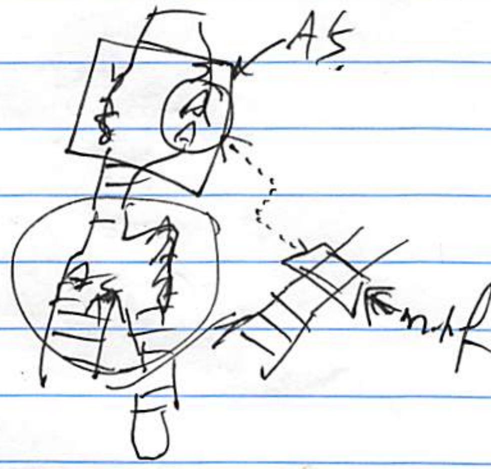
Revisiting Secondary Structure with non-canonical motifs.

25 SEP 2015

See: p-5 & p-10 FOR CLEANER NOTES

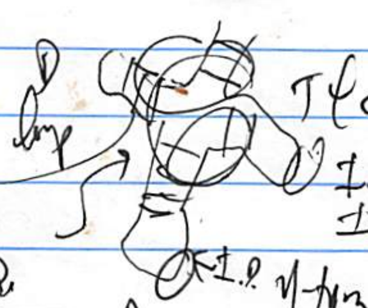
glycine Riboswitch A-minor:



should be able to
(1) align bases
(2) compute: 2° struct → 3° pred. (convolution)

o.k.: But what's first step where we can evaluate success?

Consider: tRNA



Triple stack

D-loop
triple junct

Anticodon

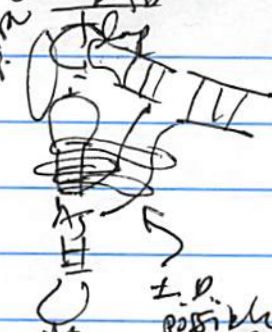
1st PASS

2nd PASS

identify things
U-turns as candidate submotif

summarize possible intercalators — estimate c.o.ff

3rd PASS



I.P. possible + triple junct

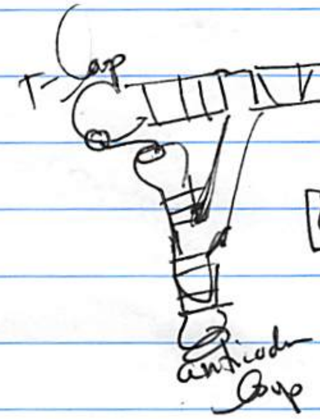
LAST PASS

triple junct

→ Could "clean up" alignments.
→ test by canonical mapping.

Input:

Sequence



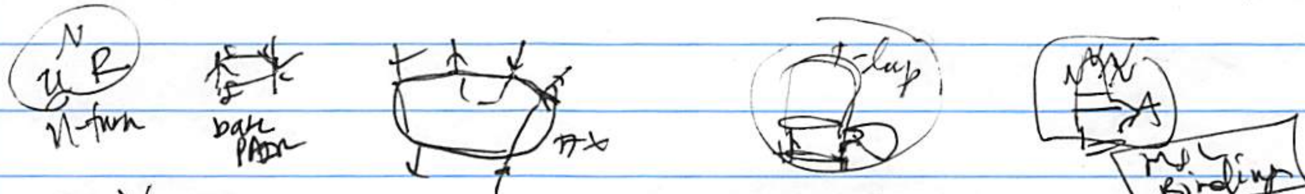
Sample of 1000 structures

graph

BASE PAIR + 3° motif list.

Should allow for discovery of tRNA motifs and T-loop interactions across transcriptome.

Need TDO detector? → Special 3D [illegible] → Define Rules for detection + scoring



Possible Recipe:

1. Scan sequence for all possible motifs.
2. Use dynamic programming [plus any motif bonuses] + calc. partition function.
3. Stochastic BACKTRACK sampling → 1000 members
4. Re-weight by 3° contacts (and remove bumpers)

As output ⇒ Compute MEA. (Some kind of consensus?) motifs detected.

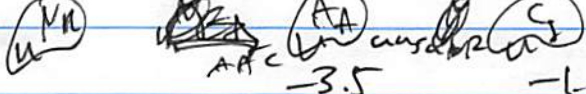
Success means: all E. coli & human tRNA have
> 1% prob of NATIVE 2° structure (?)
– then go through RNAse P, etc...

AND Beautifully Probs

Simple Int.

What are ingredients?

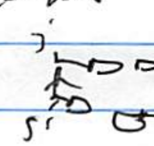
– motif identifiers & scorers [1 M standard state]



3D structures, scores

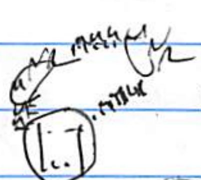
−3.5 −1.0 kcal/mol

→ −2.0 kcal/mol



← 3D structure.

– How to calculate 2d structure partition function?



⇨ Need calculator (OR rules).

⇨ Need calculator (or rules)

EACH motif has $\boxed{\text{flanking BASE PAIR}}$ { A•U, C•G, stack on pair, G•A, G•UA }

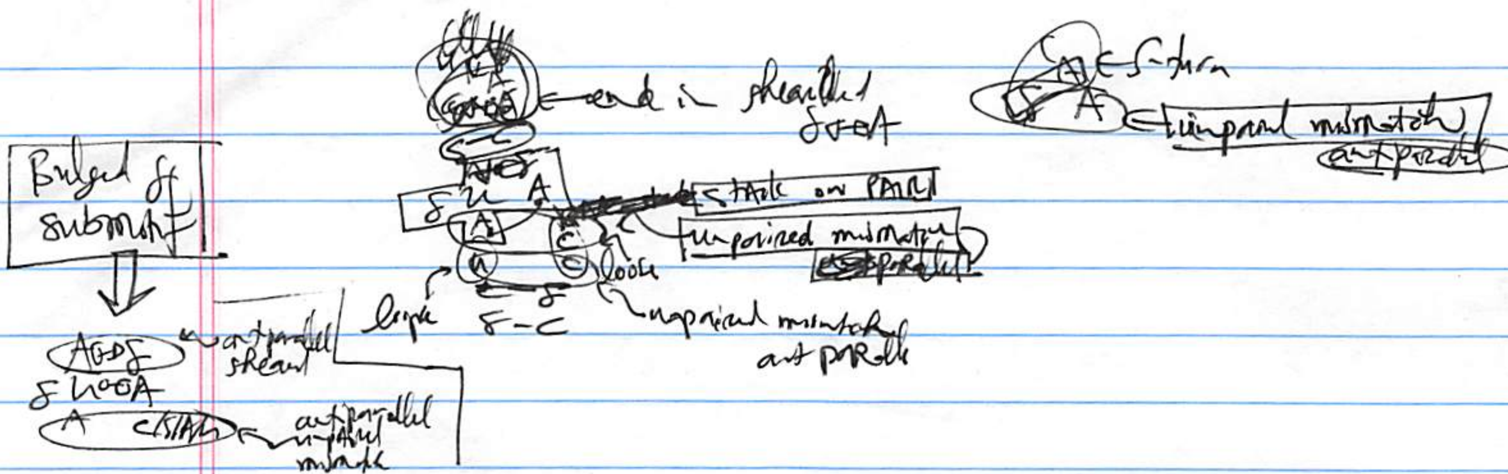
– Need rules to go between all of them.
(As default, could just use PDB stats.)

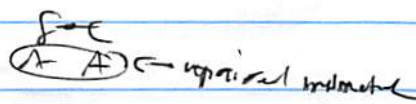
stack junction.

Need $\Delta G\left(\begin{array}{c}\text{Base doublet } i \\ \Uparrow \\ \text{Base doublet } j\end{array}\right) \leftarrow kT\log(C_{eff}/1M)$

motifs must be capped by base doublets, or

$\Delta G = +kT\log(\ldots)$





One crazy idea:

Make everything BASE PAIRED, including random coil.

BPtype : a = WC AU, WC UA, ... WC G-A, UU, coax stack

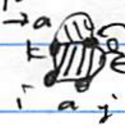
i, j → index over residues

VERY SIMPLE RECURSIONS

$$Z^{BP_a}_{(i,j)} = \cancel{\#} \frac{1}{K_d^{BP_a}} + C_{eff}$$

$$= \frac{1}{K_d^{BP_a}} \times \left[C^{loop}_{eff(i,j)} \, l + \sum_{\substack{motifs \\ m: b \to a}} C^{m(b\to a)}_{eff} Z^{BP_b}_{(k,l)} \right]$$

should also include multi-helix junction motifs.



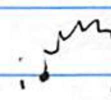
motif m(b→a)

$$C^{Loop}_{eff(i,j)} = C^{Loop}_{eff}(i, j-1) \times l$$ ← cost of closing extra nucleotide

$$+ \sum_{i < k \le j-1} C^{Loop}_{eff}(i,k) \, Z^{BP_a}_{(k,j)} \times l^{BP_a}$$

extra cost of closing loop with BP_a → Actually $C^{BP}_{init} = l^{BP_a} C_{init}$

$Z^{BP_a}_{(i,j)} C^{BP_a}_{init}$



+ ~~[crossed out]~~

[initialize $C_{eff\,(i,i)} = C_{INIT.}$]

Fill dynamic programming matrix ..

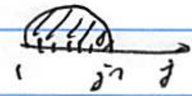
~~[crossed out]~~

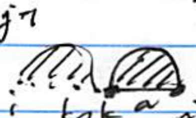
If linear,

uh wait...

Note that this can be carried out separately:

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





If circularized,

$$Z_{tot} = \frac{C_{eff}(1,N)\, l}{K_d^{ligation}} = \frac{C_{eff}(2, N+1)\, l}{K_d^{ligation}} = \ldots = \frac{C_{eff}(N, N+N-1)\, l}{K_d^{ligation}}$$

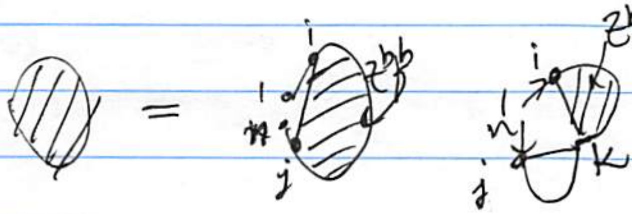
It would be powerful to have cross-checks from circular calcs... (?)

something like

~~$\sum_{i,j} Z^{BPa}_{(i,j)} C_{eff}(j,i)$~~

$$Z_{CLOSE} = C_{eff}^{N(b-a)} \, Z^{BPb}_{(i,j)} \, Z^{BPa}_{(j+1,i-1)}$$

Hofacker/Stadler style:

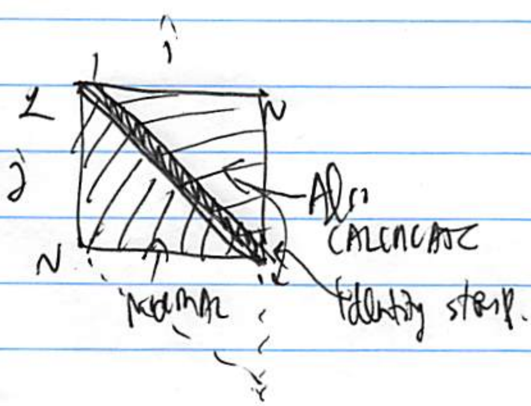


Hmm... how about:

$$Z_{circle} = C_{eff}^{(1,N)} * \left(\frac{1}{K_{\#}^{\text{LISATION}}}\right) \times l$$

provides nice cross-check.

arbitrary



TEST CASES:

Check:

$$C^{(1\to 1)} = C^{(2\to 2)} = \ldots$$

$^{1}C\ {}^{2}A\ {}^{3}A\ {}^{4}A\ {}^{5}\delta\ {}^{6}A\ {}^{7}A$

$$Z^{\text{circle}} = \frac{C^{INIT}\, \ell^4}{K_d^{CS}} \times \frac{\ell \times \ell}{K^{LIGATION}} = \frac{C^{Int}}{K^{ligation}} \frac{\ell^4}{K_d^{CS}}.$$

Check:

$^{1}C\ {}^{2}A\ {}^{3}A\ {}^{4}A\ {}^{5}\delta$

$$Z^{BP}_{(1,5)} = \frac{C^{INT}\ell^4}{K_d^{CS}} \qquad C^{Loop}_{eff}(1,5) = C^{INT}\ell^4 + \frac{C^{INT}\ell^4}{K_d^{BP}} C^{BP}_{int}$$

$^{1}C\ {}^{2}A\ {}^{3}A\ {}^{4}A\ {}^{5}\delta\ {}^{7}A$

$$C_{eff}^{(7,1\ldots 5)} = C \times \left[\frac{C^{INT}\ell^4}{K_d^{CS}}\right] \times \ell^{C-\delta} + C^{INT}\ell^5$$

$$\boxed{\ell_{INT}^{C-\delta}\, \ell\, C_{int}^{C-\delta}}$$

$^{1}C\ A\ A\ A\ {}^{5}\delta\ {}^{6}A$

$$C_{eff}^{(1,\ldots 5)} = C_{eff}^{1\ldots 5} \times \ell$$

$$= \frac{C^{INT}\ell^4 C^{C-\delta}_{int}}{K_d^{CS}} + C^{INT}\ell^5$$

Hmm Fixed.

SCRATCHWORK

Just A's loop $^{1}A\ {}^{2}A\ {}^{3}A\ {}^{4}A$

$^{2}A\ {}^{3}A$ $\qquad C_{eff}(2,3) = \ell_{int}$

$^{1}A\ {}^{2}A\ {}^{3}A$ $\qquad C_{eff}(1,3) = \ell^2_{int}$

$^{1}A\ {}^{2}A\ {}^{3}A\ {}^{4}A$ $\qquad C_{eff}(1,4) = \ell^3_{int}$

$A\ A\ A\ A$ $\qquad C_{eff}(1\to 1) = \ell^4_{int}/K_d^{LIGATION}$ ← o.K.

SCRATCHWORK

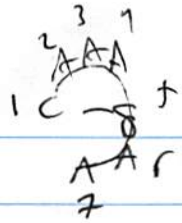
$^{1}C\ {}^{2}A\ {}^{3}A\ {}^{4}A$ $\qquad C_{eff}(1,4) = C\ell^3_{int}$

$^{1}C\ {}^{2}A\ {}^{3}A\ {}^{4}A\ {}^{5}\delta$

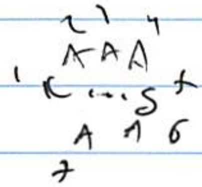
$$Z^{(1,5)}_{BP} = \left[\frac{C_{int}\ell^3}{K_d^{BP}}\right] \times \ell \rightarrow C_{eff}(1,5) = C_{int}\ell^4 + \frac{C_{int}\ell^4}{K_d^{BP}} C^{BP}_{int}$$

AHA.

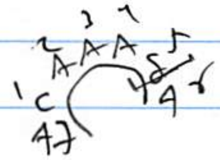
8



$$C_{\text{eff}}^{1\ldots7} = C^{\text{INIT}}\ell^{6}\left(1+\frac{C_{\text{init}}^{C-G}}{K_d^{C-G}}\right)$$



$$Z^{\text{circle}} = C^{1\to1} = \frac{C^{\text{INIT}}\ell^{7}}{K_d^{\text{lig}}}\left(1+\frac{C_{\text{init}}^{C-G}}{K_d^{C-G}}\right)$$



or closing circle differently,

$$C^{7,1\ldots5} = C^{\text{INIT}}\,{}^{6}\ell\left(1+\frac{C_{\text{init}}^{C-G}}{K_d^{C-G}}\right)$$

used $C_{\text{init}}^{C-G} = \ell\, C_{\text{init}}^{C-G}$

$$Z^{\text{circle}} = C^{7\to7} = \frac{C^{\text{INIT}}\ell^{7}}{K_d^{\text{lig}}}\left(1+\frac{C_{\text{init}}^{C-G}}{K_d^{C-G}}\right)$$

UNPAIRED CIRCLE

PAIRED CIRCLE

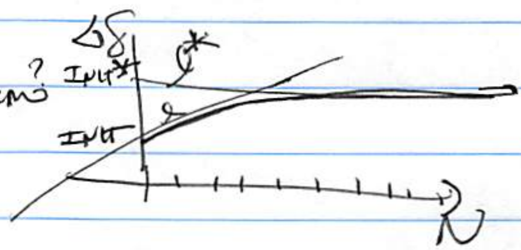
only works out so nicely due to assumption that

$$C^{\text{eff}}_{\text{CLOSURE}} = \ell^{N}, \quad \text{i.e.,}$$

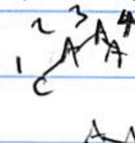
$$\Delta G_{\text{CLOSURE}} \sim N \log L$$

linear dependence on elements.

What if we assumed "two-line" form?

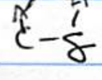


Does this work?



$C_{\text{eff}}^{1\to4} = c^{\text{init}}\ell^{3}$ ~~...~~

$$C^{*\,1\to4}_{\text{eff}} = c^{\text{init}*}\ell_*^{3}$$



$$Z^{\text{sp}}_{\text{lig}} = \frac{1}{K_d^{C-G}}\left[c^{\text{init}}\ell^{4} + c^{\text{init}*}_{*}\ell_*^{4}\right]$$

9

$$Z^{BP}_{i,1} = \frac{1}{K_d^{C-S}}\left[c^{init}l^4_{+} \quad c^{init}_{*}l^4_{*}\right]$$

$$C^{1\to4}_{eff} = c^{init}l^5 + \frac{1}{K_d^{C-S}}\left[c^{init}l^4 + c^{init}_{*}l^4_{*}\right]C^{C-S}_{init}$$

$$C^{1\to4}_{eff,*} = c^{init}_{*}l^5_{*} + \frac{1}{K_d^{C-S}}\left[c^{init}l^4 + c^{init}_{*}l^4_{*}\right]C^{C-S}_{init,*}$$

$$C^{1\to1}_{eff,*} \stackrel{?}{=} \frac{c^{init}_{*}l^7_{*}}{K_d^{lig}} + \frac{l^3_{*}}{K_d^{C-S}K_d^{lig}}\left[c^{init}l^4 + c^{init}_{*}l^4_{*}\right]C^{C-S}_{init,*}$$

$$C^{1\to1}_{eff} = \frac{c^{init}l^7}{K_d^{lig}} + \frac{l^3}{K_d^{C-S}K_d^{lig}}\left[c^{init}l^4 + c^{init}_{*}l^4_{*}\right]C^{C-S}_{init}$$

Sum $C^{1\to1}_{eff}[TOTAL] = C^{1\to1}_{eff} + C^{1\to1}_{eff,*}$

$$= \frac{c^{init}l^7}{K_d^{lig}} + \frac{c^{init}_{*}l^7_{*}}{K_d^{lig}}$$

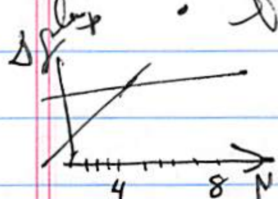
$$+ \frac{l^7 c^{init} C^{init}_{[C-S]}}{K_d^{C-S}K_d^{lig}} + \frac{l^3 l^4_{*}}{K_d^{C-S}K_d^{lig}} c^{init}_{*} C^{C-S}_{init}$$

$$+ \frac{l^3_{*} l^4 c^{init} C^{init[C-S]}_{*}}{K_d^{C-S}K_d^{lig}} + \frac{l^7_{*} c^{init}_{*} C^{C-S}_{init,*}}{K_d^{C-S}K_d^{lig}}$$

Cool, looks well defined. $\boxed{\text{Write out clearly on next page}}$

Model with

- multiple BASE PAIR TYPES $a = \frac{WC}{C-G}, \frac{WC}{G-C}, \ldots$ COAX, ...
- Nonlinear Loop model [CAPTURED through two linear approximations]



$C^{init} = 10^{2} M$; $\ell = 10$; $C^{init}_{*} = 10^{4} M$; $\ell_{*} = 8$

$K_d^{C-G} = 10^{3} M$

$C_{eff}\left(\begin{smallmatrix} C--G \\ G--C \\ \text{motif} \end{smallmatrix}\right) = \sim 10^{5} M$; $C_{eff}\left(\begin{smallmatrix} C--G \\ G--C \\ \text{motif} \end{smallmatrix}\right) = \sim 10^{5} M$

$C_{eff}\left(\begin{smallmatrix} G--C \\ C--G \\ \text{motif} \end{smallmatrix}\right) = \sim 10^{5} M$

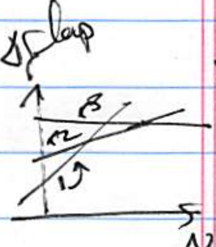
$\left[C^{init}_{CG} = c^{init} \; ; \; C^{init}_{CG*} = c^{init}_{*} \right]$

$$Z^{BP_a}_{(i,j)} = \frac{1}{K_d^{BP_a}} \cdot \left[C_{eff}(i,j-1)\ell + C^{*}_{eff}(i,j-1)\ell^{*} + \sum_{\substack{m(b\to a) \\ (k,l)(i,j)}} C_{eff}^{m(b\to a)} Z^{BP_b}_{(k,l)} \right]$$

$$C_{eff}(i,j) = C_{eff}(i,j-1)\,\ell + \sum_{i<k<j} C_{eff}(i,k-1)\,\ell \left(\frac{C^{BP_a}_{init}}{C_{init}}\right) Z(k,j)$$

$$C^{*}_{eff}(i,j) = C^{*}_{eff}(i,j-1)\,\ell^{*} + \sum_{i<k<j} C^{*}_{eff}(i,k-1)\,\ell^{*} \left(\frac{C^{BP_a *}_{init}}{C^{*}_{init}}\right) Z^{BP_a}(k,j)$$

Call this ρ^{BP_a}



In fact, could generalize to any number of lines composing loop cost.

Let q index over loop "type":

$$\boxed{\begin{aligned} Z^{BP_a}_{(i,j)} &= \frac{1}{K_d^{BP_a}} \left[\sum_q C^{q}_{eff}(i,j-1)\,\ell_q + \sum_{\substack{m(b\to a) \\ k,l \; i,j}} C_{eff}^{m(b\to a)} Z^{BP_b}_{(k,l)} \right] \\ C^{q}_{eff}(i,j) &= C^{q}_{eff}(i,j-1)\,\ell^{q} + \sum_{i<k<j} C^{q}_{eff}(i,k-1)\,\ell^{q} \rho^{BP_a} Z^{BP_a}_{(k,j)} \end{aligned}}$$

and that's it...

Well, if we need $Z^{LINEAR}(i,j) = Z^{linear}_{(i,j-1)} + \sum_{i<k<j} Z^{linear}_{(i,k-1)} Z^{BP_a}_{(k,j)}$

Open Questions

- Can we calculate base pair probabilities?
- Can we calculate derivatives?
- Can we BACTRACK?
- Double-check: will 3WJ be penalized compared to 4WJ?
- In fact... how to handle coax?
- How to prevent short apical loops? (Maybe just initialize: $C_{eff}(i,i)=\infty$, $C_{eff}(i,i+1)=\infty$, $C_{eff}(i,i+2)=C^{init}$)

Additional Notes

- SAVE All params as doublets (μ, σ)

  BEST GUESS ERROR. (FOR BOOTSTRAPPING)

-

COAXIAL STACKS thoughts

$$Z^{COAX}_{(i,j)} = Z^{BP}_{(i,k)} \, Z^{BP}_{(k+1,j)} \, K^{COAX}$$

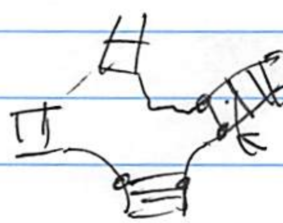
$\boxed{\text{BONUS} > 1}$

$$Z^{BP}_{(i,j)} = C_{eff}^{(i,k-1)} Z^{BP}_{(k,j)} K^{COAX} + Z^{BP}_{(i,k-1)} K^{COAX\,(i,k)} C_{eff} K^{COAX}$$

$\boxed{\text{But wait a second}}$ → 2 and 3 will form — But how about 1–4?

$\boxed{\text{Does not work}}$

INSTEAD, may need to pre-identify 4WJ motifs on possible co-ax on likely structure, treat carefully



Again – Need to work out carefully w/ model system with circularization.

0 1 2 3

CAδA

CHECK THIS

$i \backslash j$	0	1	2	3
0	C $C_{eff} = 1$ $Z = 0$	CA $C_{eff} = l$ $Z = 0$	CAδ $C_{eff} = l^2 + \left(\frac{l^2}{K_d}\right)$ [illegible] $Z = l^2/K_d$	CAδA C_{init} l
1	ASA $C_{eff} = l^3 + \frac{l^2}{K_d}$ [illegible] $Z = 0$	A $C_{eff} = 1$ $Z = 0$	Aδ $C_{eff} = l$ $Z = 0$	AδA $C_{eff} = l^2$ $Z = 0$
2	δA $C_{eff} = l^2 + \frac{l^2}{K_d}$ [illegible] C_{init} l $Z = l^2/K_d$	δA δA	δ $C_{eff} = 1$ $Z = 0$	δA $C_{eff} = l$ $Z = 0$
3	C A $C_{eff} = l$ $Z = 0$	CA A $C_{eff} = l^2$ $Z = 0$	CAδ A $C_{eff} = l^3 + \left(\frac{l^2}{K_d}\right)$ [illegible] $Z = 0$	A $C_{eff} = 1$ $Z = 0$

0	1	2	3
C	CA	CAδ	CAδA
	A	Aδ	AδA C

Rewritten on next page

	$i=0$	1	2	3			
$j=0$	C $Z_{BP}=0$ $C_{eff}=c^{INIT}$	CA $Z_{BP}=0$ $C_{eff}=C^{INIT}\ell$	$Z_{BP}=\frac{C^{INIT}\ell}{K_d}$ $C_{eff}=C^{INIT}\ell^2+\frac{C^{INIT}C^{INIT}_{BP}\ell}{K_d}$	$Z_{BP}=0$ $C_{eff}=C^{INIT}\ell^3+\frac{C^{INIT}C^{INIT}_{BP}\ell^3}{K_d}$			
1		A $Z_{BP}=0$ $C_{eff}=c^{INIT}$	A δ $Z_{BP}=0$ $C_{eff}=C^{INIT}\ell$	$Z_{BP}=0$ $C_{eff}=C^{INIT}\ell^2$	$Z_{BP}=0$ $C_{eff}=C^{INIT}\ell^3+\frac{C^{INIT2}\ell}{K_d}\ell_{BP}^{INIT}$		
2			δ $Z_{BP}=0$ $C_{eff}=c^{INT}$	$Z_{BP}=0$ $C_{eff}=C^{INIT}\ell$	$Z_{BP}=\frac{C^{INIT}\ell^2}{K_d}$ $C_{eff}=C^{INIT}\ell^2+\frac{C^{INIT}_{BP}C^{INIT}\ell}{K_d}$	$Z_{BP}=0$ $C_{eff}=C^{INIT}\ell^3+\frac{C^{INIT}_{BP}C^{INIT}\ell^3}{K_d}$	
3				A $Z_{BP}=0$ $C_{eff}=C^{INIT}$	$Z_{BP}=0$ $C_{eff}=C^{INIT}\ell$	$Z_{BP}=0$ $C_{eff}=C^{INIT}\ell^2$	$Z_{BP}=0$ $C_{eff}=C^{INIT}\ell^3+\frac{C^{INIT}C^{INIT2}\ell}{K_d}\ell^{BP}$

CIRCLE

$C_{eff}=C^{INIT}\left[\ell^4+\frac{\ell^3\ell_{BP}C_{INIT}}{K_d}\right]$

yes → Agrees with myfold.py after bug fixes.

BASE PAIR PROBABILITY

$C_{eff}(j,i)$

$$\underbrace{Z}_{(\text{force BP } i,j)} = Z^{BP}(i,j) \times \left[C_{eff}^{\text{INCLUDING BP's}}(j,i) \times \ell^{BP} \right] / K_d^{\text{ligation}}$$

$$= Z^{BP}(i,j) \times \frac{Z^{BP}(j,i) \times K_d^{BP_a}}{\ell^{BP}} \ell^{BP}$$

How derived? well

$$Z^{BP}(j,i) = \frac{1}{K_d^{BP}} \left[C_{eff}^{\text{loop}} \ell + \sum_{j<k<l<i}^{m(k,l)} C_{eff}(k,l) \right] \equiv \frac{1}{K_d^{BP}} C_{eff}^{BP} \ell$$

i.e. Boltzmann weight for dang pair from j,i side.

Identify

$$C_{eff}^{BP}(j,i) = \frac{Z^{BP}(j,i) K_d^{BP_a}}{\ell}, \text{ i.e.}$$

So:

$$BPP(i,j) = Z(\text{force } BPP_{i,j}) / Z$$

$$= \frac{Z^{BP}(i,j)\, Z^{BP}(j,i) \times K_d^{BP_a} (\ell^{BP}/\ell) / K_d^{\text{lig}}}{C_{eff}^{\text{CIRCLE}} / K_d^{\text{lig}}}$$

$$\boxed{BPP(i,j) = \frac{Z^{BP}(i,j)\, Z^{BP}(j,i)\, K_a^{BP_a} (\ell^{BP}/\ell)}{C_{eff}^{\text{circle}}}}$$

Check:

(see p. 14)

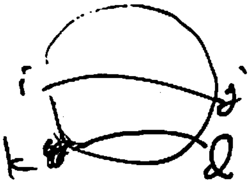
$$BPP(C\text{-}G) = \frac{Z^{BP}(1,2)\, Z^{BP}(2,1)\, K_d^{BP} (\ell^{BP}/\ell)}{C_{eff}^{\text{circle close}}}$$

$$= \frac{\left(\frac{C^{INIT}\ell^2}{K_d^{BP}}\right)\left(\frac{C^{INIT}\ell^2}{K_d^{BP}}\right) K_d (\ell^{BP}/\ell)}{C^{INIT}\ell^4 + \frac{C^{INIT^3} \ell \ell_{BP} C_{INIT}}{K_d}} = \frac{(C^{INIT})^2 \ell^3 \ell_{BP} / K_d^{BP}}{C^{INIT}\ell^4 + \frac{(C^{INIT})^2 \ell^2 \ell_{BP}}{K_d^{BP}}}$$

yes.

The other way to get BPP is through BACKTRACKING.

How does that work here?



$$BPP(i,j) = \frac{Z_{1,i-1} Z^{BP}_{ij} Z_{j+1,N}}{Z_{1,N}} \leftarrow \text{if no external PAIRS}$$

$$+ \sum_{\substack{\text{MOTIF}\\ i,j\to k,l}} \frac{Z^{BP}_{ij}}{Z^{BP}_{kl}} \, \frac{\text{eff}^{\text{MOTIF}}(\cancel{\ldots}\, i,j \to k,l)}{K_d^{BP}} \, \cancel{\ldots}\, BPP(k,l) \leftarrow \text{next external is motif}$$

~ Sierna / McCaskill

$+ \ldots$

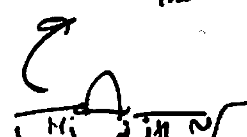
?

$\boxed{\text{Let's do non-circular case}}$.

In McCaskill:

$$BPP(i,j) = \frac{Z_{1,i-1} Z^{BP}_{ij} Z_{j+1,N}}{Z_{1,N}} + \sum_{\substack{\text{MOTIF}\\ i,j\to k,l}} \frac{Z^{BP}_{ij}\, e^{-\Delta G^{\text{MOTIF}}_{ijkl}}}{Z^{BP}_{kl}} BPP(k,l) + \ldots$$

$\frac{Z(\text{structures with } i,j)}{\cancel{\ldots}\, Z(\text{all})}$



$\frac{Z(\text{structures with } i,j \xrightarrow{\text{MOTIF}} k,l)}{Z_{1,N}}$

?

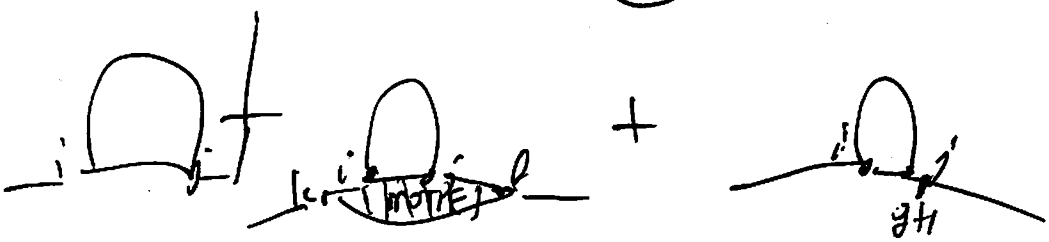
$\frac{1}{Z_{1,N}} Z^{BP}_{ij}\, e^{-\Delta G^{\text{MOTIF}}_{ijkl}} \times \frac{Z(\text{structures w/ } kl)}{Z^{BP}_{kl}}$

$\frac{Z^{BP}_{ij}\, e^{-\Delta G^{\text{MOTIF}}_{ijkl}}}{Z^{BP}_{kl}} BPP(k,l)$

All are mutually exclusive.

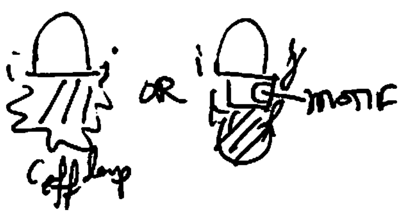
$$+ \sum_{k<i} Z^{BP}_{ij}\, e^{-(a+b)/kT} \left(BPP(k,j)\right)?$$

Hmm



$Z(i,j\ BP)$ no extn

$Z(i,j,\ BP)$ k,l next BP extn



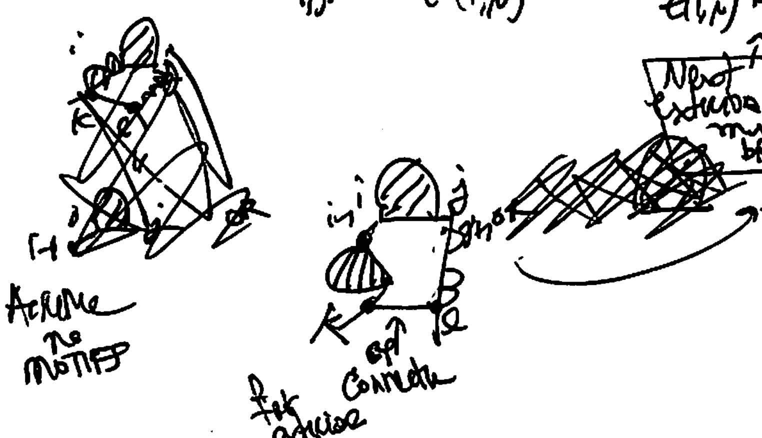
PARTITION fn of structures with extern base pair k, l

structures with no extern base pair

$$BPP(i,j) = \frac{Z(1,i-1)\, Z^{BP}(i,j)\, Z(j,N)}{Z(1,N)} + \frac{1}{Z(1,N)} \sum_{k,l} C^{\times l}_{INIT} \times \frac{C_{eff}(k,i-1)}{C_{INIT}} \times \frac{l_{BP}(j+1,l)}{C_{INIT}} \Big/ K^{BP}_{d} \times \left[\frac{Z(\text{all structures w/ } k,l^{BP})}{Z^{BP}_{kl}}\right]$$

Next extension - must bridge

"external Z" for (k,l)



$$BPP(i,j) = \frac{Z(1,i-1)\ Z^{BP}(i,j)\ Z(j,N)}{Z(1,N)} + \sum_{k,l} C_{init} \times \frac{l_{BP}}{C_{INIT}} \times \frac{C_{eff}(k,i-1)\, C_{eff}(j+1,l)\, Z^{BP}_{ij}}{Z^{BP}_{kl}}$$

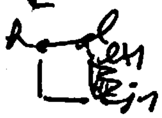
$\times\ BPP(k,l)$

(ignoring MOTIFS for now)

(27B) → SCRATCH WORK

McCaskill of (20)

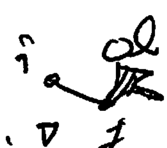
$$P_{kl} = \sum_{ij} P_{ij} \frac{Q^b_{kl}}{Q^b_{ij}} \, e^{-(a+b)/kT} \times \Big[e^{-(k-i-1)c/kT} \, Q^m_{l+1,j-1} + e^{-(j-l-1)c/kT} \, Q^m_{i+1,k-1} + Q^m_{i+1,k-1} Q^m_{l+1,j-1} \Big]$$



ab. -- enforces multiloop, i.e. ≥ 1 BRANCH (Q^M)

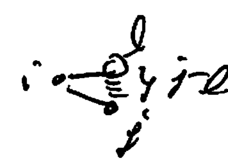
$$P^m_{il} = \sum_{j>l} \frac{P_{ij}}{Q^b_{ij}} Q^m_{l+1,j-1}$$

← note down weight by Q^b_{ij} !

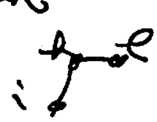


$$P^{m1}_{il} = \sum_{j>l} \frac{P_{ij}}{Q^b_{ij}} e^{-[(j-l-1)c]/kT}$$

← again note special weights.

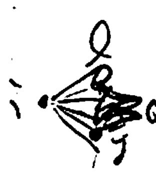


$$P_{kl} = \cdots + \sum_{i<k} Q^b_{kl} \, e^{-(a+b)/kT} \, (\ldots$$



interpretation?

$$P^m_{il} = \sum_{j>l} \frac{Q^m_{l+1,j-1}}{Q(\text{all structs ignoring } i,j \text{ BP})} \cdot \frac{Q(\text{all structs in } l+1, j \text{ BP})}{Q_{1,N}} = \ldots =$$



$Q^m_{l+1,j-1}$ ← c/f?

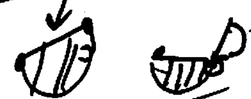
(18)

Need to further simplify to N^3 algorithm.

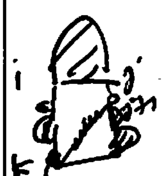
index of first exterior BASE PAIR

$C^{out}_{eff}(k,j+1)$ for all connections of k to j+1, when k is most exterior pair position (j may not be)

$$\cancel{C^{out}_{eff}(k,j) = \ell\, C^{out}_{eff}(k,j+1) + \ell Z^{BP}}$$



$$\cancel{C^{out}_{eff}(k,j) = C^{out}_{eff,BP}(k,j) + C^{out}_{eff,BP}(k,j+1) + \dots + C^{out}_{eff,BP}(k,N)}$$



First do it algebraically:

$$BPP(i,j) = \frac{Z(1,i-1)\, Z^{BP}(i,j)\, Z(j,N)}{Z(1,N)}$$

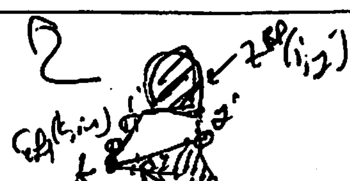
$$+ \sum_k \frac{\ell^2 \ell_{BP}}{C_{init}} C_{eff}(k,i-1)\, Z^{BP}(i,j) \sum_l \frac{C_{eff}(j+1,l)}{Z^{BP}(k,l)} BPP(k,l)$$

DEFINE Accessory Quantity $R(k,j)$:

$$R(k,j) = \sum_l \frac{C_{eff}(j,l)\, BPP(k,l)}{Z^{BP}(k,l)}$$

$$BPP(i,j) = \frac{Z(1,i-1)\, Z^{BP}(i,j)\, Z(j,N)}{Z(1,N)}$$

$$+ \sum_k \frac{\ell^2 \ell_{BP}}{C_{init}} C_{eff}(k,i-1)\, Z^{BP}(i,j)\, R(k,j+1)$$

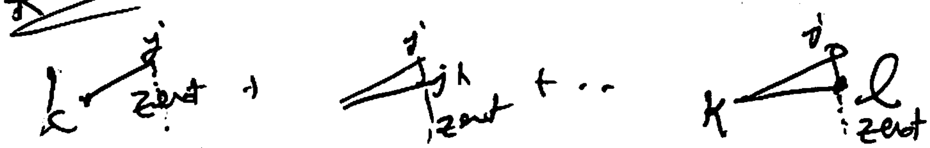


Interpretation of R ?

$$R(k,j) = \sum_{l} \mathrm{Ceff}(j,l) \frac{\mathrm{BPP}(k,l)}{Z^{BP}(k,l)}$$

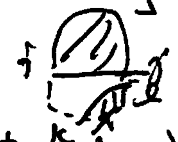
$$= \sum_{l} \mathrm{Ceff}(j,l) \frac{Z(\text{all structures with } \overset{BP}{k,l})}{Z(1,N)} \times \frac{1}{Z^{BP}_{kl}(k,l)}$$

$$= \frac{1}{Z(1,N)} \sum_{l} \left\{ \frac{Z(\text{all structures with } \overset{k,l}{BP})}{\underbrace{Z^{BP}_{kl}(k,l)}_{\text{"exterior } Z\text{" to } (k,l)}} \times \mathrm{Ceff}(j,l) \right\}$$



Let's now generalize BPP (McCaskill) expression [p.18] to circular case

[and, later convert circular expression (p.15) to linear, à la Mathews and Zuker]



$$BPP(i,j) = \frac{Z\left(\text{no "exterior" pairs}\right)}{Z_{total}} + \frac{Z\left(\text{exterior pair beginning at } k\right)}{Z_{total}}$$

$$\frac{\frac{l_{BP}}{l}\, C_{init} \times \frac{C_{eff}(1,i-1)\times l \times Z^{BP}(i,j)}{C_{INIT}} \times l \times \frac{C_{eff}(j+1,N)\times l}{C_{INIT}} \times \frac{1}{K_d^{fig}}}{C_{eff}^{CLOSE} / K_d^{fig}}$$

$$\sum_{k,l} C_{init} \times \frac{C_{eff}(k,i-1)}{C_{INIT}} \times l \times Z^{BP}(i,j) \times \frac{C_{eff}(j+1,l)}{C_{INIT}} \times l \times \left[\frac{Z\left(\text{"exterior" to } k,l\right)}{Z(1,N)}\right]^{BP}$$

$$\downarrow$$

$$BPP(k,l) / Z^{BP}_{(k,l)}$$

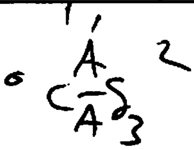
$$BPP(i,j) = \frac{C_{eff}(1,i-1)\, C_{eff}(j+1,N)\, \frac{l^2}{l_{BP}}\, Z^{BP}(i,j)}{C_{INIT}\, C_{eff}^{CLOSE}} + \sum_{k,l} \frac{l^2}{C_{init}}\, \frac{C_{eff}(k,i-1)\, C_{eff}(j+1,l)\; BPP(k,l)}{Z^{BP}(k,l)}$$

$$\boxed{BPP(i,j) = \frac{C_{eff}(1,i-1)\, C_{eff}(j+1,N)\, Z^{BP}(i,j)\, l^2\, l_{BP}}{C_{INIT}\, C_{eff}^{CLOSE}} + \sum_{k} \frac{l^2}{C_{init}}\, C_{eff}(k,i-1)\, R_{(k,j+1)}}$$

edge case
if $i=1$, $C_{eff}(1,i-1) \to C_{INIT}/l$
if $j=N$, $C_{eff}(j+1,N) \to C_{INIT}/l$

where $R(k,j) = \sum_{l} \frac{BPP(k,l)}{Z^{BP}(k,l)}\, C_{eff}(j,l)$

It's pretty hard to see how this is equal to p.15 ?

Sanity check on p(20): 

[Not particularly stringent since no use of "n"]

$$BPP(i,j) = BPP(0,2) = \frac{C_{eff}(0,0)\, C_{eff}(3,3)\, l_{BP}^{2}\, Z^{BP}(0,2)}{C_{INIT}\; C_{eff}^{CLOSE}}$$

$$= \frac{\cancel{C_{INIT}}/l \times \cancel{C_{INIT}}\, l^{2}\, l_{BP}\, C^{INIT}\, l^{2}/K_d}{\cancel{C_{INIT}}\; C^{INIT}\left[l^{4} + l\, l_{BP}\, C_{INIT}/K_d\right]}$$

$$= \frac{(C_{INIT})^{2}\, l^{3} \times l_{BP}/K_d^{BP}}{C^{INIT}\left[l^{4} + l^{3}\, l_{BP}\, C_{INIT}/K_d\right]}$$

O.K.

except probably should re-write to avoid funny business at edges... on just an edge case

$$BPP(i,j) = \frac{C_{eff}(1,i)\; C_{eff}(j+1,N)\; l^{3}\; Z^{BP}(i,j)}{C_{INIT}\; C_{eff}^{CLOSE}}$$

If $i=1$, $C_{eff}(1,i-1) \Rightarrow (\ldots)\left(\frac{1}{l}\right)$

or $j=N$, $C_{eff}(1,j+1) \Rightarrow (\ldots)\left(\frac{1}{l}\right) = \ldots$

⊛ should also work through example:

← will ~~include~~ test "R" auxiliary variable

Last calc...
Work out Z^{BPP} in "circle" style but taking into account chain break.

$$Z^{BP}(i,j) = \frac{1}{K_d^{BPa}} \times C_{\text{eff}}$$

$$= \frac{1}{K_d^{BPa}}\left(C_{\text{eff}}^{\text{Loop}}(i,j-1)\,\ell + \sum_{\text{motifs}} C_{\text{eff}}^{\text{motif}(k,l)} Z^{BP}(k,l) \right)$$

$$C_{\text{eff}}^{\text{Loop}}(i,j) = C_{\text{eff}}^{\text{Loop}}(i,j-1) \times \ell + C_{\text{INT}}\, Z^{BP}(i,j)\,(\ell^{BP}/\ell) + \sum_{i<k\leq j} C_{\text{eff}}^{\text{Loop}}(i,k-1)\, Z^{BP}(k,j)\, \ell^{BP}$$

PROBLEM... how to avoid ~~2~~ BASE PAIRS?

INITIALIZE $\quad C_{\text{eff}}(i \overset{c}{\swarrow}_{\text{CHAINBREAK}},\ i \overset{c}{\swarrow}_{\text{CHAINBREAK}}) \to 0$

add extra case to Z^{BP}...

$$Z^{BP}(i,j) = \frac{1}{K_d^{BP}}\Big(C_{\text{eff}}^{\text{Loop}}(i,j-1)\,\ell + \sum_{\text{motifs}} C_{\text{eff}}^{\text{motif}(k,l-1)} Z^{BP}(k,l)$$

$$+\; Z(i+1,c)\,Z(c+1,j-1)\,C_{\text{STD}} \Big)$$

[a la Pierce et al. - see: $\boxed{\text{Mathews?}}$]

remembering that

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

(28)

$\overset{0}{C}\cdots\overset{1}{G}$

$$Z_{BP}(0,1) = C_{STD}/K_D^{BP}$$

$$C_{eff}^{(0,0)} = C_{eff}(1,1) = 0$$

$$C_{eff}(0,1) = \left[C_{STD}/K_D^{BP}\right] \times C_{int}^{BP}$$

$$Z(0,0) = Z(1,1) = 1 \quad ?$$

$$Z(0,1) = 1 + C_{std}/K_d^{BP}$$

$$Z^{close} = Z(0,1) \rightarrow \boxed{\text{no need to close}}$$
$$= Z(1,0)$$
$$= 1 + C_{std}/K_d^{BP}$$

$\overset{0}{C}\cdots\overset{A^1}{G^2}$ (see also p. 14) note 1

~~Hmm... this is a showstopper!~~

C: $Z_{BP}=0$, $C_{eff}=C^{INIT}$, $Z=1$

CA: $Z_{BP}=0$, $C_{eff}=C^{INIT}\ell$, $Z=1$

A over C G: $Z_{BP}=C^{INIT}\ell^2/K_d$, $C_{eff}=C^{INIT}\ell^2+C^{INIT}\ell\ell\ell_{BP}/K_d$, $Z=1+C^{INIT}\ell^2/K_d$ $\rightarrow$ $Z=1+\frac{C^{INIT}\ell^2}{K_d}$

A over C–G "close":

A: $Z_{BP}=0$, $C_{eff}=C^{INIT}$, $Z=1$

AG: $Z_{BP}=0$, $C_{eff}=C^{INIT}\ell$, $Z=1$

G

$\times$ AG, C–G: $Z_{BP}=C_{STD}/K_d$, $C_{eff}=0+C_{STD}/K_d\,\ell\ell\ell C_{INIT}BP$, $Z=\,?+\,?$

C–G

close: A over C–G, $Z_{BP}=0$

A over C–G

A over C–G close

Rewrite on next page - THIS ACTUALLY WORKS.

$C\ \overset{A}{G} \rightarrow \overset{A}{CG}$ $\qquad Z=\frac{C}{K_D}\times C_{STD}/K_{\ell ig}$

$C\,~~\text{III}~~\,\overset{A}{G} \rightarrow \overset{A}{C\,\text{III}\,G}$ $\qquad Z=\left(\frac{C_{STD}C_{INIT}\ell\ell\ell_{BP}}{K_d\,C_{STD}}\right)\times\frac{\ell}{K^{LIG}}$

$$Z_{AG}=\frac{C_{std}^2}{K_{Dg}}\left[1+\frac{\ell\ell\ell}{K_d}C_{INIT}\right]$$

Hmm...

(24)

FACTORS of C^{STD}/K_{lig} take into account "cost" of chemical BRINGING MONOMERS together

C

$Z_{BP} = 0$

$C_{eff} = C^{INIT}$

$Z = 1$

$C\,A$

$Z_{BP} = 0$

$C_{eff} = C^{INIT}\ell$

$Z = \left(\frac{C^{STD}}{K_{lig}}\right)$

$C \overset{A}{\cdots} g$

$Z_{BP} = \frac{C^{INIT}\ell^2}{K_d}$

$C_{eff} = C^{INIT}\ell^2 + \frac{C^{INIT}\ell\,\ell_{BP}}{K_d}$

$Z = \left[1 + \frac{C^{INIT}\ell^2}{K_d}\right]\left(\frac{C^{STD}}{K_{lig}}\right)$

$\left(C \overset{A}{} g\right)$

THIS FACTOR CORRECTS for fact that above I didn't include floppiness of BP in K_d

A

$Z_{BP} = 0$

$C_{eff} = C^{INIT}$

$Z = 1$

$A\,g$

$Z_{BP} = 0$

$C_{eff} = C^{INIT}\ell$

$Z = \left(\frac{C^{STD}}{K_{lig}}\right)$

$C \overset{A}{\cdots} g$

$Z_{BP} = \frac{C^{std}}{K_d}\left(\frac{\ell}{\ell_{BP}}\right)$

$C_{eff} = \frac{C^{STD}}{K_d}\left(\frac{\ell}{\ell_{BP}}\right) C^{INIT}\ell_{BP}$

$Z = \left(\frac{C^{STD}}{K_{lig}}\right) + \left(\frac{C^{std}}{K_d}\right)\left(\frac{\ell_{BP}}{\ell}\right)\left(\frac{C^{STD}}{K_{lig}}\right)$

$Z \overset{?}{=} \left(\frac{C^{std}}{K_{lig}}\right) + \left(\frac{C^{std}}{K_{lig}}\right)\frac{\ell_{BP}}{\ell K_d}$

g

$Z_{BP} = 0$

$C_{eff} = C^{INIT}$

$Z = 1$

SCRATCH WORK & Resolved Later

$C \cdots g$

$Z_{BP} = \frac{C^{std}}{K_d}\left(\frac{\ell}{\ell_{BP}}\right)$

$C_{eff} = \left[\frac{C^{INIT}\ell_{BP}}{\ell_{BP}}\right]$

$Z = 1 + \frac{C^{std}}{K_d}\left(\frac{\ell}{\ell_{BP}}\right)$

$Z = \left(\frac{C^{std}}{K_{lig}}\right) + \frac{C^{std}\,\ell}{K_d\,\ell_{BP}}$ TO DO

TO DO

$C_{eff} = C^{std} + \left(\frac{C^{std}}{K_d}\right)\left(\frac{\ell}{\ell_{BP}}\right) C^{BP}_{INIT}$

STILL PROBLEMATIC

☆ New Rule for C_{eff}: $C^{std}_{ohmwork}$

$C \cdots g$

$C^{std} + \frac{C^{std}}{K_d}\left(\frac{\ell}{\ell_{BP}}\right) C^{INIT}\left(\frac{\ell}{\ell_{BP}}\right)$

$C \overset{A}{\cdots} g$

$Z_{BP} = 0$

$C_{eff} = C^{std} + \left(\frac{C^{std}}{K_d}\right)\left(\frac{\ell}{\ell_{BP}}\right) C^{INIT}\ell\,\ell_{BP}$

$= C^{std}\left(1 + \frac{\ell^2}{K_d} C^{INIT}\right)$

BUT... DO I NEED ANOTHER $\left(C^{std}/K_{lig}\right)$?

$Z = \frac{C^{LOE}}{Z} = \left(\frac{C^{std}}{K_{lig}}\right)\left(1 + \frac{\ell}{K_d} C^{INIT}\right)$

$\frac{C_{eff}\left(C \overset{A}{\cdots} g\right)_{\ell}}{K_{lig}}$

How to handle chainbreaks?

- Let's go to extreme where every linkage is breakable.
- For cyclic dinucleotide,

$$Z = 1 + 2\alpha + \alpha \cdot \frac{C_{int}\, l^2}{K_d^{lig}}$$

A S → A⌒S, A‿S   A○S

with $\alpha = \frac{C_{STD}}{K_d^{lig}}$ (>>1 if ligation favored)

Ignoring symmetry factors....

$$\to \left(\frac{C_{STD}}{K_d^{lig}}\right)\left(\frac{C_{STD}}{K_d^{lig}}\right)\frac{C_{INIT}\, l^3}{K_d^{lig}}$$ (for closed conformation)

$$= \alpha^3 \left(\frac{C_{INIT}}{C_{STD}}\right) l^3$$

$$Z = 1 + 3\alpha + 3\alpha^2 + \alpha^3 \left(\frac{C_{INIT}}{C_{STD}}\right) l^3$$

Might be easier to rewrite:

$$\alpha^{-3} Z = \underset{3\text{ cuts}}{\alpha^{-3}} + \underset{2\text{ cuts}}{3\alpha^{-2}} + \underset{1\text{ cut}}{3\alpha^{-1}} + \underset{\text{no cuts (circle)}}{\left(\frac{C_{INIT}}{C_{STD}}\right) l^3}$$

OK, to go back to my original convention:

$$\alpha^{-2} Z = \alpha^{-2} + 3\alpha^{-1} + 3 + \frac{C_{INIT}}{K_d^{lig}}\, l^3$$

How could we generate that partition function over cuts?

$$Z_{FREE}(i,j) = (\alpha^{-1}+1) Z_{FREE}(i,j-1) + \alpha^{-1} Z_{continuous}(i,j-1)$$

$$Z_{continuous}(i,j) = Z_{continuous}(i,j-1)$$

$$C_{eff}^{(continuous)}(i,j) = l \, C_{eff}(i,j-1)$$

$$\left[\; \underbrace{Z_{close}(i,j)}_{\text{sealing cut!}} = \frac{C_{eff}(i,i-1)\, l}{K_d^{cis}} \quad \text{at end.} \; \right]$$

Initialize: $Z_{FREE}(i,i) = 0$

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

$C_{eff}(i,j) = C_{INIT}$

Final $Z = Z_{FREE}(i,i-1) + Z_{close}(i,i-1) + \alpha Z_{FREE}(i,i-1) + Z_{continuous}(i,i-1)$

$Z = \alpha Z_{FREE}(i,i) + Z_{close}(i,i)$

Check on simple example A A A

A	A ⋯ A	A ⋯ A, A	~~A A A~~
$Z_{free}=0$	$Z_{free}=\alpha^{-1}$	$Z_{free}=\alpha^{-2}+2\alpha^{-1}$	~~$Z_{free}=\alpha^{-3}+3\alpha^{-2}$~~
$Z_{cont}=1$	$Z_{cont}=1$	$Z_{cont}=1$	~~$Z_{cont}=1$~~
$C_{eff}=C_{INIT}$	$C_{eff}=C_{INIT}\, l$	$C_{eff}=C_{INIT}\, l^2$	~~$C_{eff}=C_{INIT}\, l^3$~~

Nothing new.

$$Z = (1+\alpha) Z_{FREE}(0,2) + Z_{close}(i,i) + Z_{cont}(i,i-1)$$

$$= \alpha^{-2} + 3\alpha^{-2} + 3\alpha^{-1} + C_{INIT} \frac{l^3}{K_d^{loop}},$$

as expected — ~~See next page~~

27

other choice → Allow α to keep ~~[illegible]~~ Ittly toppled...

i ⋯ j-1 j i ⋯ j-1 × j

$$Z_{free}(i,j) = (1+\alpha) Z_{free}(i,j-1) + Z_{continuous}(i,j-1)$$

$$Z_{cat}(i,j) = \alpha Z_{cat}(i,j)$$

$$C_{eff}(i,j) = \ell \alpha C_{eff}(i,j-1)$$

A	A–A	A–A–A	A–A–A–A
$Z_{free} = 0$	$Z_{free} = 1$	$Z_{free} = 1+2\alpha$	$Z_{free} = 1+3\alpha+3\alpha^2$
$Z_{cat} = 1$	$Z_{cat} = \alpha$	$Z_{cat} = \alpha^2$	$Z_{cat} = \alpha^3$
$C_{eff} = Z_{I...}$			

A–A–A–A–A

$$Z_{free} = (1+3\alpha+3\alpha^2)(1+\alpha) + \alpha^3 = 1+4\alpha+6\alpha^2+4\alpha^3$$

$$Z_{cat} = \alpha^4$$

O.K.

$\alpha = C^{std}/K_d^{lig}$

Slight alternate choice that best matches prior work

α^{-1} = cost of breaks

$$Z_{free}(i,j) = (1+\alpha^{-1})\, Z_{free}(i,j-1) + \alpha^{-1} Z_{continuous}(i,j-1)$$

$$Z_{continuous}(i,j) = Z_{continuous}(i,j-1)$$

$$C_{eff}(i,j) = l\, C_{eff}(i,j-1)$$

Initialize:

$Z_{free}(i,i) = 0$

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

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

At end:

$$Z_{final} = Z(i,i) = \text{"}Z_{free}(i,i)\text{"} + \alpha^{-1}\text{"}Z_{close}(i,i)\text{"}$$

$$= (1+\alpha^{-1}) Z_{free}(i,i-1) + \alpha^{-1} Z_{cont}(i,i-1) + \frac{\alpha^{-1} C_{eff}(i,i-1)\, l}{K_d^{lig}}$$

$$= (1+\alpha^{-1})\left[Z_{free}(i,i-1) + \alpha^{-1} Z_{cont}(i,i-1)\right] + \frac{C_{eff}(i,i-1)\, l}{C^{std}}$$

(K_d^{lig} drops out)

Check

A ↔ A ↔ A

$Z_{free} = \alpha^{-2} + 2\alpha^{-1}$

$Z_{ext} = 1$

$C_{eff} = C_{INT}\, l^2$

$$Z_{final} = Z(0,0) = (1+\alpha^{-1})(\alpha^{-2} + 2\alpha^{-1}) + \alpha^{-1} + \frac{C_{INT}\, l^3}{C_{STD}}$$

$$= \alpha^{-3} + 3\alpha^{-2} + 3\alpha^{-1} + \frac{C_{INT}\, l^3}{C_{STD}}$$

O.R.

interesting that Z is scaled so that K_d^{lig} drops out.

Again ignoring <u>symmetry factors</u>.

30

cf. p. 8, p. 28

Write out recursions allowing for chainbreaks with probability $\alpha^{-1} = K_d^{lig}/C^{std}$

$$Z^{BP}_{(i,j)} = \frac{1}{K_d^{BP}} C_{eff}(i,j-1)\,l + \sum_{i<k<j} \frac{l\,C_{init}}{K_d^{BP}} Z_{cut}(i,k) Z_{cut}(k+1,j) \frac{1}{K_d^{BP}}$$

$$C_{eff}(i,j) = \underbrace{C_{eff}(i,j-1)\,l}_{[0 \text{ if cutpoint forced}]} + Z^{BP}(i,j)\, C^{BP}_{INIT} + \sum_{i<k<j} C_{eff}(i,k-1) Z^{BP}(k,j)\, l^{BP}$$

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

$$Z_{cut}(i,j) = (\underbrace{1}_{\text{If } i \text{ not cutpoint}} + \underbrace{\alpha^{-1}}_{\text{If } i \text{ cutpoint}}) Z_{cut}(i,j-1) + \alpha^{-1} \underbrace{Z_{cont}(i,j-1)}_{\text{If } i \text{ cutpoint}}$$

$$\alpha^{-1} Z_{FINAL} = \alpha^{-1} Z(i,i) = \overset{(1+\alpha^{-1})}{Z_{cut}}(i,j-1) + \alpha^{-1} Z_{continuous}(i,i-1) + \frac{C_{eff}(i,i-1)\,l}{C^{std}}$$

Should dominate for cyclic

Standard linear partition function, with α^{-1} factor

OR,

$$Z_{FINAL} = Z(i,i) = \underbrace{Z_{continuous}(i,i-1)}_{\text{If } i \text{ cutpoint}} + \underbrace{(\alpha+1)}_{\text{If } i \text{ not cutpoint}} Z_{cut}(i,i-1) + \frac{C_{eff}(i,i-1)\,l}{K_d^{lig}}$$

If i not cutpoint; [If closure allowed at cutpoint]

(31)

$\alpha \equiv C^{std}/K_d^{lig}$

C

$Z^{BP} = 0$

$C_{eff} = C_{INIT}$

$Z = 1$ continue

$Z_{cut} = 0$

cA

$Z^{BP} = 0$

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

$Z_{cont} = 1$

$Z_{cut} = 0$

c…g (A)

$Z^{BP} = C_{INIT}\, l^2 / K_d^{BP}$

$C_{eff} = C_{INIT}\, l^2 + \frac{C_{INIT} C_{INIT}^{BP} l^2}{K_d^{BP}}$

$Z_{cont} = 1 + C_{INIT}\, l^2 / K_d^{BP}$

$Z_{cut} = 0$

$\Rightarrow Z_{final} = \left(1 + \frac{C_{INIT}\, l^2}{K_d^{BP}}\right) \leftarrow Z_{cont}$

$+ \cancel{\frac{C_{INIT}\, l^2}{K_d^{BP}}\left(1 + \frac{C_{INIT}^{BP}}{K_d^{BP}}\right)} \leftarrow$ closure NOT ALLOWED at cutpoint

A

$Z^{BP} = 0$

$C_{eff} = C_{INIT}$

$Z_{continue} = \alpha$

$Z_{cut} = 0$

A g

$Z^{BP} = 0$

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

$Z_{cont} = \alpha$

$Z_{cut} = 0$

c…g (A)

$Z^{BP} = 0$

$C_{eff} = C_{INIT} \times \frac{l\, C_{std}}{K_d^{BP}} \times l^{BP}$

$Z_{cont} = \alpha + l\, C_{std} / K_d$

$Z_{cut} = \alpha^{-1}\alpha = 1$

$\Rightarrow Z_{final} = Z_{cont}$

$+ \alpha$

$+ \frac{C_{INIT}\, l\, C_{std}\, l}{K_d^{lig} K_d^{BP}}$

$= \text{CLOSE} = \alpha\left(1 + \frac{C_{INIT}\, l^2}{K_d^{BP}}\right)$

g

$Z^{BP} = 0$

$C_{eff} = C_{INIT}$

$Z_{continue} = \alpha$

$Z_{cut} = 0$

c…g

$Z^{BP} = \frac{l\, C_{std}}{K_d^{BP}} \times \left(\frac{1}{l_{BP}}\right)$

$C_{eff} = l\, \frac{C_{std} C_{INIT}^{BP}}{K_d^{BP}} \left(\frac{1}{l_{BP}}\right)$

$Z_{continue} = \alpha + l\, C_{std} / K_d^{BP}$

$Z_{cut} = \alpha^{-1}\alpha = 1$

c…g (A)

$Z^{BP} = 0$

$C_{eff} = l^2\, \frac{C_{std} C_{INIT}^{BP}}{K_d^{BP}} \left(\frac{1}{l_{BP}}\right)$

$Z_{cont} = ?$

$Z_{cut} = 1$

$Z_{final} = \alpha + l^2\, \frac{C_{std} C_{INIT}^{BP}}{K_d^{BP}}\, \frac{l}{l_{BP}\, K_d^{lig}}$

$= \alpha\left[1 + \frac{l^2 C_{INIT}^{BP}\, l}{K_d\, l_{BP}}\right]$

$= \alpha\left[1 + \frac{l^2 C_{int}}{K_d}\right]$

we $C_{INIT}^{BP}\, \frac{l}{l_{BP}} = C_{int}$

TODO:
1. Rewrite Recursion Relations clearly at cutpoints
2. Track down α factor mismatch in $Z_{FINAL}(0)$ vs. $Z_{FINAL}(1)$, $Z_{FINAL}(1)$

Scratch:

A / c g

A / c…g

$\frac{C_{INIT}\, l^2}{K_d}$

A / c…g

$C_{INIT}\, l \left(\frac{C_{std}}{K_d^{BP}}\right) \times \frac{1}{l_{BP}} \times \frac{1}{K_d^{lig}} \times \alpha$

correction

correct for the ref state being already ligated

$= \frac{C_{INIT}\, l^2}{K_d^{BP}}$

Let's allow for each segment to be broken/ligate

A A) $\Rightarrow Z = 1 + C_{INIT}\,l/K_{lig}$

A A A⁀A A A (A A) $\Rightarrow Z = 1 + 2\frac{C_{std}}{K_{lig}} + \frac{C_{std}}{K_{lig}}\left(\frac{C_{INIT}\,l}{K_{lig}}\right)^2$

and if we include BASE PAIRING:

$Z =$

C g 1

C g $+ C_{std}/K_{lig}$

C g $+ C_{std}/K_{lig}$

C g $+ (C_{std}/K_{lig})\left(\frac{C_{INIT}\,l^2}{K_{lig}}\right)$

C‖g $+ \left(C_{std}/K_d^{BP}\right)\left(\frac{l}{l_{BP}}\right)$

C‖g $+ \left(C_{std}/K_{lig}\right)\left(\frac{C_{INIT}\,l}{K_d^{BP}}\right) \rightarrow C_{INIT}\left(\frac{l_{BP}}{l}\right)$

C‖g $+ \left(\frac{C_{std}}{K_d^{BP}}\right)\left(\frac{l}{l_{BP}}\right) \times \frac{\cancel{BP}}{C_{INIT}} \times \frac{l}{K_d}\,lig$

C‖g $\left[\left(\frac{C_{std}}{K_d^{BP}}\right)\left(\frac{l}{K_d^{lig}}\right)C_{INIT}\right]\left[\frac{C_{INIT}\,l}{K_d^{lig}}\right]$

$\left[\frac{C_{std}}{K_d^{BP}}\;\frac{C_{INIT}\,l}{K_d^{lig}}\right]$

$Z = 1 + 2\left(\frac{C_{std}}{K_d^{lig}}\right) + \left(\frac{C_{std}}{K_d^{lig}}\right)\frac{C_{INIT}\,l^2}{K_d^{lig}} + \frac{C_{std}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right) + 2\left(\frac{C_{std}}{K_d^{lig}}\right)\frac{C_{INIT}\,l}{K_d^{BP}} + \left(\frac{C_{std}}{K_d^{lig}}\right)\left[\frac{C_{INIT}^2\,l^2}{K_d^{BP}\,K_d^{lig}}\right]$

$= 1 + \cancel{\ldots}\;\left[\alpha \equiv C_{std}/K_d^{lig}\right]$

$= 1 + 2\alpha + \alpha^2\left(\frac{C_{INIT}\,l^2}{C_{std}}\right) + \frac{C_{std}}{K_d^{BP}}\frac{l}{l_{BP}} + 2\alpha\frac{C_{INIT}\,l}{K_d^{BP}} + \alpha^2\frac{(C_{INIT})^2\,l^2}{C_{std}\,K_d^{BP}}$

(32B)

c^{iii} ← what if 5′ & 3′ "termini" lower chemical potential for ~~[illegible]~~ 3′→5′ connection to $\beta \ll \alpha$, i.e., $K_d^{\text{leg}(3'\text{-}5')} \gg K_d^{\text{leg}}/(\ldots)$

c s c s

$$Z_{\text{full}} = 1 + \alpha + \beta + \underbrace{\left(\frac{c_{\text{std}}}{K_d^{\text{leg}}}\right)\left(\frac{c_{\text{std}}\, l^2}{K_d^{\text{leg}'}}\right)}_{\alpha\beta\left(\frac{c_{\text{INIT}}}{c_{\text{std}}}\, l^2\right)}$$

$$+ \frac{c_{\text{std}}}{K_d^{BP}}\frac{l}{l_{BP}} + \alpha\frac{c_{\text{INIT}}\, l}{K_d^{BP}} + \beta\frac{c_{\text{INIT}}\, l}{K_d^{BP}} + \alpha\beta\left(\frac{c_{\text{INIT}}}{c_{\text{std}}}\right)^2\frac{l^2}{K_d^{BP}}$$

~~Note that ... still ... "divide through" ...~~

Try "divide through" by $\alpha\beta$:

$$(\alpha\beta)^{-1} Z_{\text{full}} = (\alpha\beta)^{-1} + \beta^{-1} + \alpha^{-1} + \frac{c_{\text{INIT}}}{c_{\text{std}}} l^2$$

$$+ (\alpha\beta)^{-1}\frac{c_{\text{std}}}{K_d^{BP}}\frac{l}{l_{BP}} + \beta^{-1}\frac{c_{\text{INIT}}\, l}{K_d^{BP}} + \alpha^{-1}\frac{c_{\text{INIT}}\, l}{K_d^{BP}} + \left(\frac{c_{\text{INIT}}}{c_{\text{std}}}\right)^2\frac{l^2}{K_d^{BP}}$$

Imagine taking linear limit $\alpha \gg 1$, $\beta \ll 1$: c s (linear)

$$(\alpha\beta)^{-1} Z_{\text{full}} = \beta^{-1}\left[1 + \frac{c_{\text{INIT}}\, l}{K_d^{BP}}\right]$$

Imagine taking circular limit $\alpha, \beta \gg 1$:

$$(\alpha\beta)^{-1} Z_{\text{full}} = \left(\frac{c_{\text{INIT}}}{c_{\text{std}}}\right)\left[1 + \frac{c_{\text{INIT}}\, l^2}{K_d^{BP}}\right]$$

$\boxed{\text{Ah, OK!}}$

- For typical calc, I know number of connections, N and BROKEN linkages.
- Decide ahead of time to compute $Z = \boxed{\alpha^{-N} Z_{\text{full}}}$
  If starting at transient chainbreak, suppress by ? } → next page

How to handle transient chainbreaks

$C'''g \Leftrightarrow C'''g$ with $\left\{ \begin{matrix} \alpha \gg 1 \\ \beta \ll 1 \end{matrix} \right\}$

β a "full" partition function allows transient pair/close of each domain.

$$Z \equiv \alpha^{-1} Z_{full}$$

$$= \alpha^{-1} + 1 + \frac{\beta}{\alpha} + \beta\left(\frac{C_{INIT}}{C_{std}} l^2\right)$$

$$+ \alpha^{-1}\frac{C_{std} l}{K_d^{BP} l_{BP}} + \frac{C_{INIT} l}{K_d^{BP}} + \alpha^{-1}\beta\frac{C_{INIT} l}{K_d^{BP}} + \beta\left(\frac{C_{INIT}^2}{C_{std}}\right)\frac{l^2}{K_d^{BP}}$$

$$\rightarrow 1 + \frac{C_{INIT} l}{K_d^{BP}} \quad \left(\text{Taking limits } \begin{matrix}\alpha \gg 1 \\ \beta \gg 1\end{matrix}\right)$$

But let's imagine build up path:

$g \Rightarrow C'''g \xrightarrow{\text{ligate}} C'''g$ Do it all in "full" partition function

$Z_{BP} = 0$	$Z_{BP} = \frac{C_{std}}{K_d^{BP}} \cdot \left(\frac{l}{l_{BP}}\right)$
$C_{eff} = C_{INIT}$	$C_{eff} = C_{INIT} \cancel{\left(\frac{l_{BP} C_{std}}{K_d^{BP} l_{BP}}\right)}$
$Z_{cont} = 1$	$Z_{cont} = \frac{C_{std}}{K_d^{BP}}$
$Z_{cut} = 0$	$Z_{cut} = 1$

$$Z_{full} = \alpha + \frac{C_{std}}{K_d^{BP}} + \frac{C_{INIT} \cdot l}{K_d^{BP}} \cdot \frac{C_{std}}{K_d^{BP}}$$

$$= \alpha + \frac{C_{std}}{K_d^{BP}} + \frac{C_{INIT} l}{K_d^{BP}} \alpha$$

$$\rightarrow \alpha\left[1 + \frac{C_{INIT} l}{K_d^{BP}}\right]$$

$\boxed{\text{O.K.}}$

works but a little complicated —
better to encode simple rules,
rather than track α's

In following $\alpha(i, i+1) = \begin{cases} 0 & \text{if } i \text{ is cutpoint } j \\ \frac{c_{std}}{K_d^{dig}} & \text{if } i \text{ is connection} \end{cases}$

$$Z^{BP}(i,j) = \frac{1}{K_d^{BP}} C_{eff}(i,j-1)\,l + \sum_{i<k<j} \frac{c_{std}}{K_d^{BP}} Z_{cont}(i,k)\, Z_{cont}(k+1,j) \cdot \frac{l}{l_{BP}}$$

$$C_{eff}(i,j) = C_{eff}(i,j-1)\,l + Z^{BP}(i,j)\, C_{INIT}\left(\frac{l_{BP}}{l}\right)$$
$$+ \sum_{i<k<j} C_{eff}(i,k-1)\, Z^{BP}(k,j)\, l^{BP}$$

$$Z_{cont}(i,j) = Z_{cont}(i,j-1)\,\alpha(j-1,j) + Z^{BP}(i,j)$$
$$+ \sum_{i<k<j} Z(i,k-1)\, \cancel{\alpha(\ldots)}\, Z^{BP}(k,j)$$
?

$$Z_{cut}(i,j) = [\alpha(j-1,j)+1]\, Z_{cut}(i,j-1) + Z_{cont}(i,j-1)$$

$$Z_{Full} = Z(i,i) = [\alpha(i-1,i)+1]\, Z_{cut}(i,i-1) + Z_{cont}(i,i-1)$$
$$+ \frac{C_{eff}(i,i-1)\,l}{K_d^{dig}}$$

- Conventionally Report $Z = Z_{Full}(\alpha^{-n}$ connections.)
- Actually, conventionally choose $c_{std} = 1M$, $K_d^{lig} = 1M$ and "stuff cancels"
- work through some examples explicitly
- Read Dirks et al. "interacting strands" carefully.

38

$\to \alpha_i = \begin{cases} 1 & i\cdots i{+}1 \text{ is connected} \\ 0 & i\cdots i{+}1 \text{ is cut} \end{cases}$

Assume cutpoints are pre-defined ... i.e., we're <u>not</u> trying to get partition function summed over states with broken or connected linkages.

(*) Needed?

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

Wrong! Need special handling to ensure i,j not already BASE PAIRED

(*) Consider folding into K_d^{BP}

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

Redundant? If $j-1$ is <u>not</u> chainbreak

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

[Optional: ~~...~~] could be ...

$$Z_{Final} = Z(i,i) = \underbrace{(1-\alpha_i)}_{\text{cutpoint}} Z_{linear}(i,i-1) + \underbrace{(\alpha_i)}_{\text{connection}} \frac{C_{eff}(i,i-1)\, l}{C_{std}} + \alpha_i \left[\sum_{i\le c<j} Z_{linear}(i,c)\, Z_{linear}(c+1,i)^{i-1} \right]$$

or ...

Would be great to work out how to hide stray factors of C_{std}, l/l_{BP} to better match Dirks, Bois, ... & Mathews

C

$Z^{BP} = 0$

$C_{eff} = C^{INIT}$

$Z_{linear} = 1$

C–A

$Z^{BP} = 0$

$C_{eff} = C^{INIT}\ell$

$Z_{linear} = 1$

C...A...δ

C...A...δ

$Z^{BP} = C^{INIT}\ell^2 / K_d^{BP}$

$C_{eff} = C^{INIT}\ell^2 + C^{INIT^2}\dfrac{\ell}{K_d^{BP}}\dfrac{\ell_{BP}}{\ell}$

$Z_{linear} = 1 + C^{INIT}\ell^2 / K_d^{BP}$

C A ⧺ δ

$Z_{final} = Z_{linear} = 1 + C^{INIT}\ell^2 / K_d^{BP}$

A

$Z^{BP} = 0$

$C_{eff} = C^{INIT}$

$Z_{linear} = 1$

A–δ

$Z^{BP} = 0$

$C_{eff} = C^{INIT}\ell$

$Z_{linear} = 1$

+A

C...δ

$Z^{BP} =$ [illegible] 0

$C_{eff} = (C^{INIT}\ell)\left(\dfrac{C_{std}}{K_d^{BP}} \cdot \dfrac{\ell}{\ell_{BP}}\right)\cancel{\ell_{BP}}\cancel{\ell}$

$Z_{linear} = \dfrac{C_{std}}{K_d^{BP}} \cdot \dfrac{\ell}{\ell_{BP}} \cdot \ell_{BP}$

C...A...δ

$Z_{FINAL} = \dfrac{C^{INIT} \cdot \cancel{C_{std}}\ell^2}{\cancel{C_{std}} K_d^{BP}} + 1$

δ

$Z^{BP} = 0$

$C_{eff} = C^{INIT}$

$Z_{linea} = 1$

C...δ

$Z = \dfrac{C_{std}}{K_d^{BP}} \cdot \dfrac{\ell}{\ell_{BP}}$

$C_{eff} = \dfrac{C_{std}\cancel{\ell}}{K_d^{BP}\ell_{BP}} C_{INIT} \cancel{\ell_{BP}}$

$Z_{linear} = Z_{BP} = \dfrac{C_{std} \cdot \ell}{K_d^{BP}\ell_{BP}}$

C...A...δ

$Z^{BP} = 0$

$C_{eff} = \dfrac{C_{std} C_{init}\ell}{K_d^{BP}}$

$Z_{linear} = \dfrac{C_{std}\ell^2}{K_d^{BP}\ell_{BP}}$

C...A...δ

$Z_{FINAL} = \dfrac{C_{INIT}\ell^2}{K_d^{BP}} + 1$

37

O.K.

C
$Z^{BP} = 0$
$C_{eff} = C^{INIT}$
$Z_{linear} = 1$

C–A
$Z^{BP} = 0$
$C_{eff} = C^{INIT}\ell$
$Z_{linear} = 1$

C…A…δ
$Z^{BP} = C^{INIT}\ell^2 / K_d^{BP}$
$C_{eff} = C^{INIT}\ell + \frac{C^{INIT}\ell^2 \ell_{BP}}{K_d^{BP}\,\ell}$
$Z_{linear} = 1 + \frac{C^{INIT}\ell^2}{K_d^{BP}}$

C…A…δ A
$Z^{BP} = 0$
$C_{eff} = \ell\left[C^{INIT}\ell + \frac{C^{INIT}\ell^2 \ell_{BP}}{K_d^{BP}}\right]$
$Z_{linear} = 1 + \frac{C^{INIT}\ell^2}{K_d^{BP}}$

C…A…δ A

C…A…δ A

$\rightarrow Z_{FINAL} = Z_{linear} = 1 + C^{INIT}\ell^2 / K_d^{BP}$

A
$Z_{BP} = 0$
$C_{eff} = C^{INIT}$
$Z_{linear} = 1$

A–δ
$Z_{BP} = 0$
$C_{eff} = C^{INIT}\ell$
$Z_{linear} = 1$

A–δ A
$Z_{BP} = 0$
$C_{eff} = C^{INIT}\ell^2$
$Z_{linear} = 1$

C…A…δ ×A
$Z_{BP} = 0$
$C_{eff} = C^{INIT} \times \left(\frac{C_{std}\,\ell}{K_d^{BP}\,\ell_{BP}}\right)$
$Z_{linear} = \frac{C_{std}\,\ell}{K_d^{BP}\,\ell_{BP}}$

C…A…δ ×A
$Z_{final} = \frac{C^{INIT}\cancel{C_{std}}\,\ell}{K_d^{BP}\,\ell\,\cancel{C_{std}}} + 1 = 1 + \frac{C^{INIT}\ell^2}{K_d^{BP}}$

(38)

δ
$Z_{BP} = 0$
$C_{eff} = C^{INIT}$
$Z_{linear} = 1$

A–δ
$Z_{BP} = 0$
$C_{eff} = C^{INIT}\ell$
$Z_{linear} = 1$

C…δ ×A
$Z_{BP} = \frac{C_{std}\,\ell}{K_d^{BP}\,\ell_{BP}}$
$C_{eff} = \frac{C_{std}\,C_{INIT}}{K_d^{BP}}\ell$
$Z_{linear} = \frac{C_{std}\,\ell}{K_d^{BP}\,\ell_{BP}}$

C…A…δ A
$Z_{BP} = 0$
$C_{eff} = \ell\,\frac{C_{std}\,C_{INIT}}{K_d^{BP}}$
$Z_{linear} = \frac{C_{std}\,\ell}{K_d^{BP}\,\ell_{BP}}$

C…A…δ A
$Z_{final} = \ell\,\frac{C_{std}\,C_{INIT}}{K_d^{BP}}\,\frac{\ell}{C_{std}} + 1 = 1 + \frac{C^{INIT}\ell^2}{K_d^{BP}}$

A
$Z_{BP} = 0$
$C_{eff} = C^{INIT}$
$Z_{linear} = 1$

C×A
$Z_{BP} = 0$
$C_{eff} = 0$
$Z_{linear} = 0$

C×A A
$Z_{BP} = 0$
$C_{eff} = 0$
$Z_{linear} = 0$

C…A…δ ×A
$Z^{BP} = 0$
$C_{eff} = 0$
$Z_{linear} = 0$

C…A…δ A
$Z_{final} = 1 + \frac{C^{INIT}\ell^2}{K_d^{BP}}$

SCRATCH WORK

(39)

$z^{BP} \Rightarrow 0$

$C_{eff} = C_{INIT}$

$z^{linear} = 1$

$z^{BP} = \frac{C_{std}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)$

$C_{eff} = C_{std} / K_d^{BP} C_{INIT}$

$z^{BP} \Rightarrow$

$C_{eff} = C_{INIT}$

$= 1$

z^{linear}

$z^{BP} = \frac{C_{std}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)$

$C_{eff} = C_{std} / K_d^{BP} C_{INIT}$

$z^{linear} = z^{BP} \frac{C_{std}\, l}{K_d^{BP}\, l_{BP}}$

$z_{final} = \left(1 + \frac{C_{std}\, l}{K_d^{BP}\, l_{BP}}\right)$

actually do not include dissociated strands?

$z_{BP} \Rightarrow$

$C_{eff} = C^{INIT}$

$z^{linear} = 1$

$z_{BP} \Rightarrow$

$C_{eff} = C^{INIT}\, l$

$z^{linear} = 1$

$z_{BP} = 0$

$C_{eff} = C \frac{C_{std}}{K_d^{BP}} \left(\frac{l}{l_{BP}}\right)$

$z^{linear} = \frac{C_{std}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)$

$z^{BP} \Rightarrow 0$

$C_{eff} = C^{INIT}$

$z^{linear} = 1$

$z^{BP} = \frac{C_{std}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)$

$C_{eff} = C_{INIT}$

$z^{linear} = \frac{C_{std}\, l}{K_d^{BP}\, l_{BP}}$

$z^{BP} \Rightarrow$

$C_{eff} = \frac{C_{std}}{K_d^{BP} l_{BP}}$

$z^{linear} = \frac{C_{std}\, l}{K_d^{BP}\, l_{BP}}$

$z^{BP} = 0$

$C_{eff} =$

$z^{linear} = \frac{C_{std}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)\left[\frac{C_{std}}{K_d}\frac{l}{l_{BP}}\right]$

$\Downarrow$

$z^{final}_{complex} = \left(\frac{C_{std}\, l^2}{K_d^{BP^2}\, l_{BP}^2}\right) \cdot \frac{1}{2}$

Symm Factor $\left(\frac{1}{2}\right)$ $+ 2 \times \frac{C_{std}\, l}{K_d\, l_{BP}} + 1$

SEPARATE complex to complex

oh! whoa - made some kind of mistake.

δ

$Z^{BP}=0$

$C_{eff}=C_{INIT}$

$Z^{linear}=1$

$Z^{BP}=0$

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

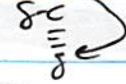
$Z^{linear}=1$

c

$Z^{BP}=0$

$C_{eff}=C_{INIT}$

$Z^{linear}=1$



$Z^{BP}=0$

$C_{eff}= C^{INIT} \cancel{x} \frac{C^{std}}{K_d^{BP}}\frac{l}{l_{BP}} = \frac{C^{std}C^{INIT}l}{K_d^{BP}}$

$Z^{linear} = \frac{C^{std}\,l}{K_d^{BP}\,l_{BP}}$

$Z^{BP}=0 \quad \frac{C^{std}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)$

$C_{eff}= \frac{C^{std}}{K_d^{BP}}C_{INIT}$

$Z^{linear}= \frac{C^{std}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)$

$Z^{BP} = \frac{C^{std}C^{INIT}l}{K_d^{BP}}\frac{l}{l_{BP}} + \frac{C^{std}}{K_d^{BP}}\frac{l}{l_{BP}}$

$C_{eff} = \frac{C^{std}C^{INIT}l^2}{K_d^{BP}} + \frac{C^{std}C^{INIT}l^2}{K_d^{BP\,2}}$

$Z_{linear} = \frac{C^{std}}{K_d^{BP}}\frac{l}{l_{BP}} + \frac{C^{std}C^{INIT}l^2}{K_d^{BP\,2}} + \frac{C^{std}}{K_d^{BP}}\frac{l}{l_{BP}}$

$Z_{FINAL} = Z_{linear}$

$= \frac{C^{std}}{K_d^{BP}}\left[\frac{l}{l_{BP}} + \frac{l^2}{l_{BP}}\frac{C^{INIT}}{K_d^{BP}}\right]$

41

$\delta \because \frac{c}{c}$

g

$Z_{BP} = 0$

$c_{eff} = C_{init}$

$Z_{linear} = 1$

$\delta \because c$

$Z_{BP} = \frac{C_{std}}{K_d^{BP}} \left(\frac{l}{l_{BP}} \right)$

$C_{eff} = \frac{C_{std}}{K_d^{BP}} C_{init}$

$Z_{linear} = \frac{C_{std}}{K_d^{BP}} \left(\frac{l}{l_{BP}} \right)$

$\delta \because \frac{a}{b}$ How to enforce mutually exclusive?

$Z_{BP} = \frac{C_{std}}{K_d^{BP}} \left(\frac{l}{l_{BP}} \right) \sim_{BP} b$

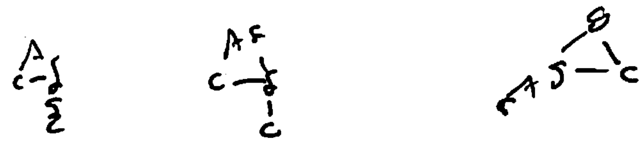
$C_{eff} = \frac{C_{std}}{K_d^{BP}} C_{init}\, l + \frac{C_{std}}{K_d^{BP}} \left(\frac{l}{l_{BP}} \right) C_{init} \left(\frac{l_{BP}}{l} \right)$

$Z_{linear} = \frac{C_{std}}{K_d^{BP}} \left(\frac{l}{l_{BP}} \right) + \frac{C_{std}}{K_d^{BP}} \left(\frac{l}{l_{BP}} \right)$

41

CAδgc

CAδδC



$$Z_{\text{final}} = 1 + \frac{C_{INIT}\,l^2}{K_d^{BP}} + \frac{C_{INIT}\,l^2}{K_d^{BP}} + \frac{C_{INIT}\,l^2}{K_d^{BP}}$$

?

$\overrightarrow{A\delta g}$ $\quad$ $A\delta\ldots C$

$Z^{BP} = 0$

$C_{eff} = l^{2}$

$Z_{linear} = 1$

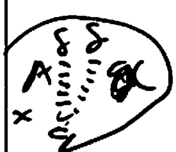
$Z^{BP} = 0$

$C_{eff} = l^3 + C_{INIT}\left(\frac{C_{INIT}\,l^2}{K_d^{BP}}\right) l^{BP}$

$Z_{linear} = 1 + \left(\frac{C_{INIT}\,l^2}{K_d^{BP}}\right)$

$$\left[\ \delta \text{ inc} \quad Z_{BP} = \frac{C_{\delta\,t\,\delta}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)\right]$$

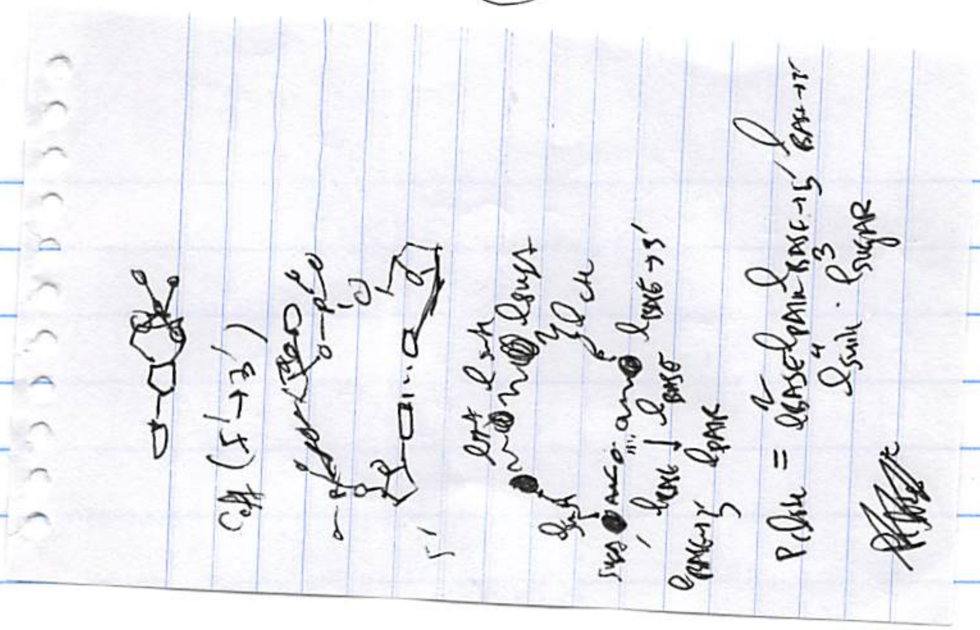
$$\left[\begin{array}{l} \delta \cdots c \\ Z_{BP} = \dfrac{C_{INIT}\,l^2}{K_d^{BP}} \\ C_{eff} = C^{INIT}\,l^2 + \left(\dfrac{C^{INIT}\,l^2}{K_d^{BP}}\right) \times \left(C_{INIT}\,l^{B}/l\right) \\ Z_{linear} = 1 + \dfrac{C_{INIT}\,l^2}{K_d^{BP}} \end{array}\right]$$



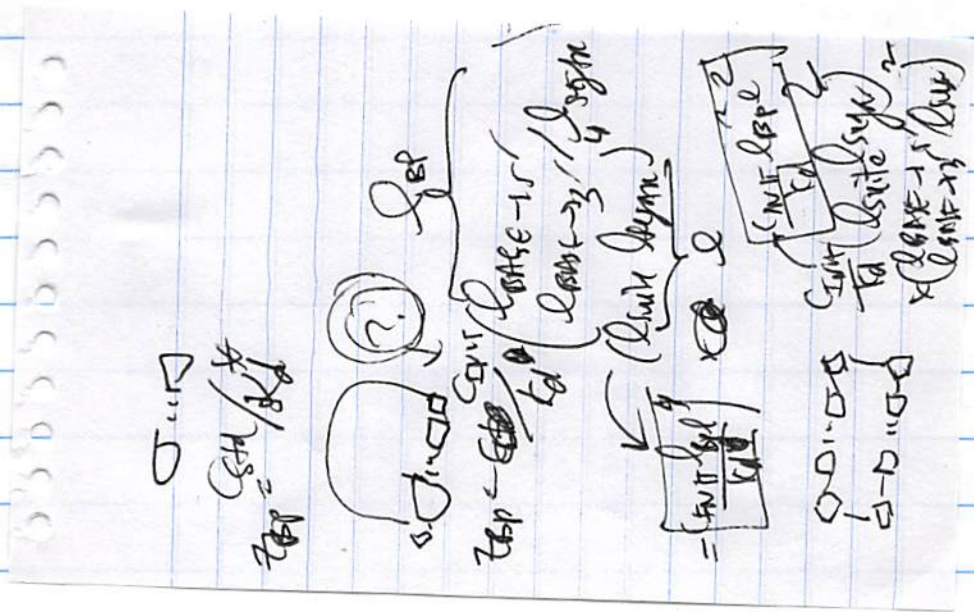
$Z^{BP} = 0$

$$C_{eff} = \underbrace{\frac{C_{\delta\,t\,\delta}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)}_{Z^{BP}[(\cdots)]} l^{BP} \times [C_{INIT}][l + l^2]$$

$$Z_{linear} = 2 \times \frac{C_{\delta\,t\,\delta}}{K_d^{BP}}\left(\frac{l}{l_{BP}}\right)$$



Simpler way to deal with ℓ/ℓ_{BP}?

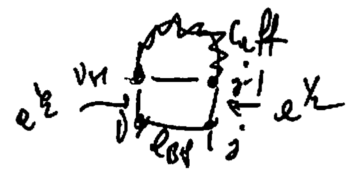


(44)

potentially simpler recursion relations $\quad \frac{f_{BP}}{K_d} = \frac{f_{BP}}{\ell} \frac{f}{K_d^{BP}}$

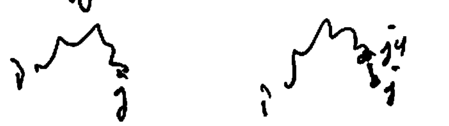
See p.(36), But substitute: $K_d^{BP*} = K_d^{BP}\, l_{BP}/\ell$]
"true" K_d if BASES are floating around.
$C^*_{init} = c_{init}\,\ell$

$$Z^{BP}(i,j) = \underbrace{\alpha_{j-1}}_{\text{connected at } j-1} \frac{1}{K_d^{BP*}} C_{eff}(i+1, j-1)\, \underbrace{\ell\, l_{BP}}_{\ell^{1/2}\ell^{1/2}} + \sum_{i \le c < j} \frac{C_{std}}{K_d^{BP*}} Z_{linear}(i+1, c) Z_{linear}(c+1, j-1)$$



define
$Z_{linear}(i+1, i) = 1$
$Z_{linear}(j-1, j) = 1$

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



define $C_{eff}(i,i-1) \equiv C^{BP}_{INIT}/\ell$
$= C_{INIT}\, l_{BP}/\ell^2 = C^*_{init}\left(\frac{l_{BP}}{\ell}\right)$ (?)

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

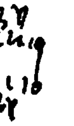
define
$Z_{linear}(i,i-1) = 1$

Initialize:
$Z_{linear}(i,i) = 1$
$C_{eff}(i,i) = C_{init} = C^*_{init}/\ell$

Finalize: $Z^{(i,j)}_{final} = \begin{cases} C_{eff}(i, i-1)\frac{\ell}{K_d^{BP}} \times \left(\frac{K_d\, l_{BP}}{C_{init}}\right) + \sum_{i \le c < j} Z_{linear}(i,c) Z_{linear}(c+1, j-1) & \text{if connected at } i \\ Z_{final}(i, i+1) & \text{if not } i \end{cases}$ (converting)

checks:

$$Z = \frac{C_{INIT}\,\ell^4}{K_d^{BP}} = \frac{C_{INIT}\,\ell^2 l_{BP}}{K_d^{BP*}} = \frac{C^*_{INIT}\, \ell^4 l_{BP}}{K_d^{BP*}} \quad \text{OK?}$$ not intuitive.



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

Write out original/current alphafold recursions explicitly

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

$\hookrightarrow Z_{linear}(i,i-1) \equiv 1.$

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

$\hookrightarrow C_{eff}(i,i-1) \equiv C_{init}\,\ell_{BP}/\ell^2$

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

$$Z_{final}(i,i) = \begin{cases} C_{eff}(i,i-1)\,\ell/C^{std} + \sum_{i\le c<i+N-1} Z_{linear}(i,c)\, Z_{linear}(c+1,i-1) \\ \quad \text{if } \alpha_i\ [i \cdots i+1 \text{ connected}] \\ Z_{linear}(i,i-1) \quad \text{if cut at } i \end{cases}$$

Initialize:

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

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

$Z^{BP}(i,i) = 0$

How do we take a derivative? e.g.,

$$\frac{\partial \log Z_{final}(i,i)}{\partial \log K_d^{BP}} = \frac{\partial \log Z_{linear}(i,i-1)}{\partial \log K_d^{BP}} \quad \text{if cut at } i$$

$$= \frac{\partial}{\partial \log K_d^{BP}} \log \Big[\alpha_{j-1} Z_{linear}(i, i-2) + \sum_{i \le k \le j} Z_{linear}(i,k-1)\, Z^{BP}(k,j) \Big]$$

$$= \cancel{\frac{\partial}{\partial \log K_d^{BP}}} \frac{\alpha_{j-1} \frac{\partial}{\partial \log K_d^{BP}} Z_{linear}(i,i-2) + \sum_{i \le k \le j} \Big[\frac{\partial Z_{linear}(i,k-1)}{\partial \log K_d^{BP}} Z^{BP}(k,j) + Z_{linear}(i,k-1) \frac{\partial Z^{BP}(k,j)}{\partial \log K_d^{BP}} \Big]}{Z_{linear}(i,i-1)}$$

$$\cancel{\boxed{\alpha_{j-1} \frac{C_{eff}(i+1,j-1)}{K_d^{BP}} \ell^2}}.$$

$$\frac{\partial Z^{BP}(k,j)}{\partial \log K_d^{BP}} = - Z^{BP}(k,j) + \frac{\alpha_{j-1}}{K_d^{BP}} \frac{\partial C_{eff}(i+1,j-1)\ell^2}{\partial \log K_d}$$

$$+ \sum_{i \le c \le j} \frac{C_{std}}{K_d^{BP}} \frac{\ell}{\ell_{BP}} \Big[\frac{\partial}{\partial \log K_d} Z_{linear}(i+1,c)\, Z_{linear}(c+1,j-1) + Z_{linear}(i+1,c) \frac{\partial}{\partial \log K_d} Z_{linear}(c+1,j-1) \Big]$$

...

Hmm...

Call: $\log K_d^{BP} \equiv g$ (free energy of formation)

Well let's FORWARD PROPAGATE (not backpropagate derivs)

$$\frac{\partial}{\partial g} Z^{BP}_{(i,j)} = -Z^{BP}(i,j) + \alpha_{j+1}\frac{C_{eff}}{K_d^{BP}}\,\frac{\partial C_{eff}(i+1,j+1)}{K_d^{BP}} + \sum \frac{C_{eff}}{K_d^{BP}}\frac{1}{e_{ff}} \times \left[\frac{\partial Z_{linear}(i+1,i)}{\partial g} Z_{linear}(i+1,j) + Z_{linear}(i+1,i)\frac{\partial}{\partial g} Z_{linear}(i+1,j)\right]$$

Could just code these up....

- PROPAGATE FORWARD AT THE SAME TIME as Z^{BP}, C_{eff}, Z_{linear}. ← very similar to Rosetta.

- Wait, why was I thinking about backpropagation?

In neural nets

$$y = f(Wx + b)$$

$$z = f(W'y + b')$$

Then to get derivs w.r.t. b: $\text{Loss} = (z - z_0)^2$, for example

$$\frac{\partial \text{Loss}}{\partial b_i} = 2(z - z_0)\frac{\partial z}{\partial b_i}$$

$$\frac{\partial z}{\partial b_i} = \frac{\partial f}{\partial q}\cdot W' \frac{\partial y}{\partial b_i}$$

$$\frac{\partial y}{\partial b_i} = \left(\frac{\partial f}{\partial q}\right)_i$$

Oh... in neural nets there are a bajillion PARAMS. So major issue for "just" memory, then.

Cool...

- let's code up $\frac{\partial Z}{\partial K_d^{BP}}$.

- ~~should~~ Following give a nice cross-check:

$$\frac{\partial \log Z}{\partial \log K_d^{BP}} = \frac{1}{Z}\frac{\partial Z}{\partial \log K_d^{BP}} = \left(\frac{K_d^{BP}}{Z}\right)\left(\frac{\partial Z}{\partial K_d^{BP}}\right)$$

Compare to:

$$\frac{\partial}{\partial \log K_d^{BP}} \log Z = \frac{1}{Z}\frac{\partial Z}{\partial \log K_d^{BP}} = \frac{1}{Z}\sum_{\text{structures}} (-\#\text{BASEPAIRS})\, e^{-\Delta G}$$

$$Z = \sum_{\text{structures}} e^{-\Delta G(\text{structure})}$$

$$= -\langle \#\text{BASEPAIRS} \rangle$$

$$= -\sum_{i,j} BPP(i,j)$$

$$\Delta G(\text{structure}) = \cdots (\#\text{BASEPAIRS}) \times \log K_d^{BP} \cdots$$