going back to p 44, simpler recursion relations with

matches mccaskill (eq?) $\rightarrow$ B. $\qquad K_d^{BP*} = K_d^{BP}\, l_{BP}/l$

$$Z^{BP}(i,j) = \alpha_{i}\alpha_{j-1}\, \frac{C_{eff}(i+1,j-1)\, l_{BP}}{K_d^{BP*}} + \sum \frac{C_{std}}{K_d^{BP*}}\, Z_{linear}(i+1,\ldots)\, Z_{lin}(\ldots,j-1)$$

2 (?)

$$C_{eff}(i,j) = \alpha_{j-1}\, C_{eff}(i,j-1)\, l + \sum_{i\le k<j} C_{eff}\, Z^{BP}(k,j)\, l_{BP} + Z^{BP}(i,j)\, l_{BP}\, l ?$$

$$Z_{linear}(i,j) = \alpha_{j-1}\, Z_{linear}(i,j-1) + \sum_{i\le k<j} Z_{linear}(i,k-1)\, Z^{BP}(k,j) + Z^{BP}(i,j)$$

Can we identify $\left.\begin{array}{l} l_{BP} \Rightarrow e^{-b} \\ l \Rightarrow e^{-c} \\ C_{INT} \Rightarrow e^{-a} \end{array}\right\}$ in multiloop model

$\Delta G = a + b(\#\text{nts}) + c(\#\text{helices})$?

 $\rightarrow Z = C_{INT}\, l^3\, l_{BP}/K_d^{BP*}$

 $\rightarrow Z = C_{std}\, \frac{l_{BP}^{*2}\, C^{INT}}{K_d^{BP*2}}$

need to look up Mccaskill.

Define(?) $l_{BP}' = l_{BP}/l$) $\qquad l_{BP} = l_{BP}'\, l \rightarrow$

try on next page.

(50)

$K_d^{BP} = \frac{l}{l_{BP}} K_d^{BP*} = \frac{K_d^{BP*}}{l_{BP}^*}$

Further revisiting of "simplified" recursion: $l_{BP} = l_{BP}^* l$

$$Z^{BP}(i,j) = \alpha_i \alpha_{j-1} \frac{C_{eff}(i+1,j-1)\, l^2 l_{BP}^*}{K_d^{BP*}} + \sum_{c=\text{cutpoint}} \frac{C_{std}}{K_d^{BP*}} Z_{linear}(i+1,c)\, Z_{linear}(c+1,j-1)$$

$$C_{eff}(i,j) = \alpha_{j-1} C_{eff}(i,j-1)\, l + C_{init} Z^{BP}(i,j)\, l_{BP}^* + \sum_{i<k<j} C_{eff}(i,k-1)\, l\, Z^{BP}(k,j)\, l_{BP}^*$$

$$Z_{linear}(i,j) = \alpha_{j-1} Z_{linear}(i,j-1) + Z^{BP}(i,j) + \sum_{i<k<j} Z_{linear}(i,k-1)\, Z^{BP}(k,j)$$

Initialize:

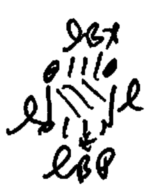
$Z_{linear}(i,i) = 1$

$C_{eff}(i,i) = C_{init}$

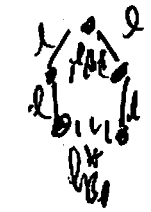
Finalize

$$Z_{final}(i,i) = \begin{cases} C_{eff}(i,i-1)\, l / C^{std} + \sum_{i \le c \le i+N-1} Z_{linear}(i,c)\, Z_{linear}(c+1,i-1) & (\text{if } i-1 \ldots i \text{ connected.}) \\ Z_{linear}(i,i-1) & (\text{if } i-1, \ldots i \text{ cutpoint}) \end{cases}$$

$$Z = \frac{C_{init}\, l^4}{K_d^{BP}} = C_{INIT} \frac{l^4 l_{BP}^*}{K_d^{BP*}}$$



$$Z = \left(\frac{C_{std}\, l}{K_d^{BP}}\right) \times \frac{l\, C_{INIT}}{K_d^{BP}} = C_{INIT} \frac{l^2 l_{BP}^{*2} C_{INIT}}{K_d^{BP*2}}$$



$$Z = C_{INIT}^2 \left(\frac{l_{BP}}{l}\right) \frac{l^4}{K_d^{BP2}} = C_{INIT}^2 \frac{l^4 l_{BP}^{*3}}{K_d^{BP2}} = \left(\frac{C_{INIT}\, l^2 l_{BP}^*}{K_d^{BP*}}\right)\left(\frac{C_{INIT}\, l^2 l_{BP}^*}{K_d^{BP*}}\right)$$

Nicely matches tz-fold framework → cool!

(51)

How to add coaxial stack [Rough]

$$Z^{coax}(i,j) = \sum_{i<k<j} Z_{BP}(i,k-1) Z_{BP}(k,j)\, K_{coax}$$

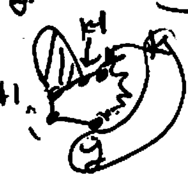
new STATE

$$C_{eff}(i,j) \mathrel{+}= Z^{coax}(i,j)\, l_{coax} + \sum_{i<k<j} C_{eff}(i,k-1)\, l\, Z^{coax}(k,j)\, l_{coax}.$$

$$Z_{linear}(i,j) \mathrel{+}= Z^{coax}(i,j) + \sum_{i<k<j} Z_{linear}(i,k-1)\, Z_{coax}(k,j)$$

?

$$Z^{BP}(i,j) \mathrel{+}= \alpha_i \sum Z^{BP}(i+1,k-1)\, l\, C_{eff}(k,j-1)\, l\, l_{BP}^{*}\, K_{coax}$$

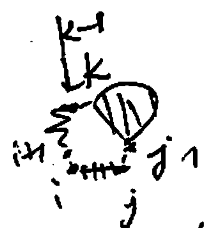


K_d^{BP*}

$\boxed{k<j-1}$ [or cutpoint]

$$+ \alpha_{j-1} \sum l\, C_{eff}(i+1,k-1)\, l\, Z_{BP}(k,j-1)\, l_{BP}$$

K_d^{BP*} K_{coax}



$\boxed{k>i+1}$ [or cutpoint]

Write clearly on next page.

(52)

If no constraint on coax ⟺ BP, i.e. if bulgeless 3WJ allowed,

$$Z^{coax}(i,j) = \sum_{i<k,\,k-1} \alpha\, Z_{BP}(i,k-1)\, Z_{BP}(k,j)$$

(?) k^{coax} $2>1$ $\boxed{\text{Bonus}}$

$$+ \sum_{i<c\leq k-1} \frac{C_{std}}{K_d^{coax}} Z_{BP}(i,k-1)\, Z_{BP}(k,j)$$

Supplement expression on (p.50) →

$$Z^{BP}(i,j) \mathrel{+}= \frac{Z^{BP}(i+1,j-1)\, \ell\, \ell_{BP}}{K_d^{BP\,coax}} \quad [C_{init}?] + \sum_{i+1<k<j} \frac{C_{eff}(i,k-1)\, \ell\, [Z_{BP}(k,j-1)\, \ell\, \ell_{BP}]}{K_d^{BP}\, K^{coax}}$$

$$+ \sum_{i+1<k<j} \frac{[\ell\, Z_{BP}(i+1,k)\, \ell\, C_{eff}(k+1,j-1)\, \ell]}{K_d^{BP}\, K^{coax}}$$

$$+ \sum_{c=\text{cutpoint}} \frac{C_{std}}{K_d^{BP}} \left\{ Z_{BP}(i+1,c)\, Z_{linear}(c+1,j-1) + Z_{linear}(i+1,c)\, Z_{BP}(c+1,j-1) \right\} K_{coax}$$

$$C_{eff}(i,j) \mathrel{+}= C_{init} Z^{coax}(i,j)\, \ell_{coax} + \sum_{i<k<j} C_{eff}(i,k-1)\, \ell\, Z^{coax}(k,j)\, \ell_{coax}.$$

Need to check through some examples.

$$Z_{linear}(i,j) \mathrel{+}= Z^{coax}(i,j) + \sum_{i<k<j} Z_{linear}(i,k-1) \times Z^{coax}(k,j)$$

53

$C\text{-}\delta\ \ \delta^*\text{-}C^*$ (disallow $C\text{-}\delta^*$, $\delta\text{-}C^*$ for now)

C

$Z_{BP} = 0$

$Z_{LOAX} = 0$

$C_{eff} = C_{INT}$

$Z_{linear} = 1$

CA

$Z_{BP} = 0$

$Z_{LOAX} = 0$

$C_{eff} = C_{INT}\,l$

$Z_{linear} = 1$

$C\text{-}\delta$

$Z_{BP} = \dfrac{C_{INT}\, l\, l_{BP}^{2*}}{k_d^{BP*}}$

$Z_{LOAX} = 0$

$C_{eff} = C_{INT}\, l^2 + \dfrac{C_{INT}^2\, l^2\, l_{BP}^{2*2}}{k_d^{BP*}}$

$Z_{linear} = 1 + \dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP*}}$

$C\text{-}\delta\text{-}\delta^*$

$Z_{BP} = 0$

$Z_{LOAX} = 0$

$C_{eff} = C_{INT}\, l^3 + \dfrac{C_{INT}^2\, l^3\, l_{BP}^{*2}}{k_d^{BP}}$

$Z_{linear} = 1 + \dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP}}$

$C\text{-}\delta\text{-}\delta^*$

$Z_{BP} = 0$

$Z_{LOAX} = 0$

$C_{eff} = C_{INT}\, l^4 + \dfrac{C_{INT}^2\, l^4\, l_{BP}^{*2}}{k_d^{BP*}}$

$Z_{linear} = 1 + \dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP*}}$

$C\text{-}\delta\ \ \delta^*\text{-}C^*$

$Z_{BP} = 0$

$Z_{LOAX} = \left[\dfrac{C_{INT}\, l\, l_{BP}^{2*}}{k_d^{BP*}}\right]^2 \underbrace{l_{BP}^{*2}\, l\, K^{LOAX}}_{?}$

$C_{eff} = C_{INT}\, l^5 + \dfrac{C_{INT}^2\, l^5\, l_{BP}^{*2}}{k_d^{BP*}} +$

$\left[C_{INT}\, l^2 + \dfrac{C_{INT}^2\, l^2\, l_{BP}^{*2}}{k_d^{BP}}\right] l\, \dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP*}}$

$+ C_{INT}\, l_{LOAX}\, l \left[\dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP}}\right]^2 \underbrace{l_{BP}^{*2}\, l\, K^{LOAX}}_{?}$

~~$Z_{linear} = 1 +$~~

$Z_{linear} = 1 + \dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP*}} +$

$\left(1 + \dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP*}}\right) \dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP*}}$

$+ \left[\dfrac{C_{INT}\, l^2\, l_{BP}^{*}}{k_d^{BP*}}\right]^2 \underbrace{l_{BP}^{*2}\, l\, K^{LOAX}}_{?}$

Need to calculate

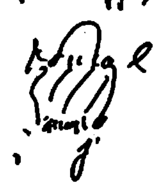
Z_{final} $\boxed{6 \text{ ways}}$

Return to MOTIFS...

Same recursion as on p-53, but add motif term to Z^{BP}:

$$Z^{BP}(i,j) = \alpha_i \alpha_{j-1} \frac{C_{eff}(i+1, j-1)\, l_{BP}^{2*}}{K_d^{BP*}} + \sum_{c = \text{cutpoint}} \frac{C_{std}}{K_d^{BP*}} Z_{linear}(i+1, c) \times Z_{linear}(c+1, j)$$

$$+ \alpha_i \alpha_{j-1} \sum_{k,l} C_{eff}^{motif(k,l;i,j)} / K_d^{BP*}\, Z^{BP}(k,l)$$



pretty simple.

oh WAIT:

$$Z^{final}(i,i) = \alpha_{i-1} \Big[\frac{C_{eff}(i, i-1)\, l}{C^{std}} + \underbrace{\sum_{i \le c < (i-1)+N} Z_{linear}(i,c)\, Z_{linear}(c+1, i-1)\, \alpha_c}_{Z_{cut}(i-1,\, i-1)} \Big]$$

[see p. 52 for definition]

$$+ (1-\alpha_{i-1}) \left[Z_{linear}(i, i-1) \right]$$

$$+ \sum_{i < j < (i-1)+N} \alpha_j Z_{BP}(i,j)\, Z_{BP}(j+1, i-1)\, C_{eff}^{motif\ j,i;\ j+1,i-1}$$

55

15 OCT 2018

Let's ask a question... $\boxed{\text{Questions in inference}}$

- Fix $\ell = \ell_{BP} = 1$

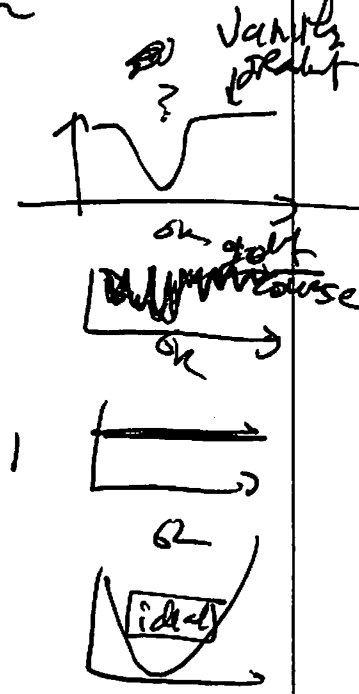
$C_{INT} K_d^{BP} = \exp[-4.1 \text{ kcal/mol} / kT]$?

~~Sim~~
- Simulate a bunch of K_d measurements to a motif like

or "MS2" hairpin.

- What's the "spectrum" of K_d?

[TEST 1] - If we have a bunch of measurements of $\log(K_d)$ with 1 kcal/mol accuracy, could we infer ~~log~~ $\log(C_{INT} K_d^{BP})$?



[TEST 2] - then add in C_{eff}^{MOTIF} for base pair step (assume constant, Fix (GC/CG) (CC/GG) (GC/CG))

Same question is $\boxed{\text{2D SCAN}}$

[TEST 3] suppose we could measure ΔG for a bunch of designed 2 state constructs — infer $\log(C_{INT} K_d^{BP})$?

[TEST 4] suppose we could design a bunch of sequences with low ensemble defect, what if we optimize

$$\sum_{\text{sequences}} \Delta G(\text{MFE}_{\text{target structure}}) - \Delta G(\text{all structures})$$

[TEST 5] what if we simulate "chemical mapping" profiles — could we make inference?

- Can we do <u>de novo</u> inference?

(1) New kinds of base pairs that build off W/C?

$\left.\begin{array}{l} A\text{-}U \\ G\text{-}C \\ G\text{-}U \end{array}\right\} \Rightarrow$ add: G-A wk

$4 \times \underset{e_{G\text{-}A}}{3} = 12$ MOTIF Coff paramtrs

$+ \ K_d^{G\text{-}A} = 1$ Coff PARAMETERS

$\overline{13 P}$

$\boxed{\text{Could scan all } 4 \times 4 = 16 \text{ VARIANTS}}$

(2) Look for UNR turns → $\left\{\begin{array}{l} K^{UUCA} \\ \ell_{UCA} \end{array}\right.$ TREAT similar to "BP"s.

Could <u>FREEZE</u>

In fact, could scan all $\boxed{64 \text{ VARIANTS}}$

Need <u>PRIORS</u> to constrain searches,

either naive ones, eg, $\ell_{BP} \sim \ell_{G\text{-}C}$

or sophisticated ones for, eg,

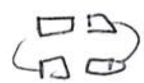
$Coff\begin{pmatrix} G\text{-}C \\ | \ \ | \\ G\text{-}A \end{pmatrix}$,

FROM Rosetta.

pre - 57A

(1)

THE FARMINGTON INN
827 Farmington Avenue, Farmington, CT 06032
860-677-2821 www.FarmingtonInn.com

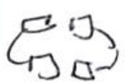


$\frac{C_{std}}{K_d^{BP}} \times \frac{l_{int}^2 l_{BP}^2}{K_d^{BP}}$

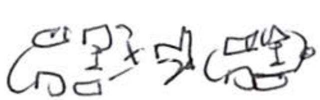
$\left(\frac{C_{std}}{K_d^{BP}}\right)^2 \times \frac{l_{int}^2 l_{BP}^2 C_{int}}{K_d^{l_{eq}}}$

$\times \left(K_d^{l_{eq}}/C_{std}\right)$ ← conventional x-lit factor for reality unk.

$= \frac{C_{std}\, C_{int}\, l_{int}^2 l_{bp}^2}{K_d^{BP\,2}}$



$\frac{C_{std}}{K_d^{BP}} \cdot \frac{C_{eff}^{BPS}}{K_d^{BP}}$



$\left\{\left(\frac{C_{std}}{K_d^{BP}}\right)^2 K_{COAX}\right\} \times \frac{C_{INIT}\, l\, l_{COAX}}{K_d^{l_{eq}}}$

$\times \left(K_d^{l_{eq}}/C_{std}\right)$

(2)

THE FARMINGTON INN
827 Farmington Avenue, Farmington, CT 06032
860-677-2821 www.FarmingtonInn.com

K_{coax}

$C_{eff}^{BPS\,2}$

$$\boxed{C_{EFF}^{BPS} \stackrel{?}{=} \left[C_{INIT}\, l_{COAX}\right] l\, K_{COAX}}$$

Boost

yes??

Now imagine bulge:

$Z = \left(\frac{C_{std}}{K_d^{BP}}\right)^2 K_{COAX} \frac{C_{INIT}\, l_{COAX}\, l^4}{\cancel{K_d^{l_{eq}}}} \times \frac{\cancel{K_d^{l_{eq}}}}{\cancel{C_{std}}}$

Alternatively: $= C_{std}\, C_{INIT}\, l_{COAX}\, l^4 K_{COAX} / K_{d\,BP}^2$

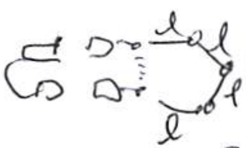
$Z = \left[C_{INIT}\, \frac{C_{std}\, l^5\, \cancel{l_{BP}}}{K_{d\,BP}^2}\right] \times \left[\frac{C_{EFF}^{BPS}}{C_{INIT}\, \cancel{l_{BP}}\, l}\right]^2$

$= \frac{C_{std}\, C_{EFF}^{BPS}\, l^3}{K_d^{BP\,2}}$

(3)

THE FARMINGTON INN
827 Farmington Avenue, Farmington, CT 06032
860-677-2821 www.FarmingtonInn.com

$$\left[\left(\frac{C_{std}}{K_d^{BP}}\right)^2 K_{conx}\right]$$

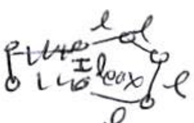


$$\left[\left(\frac{C_{std}}{K_d^{BP}}\right)^2 K_{conx}\right]\left[C_{init}\,\ell^4 \; \ell_{conx}\right]$$

$$\times K_{d,l}/C_{std}$$

$$\stackrel{?}{=} \frac{C_{std}\,K_{conx}\,C_{init}\,\ell^4\,\ell_{conx}}{K_d^{BP\,2}}$$

$\boxed{\text{yes!}}$



$$Z = \left(\frac{C_{std}}{K_d^{BP}}\right)^2 K_{conx} \quad \& \quad \boxed{\ell_{conx} \sim \frac{\overset{\text{BASE PAIR STEP in THE}}{C_{eff}}}{C_{init}\,\ell\,K_{conx}}}$$

unitless equilibrium constant!

(4)

THE FARMINGTON INN
827 Farmington Avenue, Farmington, CT 06032
860-677-2821 www.FarmingtonInn.com

10^5 M

ok: choose ℓ_{conx} first, the BASE PAIR STEP MOTIF

$$K_{conx} \sim \left[\frac{C_{eff}}{C_{init}\,\ell}\right] \sim 10^5 ?$$

$\ell \sim 1$ m?

& no connection $\left(\frac{C_{std}}{K_d}\right)^2 \frac{C_{std}}{K_d^{conx}}$

$$\left(\frac{C_{std}}{K_d}\right)^2 \frac{\cancel{C_{std}}}{K_d^{conx}} \; \frac{C_{init}\,\ell_{conx}}{K_d^{dig}} \; \frac{K_{dlg}}{C_{std}}$$

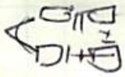
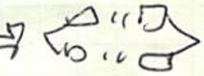
$$= \left(\frac{C_{std}}{K_d}\right)^2 K_{conx}$$

So... $\boxed{K_d^{conx} = \frac{C_{init}\,\ell_{conx}}{K_{conx}}}$

PTC-57B

20 Oct 2018

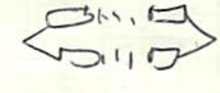
Back-of-the-envelope calculations for coaxial stacking.

 ⇒ 

$$\left(\frac{C_{std}}{K_{d}^{BP}}\right)^2 K_{coax} \qquad \left(\frac{C_{std}}{K_d^{coax}}\right)^2 C_{init}\,\ell_{coax}\,\ell \qquad K_d^{lig} \qquad \left(\frac{K_d^{lig}}{C_{std}}\right)$$

$$= \frac{C_{std}\,C_{init}\,\ell_{coax}\,\ell\,K_{coax}}{K_d^{BP\,2}}$$

Alternatively,

 ⇒ $C_{std}\dfrac{C_{eff}^{BPS}}{K_d^{BP\,2}}$ ← should be the same

So... identify: $C_{eff}^{BPS} = c_{init}\,\ell_{coax}\,\ell\,K_{coax}$

So, in $\boxed{K_{coax} = \dfrac{C_{eff}^{BPS}}{C_{init}\,\ell_{coax}\,\ell}}$ ← 10^5 M $= 200$ ← was that miss?

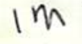
But very similar to 10^2–10^3 in Walter et al, PNAS 1994

$\frac{1}{2}$?

10^3 ?

1 M ?

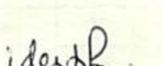
★ ℓ_{coax} may be quite high!

 $\approx C_{eff} \times \left(\frac{0.3}{1\,M}\right)^2 \cdot 1 \sim 10^3$ nM

0.05 nM

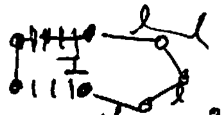
?

0.5 M

 ⇒ 

$$\left(\frac{C_{std}}{K_d}\right)^2 \left(\frac{C_{std}}{K_d^{coax}}\right) \times \frac{C_{init}^{*}\,\ell_{coax}\,\ell}{K_d^{lig}} \cdot \frac{K_d^{lig}}{C_{std}} \Leftrightarrow \left(\frac{C_{std}}{K_d^{BP}}\right)^2 K_{coax}$$

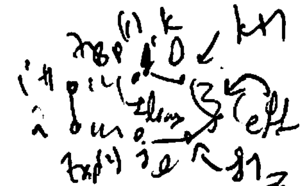
$$K_d^{coax} \sim \frac{C_{init}^2\,\ell^2\,\ell_{coax}\,\ell^{*}_{coax}}{C_{eff}^{BPS}} \longleftarrow K_d^{coax} = \frac{\overset{1M}{C_{init}}\ \overset{10^3}{\ell_{coax}}\ \overset{\frac{1}{2}}{\ell}}{\underset{200}{K_{coax}}} \sim \frac{1M\cdot\frac{1}{2}\cdot 10^3}{200} \sim \boxed{2.5\,M}$$



"seed" for C_{eff} — ligation scalefactor

$$Z = \left[\frac{C_{std}}{K_d^{BP}}\right]^2 K_{coax} \times \frac{C_{int}\, l_{coax}\, l^4}{K_d^{lig}} \cdot \left[\frac{K_d^{lig}}{C_{std}}\right]$$

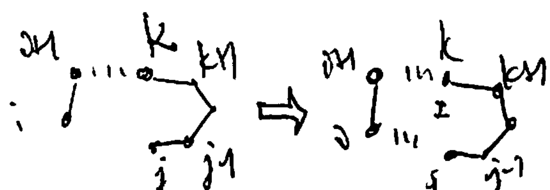
$$= C_{int}\, C_{std}\, K_{coax}\, l^4\, l_{coax} / K_d^{BP\,2}$$



$C_{eff} = C_{int}\, l^2$ — ligation scalefactor

$$Z = \underbrace{\left[\frac{C_{std}}{K_d^{BP}}\right]^2}_{Z_{BP}^{(1)}\, Z_{BP}^{(2)}} \cdot K_{coax} \cdot l \cdot \frac{[C_{int}\, l^2] \cdot l \cdot l_{coax}}{K_d^{lig}} \cdot \left[\frac{K_d^{lig}}{C_{std}}\right]$$

$$= Z_{BP\,(i,j)}\, Z_{BP\,(m,k)} \cdot l_{coax}\, l\, C_{eff}(k+1, j-1)\, l \,/\, C_{std}$$ ← [to match convention?]



$$Z = \left(\frac{C_{std}}{K_d^{BP}}\right) \qquad Z_{(i,j)} = \left(\frac{C_{std}}{K_d^{BP}}\right) \times \frac{C_{int}\, l_{coax}\, l^4}{K_d^{BP}} \times K_{coax} = \frac{Z_{BP}(i+1, k+1) \times l \times C_{eff}(k+1, j-1) \times l \times [l_{coax}\, K_{coax}]}{K_d^{BP}}$$

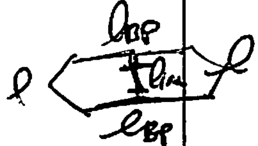
See later — more intuitive to define K_d^{coax}

Alternative scalings

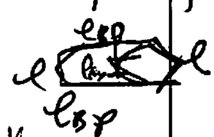
define K_d^{coax*} by $K_{coax} = \dfrac{C_{INIT}\, l\, l_{BP}^2\, l_{coax}}{K_d^{coax*}}$?

then,

$$K_{coax} = \frac{C_{INIT}\, l\, l_{BP}^2\, l_{coax}}{K_d^{coax*}} \stackrel{?}{=} \frac{C_{eff}^{BPS}}{C_{init}\, l_{coax}\, l}$$



$$\frac{1}{K_d^{coax*}} = \left[\frac{C_{eff}^{BPS}}{C_{init}\, l\, l_{BP}^2}\right]\left(\frac{1}{l_{coax}^2\, C_{INIT}}\right) = \frac{C_{eff}^{BPS}}{(C_{INIT}\, l\, l_{BP}^2\, l_{coax})(C_{init}\, l\, l_{coax})}$$



and

$$K_d^{coax} = \frac{C_{INIT}\, l_{coax}^*\, l}{K_{coax}} = \frac{C_{INIT}\, l_{coax}^*\, l}{C_{INIT}\, l\, l_{BP}^2\, l_{coax}}\, K_d^{coax*}$$

$$= \left[\frac{l_{coax}^*}{l_{coax}\, l_{BP}^2}\right] K_d^{coax*} = \left(\frac{l_{coax}^*}{l_{coax}}\right) \frac{K_d^{coax*}}{l_{BP}^2}$$

Alternatively define

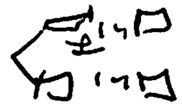
$$K_{coax} = \frac{C_{INIT}\, l\, l_{coax}}{K_d^{coax}}$$

$$\frac{C_{init}\, l\, l_{coax}}{K_d^{coax}} = \frac{C_{eff}^{BPS}}{C_{init}\, l_{coax}\, l}$$

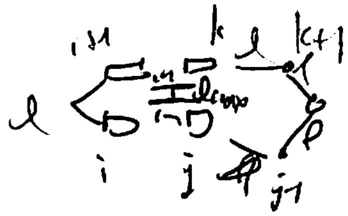
$$\boxed{\frac{1}{K_d^{coax}} = \frac{C_{eff}^{BPS}}{(C_{init}\, l\, l_{coax})^2}}$$

and

$$K_d^{coax} = \frac{C_{INIT}\, l_{coax}^*\, l}{C_{INIT}\, l\, l_{coax}}\, K_d^{coax\prime} = \left[\frac{l_{coax}^*}{l_{coax}}\right] K_d^{coax\prime}$$



$$Z = \left(\frac{C_{std}}{K_d^{BP}}\right)^2 K_{coax} = \left(\frac{C_{std}}{K_d^{BP}}\right)^2 \frac{C_{INIT}\, l\, l_{coax}}{K_d^{coax}}$$



$$Z = C_{init}\, C_{std}\, K_{coax}\, l^4\, l_{coax} / K_d^{BP^2}$$

$$= C_{INIT}\, C_{std} \frac{C_{INIT}\, l\, l_{coax}}{K_d^{coax}}\, l^4\, l_{coax} / K_d^{BP^2}$$

$$= \left(\frac{C_{std}}{K_d^{BP}}\right) \times \left(\frac{C_{INIT}\, l\, l_{coax}}{K_d^{BP}}\right)\left(\frac{C_{INIT}\, l\, l_{coax}}{K_d^{coax}}\right)$$

$Z(i+1, l-1)$ — FROM BASE PAIR (i,j) — cost of doing extra loop?

$\boxed{\text{Even clearer on next page (62)}}$

Let's try to write out recursions cleanly:

i.e. force cutpoint, $[1-\alpha_c]$

$$Z^{BP}(i,j) = \alpha_i \alpha_{j-1} \frac{C_{eff}(i+1,j-1)\, l_i^2 l_{BP}}{K_d^{BP}} + \sum_{i<c<j-1} \frac{C_{std}}{K_d^{BP}} Z_{linear}(i+1,c)\, Z_{linear}(c+1,j-1)$$

$$+ \alpha_i \alpha_{j-1} Z_{BP}(i+1,j-1) \left(\frac{C_{std}}{K_d^{BP}}\right) K_{coax}$$

cutpoint

$\frac{C_{std}}{K_d^{BP}} Z_{cut}(i+1,j-1)$

Could combine with term below if $Z_{cut}(j,j-1) \equiv 1$

SWITCH ORDER

$$+ \sum_{i+1<k<j-1} \alpha_i \alpha_k \alpha_{j-1} Z_{BP}(i+1,k) \frac{C_{eff}(k+1,j-1)\, l^2 l_{coax}}{K_d^{BP}}$$

$$+ \sum_{i+1<k<j-1} \alpha_i \alpha_{j-1} Z_{BP}(i+1,k)\, K_{coax} \frac{C_{std}}{K_d^{BP}} \times \sum_{k<c<j-1} (1-\alpha_c) Z_{linear}(k+1,c)\, Z_{linear}(c+1,j-1)$$

Replace with:

$$Z_{cut}(k+1,j-1) \equiv \sum_{k<c<j-1} (1-\alpha_c) Z_{linear}(k+1,c)\, Z_{linear}(c+1,j-1)$$

$$+ (1-\alpha_i)\alpha_{j-1} Z_{BP}(i+1,j-1) \frac{C_{std}}{K_d^{BP}} K_{coax}$$

$$+ \sum_{i+1<k<j-1} \alpha_i \alpha_{k-1} \alpha_{j-1} Z_{BP}(k,j-1) \frac{C_{std}}{K_d^{BP}} K_{coax} Z_{cut}(i+1,k-1)$$

$$+ \sum_{i+1<k<j-1} \alpha_i \alpha_{k-1} \alpha_{j-1} Z_{BP}(k,j-1) \frac{C_{eff}(i+1,k-1)\, l^2 l_{coax} K_{coax}}{K_d^{BP}}$$

$$+ \alpha_i \alpha_{j-1} \frac{C_{eff}^{BPS}(i,j;i+1,j-1)}{K_d^{BP}}$$

$$\boxed{C_{eff}^{BPS} \sim C_{init}\, l_{coax} K_{coax}}$$

← IN FACT USE THIS AS FIRST ESTIMATE OF K_{coax}

Let's REORDER and simplify Z^{BP} recursion.

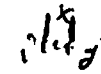
First Redefine:

$$Z_{cut}(i,j) = \sum_{i\le c<j} Z_{linear}(i+1,c)\, Z_{linear}(c+1,j-1)\, \underbrace{[1-\alpha_c]}_{c \to \text{cutpoint}}$$

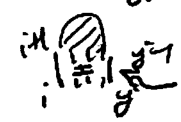
[Single cut] DEFINING $Z_{linear}(i+1,i) \equiv 1$. (Check: $Z_{cut}(i,i+1) = 4$]

So ...

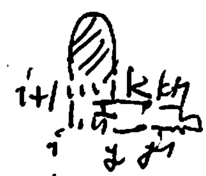
$$Z^{BP}(i,j) = \alpha_i \alpha_{j-1} C_{eff}(i+1,j-1)\, e^{2} l_{BP}/K_d^{BP} + \frac{C_{std}}{K_d^{BP}} Z_{cut}(i,j)$$



$$+ \alpha_i \alpha_{j-1} C_{eff}^{BPS}(i,j;i+1,j-1)/K_d^{BP}$$

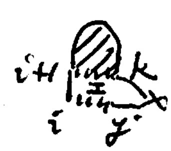


$$+ \sum_{i+1<k<j-1} \alpha_i \alpha_k \alpha_{j-1} Z_{BP}(i+1,k)\, C_{eff}(k+1,j-1)\, e\, l_{coax} K_{coax}^{(i,j;i+1,k)}/K_d^{BP}$$

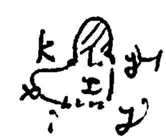


$$+ \sum_{i+1<k<j-1} \alpha_i \alpha_{k-1} \alpha_{j-1} C_{eff}(i+1,k-1)\, Z_{BP}(k,j-1)\, e\, l_{coax} K_{coax}^{(i,j;k,j-1)}/K_d^{BP}$$

$$+ \sum_{i+1<k<j} \alpha_i Z_{BP}(i+1,k)\, C_{std} K_{coax}^{(i,j;i+1,k)} Z_{cut}(k,j)/K_d^{BP}$$



$$+ \sum_{i<k<j-1} \alpha_{j-1} Z_{cut}(i,k)\, Z_{BP}(k,j-1)\, C_{std} K_{coax}^{(i,j;k,j-1)}/K_d^{BP}$$



$K_{coax}^{(BP1,BP2)} = [C_{eff}^{BPS}(BP1,BP2)/C_{INIT}\, l_{coax}]$ to first approximation.

OVER

Rest of the recursions...

$Z_{cut} \rightarrow$ see prev. page 62

$Z_{BP} \rightarrow$ see prev. page 62

$$Z^{coax}(i,j) = \sum_{i<k<j-1} \alpha_k Z_{BP}(i,k) Z_{BP}(k+1,j) K_{coax}(k,i;k+1,j)$$

$$C_{eff}(i,j) = \alpha_{j-1} C_{eff}(i,j-1) l + C_{init} l_{BP} Z^{BP}(i,j) + \sum_{i<k<j} C_{eff}(i,k-1) l Z^{BP}(k,j) l_{BP}$$
$$+ C_{init} l_{coax} Z^{coax}(i,j) + \sum_{i<k<j} C_{eff}(i,k-1) l Z^{coax}(k,j) l_{coax}$$

$$Z_{linear}(i,j) = \alpha_{j-1} Z_{linear}(i,j-1) + Z^{BP}(i,j) + \sum_{i<k<j} Z_{linear}(i,k-1) Z^{BP}(k,j)$$
$$+ Z^{coax}(i,j) + \sum_{i<k<j} Z_{linear}(i,k-1) Z^{coax}(k,j)$$

$$Z_{final}(i,i) = \begin{cases} Z_{linear}(i,i-1) & \text{if } i-1 - i \text{ is cutpoint} \\ \dfrac{C_{eff}(i,i-1) l}{C_{std}} + \cancel{\sum_{i \le c < (i-1)+N} Z_{linear}(i,c) Z_{linear}(c+1,j-1)} + \sum_{i<j<(i-1)+N} \alpha_j Z_{BP}(i,j) Z_{BP}(j+1,i-1) C_{eff}^{BPS}(j,i;j+1,i-1) \\ \quad + \sum_{i<j<k<(i-1)+N} Z_{BP}(i,j) Z_{BP}(k,i-1) K_{coax}(j,i;j+1,i-1) Z_{cut}(j,k) \\ \quad + \sum_{\substack{i<j \\ j+1<k}} Z_{BP}(i,j) Z_{BP}(k,i-1) K_{coax}(j,i;j+1,i-1) C_{eff}(j+1,k) \alpha_j \alpha_{k-1} l_{coax}^2 \end{cases}$$

?

more $\boxed{\text{TO DO}}$:
- check $\alpha_j \alpha_{k-1}$ in last term of Z_{FINAL}
- Code up $C_{eff}^{\text{no coax single}}$
  - ↳ $C_{eff}^{\text{no BP single}}$
- FARFAR modeling rule — is okay?

$\boxed{\text{TO DO}}$
- Z_{FINAL} writeout (MOTIF) p 54
- Z_{FINAL} written here p 63
- Note on how to disallow p 58
- Code up Z_{cut}
- Try coax.

(64)

NOTES ON HOW TO DISALLOW: 

In expression for Z^{BP} [p. (62)], make following replacements:

$$Z^{BP} = \alpha_i \alpha_{j-1} \, C_{eff}^{\boxed{\text{no coax single}}}(i+1, j-1)\, l^2 l_{BP} / K_d^{BP} + \ldots$$

$$+ \sum_{i+1<k\leq j-1} \alpha_i \alpha_k \alpha_{j-1} Z_{BP}(i+1,k)\, C_{eff}^{\boxed{\text{no BP single}}}(k+1,j-1)\, l^2 l_{coax} K_{coax}(i,j;i+1,k) / K_d^{BP}$$

$$+ \sum_{i+1<k<j-1} \alpha_i \alpha_{k-1} \alpha_{j-1}\, C_{eff}^{\boxed{\text{no BP single}}}(i+1,k-1)\, Z_{BP}(k,j-1)\, l^2 l_{coax} K_{coax}(i,j;k,j-1) / K_d^{BP}$$

where: see (63)

$$C_{eff}^{\boxed{\text{no coax single}}}(i,j) = C_{eff}(i,j) - C_{INIT} l_{coax} Z^{coax}(i,j)$$

and

$$C_{eff}^{\boxed{\text{no BP single}}}(i,j) = C_{eff}(i,j) - C_{INIT}\, l_{BP} Z^{BP}(i,j)$$

"subtraction" expressions.

Pretty easy. might actually just track

$$C_{eff}^{loop}(i,j) = \alpha_{j-1} C_{eff}(i,j-1) l + \sum_{i<k<j} C_{eff}(i,k-1) l Z^{BP}(k,j) l_{BP} + \sum_{i<k<j} C_{eff}(i,k-1) l Z^{coax}(k,j) l_{coax}$$

$$C_{eff}^{\text{no coax single}}(i,j) = C_{eff}^{loop}(i,j) + C_{INIT} l_{BP} Z^{BP}(i,j)$$

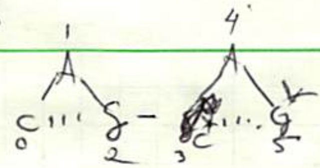
$$C_{eff}^{\text{no BP single}}(i,j) = C_{eff}^{loop}(i,j) + C_{INIT} l_{coax} Z^{BP}(i,j)$$

$$C_{eff}(i,j) = C_{eff}^{loop}(i,j) + C_{INIT} l_{BP} Z^{BP}(i,j) + C_{INIT} l_{coax} Z^{BP}(i,j)$$

Would allow for simpler back tracking than "subtraction" expressions above

Later realized that I do need to subtract out contribution from ligated BP single + coax single in Z_{FINAL}

$\Rightarrow$ $\boxed{\text{see p. (68)}}$

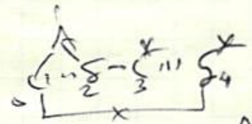


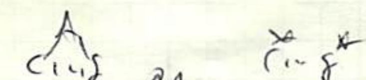
Revisiting

(trying to debug implementation in alphafold)

Start out

$Z_{BP}=0$	$Z_{BP}=0$	$Z_{BP}=0$	$Z_{BP}=0$	$Z_{BP}=0$
$Z_{coax}=0$	$Z_{coax}=0$	$Z_{coax}=0$	$Z_{coax}=0$	$Z_{coax}=0$
$C_{eff}=C_{INIT}$	$C_{eff}=C_{INIT}$	$C_{eff}=0$	$C_{eff}=0$	$C_{eff}=0$
$Z_{linear}=1$	$Z_{linear}=1$	$Z_{linear}=0$	$Z_{linear}=0$	$Z_{linear}=\frac{C_{std}}{K_d^{BP}}\left(1+\frac{C_{INIT}l^2_{BP}}{K_d^{BP}}\right)$





$$Z_{BP}=\frac{C_{INIT}l_{BP}}{K_d^{BP}} \qquad Z_{BP}=\frac{C_{std}}{K_d^{BP}}$$

$$Z_{BP}=\frac{C_{std}}{K_d^{BP}}Z_{linear}[\cdots]Z_{linear}[G]+\frac{C_{std}K_{coax}}{K_{d\,BP}}Z_{BP}[\cdots]Z_{coax}[\cdots]$$

$$=\frac{C_{std}}{K_d^{BP}}\times\left(1+\frac{C_{INIT}l^2_{BP}}{K_d^{BP}}\right)+\frac{C_{std}K_{coax}C_{INIT}l^2_{BP}}{K_d^{BP}\,K_d^{BP}}$$

$$\left\{\left[\frac{C_{INIT}l^2_{BP}}{K_d^{BP}}\right]\times\left[\frac{C_{std}}{K_d^{BP}}\right]\times K_{coax}\right\}$$

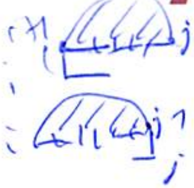
(Ah, was missing this term.)

legal

① OCT 21, 2018

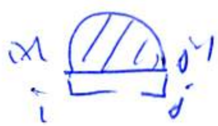
WRITE IT ALL DOWN

Stuck in Windsor locks



$Z_{5'}$

$Z_{3'}$



$Z = Z_{5'} + Z_{3'}$ ← allows only for single BASE PAIR to stack → Role of 5' or 3'.

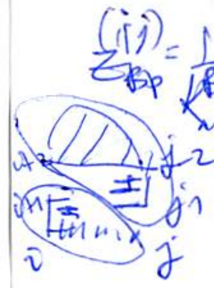
$C_{eff} = C_{INT}(l_{5'} + l_{3'})$

$Z_{BP} = \frac{1}{K_d^{BP}}\left(K_{3'}(i,j;i+1)\, l_{3'}\, C_{eff}\, [Z_{BP}(i+2,j-1) + Z_{3'}(i+1,j-2;j)\, l_{3'}]\right)$

②

WRITE IT ALL DOWN

$Z_{BP}^{(i,j)} = \frac{1}{K_d^{BP}}\left(K_{3'}(i,j;i+1)\, l_{3'}\, l_{bp}\, l\, C_{INT}\left[Z_{BP}(i+2,j-2) \times K_{3'}(i+1,j-2;j-1)\right]\right)$



- Interesting... I think I could get all the equations to work out self consistently with just 3' DANGLES

- But what if we have both 5' & 3' dangles?

Either one? •

$Z_{BP}(i,j) += \frac{1}{K_d^{BP}}\left(K_{5'}(i,j;j-1)\, l_{5'}\, l_{bp}\, l\, C_{INT}\left[Z_{BP}(i+1,j-4)\, K_{5'}(i+1,j-2;j-3)\right]\right)$

$+ K_{5'}(j-2;i+1,j-1))$

CANDLEWOOD SUITES
AN IHG HOTEL
CANDLEWOODSUITES.COM 1 888 CANDLEWOOD

CANDLEWOOD SUITES
AN IHG HOTEL
CANDLEWOODSUITES.COM 1 888 CANDLEWOOD

A-99

WRITE IT ALL DOWN

(3)

How about simultaneous 5', 3'?

→ $\boxed{\text{TREAT AS A BASEPAIR}}$

Check: only one dangle can stack, right?

can't have both, right?

$$\underbrace{Z_{BP}(i+1,j-1)\,K_{5'}\,\ell}_{Z_{5'}\text{ term in } C_{eff}(i,j-1)} + \underbrace{\ell\,Z_{BP}(i+1,j-1)\,K_{3'}}_{Z_{3'}\text{ term}}$$

BP-mismatch

$\frac{C_{eff}}{K_d^{BP}}$ or $\frac{K_{5'}K_{3'}C_{init}\,\ell\text{ mismatch}}{K_d^{BP}}$ (?)

So: $K_{5'}\ell_{3'} = K_{3'}\ell_{5'}$ or $\frac{K_{5'}C_{init}\ell_{3'}}{K_d}$ or $\frac{K_{3'}C_{init}\ell_{5'}}{K_d}$

$= K_{5'}K_{3'}C_{init}\ell_{\text{mismatch}}$
BP/mismatch
$\propto C_{eff}$

CANDLEWOOD SUITES
AN IHG HOTEL
CANDLEWOODSUITES.COM 1 888 CANDLEWOOD

WRITE IT ALL DOWN

(4)

THINK THROUGH MISMATCHES ⇒

$\boxed{\text{should be able to get asymm. correctly.}}$

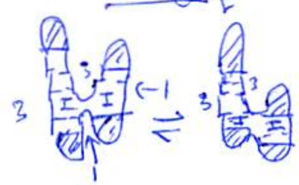
include coaxial stack for mismatch

☆ In fact, should be able to handle multiloop asymmetries?

$\boxed{\text{Cool!}}$

must break & use dissociation
single BP - coax closure

For multiloop:



If add extra nucleotide, there will be defect(s)

$\boxed{\text{Finally – I have the complete model!}}$

66-B

CANDLEWOOD SUITES
AN IHG HOTEL
CANDLEWOODSUITES.COM 1 888 CANDLEWOOD

Exactly same as Mathews 2004 model. → ✩ NOTE THAT A NUCLEOTIDE CANNOT BE BOTH 5' & 3' DANGLE BPs

✩ NOTE THAT DANGLES act as "caps" → cannot continue into mismatches, only into Ceff loops.

Model in dangles's terms to add into recursions on p(62)–(63)

$$Z^{BP}(i,j) \mathrel{+}= \alpha_i\, \alpha_{i+1}\, \alpha_{j}\, K_{3'dangle}^{(j;i;i+1)}\, C_{eff}(i+2, j-1)\, l\, l^2_{3'dangle}\, l_{BP} / K_d^{BP}$$

$$+ \; K\, \alpha_{i+1}\, \alpha_j\, C_{std}\, K_{3'dangle}(i;j;i+1)\, Z_{cut}(i+1, j) / K_d^{BP}$$

$$+ \; \alpha_i\, \alpha_{j-2}\, \alpha_{j-1}\, K_{5'dangle}(i;j;j-1)\, C_{eff}(i+1, j-2)\, l^2\, l_{5'dangle}\, l_{BP} / K_d^{BP}$$

$$+ \; \alpha_i\, \alpha_{j-2}\, \alpha_{j-1}\, K_{5'dangle}(i;j;j-1)\, Z_{cut}(i, j-1) / K_d^{BP}$$

Could also write: $K_3' = C_{INIT}\, l\, l_{3'} / K_d^{3'stack}$ $\qquad K_5' = C_{INIT}\, l\, l_{5'} / K_d^{5'stack}$

(REVISIT ON NEXT PAGE)

$$Z_{5'dangle}(i,j) \equiv K_{5'dangle}(j;i+1;i) \times Z_{BP}(i+1, j)$$

$$Z_{3'dangle}(i,j) = K_{3'dangle}(j-1;i;j)\, Z_{BP}(i, j-1)$$

$$C_{eff}(i,j) \mathrel{+}= C_{INIT}\, l_{5'dangle}\, Z_{5'dangle}(i,j) + \sum_{i<k<j} C_{eff}(i,k-1)\, l\, Z_{5'dangle}(k,j)\, l_{5'dangle}$$

$$+ \; C_{INIT}\, l_{3'dangle}\, Z_{3'dangle}(i,j) + \sum_{i<k<j} C_{eff}(i,k-1)\, l\, Z_{3'dangle}(k,j)\, l_{3'dangle}$$

$$Z_{linear}(i,j) \mathrel{+}= Z_{5'dangle}(i,j) + \sum_{i<k<j} Z_{linear}(i,k-1)\, Z_{5'dangle}(k,j)$$

$$+ \; Z_{3'dangle}(i,j) + \sum_{i<k<j} Z_{linear}(i,k-1)\, Z_{3'dangle}(k,j)$$

$$Z_{final}(i,i) \mathrel{+}= \sum_{i<j<(i-1)+N} \alpha_j\, \alpha_{i-2}\, K_{5'dangle}(j;i;i-1)\, l\, l^2_{5'dangle}\, C_{eff}(j+1, i-2)\, Z_{BP}(i,j)$$

$$+ \sum_{i<j<(i-1)+N} K_{5'dangle}(j;i;i-1)\, Z_{BP}(i,j)\, Z_{cut}(j, i-1)$$

$$+ \sum_{i<j<(i-1)+N} \alpha_i\, \alpha_{j-1}\, K_{3'dangle}(i-1;j;i)\, C_{eff}(i+1, j-1)\, l\, l^2_{3'dangle}\, Z_{BP}(j, i-1)$$

$$+ \sum_{i<j<(i-1)+N} K_{3'dangle}(i-1;j;i)\, Z_{cut}(i,j)\, Z_{BP}(j, i-1)$$

I ~~we think~~ was previously following this line of reasoning:

in $\boxed{I 3'}_j$ $\quad K_{3'} = C_{INIT}\, l_{3'} / K_d^{3'}$

in $\boxed{I 3', I 5'}$ $\quad Z = K_{3'} \times C_{INIT}\, l_{5'} / K_d^{BP} \approx C_{eff}^{\text{MISMATCH/BASE PAIR STACK}} / K_d^{BP}$, in approximation (likely incorrect) that 3' and 3' stacks are not conflicting.

So identify $C_{INIT} K_{3'}' l_{5'}' = C_{eff}^{\text{MISMATCH/BASE PAIR STACK}}$

and analogously $C_{INIT} l_3' K_5' = C_{eff}^{\text{MISMATCH/BASE PAIR STACK}}$

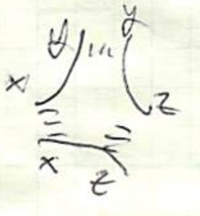
$K_3' l_5' = l_3' K_5'$?

$l_{5'} = K_{5'}\, l_{3'} / l_{5'}$ O.K.

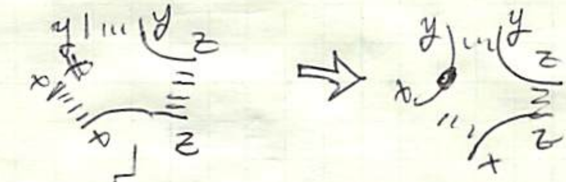
TO DO:

68

OCT 22, 2018



Seeing Z_{final} ERROR when closing this test case and excluding "strained 3WJ".

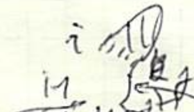


$\boxed{\text{should not allow closure!}}$

May even need to manually subtract these two rare cases off.

How ugly!

$\boxed{\text{But I kept it}}$



(A)

$$Z_{final}(i,i) \mathrel{-}= \sum_{i<j<i-1+N} \alpha_j \left[Z_{BP}(i,j)\, Z_{coax}(j+1,i-1) + Z_{coax}(i,j)\, Z_{BP}(j+1,i-1) \right] l_{BP\,coax}$$

[too complicated] Eventually, in case of BACKTRACKING, could just "divert" to all contributions to $C_{eff}(0,i-1)$ and exclude problem cases.

- I in principle could keep track of another matrix like $C_{eff}[\text{no single BP-COAX}]$

So

But NOTE: this only happens in Z_{final} when closing a circle.

$$C_{eff}(i,j) = \ldots \sum_{i<k<j} C_{eff\,no\,single\,coax}(i,k-1)\, Z_{BP}(k,j) + C_{INIT}\, Z_{coax}(i,k-1)\, Z_{BP}(k,j)$$

$$+ C_{eff\,no\,single\,coax}(i,k-1)\, Z_{BP}(k,j) + C_{INIT}\, Z_{BP}(i,k-1)\, Z_{coax}(k,j)$$

then add these

yes! Could add these FIRST $\rightarrow$ $C_{eff}[\text{no BP-coax doublet}]$

Ugly.

LAST IDEA: THIS WORKED

update C_{eff}, Recognize offset $= N-1, d_j$, switch to $\left\{ \begin{array}{l} C_{eff\,no\,single\,coax} \text{ in BP loop} \\ C_{eff\,no\,single\,BP} \text{ in loop} \end{array} \right.$

yes! yes!

27 OCT 2015

(0) ASCII Art

(1) CREATE class with periodic Boundary Condition → check tests

(2) Then, do derivs → check tests for KdP

(3) Then, do accumulations of contributions → {Enumerative BACKTRACK, Stochastic BACKTRACK, MFE}

planned

(4) Then, port parameters from Vienna 1.8.5 / ~~RNAstructure~~ RNAsoft

i.e., Turner-1999

check on tRNA, incl. F-matrix; ~~log~~ Δ G NATIVE

(A) Do we want chemical mapping data for every tRNA?

(5) Excellent ~~to~~ takeoff point for ~~optimization(?)~~ testing 3° ~~[illegible]~~ for TRNA motifs. { A Base PAIR, U-turns, T-loop, ~~[illegible]~~ D-loop, 3° contact

(6) then ~~log (~~ logN - loops

(7) then code up in C++ to optimize, [w/ python Bindings] [BFGS/SGD]

25 Oct 2015

Peripheral [GITHUB "issues"]

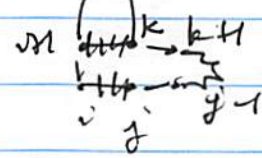
(8) Dangles

(9) Hide derivs/contribs into class, and write "compiler" to translate into FAST python or C++ code.

28 OCT, 2018 -A

A way to make the code even more readable...

Consider a term:



$$Z_{BP}(i,j) = \sum_{i+1<k<j-1} \underbrace{Z_{BP}(i+1,k)\, C_{eff}(k+1,j-1)\, l^3 l_{BP}^2 K_{coax}/K_d^{BP}}_{Z_{BP}^{CONTRIB}}$$

Then derivative w.r.t. any parameter λ is given by chain rule:

$$\frac{1}{Z^{CONTRIB}_{BP}(i,k,j)} \frac{\partial Z^{CONTRIB}_{BP}(i,k,j)}{\partial \lambda} = \frac{1}{Z_{BP}(i+1,k)} \frac{\partial Z_{BP}(i+1,k)}{\partial \lambda} + \frac{1}{C_{eff}(k+1,j-1)} \frac{\partial C_{eff}}{\partial \lambda}$$

$$+ 3 \frac{1}{l}\frac{\partial l}{\partial \lambda} + 2 \frac{1}{l_{BP}} \frac{\partial l_{BP}}{\partial \lambda} + \frac{1}{K_{coax}} \frac{\partial K_{coax}}{\partial \lambda}$$

$$- \frac{1}{K_{coax}} \frac{\partial K_d^{BP}}{\partial \lambda}.$$

and $$\frac{\partial Z_{BP}(i,j)}{\partial \lambda} = \sum_{i+1<k<j-1} \partial Z^{CONTRIB}_{BP}/\partial \lambda\,(i,k,j)$$

Could also write in terms of derivatives:

$$\frac{\partial}{\partial \lambda} \log Z_{BP}(i,j) = \frac{1}{Z_{BP}(i,j)} \frac{\partial Z_{BP}(i,j)}{\partial \lambda}$$

$$= \frac{1}{Z_{BP}(i,j)} \Big[\sum_{i+1<k<j-1} Z^{contrib}_{BP}(i,k,j) \times \Big[\frac{\partial}{\partial \lambda} \log Z_{BP}(i+1,k) + \frac{\partial}{\partial \lambda} \log C_{eff}(k+1,j) \; \ldots \Big] \Big]$$

Probably easier to write out chain rule for product:

$$Z_{BP}^{*}(i,j) = \ldots \sum_{i+1<k\leq j-1} Z_{BP}(i+1,k)\, C_{eff}(k+1,j-1)\, l\, l l l_{BP}\, l_{BP}\, K_{coax} K_{a}^{BP} \ldots$$

$$\frac{\partial}{\partial \lambda} Z_{BP}(i,j) = \ldots \sum_{i+1<k\leq j-1} \Big[\frac{\partial Z_{BP}(i+1,k)}{\partial \lambda} C_{eff}(k+1,j-1)\, l l l l_{BP}\, l_{BP}\, K_{coax} K_{a}^{BP} +$$

$$+ Z_{BP}(i+1,k) \frac{\partial C_{eff}(k+1,j-1)}{\partial \lambda}\, l l l l_{BP}\, l_{BP}\, K_{coax} K_{a}^{BP} +$$

$$+ Z_{BP}(i+1,k)\, C_{eff}(k+1,j-1) \frac{\partial l}{\partial \lambda} l l l_{BP}\, l_{BP}\, K_{coax} K_{a}^{BP}$$

$$\ldots$$

$$+ Z_{BP}(i+1,k)\, C_{eff}(k+1,j-1)\, l l l l_{BP}\, l_{BP}\, K_{coax} \frac{\partial K_{a}^{BP}}{\partial \lambda} \Big]$$

So we could actually keep track of objects

$$\begin{pmatrix} Z_{BP} \\ \frac{\partial}{\partial \lambda} Z_{BP} \end{pmatrix} \quad \begin{pmatrix} C_{eff} \\ \frac{\partial}{\partial \lambda} C_{eff} \end{pmatrix} \quad \begin{pmatrix} l \\ \frac{\partial}{\partial \lambda} l \end{pmatrix}$$

and when we take product, use product rule:

$$\begin{pmatrix} A \\ \frac{\partial A}{\partial \lambda} \end{pmatrix} * \begin{pmatrix} B \\ \frac{\partial B}{\partial \lambda} \end{pmatrix} = \begin{pmatrix} A \cdot B \\ \frac{\partial A}{\partial \lambda} B + A \frac{\partial B}{\partial \lambda} \end{pmatrix}$$

<u>Problem</u>: I am seeing a lot of overhead in Python when I overload even the __getitem__ class, e.g. to automatically wrap

$$Z_{BP}(i+1,k) \Rightarrow Z_{BP}(\mathrm{mod}(i+1,N), \mathrm{mod}(k,N))$$

Can't have these in inner loops of code.

Needed FOR MFE, stochastic BACKTRACKING

Could also accumulate separate contributions to each partition function through overloading addition:

$$Z = AB + CD$$

list of contributions

$$\begin{pmatrix} Z \\ \partial Z/\partial 1 \\ Z^{CONTRIB} \end{pmatrix} = \begin{pmatrix} A \\ \partial A/\partial 1 \\ A^{CONTRIB} \end{pmatrix} \begin{pmatrix} B \\ \partial B/\partial 1 \\ B^{CONTRIB} \end{pmatrix} + \begin{pmatrix} C \\ \partial C/\partial 1 \\ C^{CONTRIB} \end{pmatrix} \begin{pmatrix} D \\ \partial D/\partial 1 \\ D^{CONTRIB} \end{pmatrix}$$

$$= \begin{pmatrix} AB \\ \frac{\partial A}{\partial 1} B + A \frac{\partial B}{\partial 1} \\ [AB] \end{pmatrix} + \begin{pmatrix} CD \\ \frac{\partial C}{\partial 1} D + C \frac{\partial D}{\partial 1} \\ [CD] \end{pmatrix}$$

$$= \begin{pmatrix} AB + CD \\ \frac{\partial A}{\partial 1} B + A \frac{\partial B}{\partial 1} + \frac{\partial C}{\partial 1} D + C \frac{\partial D}{\partial 1} \\ [AB, CD] \end{pmatrix}$$

- Again, don't want the accumulation in inner loops of code, due to cost of overloading.

- In addition, keeping track of contributions will cost N^3 memory (will probably limit $N \ll 1000$) and incur overhead in dynamic updating of lists.

  $\Rightarrow$ instead only do it if we're backtracking and perhaps even repeat calculation at those steps.

To solve code readability problems, and code copying,
could simply write a little "compiler" py
that goes from a text file recursions.txt
of commented recursions to a
single python file:
derived_recursions.py [for alphafold]
and/or
derived_recursions.h [for C++ alphafoldplus]
• (Alternative would be MACROs but I find those cryptic)
• In fact, as stepping stone (or sanity check),
could also code up objects to ~~have~~
"magically" do derivs and contributions too,
even at the expense of speed.

E.

Thoughts on convexity

$$Z = \sum_{\substack{\text{STRUCTURES}\\ s}} e^{\sum_{\substack{\text{FEATURE}\\ j}} w_j n_j(s)}$$

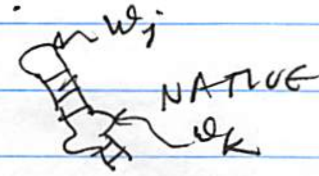
$$\frac{\partial}{\partial w_j}\log Z = \frac{1}{Z}\sum_{\substack{\text{STRUCTURES}\\ s}} n_j e^{-\sum w_j n_j(s)} = \langle n_j \rangle$$

$$\frac{\partial^2}{\partial w_j \partial w_k}\log Z = \left(-\frac{1}{Z^2}\frac{\partial Z}{\partial w_k}\right)\sum_{\substack{\text{STRUCTURES}\\ s}} n_j e^{-\sum w_j n_j(s)}$$

$$+\left(\frac{1}{Z}\right)\sum_{\substack{\text{STRUCTURES}\\ s}} n_j n_k e^{-\sum w_j n_j(s)}$$

$$= \langle n_j n_k \rangle - \langle n_j \rangle \langle n_k \rangle$$

Suppose I have maximum likelihood [or minimal free energy gap] target:



$$\Delta G_{NATIVE} = -\log Z_{NATIVE} + \log Z$$

$$\frac{\partial^2}{\partial w_j \partial w_k}\Delta G_{NATIVE} = -\cancel{\left[n_j^{NATIVE} n_k^{NATIVE} - n_j^{NATIVE} n_k^{NATIVE}\right]}^{\to 0} + \left[\langle n_j n_k \rangle - \langle n_j \rangle \langle n_k \rangle\right]$$

$$= \langle n_j n_k \rangle - \langle n_j \rangle \langle n_k \rangle$$

Is this covariance matrix always positive definite?

Notes on BACKTRACKING

Return to "simple recursions"

$$Z_{BP}^{(i,j)} = C_{eff}(i+1, j-1) l_{BP} e^{2} / k_{BP} + \sum_{i \le c < j} Z_{linear}(i+1, c) Z_{linear}(c+1, j-1) e_{BP}$$

$$C_{eff}^{(i,j)} = C_{eff}(i, j-1) l + C_{init} Z_{BP}(i,j) l_{BP} + \sum_{i < k < j} l_{BP} C_{eff}(i, k-1) Z_{BP}(k,j)$$

$$Z_{linear} = Z_{linear}(i, j-1) + Z_{BP}(i,j) + \sum_{i < k < j} Z_{linear}(i, k-1) Z_{BP}(k,j)$$

$$Z_{final}^{(i)} = Z_{linear}(i, i-1) \quad \text{if} \quad i \text{ is cutpoint}$$

7/8

How to backtrack? MFE?

Start with $Z_{final}(0)$

↓

$Z_{linear}(0, N-1)$

$Z_{linear}(0,N-2)$ $\quad$ $Z_{BP}(0,N)$ $\quad$ $Z_{linear}(0,k-1)\, Z_{BP}(k,N-1)$ ← several terms

choose max

Zlin $\quad$ ZBP

[$Z_{linear}(0,k-1)$ …] $\quad$ [Ceff $\quad$ Zlin Zlin]

To: For a recursion [depth-first]

- ask for each contribution
- choose max Z

Set that of each comp.

Keep: BP list, probability? $\quad$ o.k. worked

How about enumeration?

Zlin → return (BPs_all) → [P_1 {BP_1 BP_2 …}, P_2 {BP}] all alternative

Zlin → return (BPs_all) → ( ) all alternative

Then take all combinations

↓

output {P, BP}

2 NOVEMBER, 2018

alphafold: $\boxed{\text{mag to completion...}}$

PAPER 0 → { ~~Eterna~~ Rank all existing BPP algorithms based on Eterna { SHAPE + ~~Array Kd measurements~~ also m^2 [no 3' structure] }

PAPER 1 [AlphaFold premiere]

- $\mu \pm \sigma$ + Bootstrapping → still need theory
- Rosetta-estimated starter values for coax, etc.
- Evaluate success on RMDB$_{2,L}$; Turner ΔG; cloud lab MI; mg2 binding (Eterna Array measurements ?)
- All structured motifs incl. Rosetta estimates
- logarithmic loop penalty ← write out theory. ← check ΔA on tRNA
- noncanonical base pairs

prospective expts on switches

PAPER 2

- Partition Function incl. 3' structure ← Need { ΔG estimator (incl. normal modes); $\Delta\Delta G$ estimator }
- Evaluate success via m^2-data (e.g. tRNA) + prospective experiments on alternative states.

PAPER 3

- C_{eff} + 6D distributions → what's the story? Is that even "inside" alphafold?