Supplementary Information for "CplexA: a Mathematica package to study macromolecular-assembly control of gene expression"

Jose M. G. Vilar [1,2] *and Leonor Saiz* [3]

[1]*Biophysics Unit (CSIC-UPV/EHU) and Department of Biochemistry and Molecular Biology, University of the Basque Country, P.O. Box 644, 48080 Bilbao, Spain*

[2]*IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain*

[3]*Department of Biomedical Engineering, University of California, 451 East Health Sciences Drive, Davis, CA 95616, USA*

In[1]:=
```
SetDirectory@NotebookDirectory[];
```

In[2]:=
```
<< CplexA`
```

Detailed calculations for Figure 1

In[3]:=
```
(* Experimental data for the transcriptional activity
 of the PRM and PR promoters of phage lambda from ref. Dodd,
I.B.,Shearwin,K.E.,Perkins,A.J.,Burr,T.,Hochschild, A.and Egan,
J.B.(2004) Cooperativity in longrange gene regulation by the lambda CI repressor,
Genes Dev,18,344-354 *)
PRMdata := {{0.058, 66.367}, {0.373, 189.358}, {0.658, 214.233}, {0.922, 188.897},
   {1.168, 173.696}, {1.385, 155.271}, {1.875, 133.160}, {2.281, 136.384}};
PRdata := {{0.063, 1057.627}, {0.373, 585.763}, {0.658, 204.746}, {0.921, 70.508},
   {1.171, 27.119}, {1.384, 18.983}, {1.875, 12.203}, {2.278, 9.492}};
```

In[5]:=
```
(* Plot of the experimental data to use later on *)
plotPRM := ListPlot[PRMdata, PlotMarkers → {Automatic, 17}, PlotRange → {{0, 2.5}, {0, 250}},
   ImageSize → 300, BaseStyle → {FontFamily → "Helvetica", FontSize → 12}];
plotPR := ListPlot[PRdata, PlotMarkers → {Automatic, 17}, PlotRange → {{0, 2.5}, {0, 1200}},
   ImageSize → 300, BaseStyle → {FontFamily → "Helvetica", FontSize → 12}];
```

In[7]:=
```
(* Computes the CI dimer concentration as a function of CI monomer concentrations *)
NF[nn_] := 9.38419 * 10^-14 + 7.0922 * 10^-10 nn - 1.15373 * 10^-11 Sqrt[0.0000661586 + 1. nn]
```

In[8]:=
```
(* State variables *)
S := {sOL1, sOL2, sOL3, sOR1, sOR2, sOR3, sL}
```

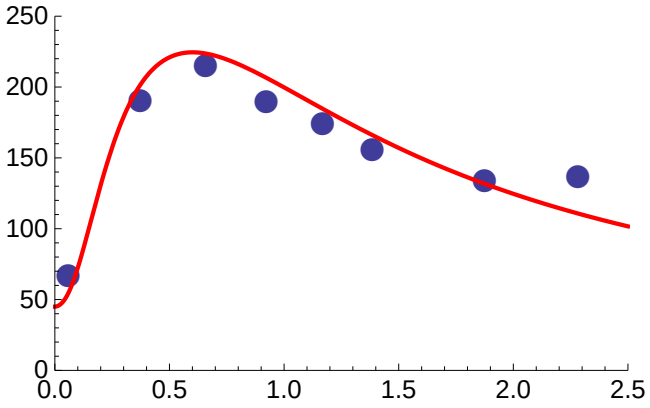
In[9]:=
```
(* Free energy *)
ΔGλ := sOR1 (-12.7 - RT Log[n]) + sOR2 (-10.7 - RT Log[n]) + sOR3 (-10.2 - RT Log[n]) -
   3 sOR1 sOR2 - 3 sOR2 sOR3 + 3 sOR1 sOR2 sOR3 + sOL1 (-13.8 - RT Log[n]) +
   sOL2 (-12.1 - RT Log[n]) + sOL3 (-12.4 - RT Log[n]) - 2.5 sOL1 sOL2 - 2.5 sOL2 sOL3 +
   2.5 sOL1 sOL2 sOL3 + sL (21 - 21.2 sOL1 sOL2 sOR1 sOR2 - 3 sOL3 sOR3);
```

In[10]:=
```
(* Transcriptional activity of the PRM promoter *)
rrm := (1 - sOR3) (45 + ((240 - 45) sL + (460 - 45) (1 - sL)) sOR2);
```

In[11]:=
```
(* Average transcriptional activity of the PRM promoter *)
Arrm = AvConf[rrm, ΔGλ, S];
```

In[12]:=
```
(* Computed average transcriptional activity of the PRM
 promoter (red line) and its experimental counterpart (blue circles)
 as a function of the normalized CI dimer concentration *)
Show[plotPRM, Plot[Arrm /. {RT → 0.6, n → NF[nn]}, {nn, 0, 2.5}, PlotStyle → {Red, Thick}]]
```

Out[12]=

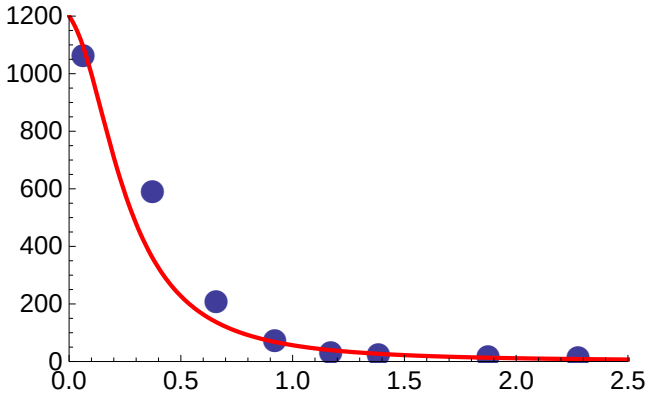


In[13]:=
```
(* Transcriptional activity of the PR promoter *)
rr := (1 - sOR1) 1200;
```

In[14]:=
```
(* Average transcriptional activity of the PR promoter *)
Arr = AvConf[rr, ΔGλ, S];
```

In[15]:=
```
(* Computed average transcriptional activity of the PR
 promoter (red line) and its experimental counterpart (blue circles)
 as a function of the normalized CI dimer concentration *)
Show[plotPR, Plot[Arr /. {RT → 0.6, n → NF[nn]},
  {nn, 0, 2.5}, PlotStyle → {Red, Thick}, PlotRange → All]]
```

Out[15]=



Sum over an exponentially large number of states in sub-exponential times

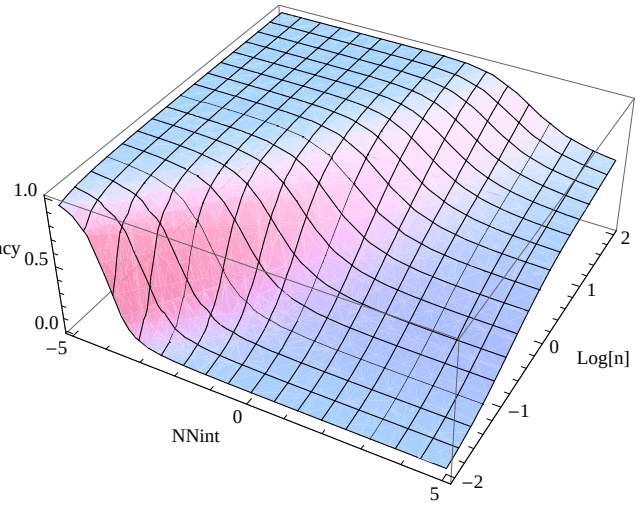
In[16]:=
```
(* Definition of a linear array of binding sites with next-
 neighbor interactions for molecules with concentration
  n. The function NNChain returns a list with the state variables,
the free energy of the chain, the average occupancy,
and the time needed for the computation. NumSites is the number of sites;
NNint is the energy of the next-neighbor interaction between bound molecules;
the free energy of binding is -1; and the thermal energy is 1. *)
NNChain[NumSites_, NNint_] := Module[{S, ΔGB, res},
  S = Table[ToExpression["s" <> ToString[i]], {i, 1, NumSites}];
  ΔGB = Simplify[Sum[(-1 - Log[n]) S[[i]], {i, 1, NumSites}] +
      NNint Sum[S[[i]] S[[i + 1]], {i, 1, NumSites - 1}]];
  res = Timing[AvConf[Mean[S], Simplify[ΔGB], S, 1]]; {S, ΔGB, res[[2]], res[[1]]}]
```

In[17]:=
```
(* Analytical expression of the average
 occupancy of a linear array of 10 binding sites *)
NNChain10 = NNChain[10, NNint][[3]]
```

Out[17]=
$$\left(e^{n}\left(5\,e^{9\,NNint}+9\,e^{1+8\,NNint}\,n+36\,e^{1+9\,NNint}\,n+12\,e^{2+7\,NNint}\,n^2+84\,e^{2+8\,NNint}\,n^2+84\,e^{2+9\,NNint}\,n^2+14\,e^{3+6\,NNint}\,n^3+126\,e^{3+7\,NNint}\,n^3+210\,e^{3+8\,NNint}\,n^3+70\,e^{3+9\,NNint}\,n^3+15\,e^{4+5\,NNint}\,n^4+150\,e^{4+6\,NNint}\,n^4+300\,e^{4+7\,NNint}\,n^4+150\,e^{4+8\,NNint}\,n^4+15\,e^{4+9\,NNint}\,n^4+15\,e^{5+4\,NNint}\,n^5+150\,e^{5+5\,NNint}\,n^5+300\,e^{5+6\,NNint}\,n^5+150\,e^{5+7\,NNint}\,n^5+15\,e^{5+8\,NNint}\,n^5+14\,e^{6+3\,NNint}\,n^6+126\,e^{6+4\,NNint}\,n^6+210\,e^{6+5\,NNint}\,n^6+70\,e^{6+6\,NNint}\,n^6+12\,e^{7+2\,NNint}\,n^7+84\,e^{7+3\,NNint}\,n^7+84\,e^{7+4\,NNint}\,n^7+9\,e^{8+NNint}\,n^8+36\,e^{8+2\,NNint}\,n^8+5\,e^{9}\,n^9\right)\right)\Big/\left(5\left(e^{9\,NNint}+10\,e^{1+9\,NNint}\,n+9\,e^{2+8\,NNint}\,n^2+36\,e^{2+9\,NNint}\,n^2+8\,e^{3+7\,NNint}\,n^3+56\,e^{3+8\,NNint}\,n^3+56\,e^{3+9\,NNint}\,n^3+7\,e^{4+6\,NNint}\,n^4+63\,e^{4+7\,NNint}\,n^4+105\,e^{4+8\,NNint}\,n^4+35\,e^{4+9\,NNint}\,n^4+6\,e^{5+5\,NNint}\,n^5+60\,e^{5+6\,NNint}\,n^5+120\,e^{5+7\,NNint}\,n^5+60\,e^{5+8\,NNint}\,n^5+6\,e^{5+9\,NNint}\,n^5+5\,e^{6+4\,NNint}\,n^6+50\,e^{6+5\,NNint}\,n^6+100\,e^{6+6\,NNint}\,n^6+50\,e^{6+7\,NNint}\,n^6+5\,e^{6+8\,NNint}\,n^6+4\,e^{7+3\,NNint}\,n^7+36\,e^{7+4\,NNint}\,n^7+60\,e^{7+5\,NNint}\,n^7+20\,e^{7+6\,NNint}\,n^7+3\,e^{8+2\,NNint}\,n^8+21\,e^{8+3\,NNint}\,n^8+2\,e^{9+NNint}\,n^9+8\,e^{9+2\,NNint}\,n^9+e^{10}\,n^{10}\right)\right)$$

In[18]:=
```
(* Plot of the analytical expression of the average
 occupancy of a linear array of 10 binding sites as a function of
 the energy of interaction and the logarithm of the concentration *)
Plot3D[NNChain10 /. n → 10^x, {NNint, -5, 5}, {x, -2, 2},
 AxesLabel → {"NNint", "Log[n]", "Occupancy", "", ""}]
```

Out[18]=

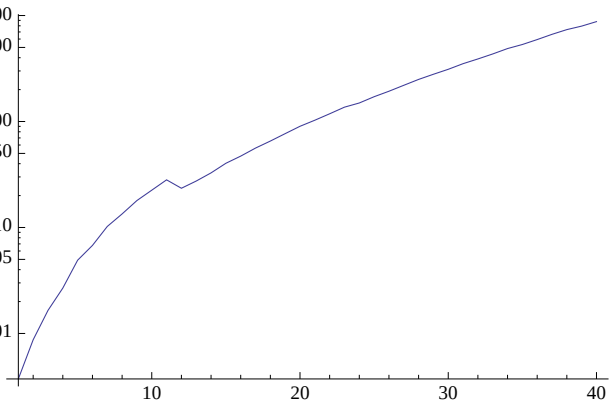


In[19]:=
```
TimesTab = Table[{NumSites, NNChain[NumSites, -1][[4]]}, {NumSites, 1, 40}]
```

Out[19]=
```
{{1, 0.003723}, {2, 0.00872}, {3, 0.016504}, {4, 0.026798}, {5, 0.049121}, {6, 0.067722},
 {7, 0.102238}, {8, 0.134266}, {9, 0.179601}, {10, 0.224848}, {11, 0.281501},
 {12, 0.234544}, {13, 0.274352}, {14, 0.327549}, {15, 0.403452}, {16, 0.471403},
 {17, 0.561451}, {18, 0.653952}, {19, 0.7683}, {20, 0.901277}, {21, 1.02825}, {22, 1.18254},
 {23, 1.36579}, {24, 1.49554}, {25, 1.71374}, {26, 1.93023}, {27, 2.19595}, {28, 2.49536},
 {29, 2.79143}, {30, 3.1087}, {31, 3.51332}, {32, 3.89292}, {33, 4.33852}, {34, 4.87941},
 {35, 5.32834}, {36, 5.93174}, {37, 6.64703}, {38, 7.36409}, {39, 7.94771}, {40, 8.75999}}
```

In[20]:=
```
(* The CPU time (in seconds) needed to compute the average occupancy is
 plotted as a function of the number of binding sites of the linear array *)
ListLogPlot[TimesTab, Joined → True, PlotRange → All]
```

Out[20]=



In[21]:=
```
TimeRatio40to20 = NNChain[40, -1][[4]] / NNChain[20, -1][[4]]
```

Out[21]= 7.64344

In[22]:=
```
NumberStatesRatio40to20 = 2^40 / 2^20
```

Out[22]= 1 048 576