1 &\rightarrow WM_0WM_1 | WM_1WM_0 \\
M_2 &\rightarrow WM_1WM_1 | WM_2WM_0 | WM_0WM_2 \\
M_3 &\rightarrow WM_0WM_3 | WM_1WM_2 | WM_2WM_1 | WM_3WM_0 \\
WM_0 &\rightarrow V_0WM_0 | V_0 \\
WM_1 &\rightarrow tWM_0 | V_0WM_1 | V_1WM_0 | V_1 | t \\
WM_2 &\rightarrow tWM_1 | V_0WM_2 | V_1WM_1 | V_2WM_0 | V_2 \\
WM_3 &\rightarrow tWM_2 | V_0WM_3 | V_1WM_2 | V_2WM_1 | V_3WM_0 | V_3
\end{aligned}
$$