

Update

Pete Rigas, Lambert Lab

February 17, 2021

introduction

- interdisciplinary work involving statistical mechanics, synthetic biology, protein-DNA interactions,
- developing models of binding dynamics (already established for Cas12a), which are inherently reliant on searching for a portion of the genome upstream of a target gene, and diffusing in one dimension along the target sequence

background

We introduce different cases of solutions of the IVP for the dimensionless second order differential equation

$$-\mathcal{A} \frac{d^2 \tau}{dx^2} + \mathcal{U}'(x) \frac{d\tau}{dx} = 1 ,$$

which can be readily obtained from a standard application of separation of variables of the diffusion PDE,

$$\frac{\partial \tau}{\partial t} = c \frac{\partial^2 \tau}{\partial x^2} ,$$

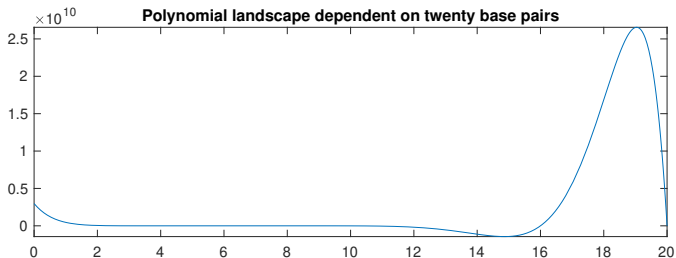
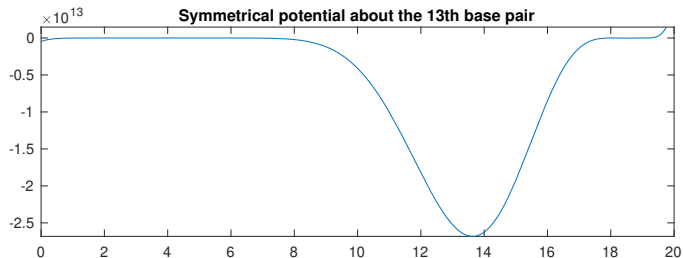
for arbitrary c .

model predictions

From exit time curves, theoretical predictions of the model can be easily confirmed in published data on Fn Cas12a, with the exit times and corresponding landscape adhering to the following qualities,

- mismatches within the first 6 base pairs of the landscape pose little detriment to the exit time that is numerically obtained from solving the second order IVP,
- mismatches at base pairs that are either consecutive, or within a one to two base pair window, after the first 6 base pairs significantly impact the development of a bound stable nuclease,
- mismatches in the last several bases of the landscape pose minimal effect on the binding process.

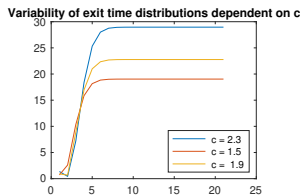
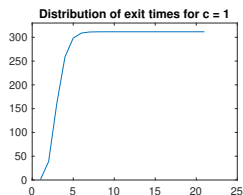
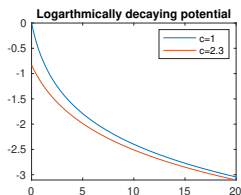
polynomial potentials $\mathcal{U}'(x)$



solutions to the IVP for different potential classes

Exit time class	Recovery formulation \mathcal{R} of u_1 terms of potential
τ_0	$\Delta_\tau + \frac{x_2(x_2 - c_1)}{c_1}$
τ_{Constant}	$\Delta_\tau - c_1 - c_1 \exp(-c_1 x) - \frac{x}{\sqrt{c_1}}$
$\tau_{\text{Asym}} \setminus \text{Sym}, \tau_{\text{Poly}}$	$\Delta_\tau - \int_0^x \int_0^x -\exp\left(\frac{\sum_{\text{degrees}} p(u)^{i+1}}{\prod_i d^i}\right) \exp\left(\frac{\sum_{\text{degrees}} p(u)^{i+1}}{\prod_i d^i}\right) du \, du -$ $\int_0^x \int_0^v \exp\left(-\frac{\sum_{\text{degrees}} p(v)^{i+1}}{\prod_i d^i}\right) \exp\left(\frac{\sum_{\text{degrees}} p(u)^{i+1}}{\prod_i d^i}\right) du \, dv - \int_0^x \exp\left(-\frac{\sum_{\text{degrees}} p(u)^{i+1}}{\prod_i d^i}\right) du$
τ_{Inc}	$\Delta_\tau - \frac{1}{\text{Deg}+1} \left((\text{Deg})^{\frac{1}{\text{Deg}}} \Gamma_{\text{Inc}} \left(\frac{1}{\text{Deg}}, \frac{y^{\text{Deg}+1}}{\text{Deg}+1} \right) - \right.$ $\left. \int_0^x \text{Deg}^{\frac{1}{\text{Deg}}} \exp\left(\frac{y^{\text{Deg}+1}}{\text{Deg}+1}\right) \Gamma_{\text{Inc}} \left(\frac{1}{\text{Deg}}, \frac{y^{\text{Deg}+1}}{\text{Deg}+1} \right) \right) + \dots$
τ_{\log}	$\Delta_\tau + \exp\left(c_1 - \log(c_1)\right) \int_0^{x_1} (u - c_1)^{-(u-c_1)} e^{-(c_1-u)} du$

variability in exit times for the exponential class of potentials



However, such a class of potentials, despite solutions existing for the IVP, is not realistic of the barriers intrinsic to the CRISPR landscape

dirac delta landscape potential

Such a class of potentials

- is resemblant of the energy barriers of binding,
- composed of step functions with a fixed "jump" in the energy landscape between neighboring base pairs,
- produces solutions \mathcal{S} that are dependent on the magnitude of the jump discontinuity of the landscape between neighboring base pairs, taking the form

$$\mathcal{S}(x) = \frac{x}{|\mathcal{U}'(x_1)|} - \frac{\exp(|\mathcal{U}'(x_1)| x)}{|\mathcal{U}'(x_1)|^2} + \frac{1}{|\mathcal{U}'(x_1)|^2} ,$$

where the landscape $\mathcal{U}'(x)$ is evaluated at the barrier corresponding to base pair x_1 in the sequence,

- is problematic to code in Matlab.

strategy

We obtain approximations of the exit time for the IVP raised in the previous slide by solving IVPs of shorter duration for the passage times across neighboring base pairs. Specifically,

- the expected exit time for passage to the **first base pair** of the sequence is obtained under the initial conditions that the exit time when the protein begins the search along the DNA strand is 0, and is subject to a constant initial unit velocity that it maintains throughout the search phase,
- passage times to all remaining base pairs in the sequence are obtained through specification of the jump discontinuity $\mathcal{U}'(x)$ at the next base pair of the sequence, from which we exhaust the procedure for all base pairs to obtain all remaining approximations of the exit time.

code

```
function [sol_array] = step_potential(U_initial, U_string)
    sol_array = [ ];
    % solve IVP for first barrier
    syms y(x)
    Dy = diff(y);
    ode = diff(y,x,2) == -1+(U_initial)*diff(y,x,1);
    cond1 = y(0) == 0;
    cond2 = Dy(0) == 1;
    conds = [cond1 cond2];
    ySol(x) = dsolve(ode,conds);

    for I = 1 : length(U_string)

        temp = matlabFunction(ySol(x));

        % formulate new IVP from exit time of previous solution
        ode_temp = diff(y,x,2) == -1+(U_string(I))*diff(y,x,1);
        cond_1 = temp(1);
        cond_2 = Dy(0)==1;
        conds_temp = [cond_1 cond_2];
        ySol(x) = dsolve(ode_temp,conds_temp);

    end

end
```