

Nearest-neighbor model in microarray hybridization

Enrico Carlon, KULeuven

Plön – May 9, 2011

Outline

- 1 Hybridization properties from controlled experiments (Agilent)

- 1 Hybridization properties from controlled experiments (Agilent)
 - 1 *Breakdown of thermal equilibrium*

- 1 Hybridization properties from controlled experiments (Agilent)
 - 1 *Breakdown of thermal equilibrium*
 - 2 *Modeling non-equilibrium effects*

- 1 Hybridization properties from controlled experiments (Agilent)
 - 1 *Breakdown of thermal equilibrium*
 - 2 *Modeling non-equilibrium effects*
 - 3 *Exploring other arrays*

- 1 Hybridization properties from controlled experiments (Agilent)
 - 1 *Breakdown of thermal equilibrium*
 - 2 *Modeling non-equilibrium effects*
 - 3 *Exploring other arrays*

- 2 Algorithms for biological data analysis (Affymetrix)

- 1 Hybridization properties from controlled experiments (Agilent)
 - 1 *Breakdown of thermal equilibrium*
 - 2 *Modeling non-equilibrium effects*
 - 3 *Exploring other arrays*

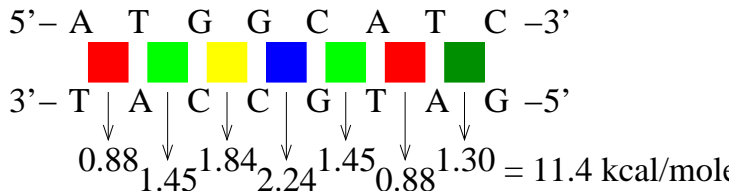
- 2 Algorithms for biological data analysis (Affymetrix)
 - 1 *AffILM (available at www.bioconductor.org)*

Nearest neighbor parameters $\Delta G_{37^\circ\text{C}}$ (DNA/DNA)

Table of nearest neighbor parameters for the hybridization free energies $\Delta G_{37^\circ\text{C}}$ in 1M NaCl expressed in kcal/mol. The orientation is 5'-3' for the upper strand and 3'-5' for the lower strand. Only 10 of the 16 parameters are independent.

$\begin{matrix} AA \\ TT \end{matrix}$	-1.00	$\begin{matrix} AT \\ TA \end{matrix}$	-0.88	$\begin{matrix} AC \\ TG \end{matrix}$	-1.44	$\begin{matrix} AG \\ TC \end{matrix}$	-1.28
$\begin{matrix} TA \\ AT \end{matrix}$	-0.58	$\begin{matrix} TT \\ AA \end{matrix}$	-1.00	$\begin{matrix} TC \\ AG \end{matrix}$	-1.30	$\begin{matrix} TG \\ AC \end{matrix}$	-1.45
$\begin{matrix} CA \\ GT \end{matrix}$	-1.45	$\begin{matrix} CT \\ GA \end{matrix}$	-1.28	$\begin{matrix} CC \\ GG \end{matrix}$	-1.84	$\begin{matrix} CG \\ GC \end{matrix}$	-2.17
$\begin{matrix} GA \\ CT \end{matrix}$	-1.30	$\begin{matrix} GT \\ CA \end{matrix}$	-1.44	$\begin{matrix} GC \\ CG \end{matrix}$	-2.24	$\begin{matrix} GG \\ CC \end{matrix}$	-1.84

Example of a calculation



... plus boundary terms (effect of the double helix ends)

Mismatches

Base pairings can occur also for non-complementary bases.

Here: possible structure of a mismatch between G (left) and A (right)
G·A forms two hydrogen bonds!

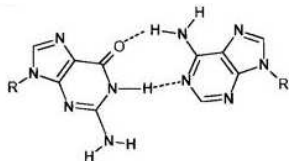


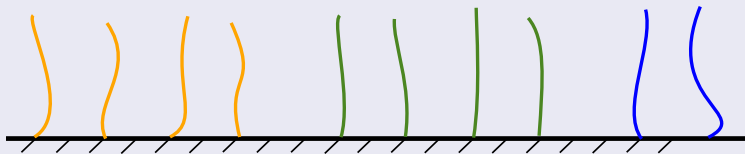
Table of nearest-neighbor parameters for AG mismatches at 37°C in 1M NaCl (unit kcal/mol)

$\frac{AA}{TG}$	0.14	$\frac{AG}{TA}$	0.02	$\frac{CA}{GG}$	0.03	$\frac{CG}{GA}$	0.11
$\frac{GA}{CG}$	-0.25	$\frac{GG}{CA}$	-0.52	$\frac{TA}{AG}$	0.42	$\frac{TG}{AA}$	0.74

Hybridization in a biological experiment

Complex system with a large set of chemical reactions

Before Hybridization



Aim

Understand/characterize hybridization in microarrays

Hybridization in a biological experiment

Complex system with a large set of chemical reactions

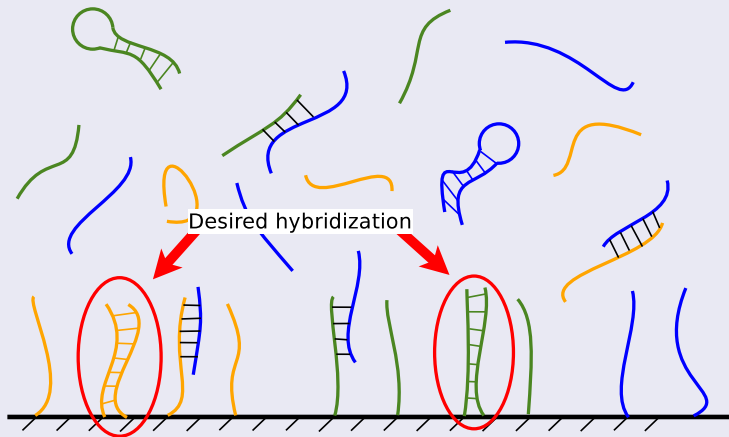


Aim

Understand/characterize hybridization in microarrays

Hybridization in a biological experiment

Complex system with a large set of chemical reactions

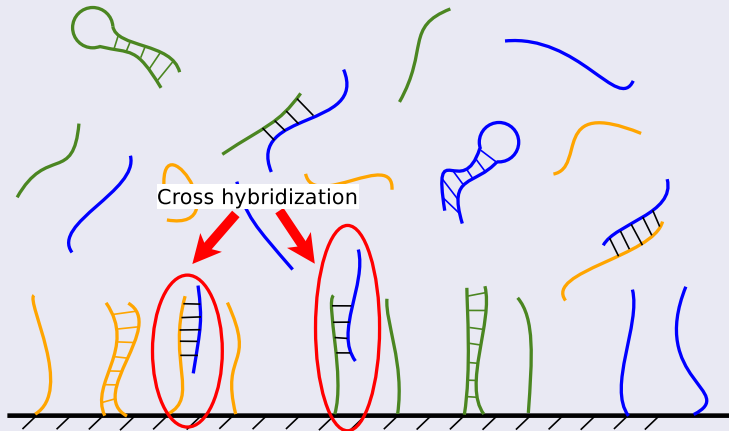


Aim

Understand/characterize hybridization in microarrays

Hybridization in a biological experiment

Complex system with a large set of chemical reactions

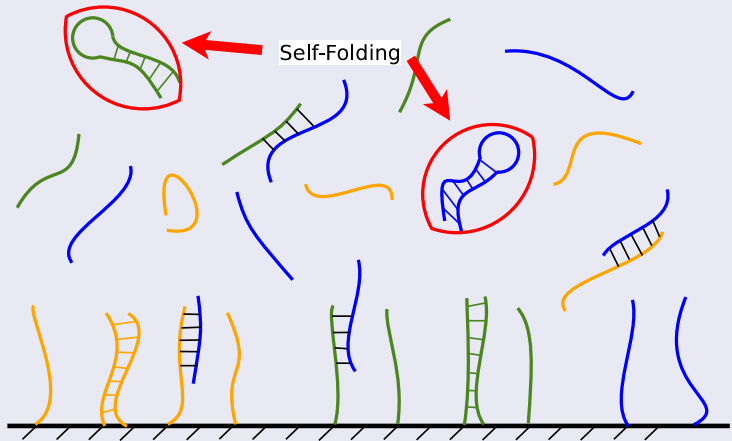


Aim

Understand/characterize hybridization in microarrays

Hybridization in a biological experiment

Complex system with a large set of chemical reactions

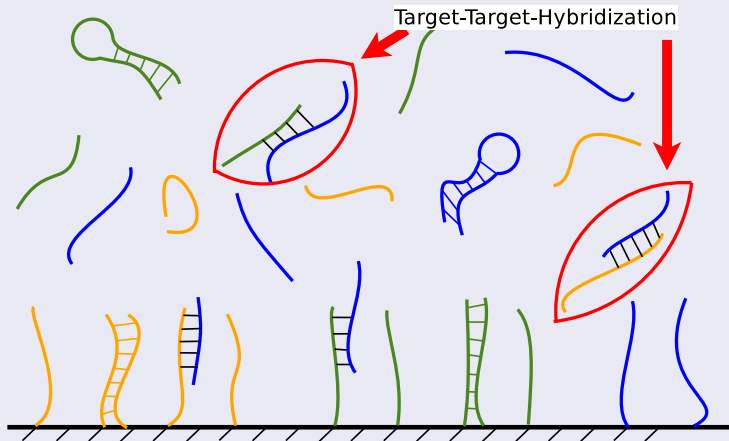


Aim

Understand/characterize hybridization in microarrays

Hybridization in a biological experiment

Complex system with a large set of chemical reactions



Aim

Understand/characterize hybridization in microarrays

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3' target in solution

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3' target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5' perfect match

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA T CTTCTAACCCACCGTGACAACATT-5'	1 mismatch

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA T CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAAG A TTCTAACCCACCGTGACAACATT-5'	1 mismatch

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA T CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAAG A TTCTAACCCACCGTGACAACATT-5'	1 mismatch
...	
3'-CAAAAGCTTCTAACCCACCGTGAC G ACATT-5'	1 mismatch

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA T CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAAG A TTCTAACCCACCGTGACAACATT-5'	1 mismatch
...	
3'-CAAAAGCTTCTAACCCACCGTGAC G ACATT-5'	1 mismatch
3'-CAAAAGCTTCTAACCCACCGTGAC T ACATT-5'	1 mismatch

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA T CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAAG A TTCTAACCCACCGTGACAACATT-5'	1 mismatch
...	
3'-CAAAAGCTTCTAACCCACCGTGAC G ACATT-5'	1 mismatch
3'-CAAAAGCTTCTAACCCACCGTGAC T ACATT-5'	1 mismatch
3'-CAAAA A CTTCT C ACCCACCGTGACAACATT-5'	2 mismatches

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA T CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAAG A TTCTAACCCACCGTGACAACATT-5'	1 mismatch
...	
3'-CAAAAGCTTCTAACCCACCGTGAC G ACATT-5'	1 mismatch
3'-CAAAAGCTTCTAACCCACCGTGAC T ACATT-5'	1 mismatch
3'-CAAAA A CTTCT C ACCCACCGTGACAACATT-5'	2 mismatches
3'-CAAAA A CTTCT G ACCCACCGTGACAACATT-5'	2 mismatches

Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA T CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAAG A TTCTAACCCACCGTGACAACATT-5'	1 mismatch
...	
3'-CAAAAGCTTCTAACCCACCGTGAC G ACATT-5'	1 mismatch
3'-CAAAAGCTTCTAACCCACCGTGAC T ACATT-5'	1 mismatch
3'-CAAAA A CTTCT C ACCCACCGTGACAACATT-5'	2 mismatches
3'-CAAAA A CTTCT G ACCCACCGTGACAACATT-5'	2 mismatches
3'-CAAAA A CTTCT T ACCCACCGTGACAACATT-5'	2 mismatches
...	
3'-CAAAAGCTTCTAACCCAC T GTGAC T ACATT-5'	2 mismatches

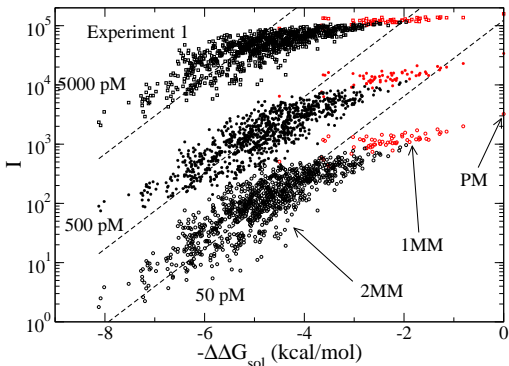
Determination of ΔG from Agilent Custom Arrays

Sufficiently simple setup to avoid “spurious” effects

5'-GTTTTCGAAGATTGGGTGGCACTGTTGTAA-3'	target in solution
3'-CAAAAGCTTCTAACCCACCGTGACAACATT-5'	perfect match
3'-CAAAA A CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA C CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAA T CTTCTAACCCACCGTGACAACATT-5'	1 mismatch
3'-CAAAAG A TTCTAACCCACCGTGACAACATT-5'	1 mismatch
...	
3'-CAAAAGCTTCTAACCCACCGTGAC G ACATT-5'	1 mismatch
3'-CAAAAGCTTCTAACCCACCGTGAC T ACATT-5'	1 mismatch
3'-CAAAA A CTTCT C ACCCACCGTGACAACATT-5'	2 mismatches
3'-CAAAA A CTTCT G ACCCACCGTGACAACATT-5'	2 mismatches
3'-CAAAA A CTTCT T ACCCACCGTGACAACATT-5'	2 mismatches
...	
3'-CAAAAGCTTCTAACCCAC T GTGAC T ACATT-5'	2 mismatches

total 1006 sequences, replicated 15 times = 15K array

Intensity vs. ΔG_{sol}



Expected: Langmuir isotherm

$$I = \frac{Ace^{\Delta G/RT}}{1 + ce^{\Delta G/RT}}$$

data for 3 different concentrations (c)

lines: slope = $1/RT$

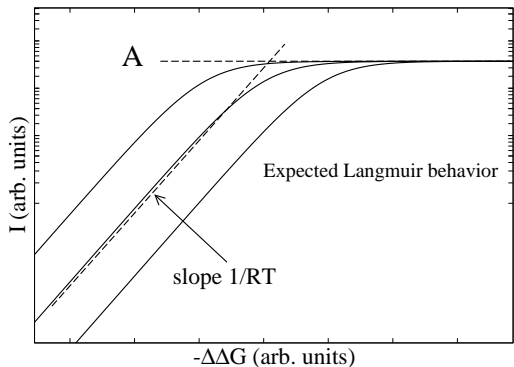
scatter of the data

deviations from $1/RT$ already far from chemical saturation

No fitting parameters!

ΔG_{sol} is computed from tabulated literature data (parameters measured in hybridization in solution)

Intensity vs. ΔG_{sol}



Expected: Langmuir isotherm

$$I = \frac{Ace^{\Delta G/RT}}{1 + ce^{\Delta G/RT}}$$

data for 3 different
concentrations (c)

lines: slope = $1/RT$

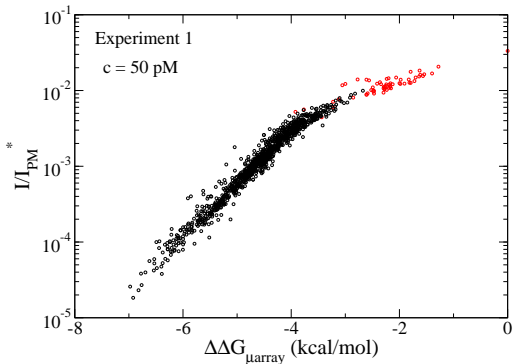
scatter of the data

deviations from $1/RT$ already
far from chemical saturation

No fitting parameters!

ΔG_{sol} is computed from tabulated literature data (parameters measured in hybridization in solution)

Fitting free energy parameters from experimental data

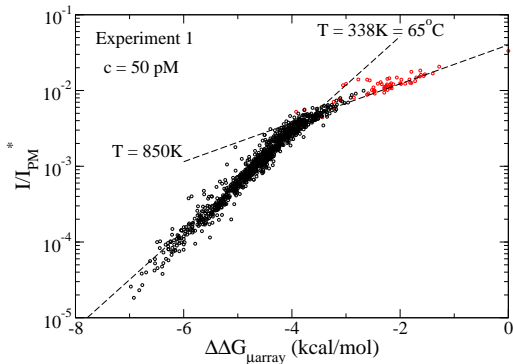


1006 data points and
58 parameters

Fit to

$$I = A \exp(\Delta G/RT)$$

Fitting free energy parameters from experimental data



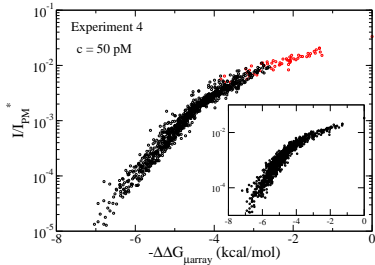
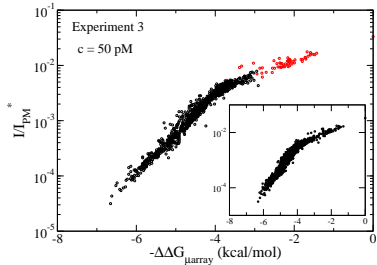
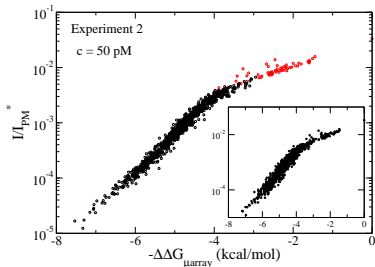
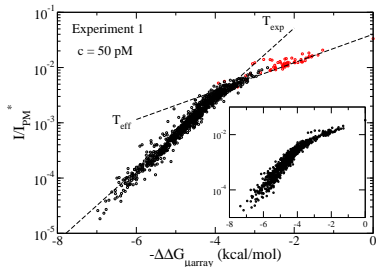
1006 data points and
58 parameters

Fit to

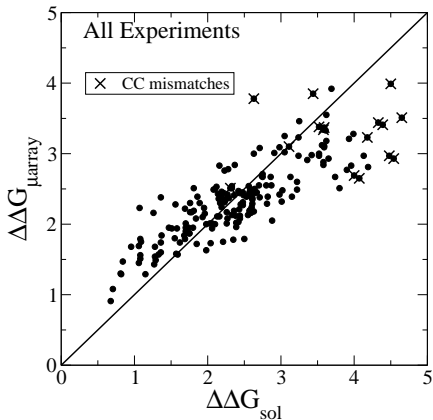
$$I = A \exp(\Delta G/RT)$$

Two temperatures!

Identical behavior for 4 different sequences



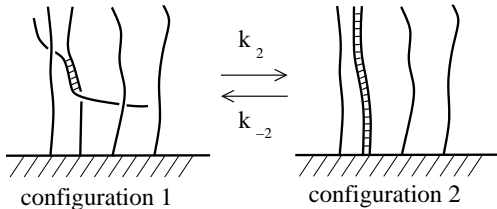
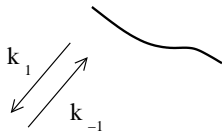
Comparison of free energy in the array and in solution



$\Delta\Delta G$: difference between perfect match hybridization and hybridization with an internal mismatch.

Moderate correlation between solution and microarray free energies

Understanding the two regimes

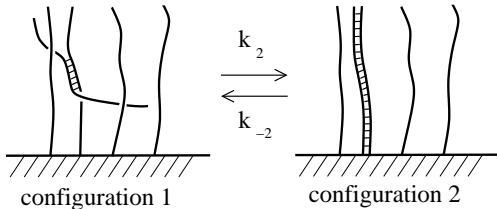
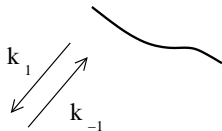


Hybridization in solution:
nucleation and fast zippering

Is zippering rapid in
microarrays?
Probably not!

The simplest description of a
long lived partially zipped
state is through a 3 state
model

Understanding the two regimes



Hybridization in solution:
nucleation and fast zippering

Is zippering rapid in
microarrays?
Probably not!

The simplest description of a
long lived partially zipped
state is through a 3 state
model

Four parameters

Hybridization rates: k_1 , k_2 (forward), k_{-1} , k_{-2} (reverse)

Thermodynamics $k_{-1}/k_1 = e^{-\Delta G'/RT}$ $k_{-2}/k_2 = e^{-(\Delta G - \Delta G')/RT}$

Two vs. three state model

Two state model
kinetics

$$\frac{d\theta}{dt} = ck_1(1 - \theta) - k_{-1}\theta$$

Equilibrium

$$\theta_{eq} = \frac{ck_1}{k_{-1} + ck_1} = \frac{cK}{1 + cK} = \frac{ce^{\Delta G/RT}}{1 + ce^{\Delta G/RT}}$$

Two vs. three state model

Two state model
kinetics

$$\frac{d\theta}{dt} = ck_1(1 - \theta) - k_{-1}\theta$$

Equilibrium

$$\theta_{eq} = \frac{ck_1}{k_{-1} + ck_1} = \frac{cK}{1 + cK} = \frac{ce^{\Delta G/RT}}{1 + ce^{\Delta G/RT}}$$

Three
state
model

$$\frac{d\theta^{(1)}}{dt} = ck_1(1 - \theta^{(1)} - \theta^{(2)}) + k_{-2}\theta^{(2)} - (k_{-1} + k_2)\theta^{(1)}$$

$$\frac{d\theta^{(2)}}{dt} = k_2\theta^{(1)} - k_{-2}\theta^{(2)}$$

Equilibrium
($c \rightarrow 0$)

$$\theta_{eq}^{(1)} = c \frac{k_1}{k_{-1}} = ce^{\Delta G'/RT}$$

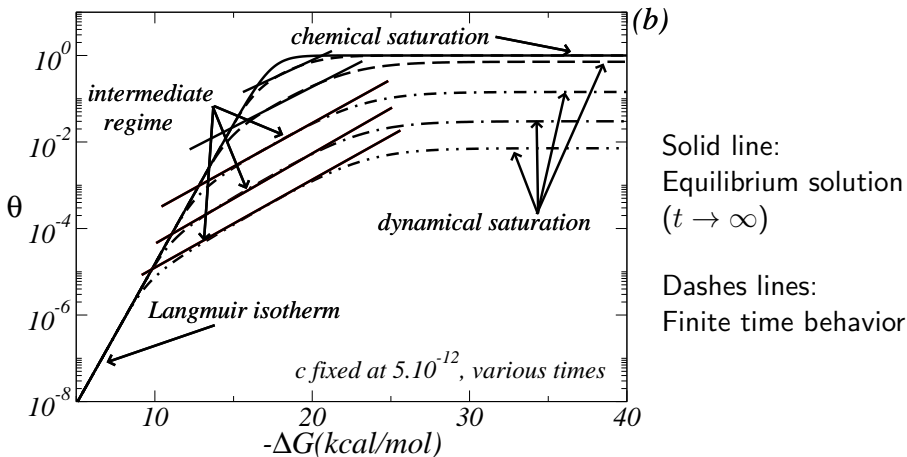
$$\theta_{eq}^{(2)} = c \frac{k_1 k_2}{k_{-1} k_{-2}} = ce^{\Delta G/RT}$$

Assumption:

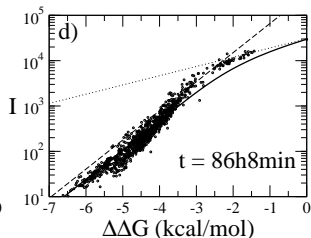
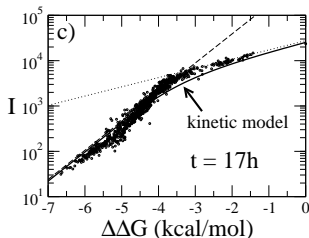
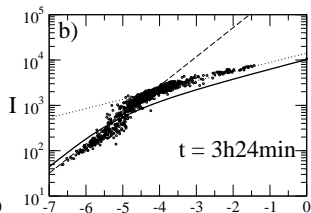
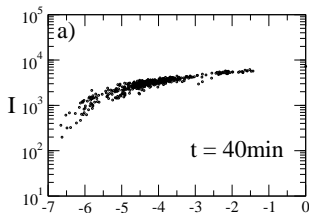
k_1, k_2 independent on ΔG

$\Delta G' \approx \gamma \Delta G$ (they should be monotonically linked)

Three state model



Experiments (Agilent arrays)



Target sequence

$L = 30$

Exper. to $\approx 86\text{h}$

Temperature 65°C

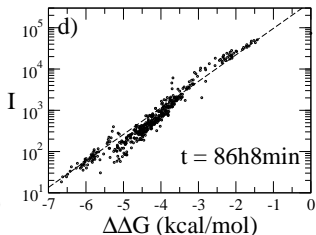
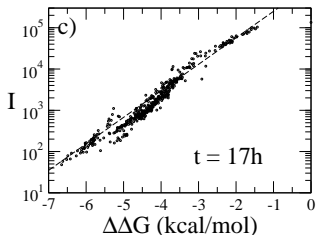
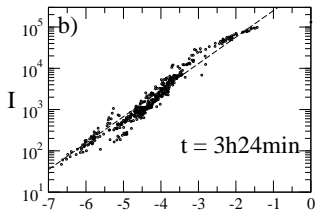
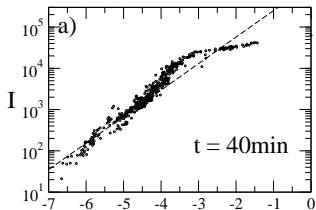
Solid line:

three state model

Dashed line:

Equilibrium isotherm

Experiments (Agilent arrays)

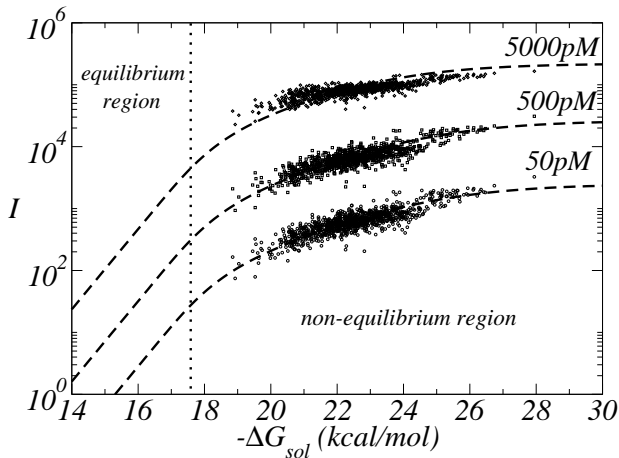


Target sequence
 $L = 25$
Exper. to $\approx 86\text{h}$

Temperature 65°C

Dashed line:
Equilibrium isotherm

Experiments (Agilent arrays)



Target sequence

$L = 30$

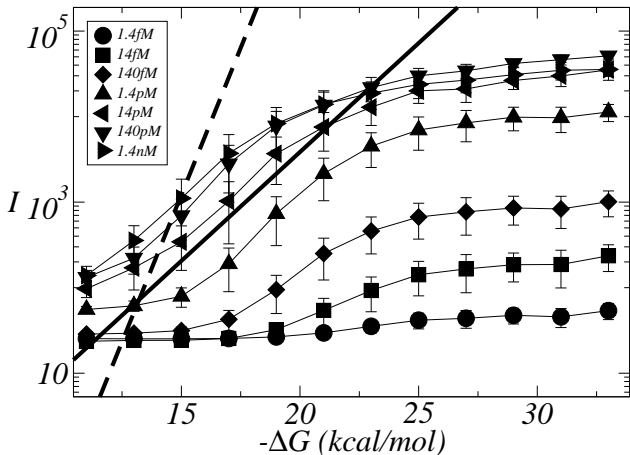
Exper. ≈ 17 h

Temperature 55°C

Dashed line:
Three state model

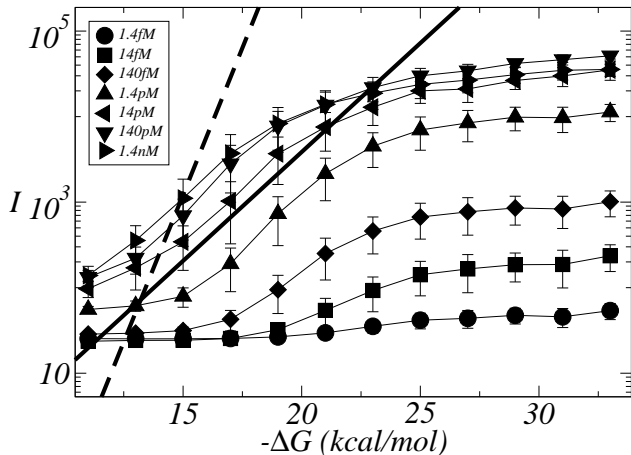
Evidence of dynamical
saturation

Experiments (High density arrays, Suzuki et al.)



140 targets and probes with 1 or 2 mismatches

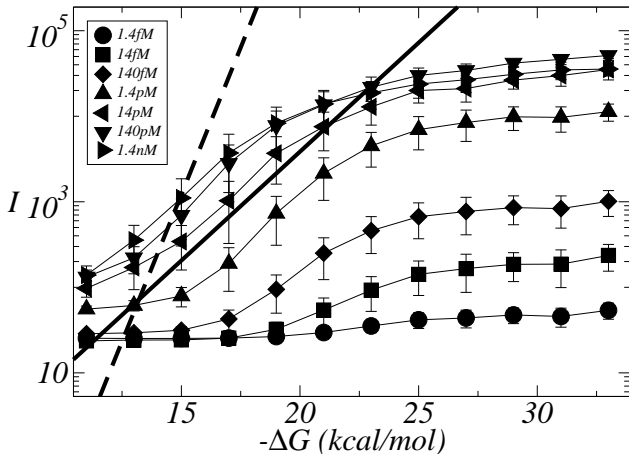
Experiments (High density arrays, Suzuki et al.)



140 targets and probes with 1 or 2 mismatches

Solid line:
 $T_{\text{eff}} \approx 1200K$

Experiments (High density arrays, Suzuki et al.)

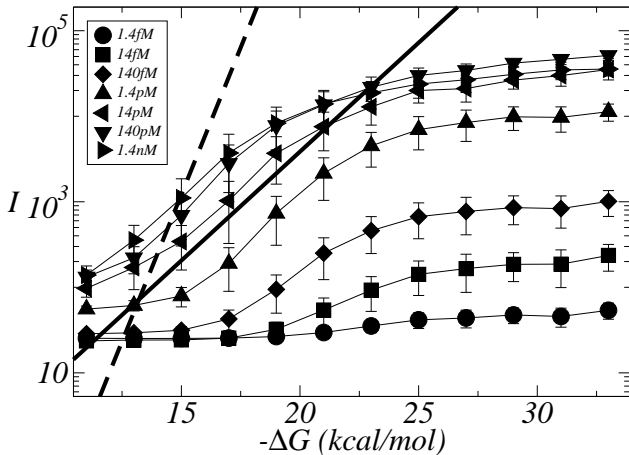


140 targets and probes with 1 or 2 mismatches

Solid line:
 $T_{\text{eff}} \approx 1200K$

Dashed line:
 $T_{\text{exp}} \approx 45C$

Experiments (High density arrays, Suzuki et al.)



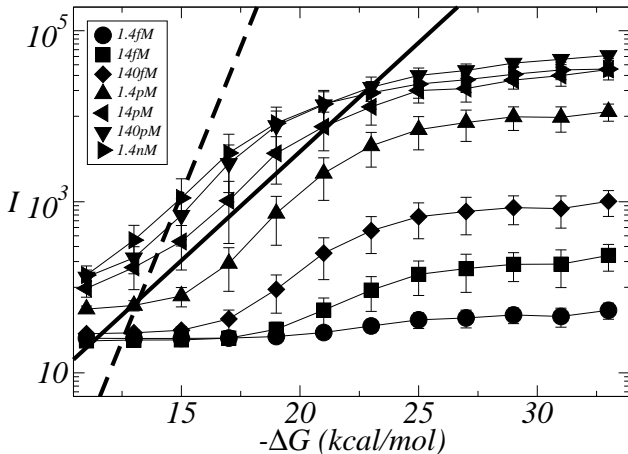
140 targets and probes with 1 or 2 mismatches

Solid line:
 $T_{\text{eff}} \approx 1200K$

Dashed line:
 $T_{\text{exp}} \approx 45C$

Dynamical saturation

Experiments (High density arrays, Suzuki et al.)



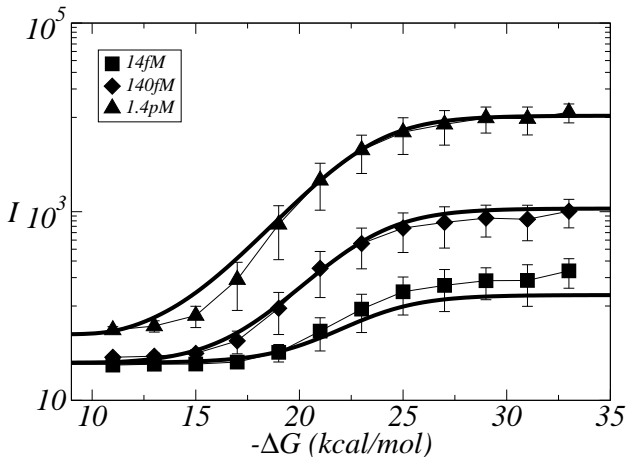
140 targets and probes with 1 or 2 mismatches

Solid line:
 $T_{\text{eff}} \approx 1200K$

Dashed line:
 $T_{\text{exp}} \approx 45C$

Dynamical saturation
... or depletion?

Experiments (High density arrays, Suzuki et al.)



140 targets and probes with 1 or 2 mismatches

Solid line:

$T_{\text{eff}} \approx 1200K$

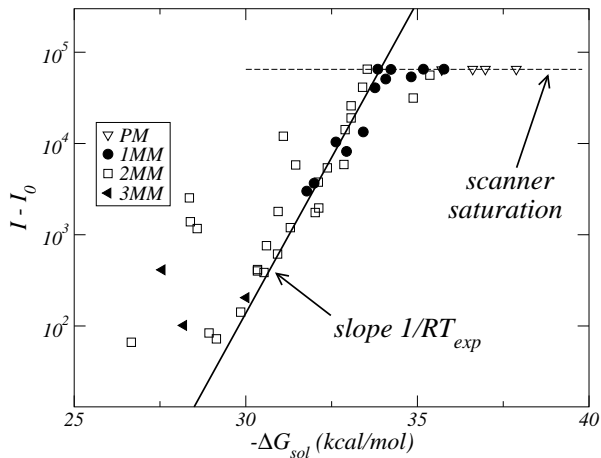
Dashed line:

$T_{\text{exp}} \approx 45C$

Dynamical saturation
... or depletion?

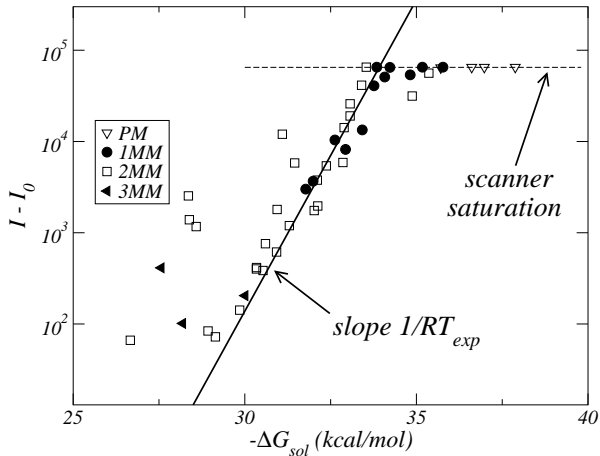
Solid line: three state model

Experiments (Spotted array, Weckx et al.)



Experiment with 4 targets and probes up to three mismatches

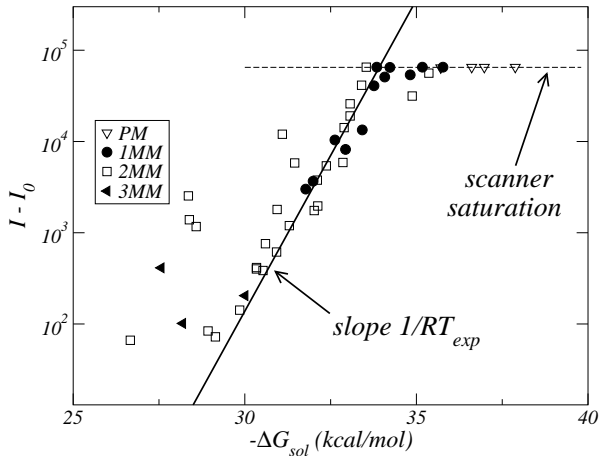
Experiments (Spotted array, Weckx et al.)



Experiment with 4 targets and probes up to three mismatches

Large range of ΔG

Experiments (Spotted array, Weckx et al.)



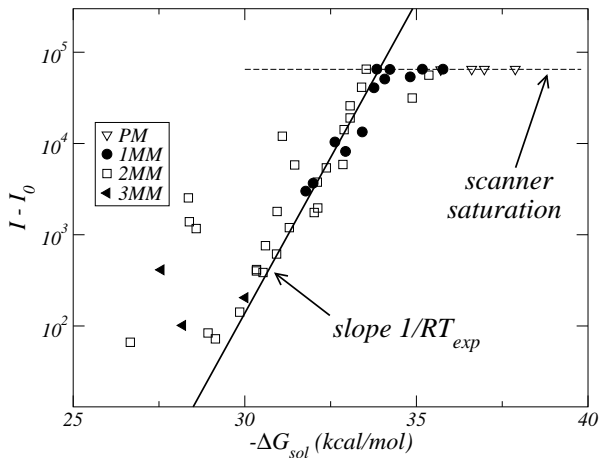
Experiment with 4 targets and probes up to three mismatches

Large range of ΔG

Solid line:

$$T_{exp} \approx 45C$$

Experiments (Spotted array, Weckx et al.)



Experiment with 4 targets and probes up to three mismatches

Large range of ΔG

Solid line:

$T_{exp} \approx 45C$

Dashed line:

scanner saturation

Practical consequences

Microarrays working in a non-equilibrium regime suffer from enhanced cross-hybridization

Practical consequences

Microarrays working in a non-equilibrium regime suffer from enhanced cross-hybridization

Example: two different sequences in solution binding to the same probe at the surface. The first is perfect match and has concentration c , the second carries mismatches and has concentration c' .

$$I = cf(\Delta G) + c'f(\Delta G')$$

Practical consequences

Microarrays working in a non-equilibrium regime suffer from enhanced cross-hybridization

Example: two different sequences in solution binding to the same probe at the surface. The first is perfect match and has concentration c , the second carries mismatches and has concentration c' .

$$I = cf(\Delta G) + c'f(\Delta G')$$

We wish to have $f(\Delta G) \gg f(\Delta G')$. In equilibrium

$$\frac{f_{eq}(\Delta G)}{f_{eq}(\Delta G')} = e^{(\Delta G - \Delta G')/RT_{exp}}$$

is maximized. . .

Practical consequences

Microarrays working in a non-equilibrium regime suffer from enhanced cross-hybridization

Example: two different sequences in solution binding to the same probe at the surface. The first is perfect match and has concentration c , the second carries mismatches and has concentration c' .

$$I = cf(\Delta G) + c'f(\Delta G')$$

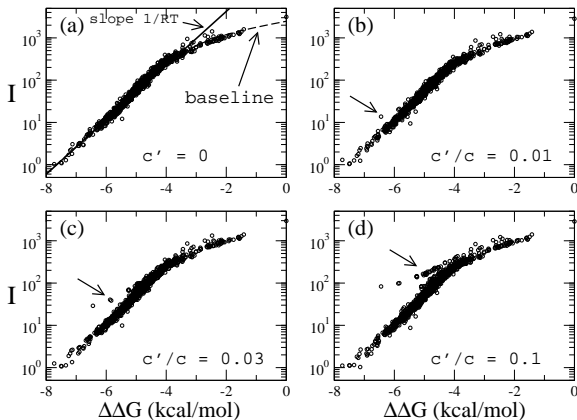
We wish to have $f(\Delta G) \gg f(\Delta G')$. In equilibrium

$$\frac{f_{eq}(\Delta G)}{f_{eq}(\Delta G')} = e^{(\Delta G - \Delta G')/RT_{exp}}$$

is maximized. . . while in non-equilibrium

$$\frac{f(\Delta G)}{f(\Delta G')} = e^{(\Delta G - \Delta G')/RT_{eff}} \ll \frac{f_{eq}(\Delta G)}{f_{eq}(\Delta G')}$$

Application: Mutant identification (Jef's talk)



Target:
Mixture of 2 sequences

Wild type at c
Mutant at c'

Detection limit
 $c'/c \approx 0.01$

Back to complex biological experiments (Affy Genechip)

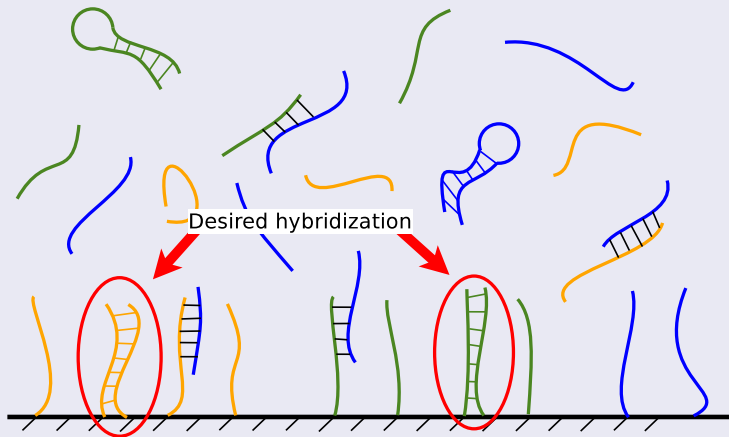
AffyILM: Algorithm available from BioConductor



AffyILM

Back to complex biological experiments (Affy Genechip)

AffyILM: Algorithm available from BioConductor

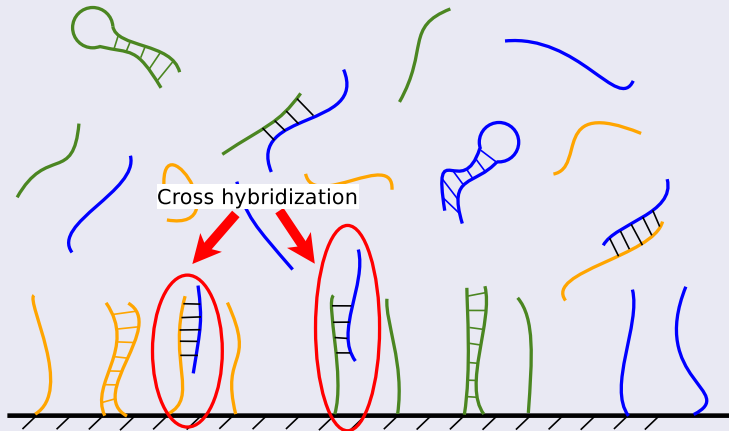


AffyILM

... Langmuir hybridization (with T_{eff})

Back to complex biological experiments (Affy Genechip)

AffyILM: Algorithm available from BioConductor

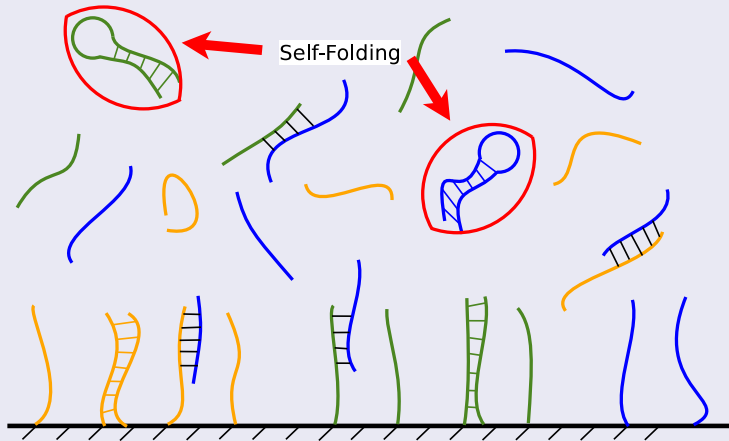


AffyILM

... background subtraction

Back to complex biological experiments (Affy Genechip)

AffyILM: Algorithm available from BioConductor

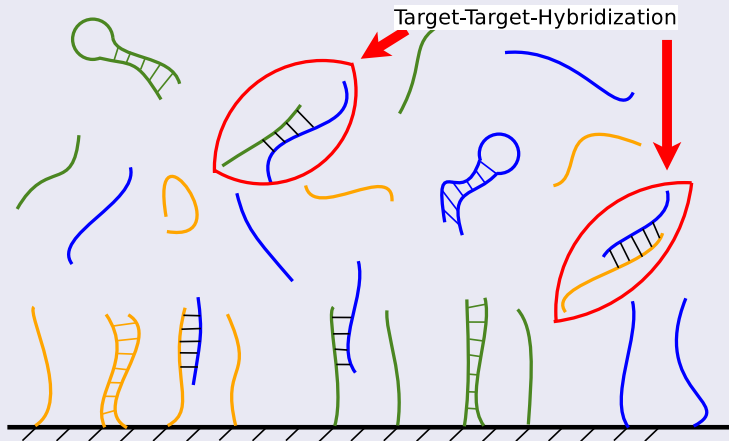


AffyILM

... not taken into account (should affect 5% of targets)

Back to complex biological experiments (Affy Genechip)

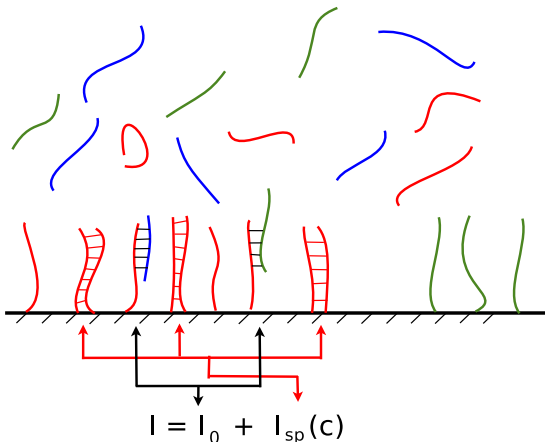
AffyILM: Algorithm available from BioConductor



AffyILM

... mean-field approximation

1st - Estimating of the Background Noise



Motivation:

- Strong signals ($I_{sp} \gg I_0$)
⇒ Gene expressed
- Weak signals ($I_{sp} \approx I_0$)
⇒ Weakly expressed gene or background noise?

Aim

Calculation of background intensity for each probe on array

2nd - Specific signal: extended Langmuir model

θ : fraction of bound probes

c : mRNA concentration

α : fraction "free" mRNA

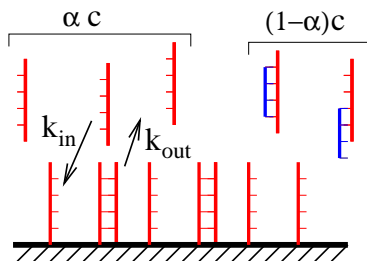
$$\theta = \frac{\alpha c e^{\Delta G/RT}}{1 + \alpha c e^{\Delta G/RT}}$$

$$\alpha = \frac{1}{1 + \tilde{c} e^{\Delta G_R/RT'}}$$

Intensity: $I = A\theta$

ΔG , ΔG_R : from data in solution

Parameters: A , T , T' , \tilde{c}

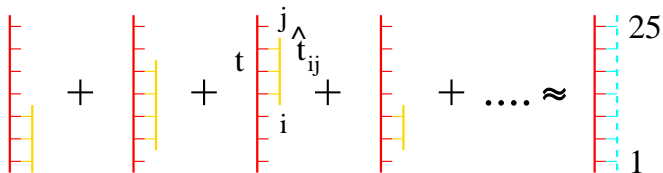


Scaling variable

$$x' = \alpha c \exp(\Delta G/RT)$$

Fitted effective temperature $T \approx 700K$

Modeling Hybridization in solution



Chemical equilibrium

$$\frac{[t][\hat{t}_{i,j}]}{[t\hat{t}_{i,j}]} = e^{\Delta G_R(i,j)/RT}$$

Target conservation

$$[t] + \sum_{i,j} [t\hat{t}_{i,j}] = c$$

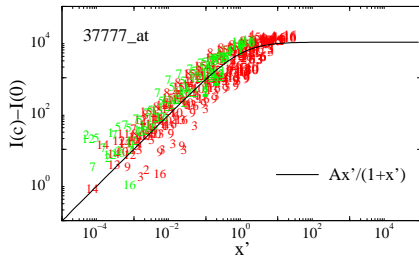
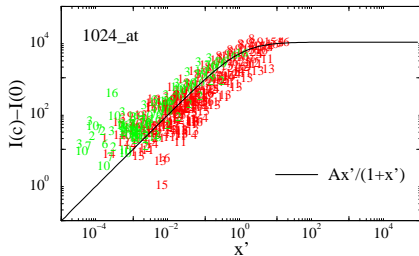
$$\alpha = \frac{[t]}{c} = \frac{1}{1 + \sum_{i,j} [\hat{t}_{i,j}] e^{-\Delta G_R(i,j)/RT}} \approx \frac{1}{1 + \tilde{c} e^{-\Delta G_R(1,25)/RT_{\text{eff}}}}$$

Calibration data (spike-in)

≈ 300 data points (concentrations $c = 0, 0.25, 0.5 \dots 1024$ pM)

Scaling collapses (4 fitting parameters)

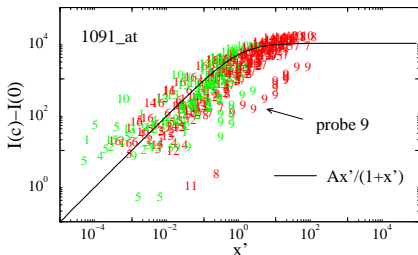
RED : PM / GREEN : MM



x-axis: scaling variable $x' = \alpha c \exp(\Delta G/RT)$

y-axis: background subtracted intensities

Outliers



1091_at9 sequence alignment (GenBank)

3'-CTG TCC TTG GTC CG C ATG GCT CGT T-5'	Affymetrix
3'-CTG TCC TTG GTC CG C ATG GCT CGT T-5'	M65066.1
3'-CTG TCC TTG GTC CG AGGCT GCT CGT T-5'	BC013368.2
3'-CTG TCC TTG GTC CG AGGCT GCT CGT T-5'	AL833563.1

Data base

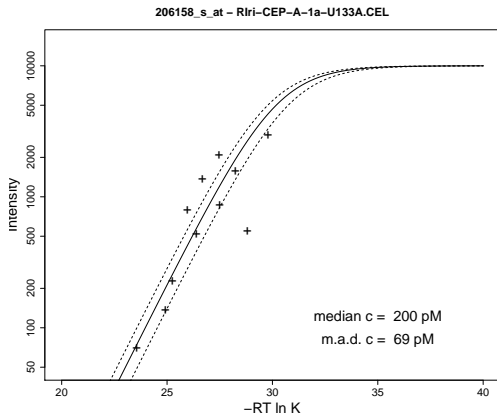
Error in the annotation:

the correct sequence was submitted to GenBank in 2003, after the microarray was designed.

affyILM (available at www.bioconductor.org)

Input: Experimental microarray data (Affymetrix)

Output: Estimate of Concentration per each probe sequence



Probe no.	conc [pM]
1	437.24
2	466.92
3	176.57
4	200.39
5	433.10
6	153.85
7	211.93
8	154.55
9	124.68
10	35.74
11	201.57
Median	200
M.A.D.	69

Collaborators

Institute for Theoretical Physics - KULeuven

M. Baiesi, F. Berger, A. Ferrantini, K. M. Kroll, J. C. Walter
J. Hooyberghs (& Flemish Institute for Technological Research, VITO)

Institute for Theoretical Physics - Utrecht University

G. T. Barkema, J. Klein Wolterink, G. Mulders

Micorarray Facility - Flemish Institute for Biotechnology (VIB)

P. Van Hummelen, S. Weckx

Financial support from

FWO (Flemish Research Council), KULeuven, VITO, VIB



Recent Papers

G. Mulders, G. Barkema and E. Carlon

Inverse Langmuir method for oligonucleotide microarray analysis
BMC Bioinformatics **10**, 64 (2009)

J. Hooyberghs, M. Baiesi, A. Ferrantini and E. Carlon

Breakdown of thermodynamic equilibrium for DNA hybridization in microarrays
Phys. Rev. E **81**, 012901 (2010).

J. Hooyberghs and E. Carlon

Hybridisation thermodynamic parameters allow accurate detection of point mutations with DNA microarrays
Biosens. and Bioel. **26**, 1692 (2010)

J. C. Walter, K. M. Kroll, J. Hooyberghs and E. Carlon

Non-Equilibrium Effects in DNA Microarrays: A Multiplatform Study
J.Phys. Chem. B (in press). DOI: 10.1021/jp2014034

Downloadable from <http://itf.fys.kuleuven.be/~enrico/>