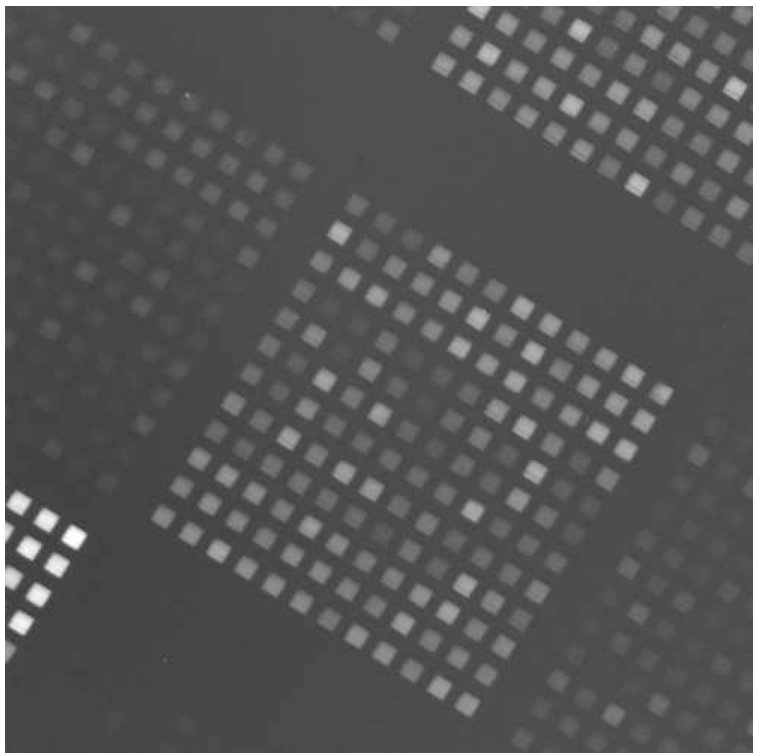
Towards the information entropy of DNA microarrays



Albrecht Ott

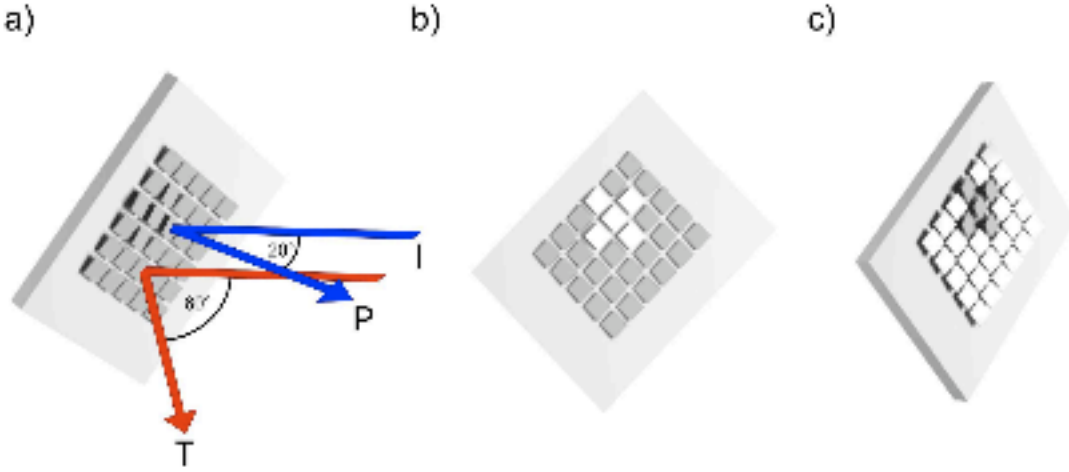
Biologische ExperimentalPhysik

Universität des Saarlandes

The DMD – a dynamic photomask

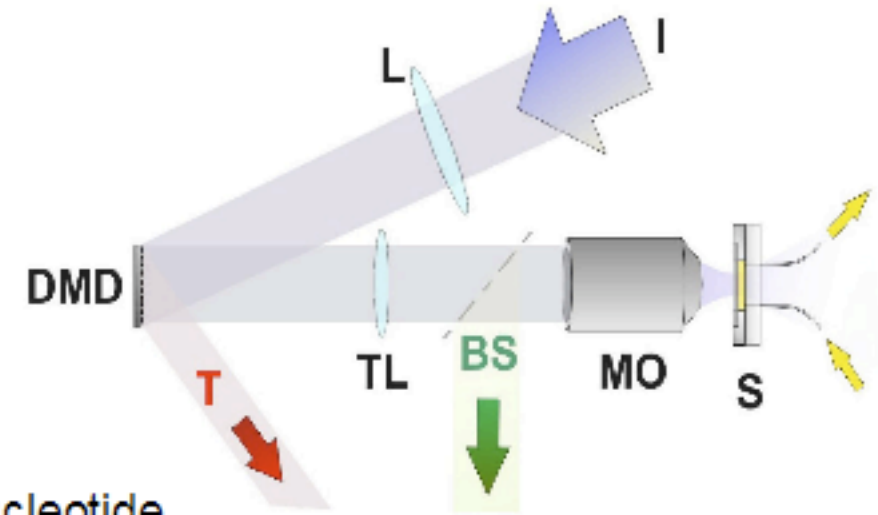
Digital Micromirror Device (DMD™)

(taken from a DLP™ video projector)



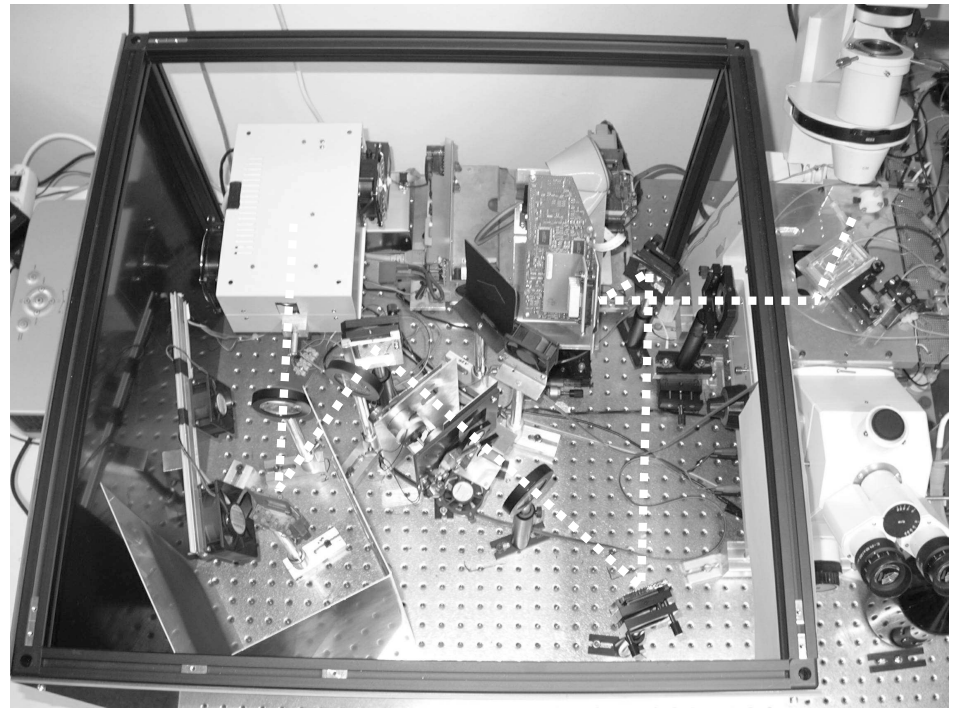
"Maskless" microscope projection photolithography

- no need for expensive chrome masks
- intrinsic mask alignment

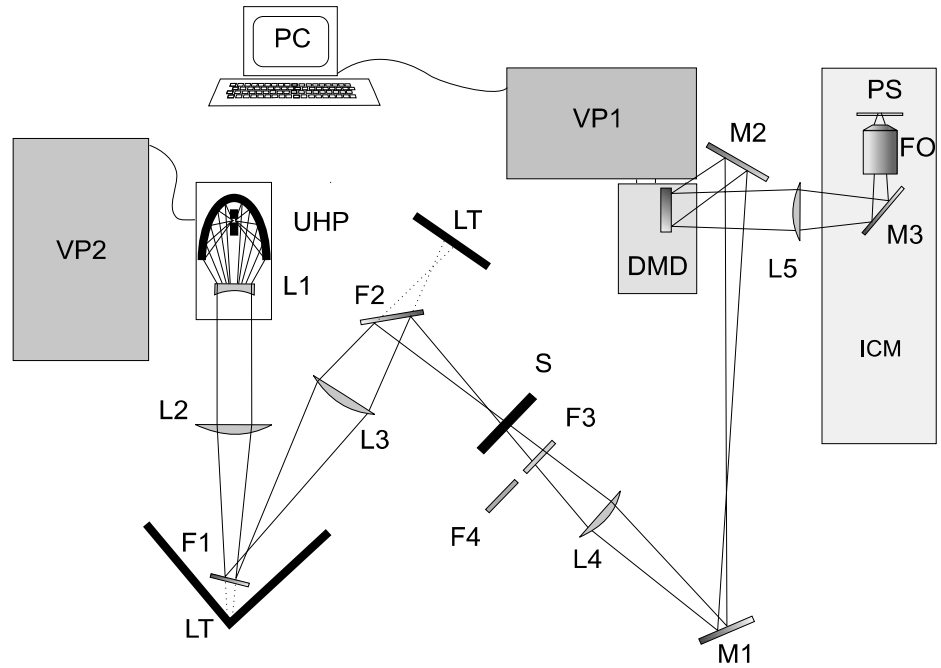


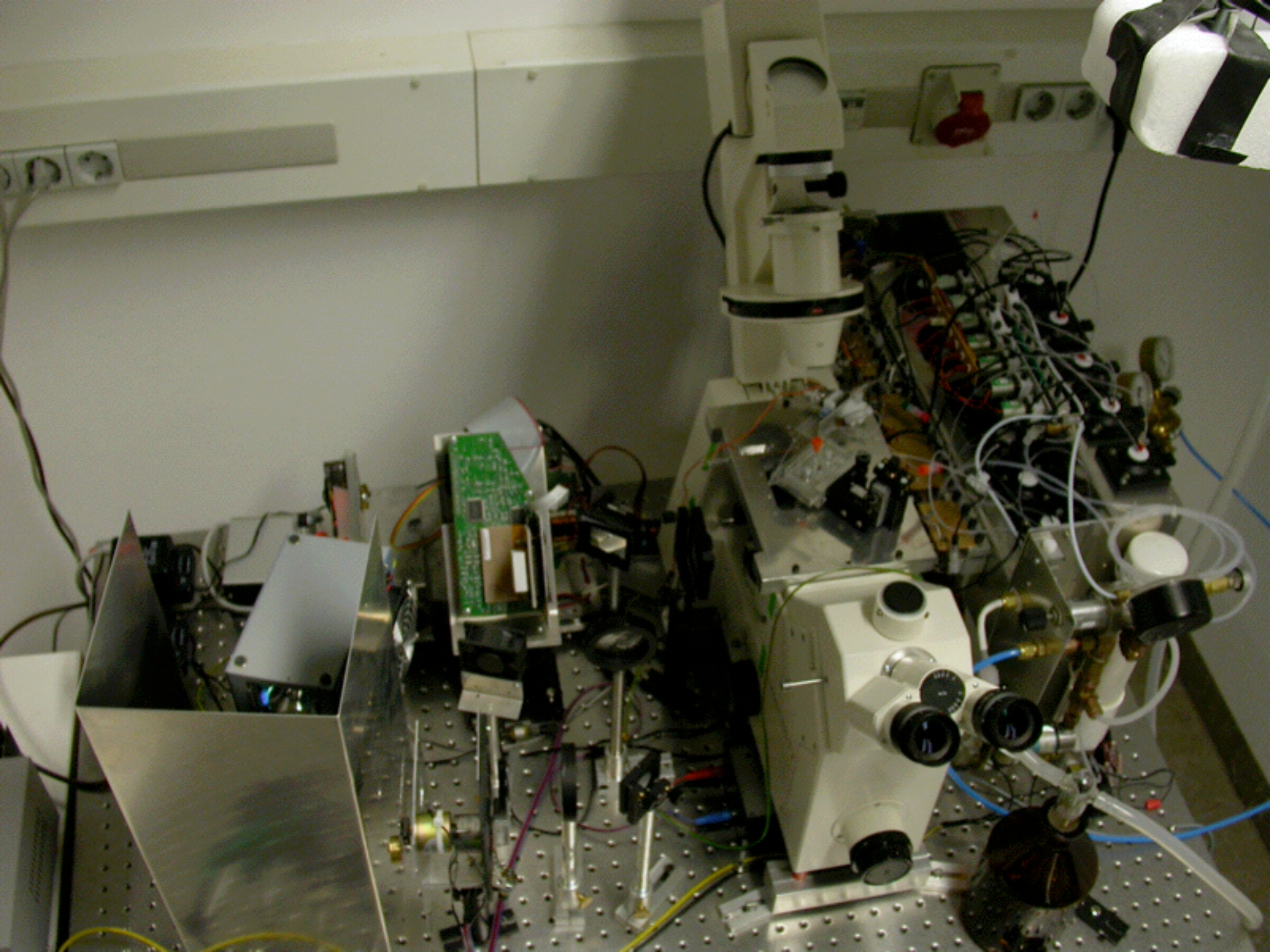
Singh-Gasson *et al.*
Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array.
Nat Biotechnol. 1999 Oct;17(10):974-978

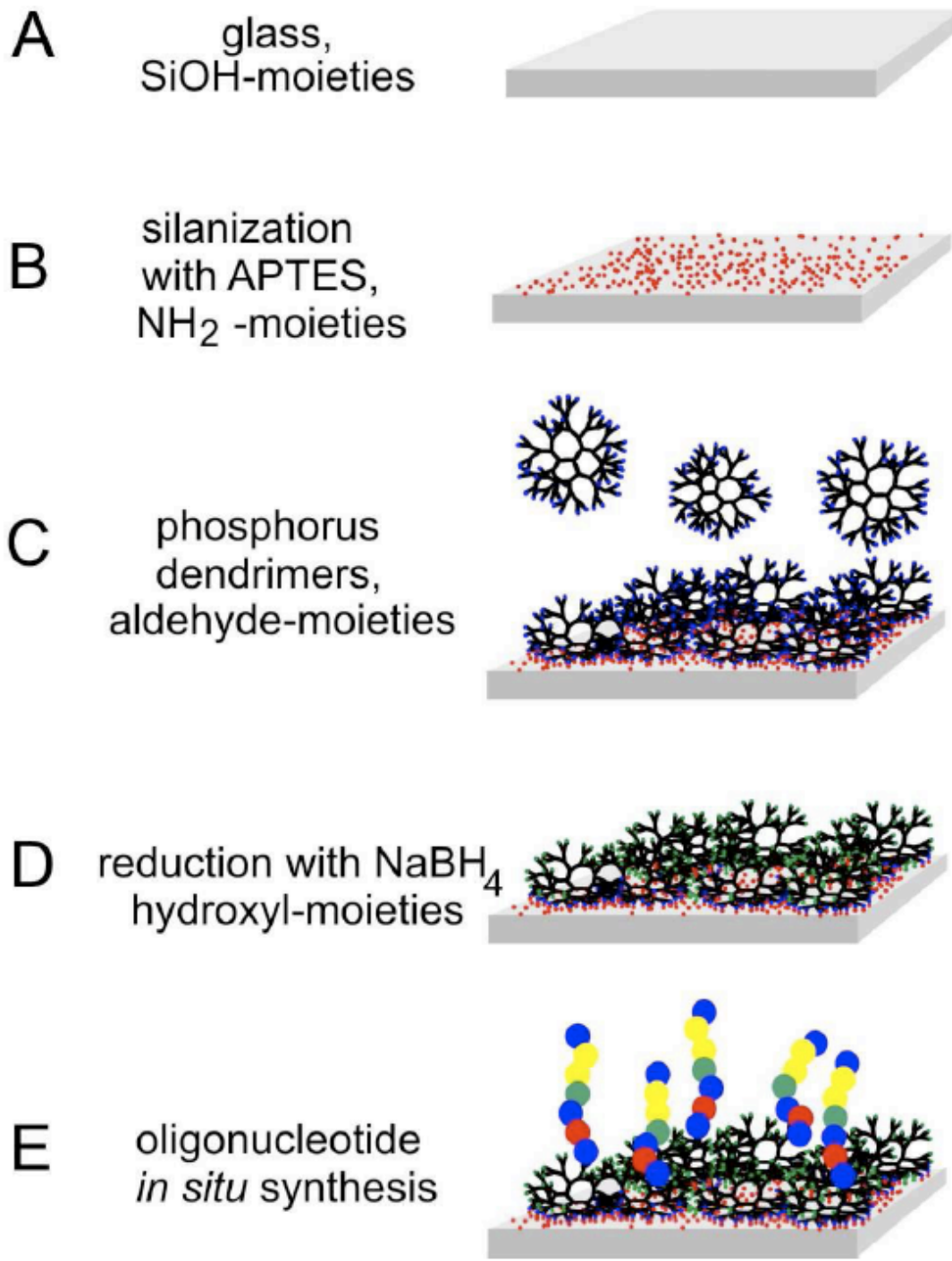
a)



b)


PC
VP1
PS
FO
M2
VP2
UHP
LT
M3
DMD
L5
L1
F2
ICM
S
L2
F3
L3
F4
L4
F1
LT
M1





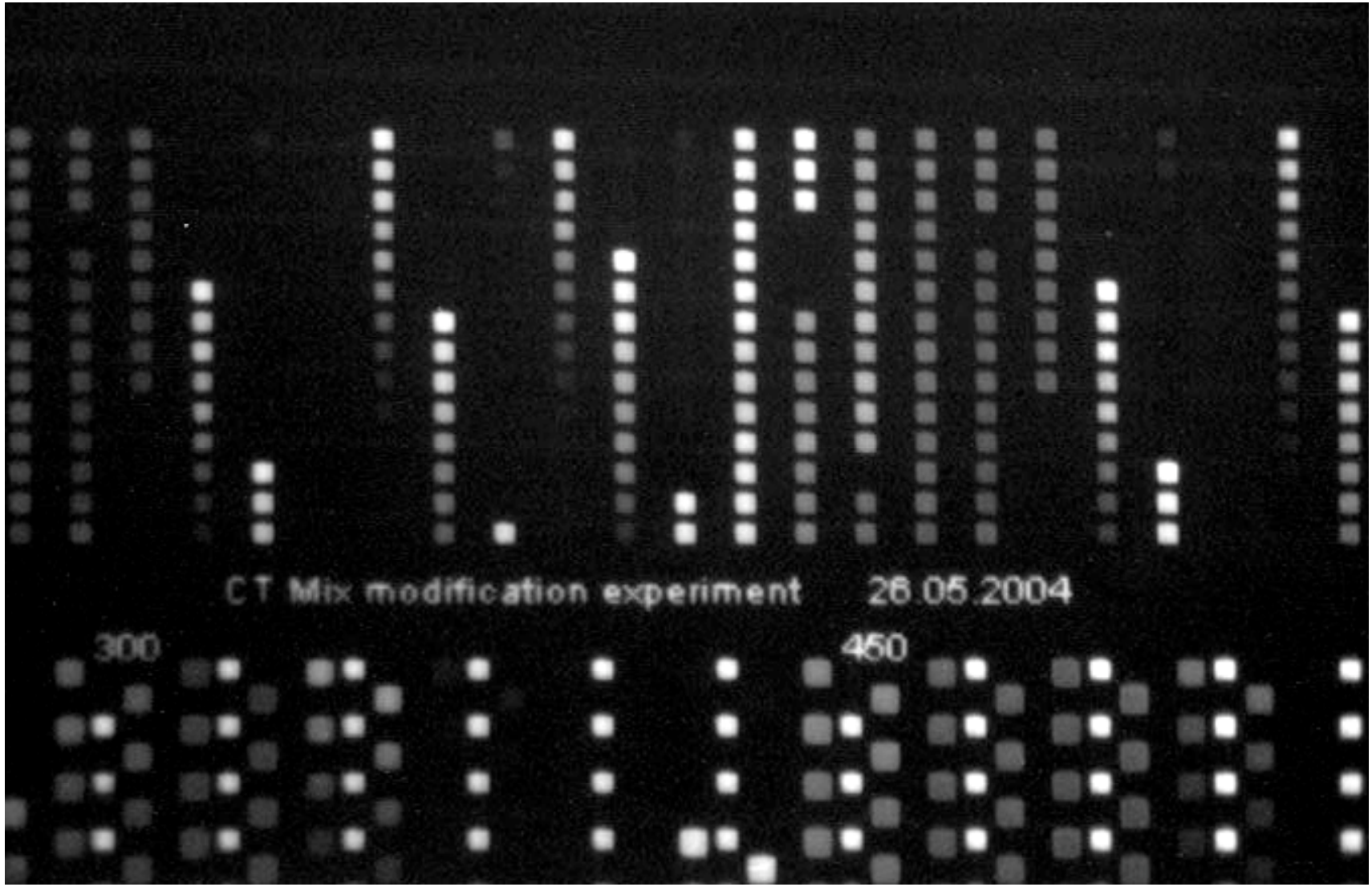
adapted from
Le Berre et al.

worked better than:

polyamidoamine (PAMAM)
dendrimers

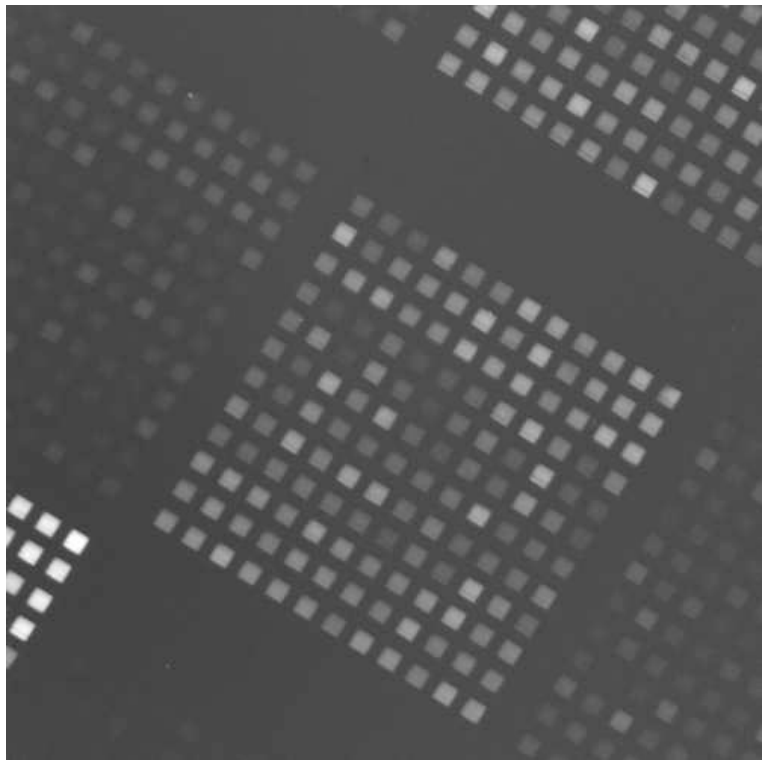
polyethylene glycol brushes
(PEG on epoxysilane - GPTS)

monohydroxsilane ....



Resolution about 3 micrometer
Reusable Array: 5€/experiment
Reproducibility better than 20 % (uncorrected)

Point defect study



Look at all the

single base mismatches

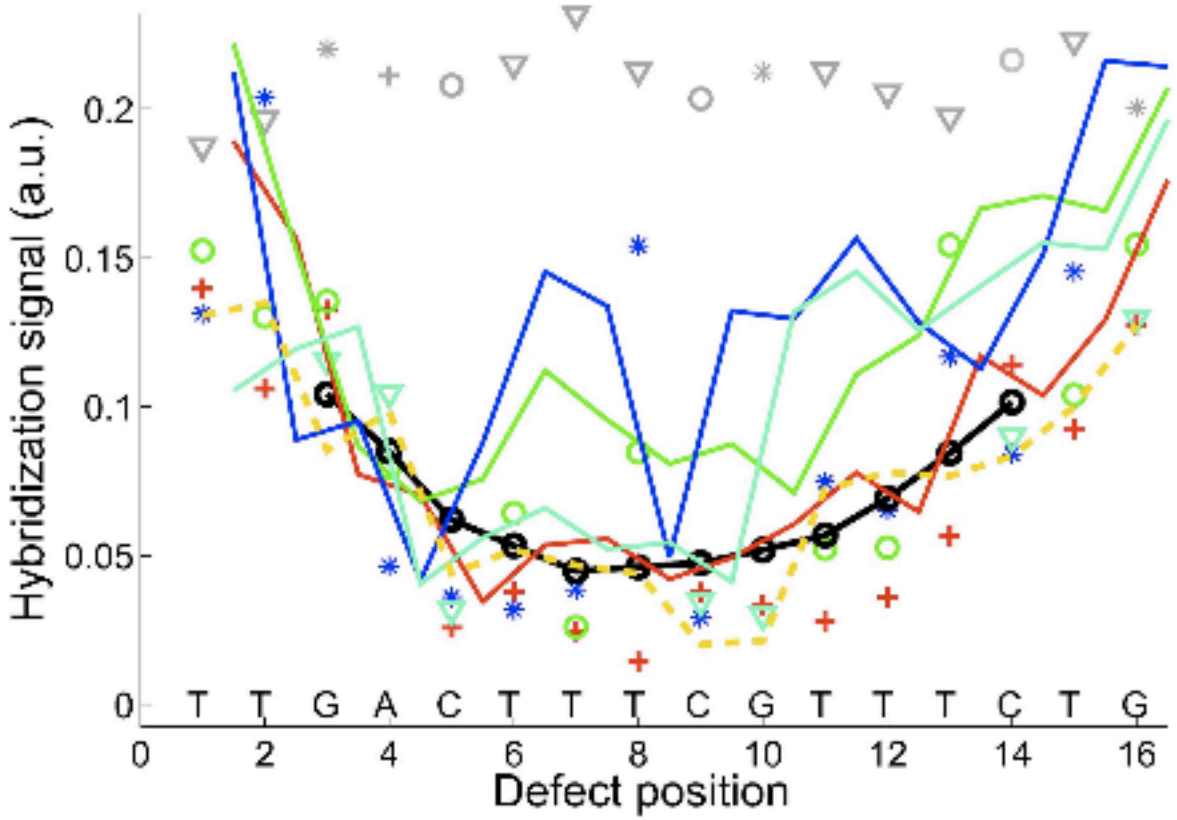
including bulges

of a given sequence

	Single base mismatch	Single base bulge
Probe	3'-TTGACTTTCGTTACTG-5'	3'-TTGACTTTCGTTTCTG-5'
Target	5'-AACTGAAAGCAAAGAC-3'	5'-AACTGAAAGCAAAGAC-3'

(A in Probe: mismatch; G in Probe: bulge)

Influence of the defect type and defect position on the hybridization signal intensity

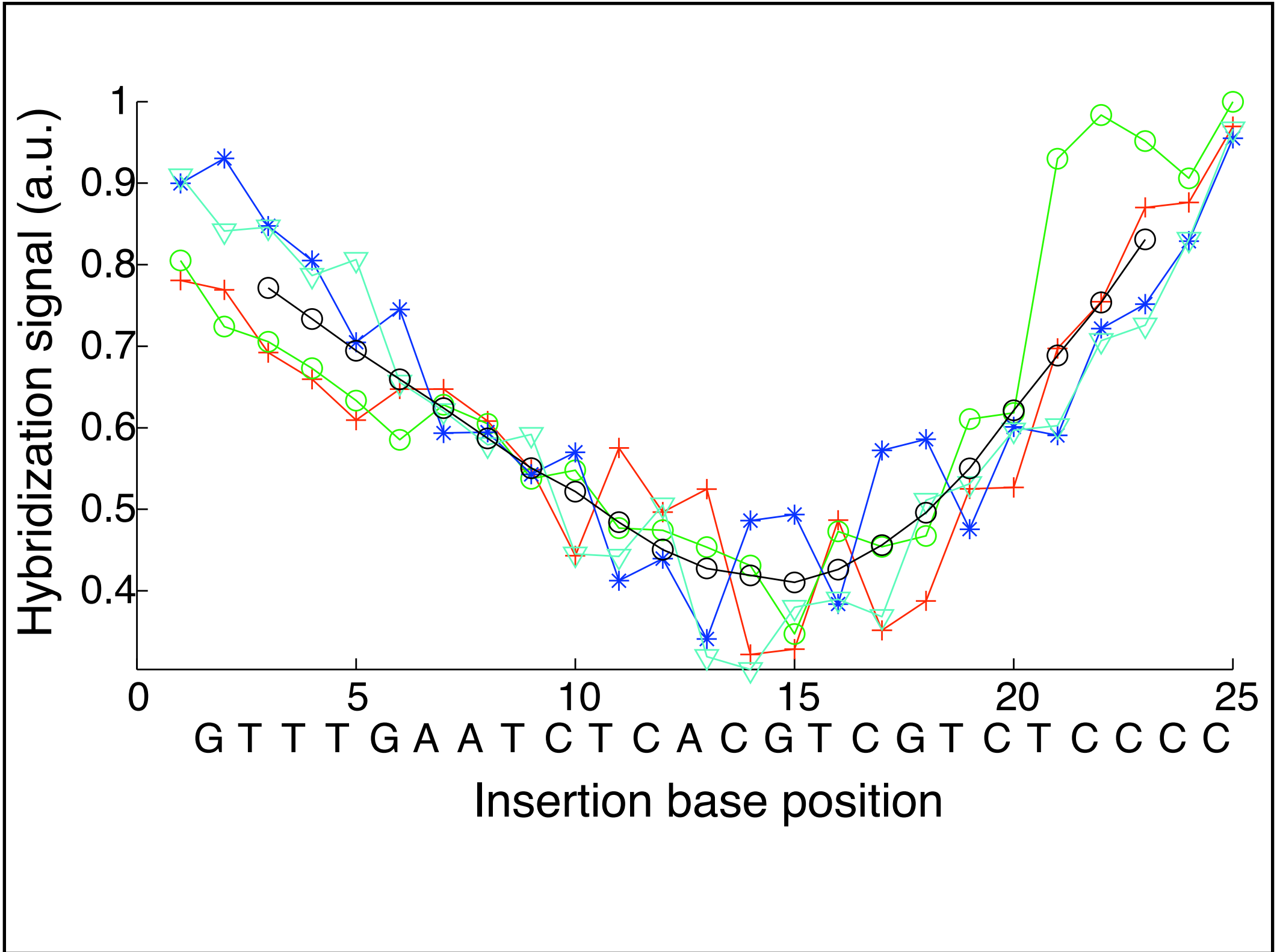


PM hybridization signals (*grey*)

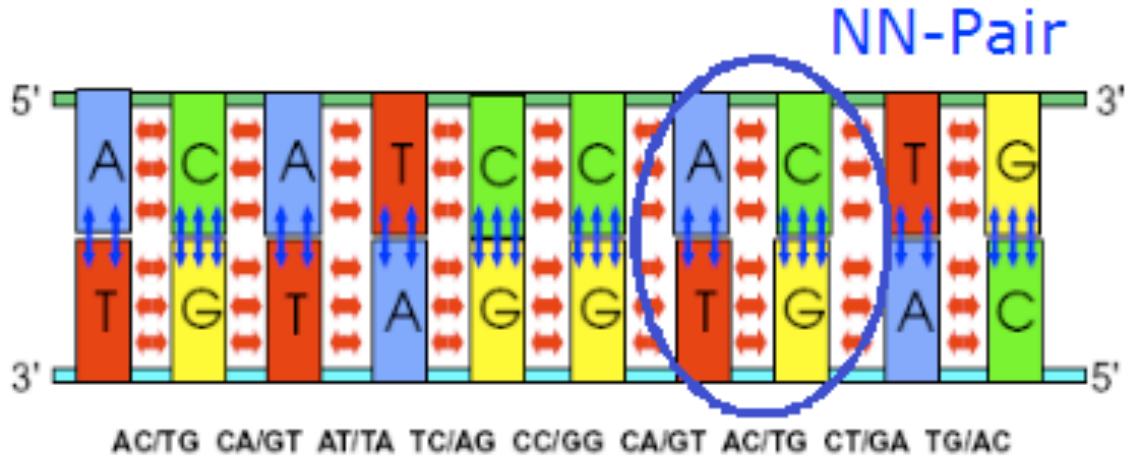
Base substitution → **MM**
- *A (red crosses)*
- *C (green circles)*
- *G (blue stars)*
- *T (cyan triangles)*

Base insertion → **bulge**
- *A (red line)*
- *C (green line)*
- *G (blue line)*
- *T (cyan line)*

Base deletion → **bulge**
- *dashed line*


Hybridization signal (a.u.)
1
0.9
0.8
0.7
0.6
0.5
0.4
0
5
10
15
20
25
G T T T G A A T C T C A C G T C G T C T C C C C
Insertion base position

Calculation of the (mismatched) nucleic acid duplex stability with the Nearest-Neighbor Model



$$\Delta G^\circ_{37}(\text{total}) =$$

$$= \Delta G^\circ_{37\,\text{init}} + \Delta G^\circ_{37\,\text{AC/TG}} + \Delta G^\circ_{37\,\text{CA/GT}} + \ldots + \Delta G^\circ_{37\,\text{TG/AC}} + \Delta G^\circ_{37\,\text{AT term}}$$

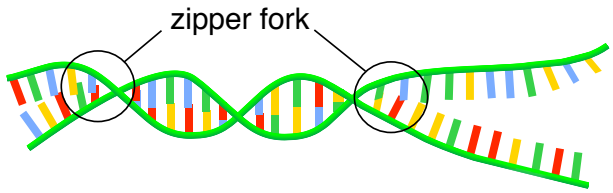
$$= (1.96 - 1.44 - 1.45 - 0.88 - 1.30 - 1.84 - 1.45 - 1.44 - 1.28 - 1.45 + 0.05)\,\frac{\text{kcal}}{\text{mol}}$$

$$= -10.52\,\frac{\text{kcal}}{\text{mol}}$$

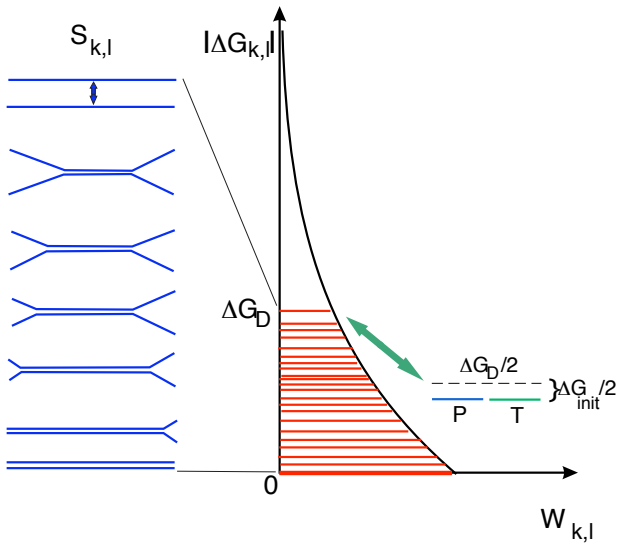
No positional influence !

Position Dependence & Molecular Zipping

Equilibrium: Double ended zipper partition function

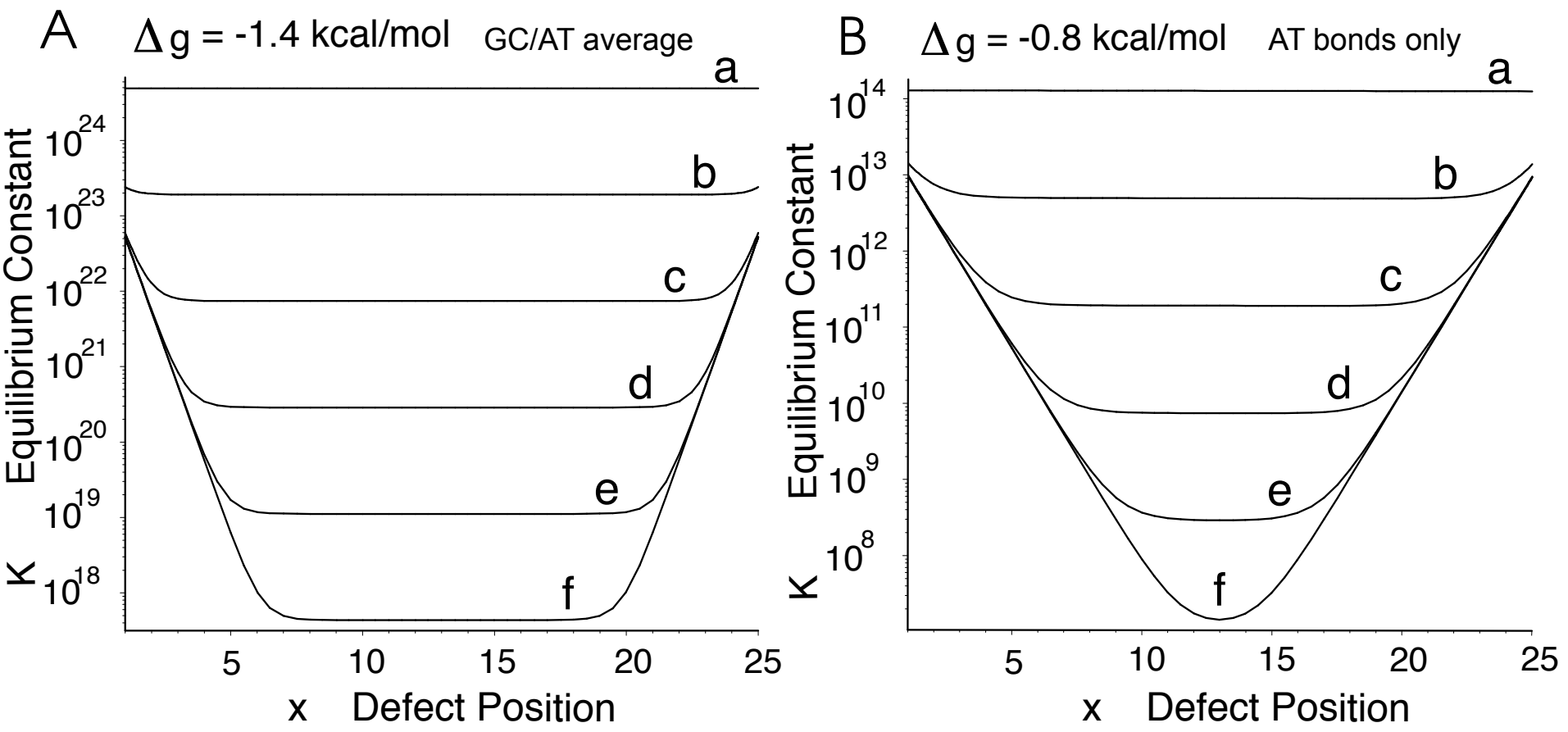


$$Z_D = \sum_{k=0}^{N-1} \sum_{l=k+1}^{N} w_{k,l} = \sum_{k=0}^{N-1} \sum_{l=k+1}^{N} e^{\Delta G^{\circ}_{k,l}/RT}$$



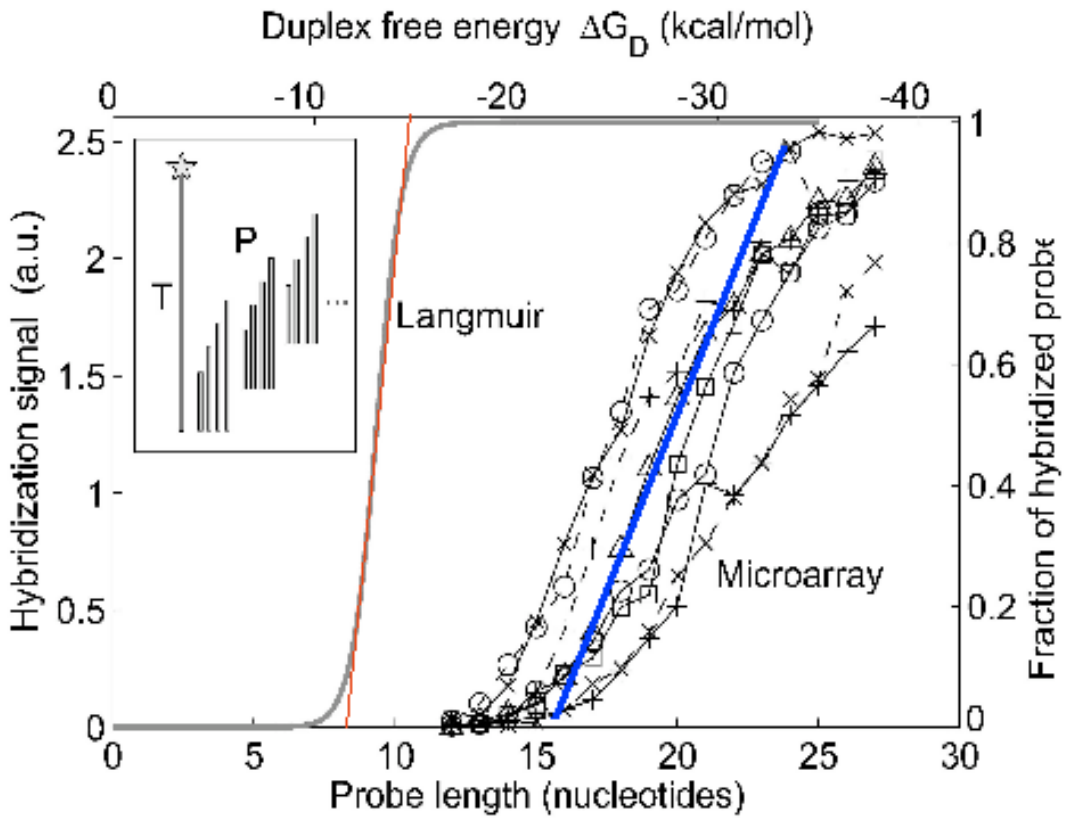
$\Delta G_{k,l}$ taken from bulk next neighbor energy tables

Binding constant for a single mismatch homopolymer



a) - f) different mismatch energies 0-5 kcal/mol

Array Signal vs Probe Length (~ΔG)



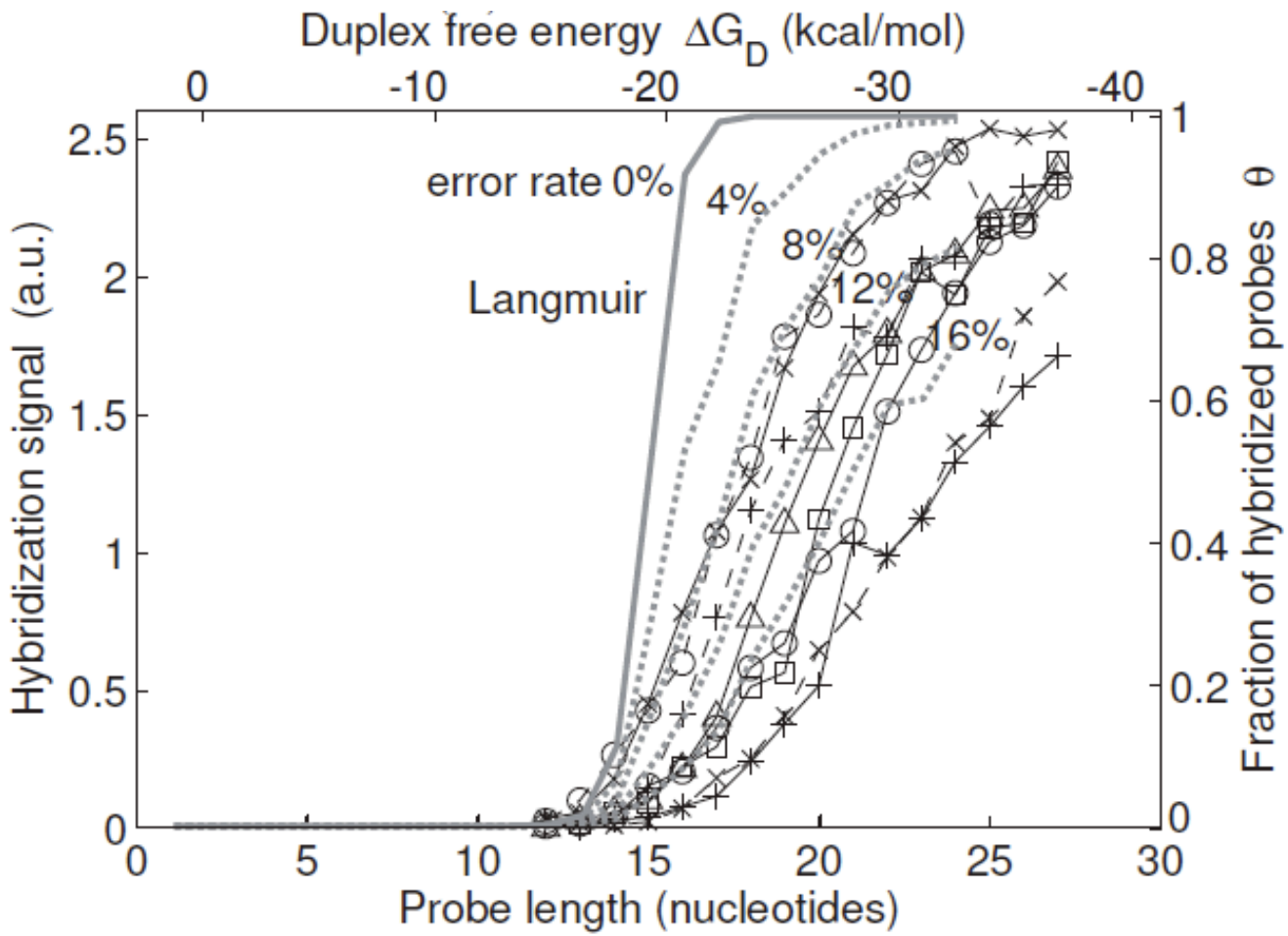
$$\theta_i = \frac{K_i \cdot [T_0]}{1 + K_i \cdot [T_0]}$$

Langmuir isotherm with a single binding affinity

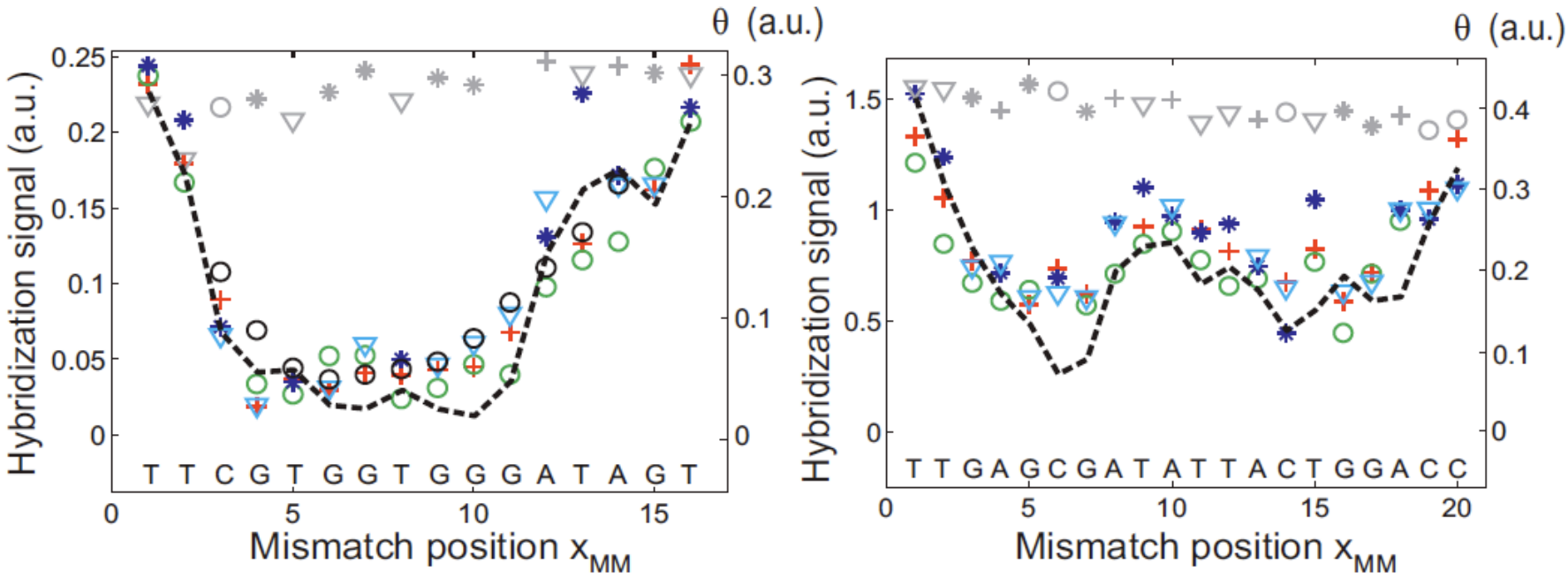
$$\theta_{total} = \sum_i^N x_i \theta_i = \sum_i^N x_i \frac{K_i \cdot [T_0]}{1 + K_i \cdot [T_0]}$$

Sips-like effective isotherm with multiple binding affinities

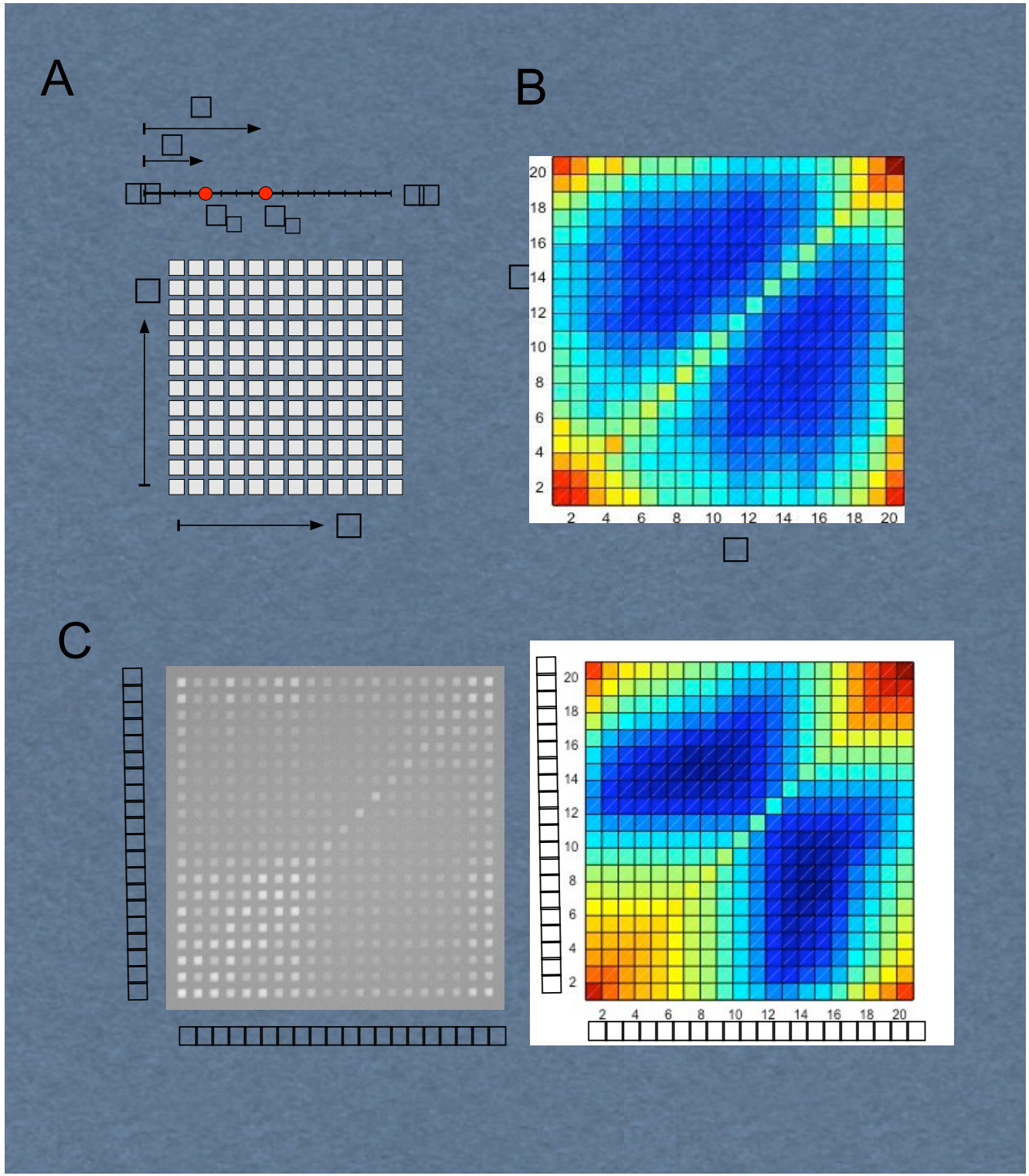
Error rate of about 10%/base & zipping dynamics: observed Sips type isotherm



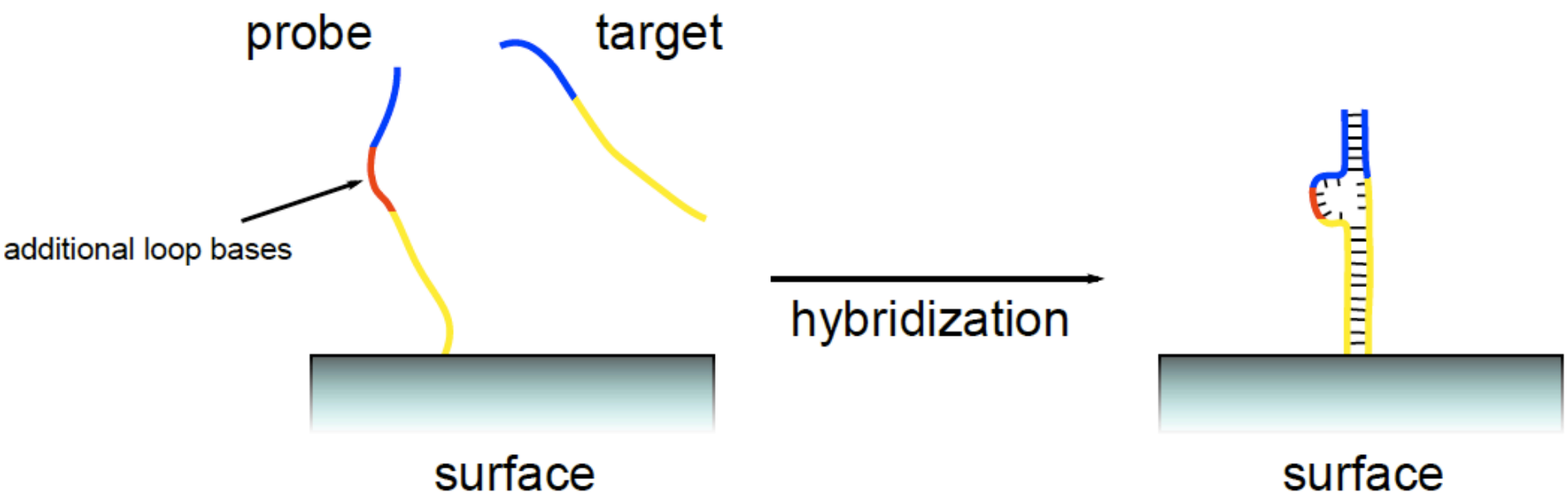
Zipping & 10% error/base



- Good agreement with experimental results
- ~1.5 kT energy penalty for the surface / bulk
- Can be mapped to Affymetrix array response and the PDNN (Position Dependent Next Neighbor) model


A
B
20
18
16
14
12
10
8
6
4
2
2 4 6 8 10 12 14 16 18 20
C
20
18
16
14
12
10
8
6
4
2
2 4 6 8 10 12 14 16 18 20

Loops


probe
target
additional loop bases
hybridization
surface
surface

Loops: Experiment and Theory

(Preliminary Results)

Signal averaged over all loop positions (loop length)

Signal averaged over all loop lengths (loop position)

Results

- **Synthesis Defects**

  ... raise the "effective temperature" of DNA hybridization

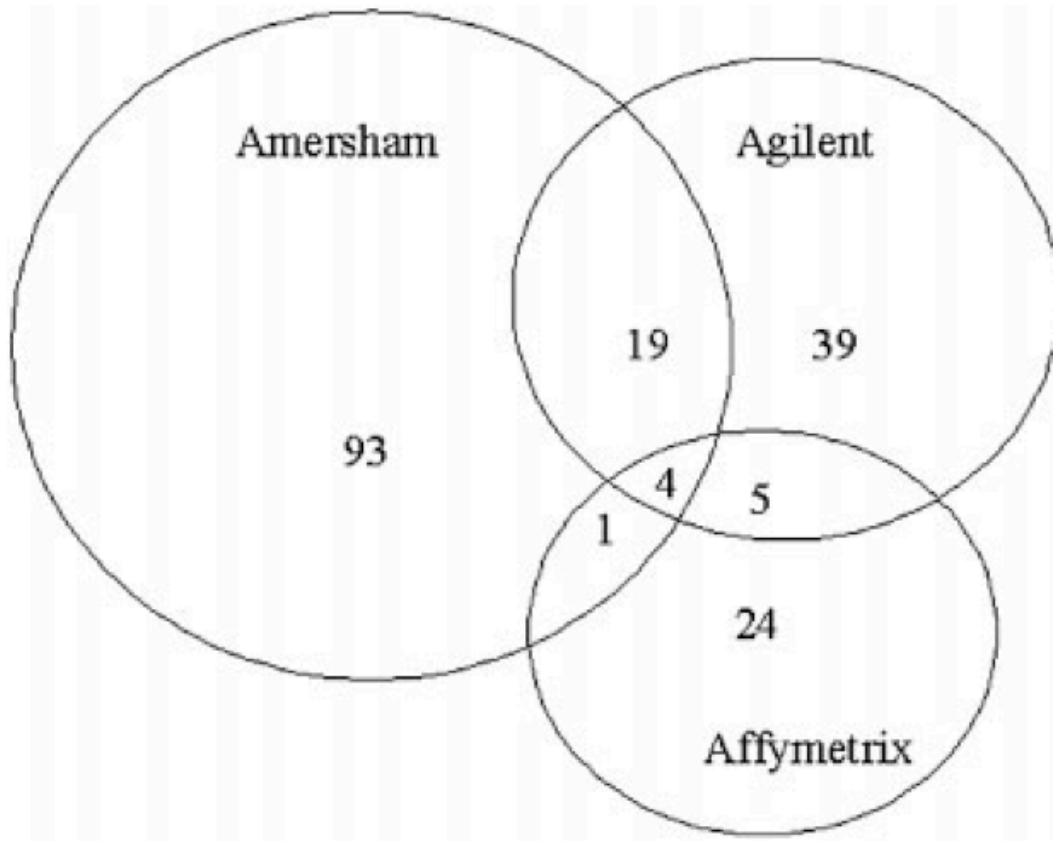
  ... compress the dynamic range &
  reduce differences in melting temperatures

  ... need to be taken into account for modelling

Results

- Equilibrium thermodynamics from bulk does work well for description
- There are (important) limitations to the array due to the physics of DNA

Differentially expressed genes: results from three commercial microarray platforms.



Tan PK, Downey TJ, Spitznagel EL Jr, Xu P, Fu D et al., 2003. Nucleic Acids Res. 31: 5676–5684

- Determine the amount of information possibly transmitted by an ideal array as a function of conditions
  (DNA is the limiting factor)
- See how close you can come with a real array...

Conclusion

- DNA microarray hybridization intensities can be predicted using parameters from bulk at thermal equilibrium (in the considered simple cases)
- DNA microarray technology is limited by the physics of DNA hybridization (among others)
- Ongoing project: determine the physical limit and compare to real arrays

Conclusion

- DNA hybridization in competition may give unexpected results

Collaborators

- Thomas Naiser (now TU Munich)
- Christian Trapp
- Marc Schenkelberger, Timo Mai

Thank you for listening !