Physico-chemical model of surface hybridisation and post-hybridisation washing

Conrad Burden

Mathematical Sciences Institute
Australian National University, Canberra

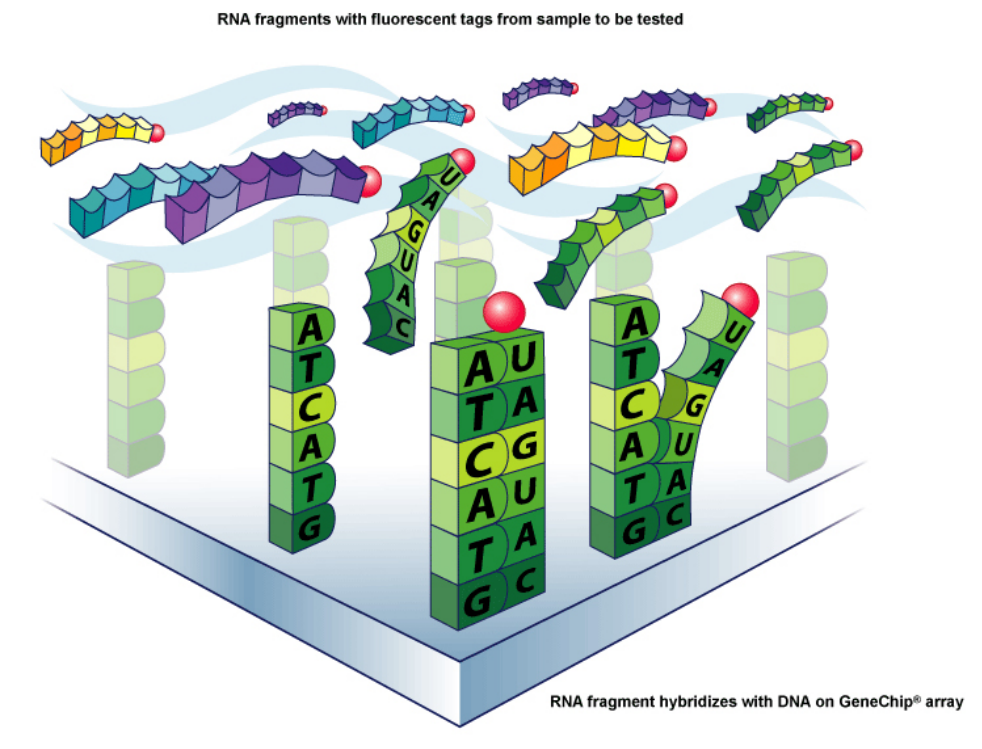

Canberra

A comprehensive model of hybridisation plus post-hybridisation washing is set out in

C.J. Burden and H. Binder: "*Physico-chemical modelling of target depletion during hybridization on oligonucleotide microarrays*", Phys. Biol. **6**, (2010) 016004

Basic idea: it's all equilibrium physical chemistry driven by the law of mass-action

$$\text{PROBE} + \text{TARGET} \underset{\text{DESORPTION}}{\overset{\text{ADSORPTION}}{\rightleftharpoons}} \text{DUPLEX}$$

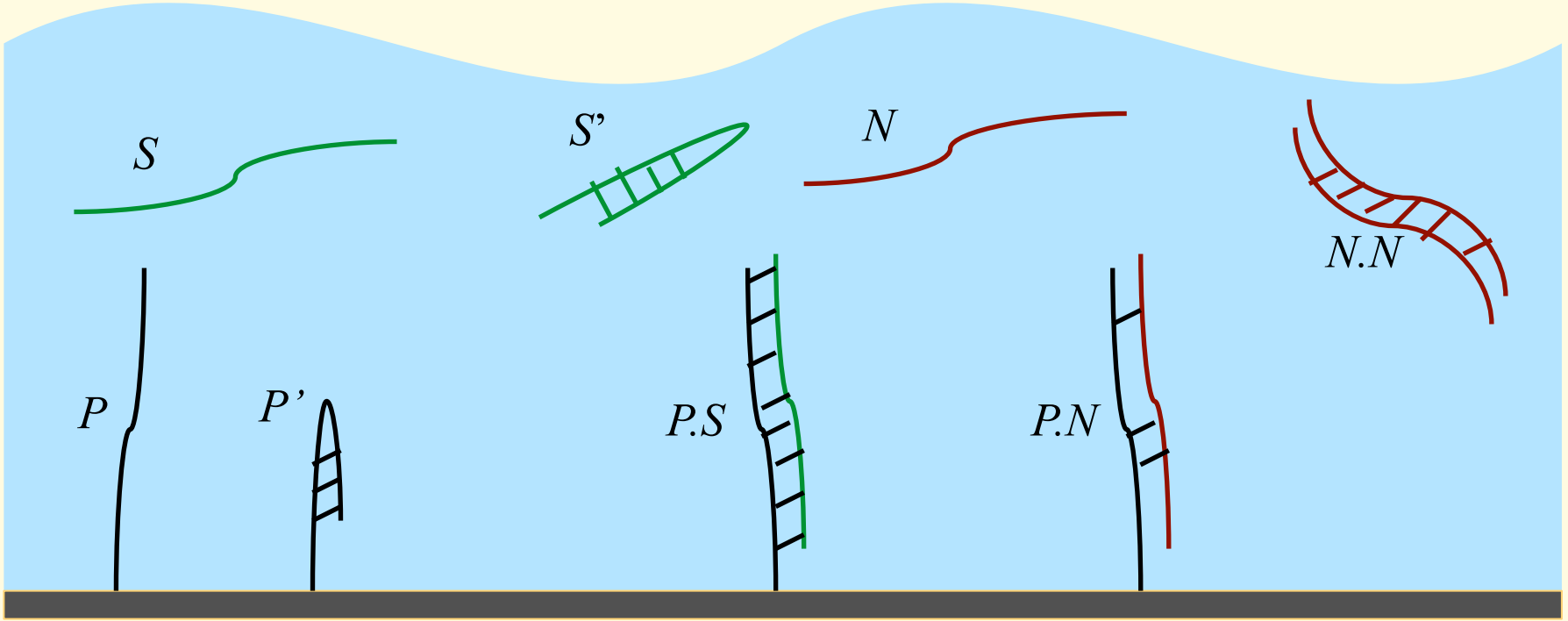


... plus all the other chemical reactions involving nonspecific hybridisation, hybridisation in bulk solution, etc.

Image courtesy of Affymetrix

Table A1. Chemical species present in the model.

	Unfolded	Folded
Specific target in the solution	S	S'
Non-specific effective target in solution	N	N'
Probe at the surface (not bound to the target)	P	P'
Duplexes in the solution	$S \cdot S$, $S \cdot N$, $N \cdot N$	
Duplexes at the microarray surface	$P \cdot S$, $P \cdot N$	



Assume a set of chemical reactions r each in equilibrium,
with equilibrium constant K_r:

Folding

$$S \rightleftharpoons S' : \qquad [S'] = K_{\text{Sfold}}[S].$$
$$N \rightleftharpoons N' : \qquad [N'] = K_{\text{Nfold}}[N].$$
$$P \rightleftharpoons P' : \qquad [P'] = K_{\text{Pfold}}[P].$$

Bulk hybridization

$$S + N \rightleftharpoons S \cdot N : \qquad [S \cdot N] = K_{\text{SN}}[S][N].$$
$$S + S \rightleftharpoons S \cdot S : \qquad [S \cdot S] = K_{\text{SS}}[S]^2.$$
$$N + N \rightleftharpoons N \cdot N : \qquad [N \cdot N] = K_{\text{NN}}[N]^2.$$

Surface hybridization

$$P + S \rightleftharpoons P \cdot S : \qquad [P \cdot S] = K_{\text{PS}}[P][S].$$
$$P + N \rightleftharpoons P \cdot N : \qquad [P \cdot N] = K_{\text{PN}}[P][N].$$

Assume a set of chemical reactions r each in equilibrium,
with equilibrium constant K_r:

Folding

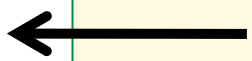
$$S \rightleftharpoons S' : \qquad [S'] = K_{\text{Sfold}}[S].$$
$$N \rightleftharpoons N' : \qquad [N'] = K_{\text{Nfold}}[N].$$
$$P \rightleftharpoons P' : \qquad [P'] = K_{\text{Pfold}}[P].$$

Bulk hybridization

$$S + N \rightleftharpoons S \cdot N : \qquad [S \cdot N] = K_{\text{SN}}[S][N].$$
$$S + S \rightleftharpoons S \cdot S : \qquad [S \cdot S] = K_{\text{SS}}[S]^2.$$
$$N + N \rightleftharpoons N \cdot N : \qquad [N \cdot N] = K_{\text{NN}}[N]^2.$$

Surface hybridization

$$P + S \rightleftharpoons P \cdot S : \qquad [P \cdot S] = K_{\text{PS}}[P][S].$$
$$P + N \rightleftharpoons P \cdot N : \qquad [P \cdot N] = K_{\text{PN}}[P][N].$$



In addition to the K_r the input parameters are

- specific target concentration

$$x_S = [S] + [S'] + [P.S] + [S.N] + [S.S]$$

(= spike-in conc. in spike-in experiments)

- effective non-specific target concentration

$$x_N = [N] + [N'] + [P.N] + [S.N] + [N.N]$$

- effective probe concentration on surface of feature

$$p = [P] + [P'] + [P.S] + [P.N]$$

The aim (in the first instance) is to determine the coverage fraction ($0 \leq \theta \leq 1$) on a particular feature of the microarray of fluorescent-dye carrying duplexes

$$\theta = \theta_S + \theta_N = \frac{[P \cdot S]}{p} + \frac{[P \cdot N]}{p}$$

in terms of the input parameters x_S, x_N, p, K_r ($r = PS, PN, SN, \ldots$).

From this the observed flourescence intensity is simply

$$I(x_S) = a + b\theta$$

physical background

saturation intensity

Solving this system for $\theta(x_S)$ is straightforward algebra:

$$\theta(x_S) = \frac{X_N + K_S(x_S - p\theta_{\text{sum}})}{1 + X_N + K_S(x_S - p\theta_{\text{sum}})}$$

where θ_{sum} is the solution to

$$\theta_{\text{sum}} = \sum_{\text{features } f} \frac{K_S^f(x_S - p\theta_{\text{sum}})}{1 + K_S^f(x_S - p\theta_{\text{sum}})}$$

with the sum taken over all features contributing to the depletion of specific target species S.

Solving this system for $\theta(x_S)$ is straightforward algebra:

$$\theta(x_S) = \frac{X_N + K_S(x_S - p\theta_{\text{sum}})}{1 + X_N + K_S(x_S - p\theta_{\text{sum}})}$$

Note:

- Three parameters:

  K_S (effective equilibrium const.) depends on x_N and K_r ($r = PS, PN, SN, \ldots$)

  X_N (non-specific binding strength) depends on x_N and K_r ($r = PS, PN, SN, \ldots$)

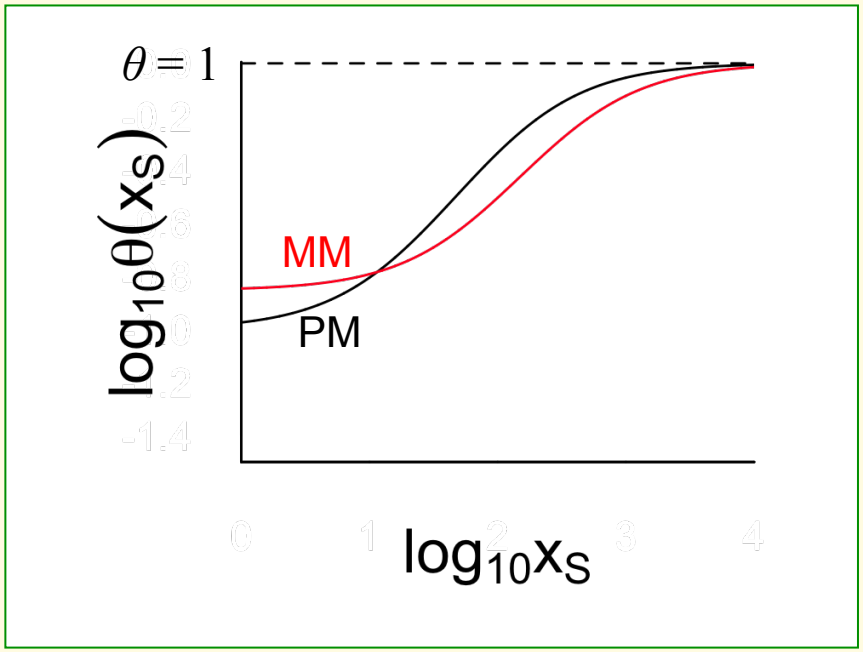
  p (probe density)

- If depletion of target from supernatant solution can be ignored, $p = 0$, and $\theta(x_S)$ is a rectangular hyperbola;
- If $x_N = 0$ then $X_N = 0$ and $\theta(0) = 0$;
- At saturation spec. target concentration $x_S \rightarrow \infty$, $\theta(x_S) \rightarrow 1$.

Saturation asymptote $\theta(\infty) = 1$ does not agree with spike-in experiments:

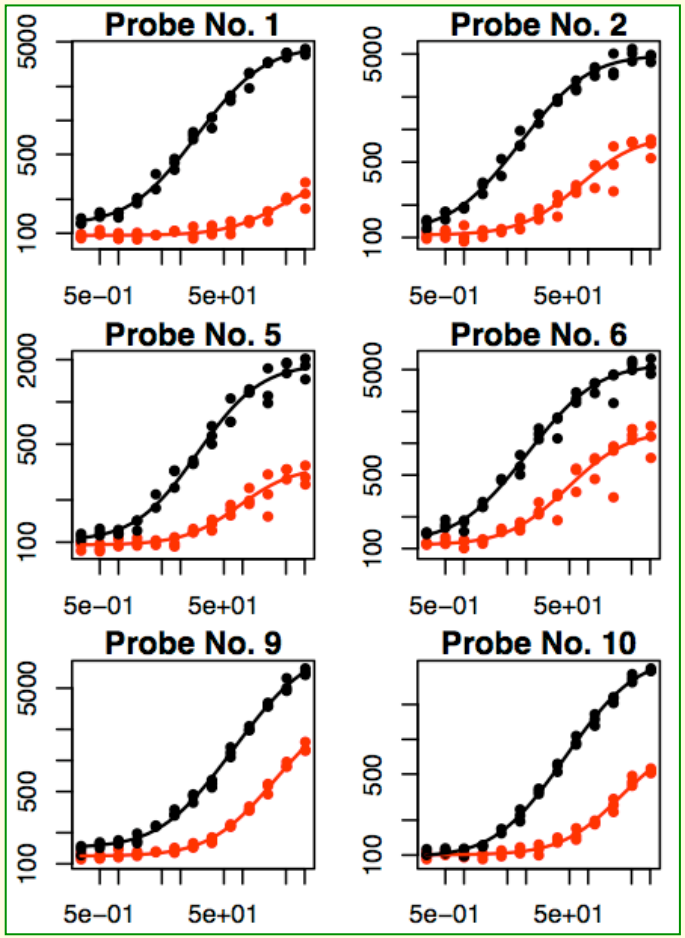
For PM/MM pair, theory gives

$$\theta^{PM}(\infty) = \theta^{MM}(\infty) = 1$$



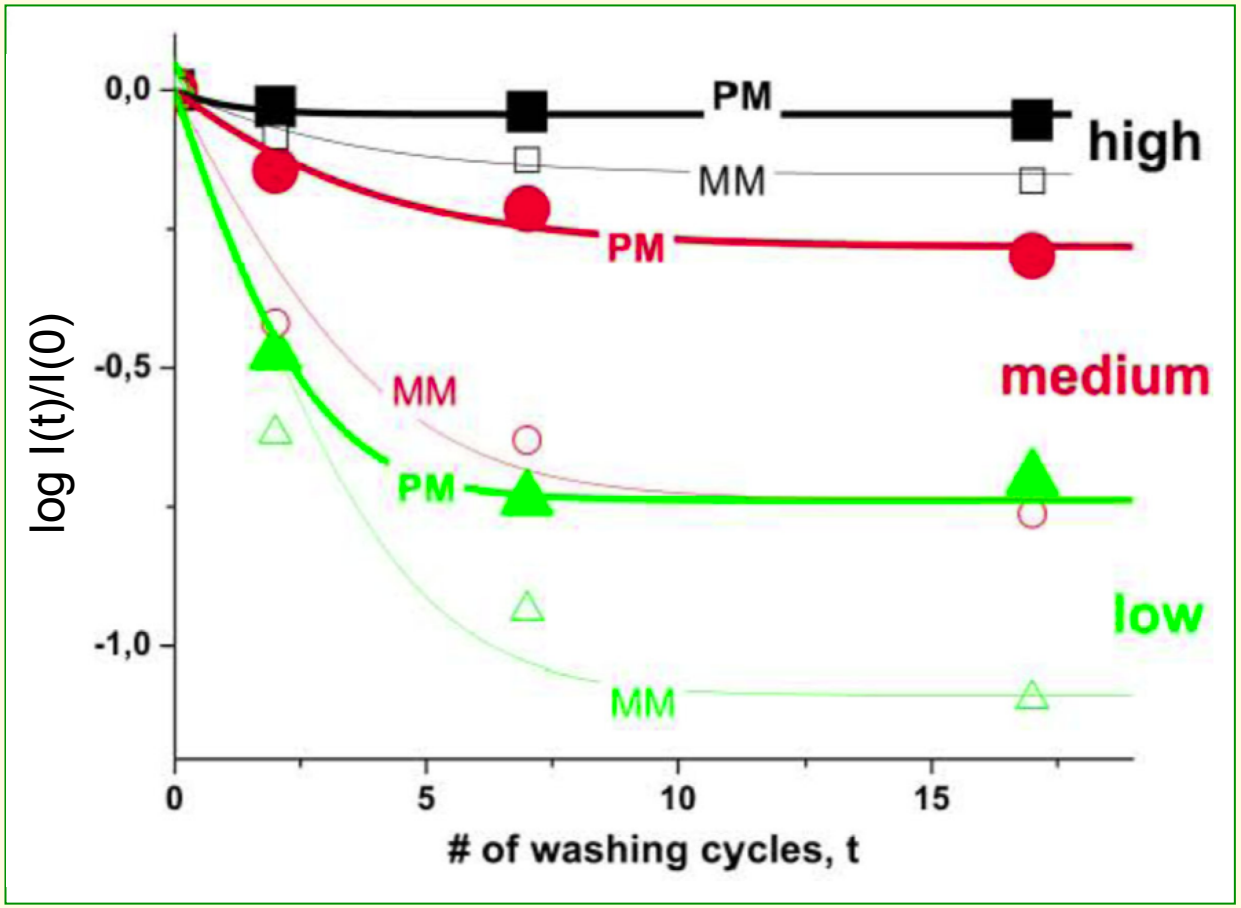
Spike-in experiments *always* give

$$\theta^{PM}(\infty) > \theta^{MM}(\infty)$$

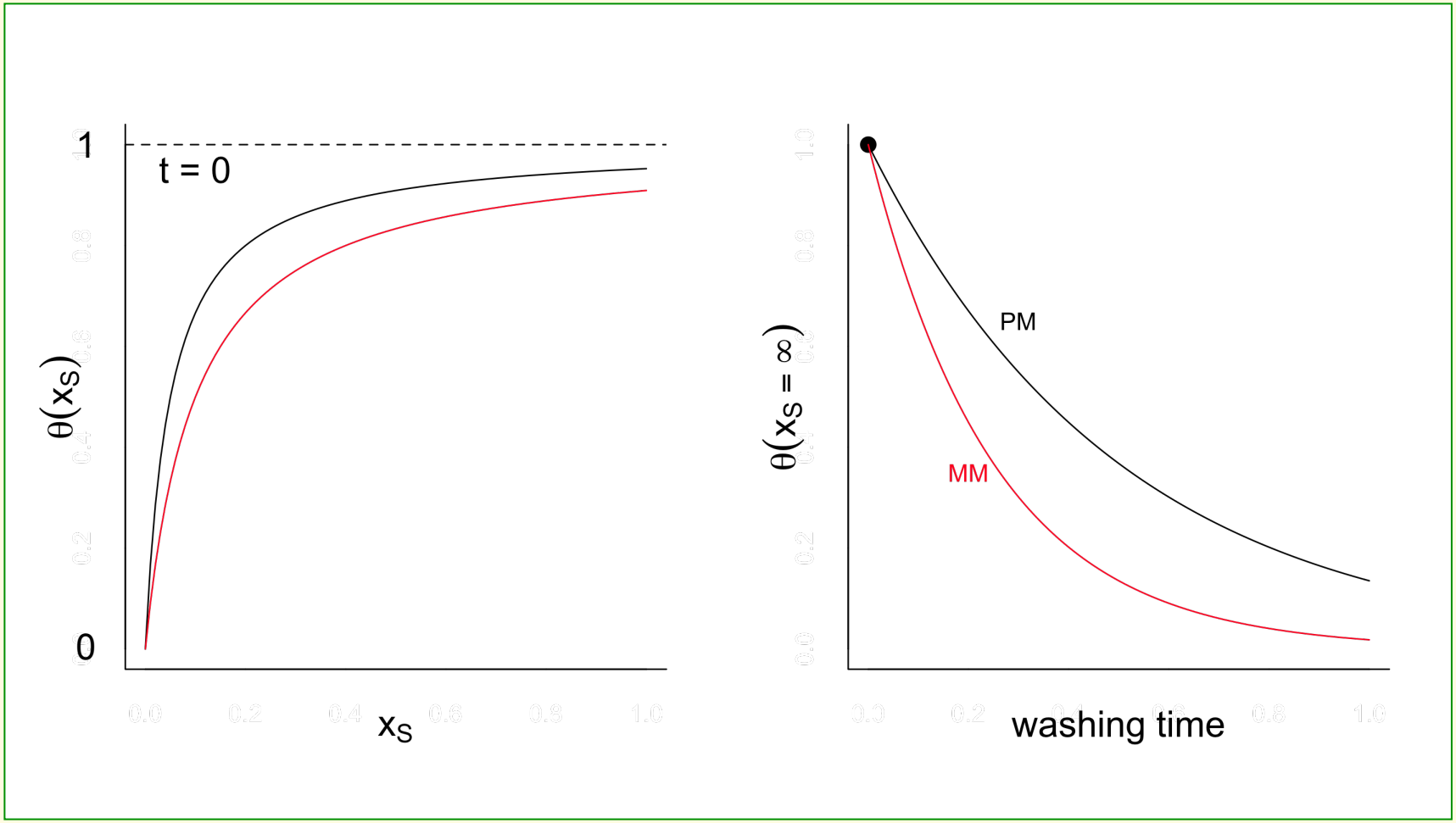


Differing asymptotes explained by post-hybridisation washing:

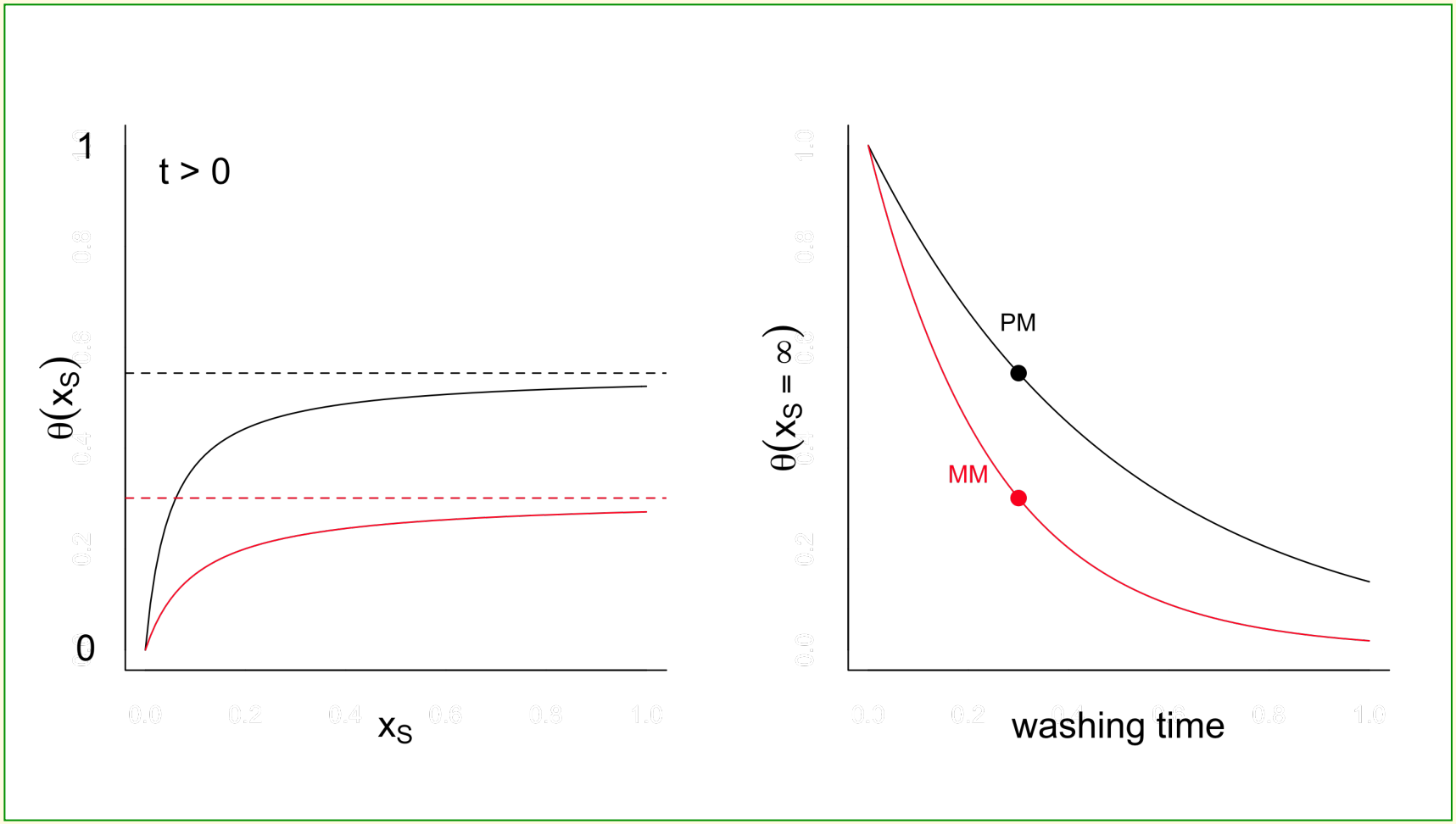
Decaying fluorescence intensity with washing is confirmed by experiment
(Binder, H. et al. BMC Bioinformatics **11**:291 (2010))



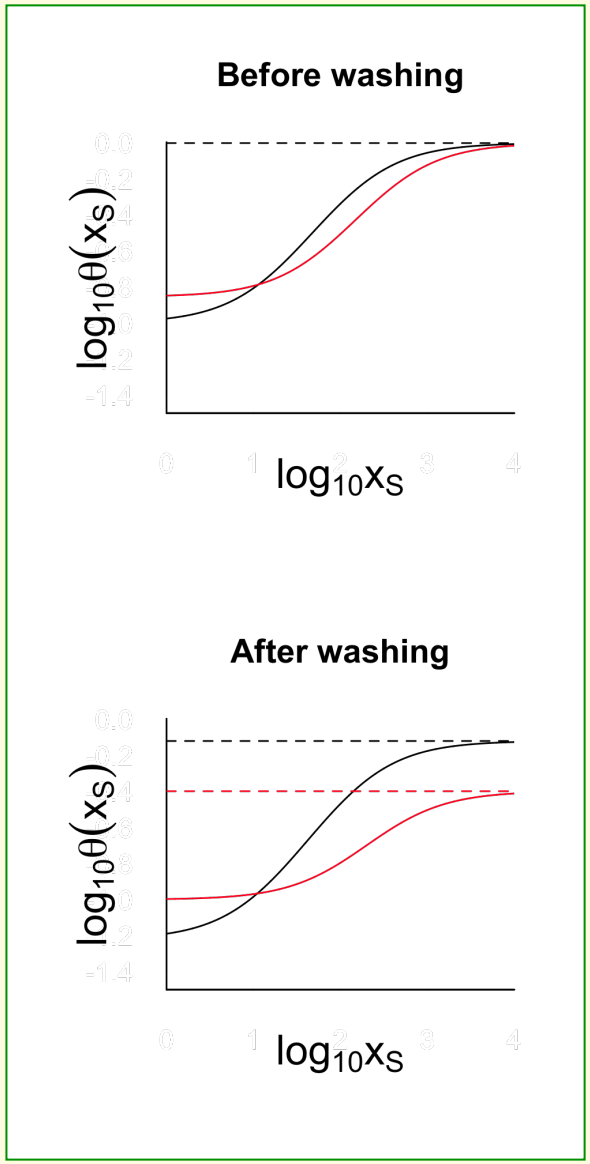
Differing asymptotes explained by post-hybridisation washing



Differing asymptotes explained by post-hybridisation washing

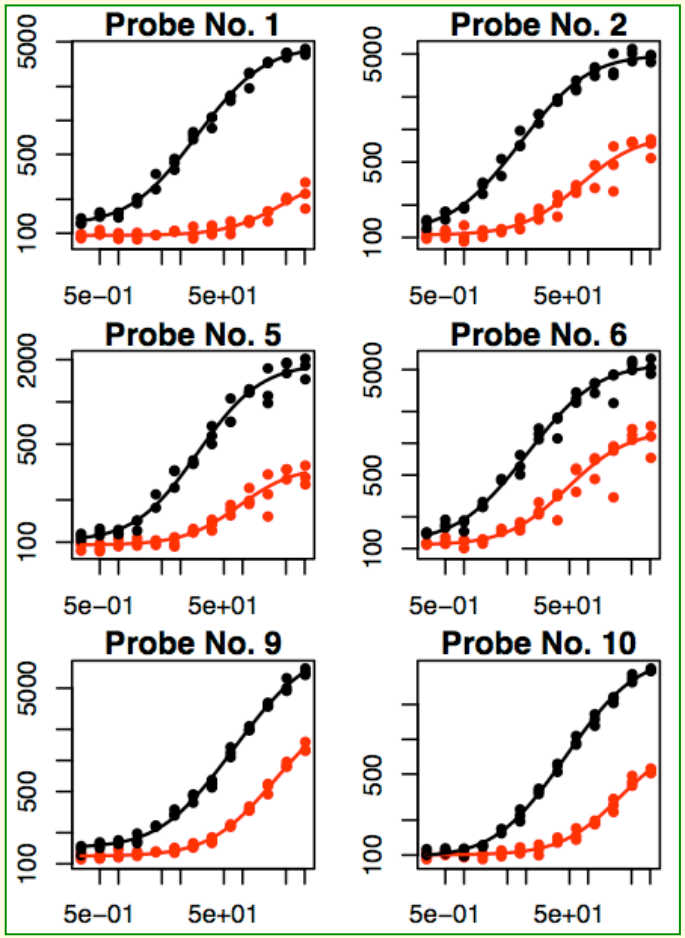


Saturation asymptote $\theta(\infty) = 1$ does not agree with experiment:



Spike-in experiments *always* give

$\theta^{PM}(\infty) > \theta^{MM}(\infty)$



Putting it all together, the fluorescence intensity on a given feature is related to the specific target concentration by a 4-parameter 'isotherm':

$$I(x_S) = A + B\frac{K(x_S - p\theta_{\text{sum}})}{1 + K(x_S - p\theta_{\text{sum}})}$$

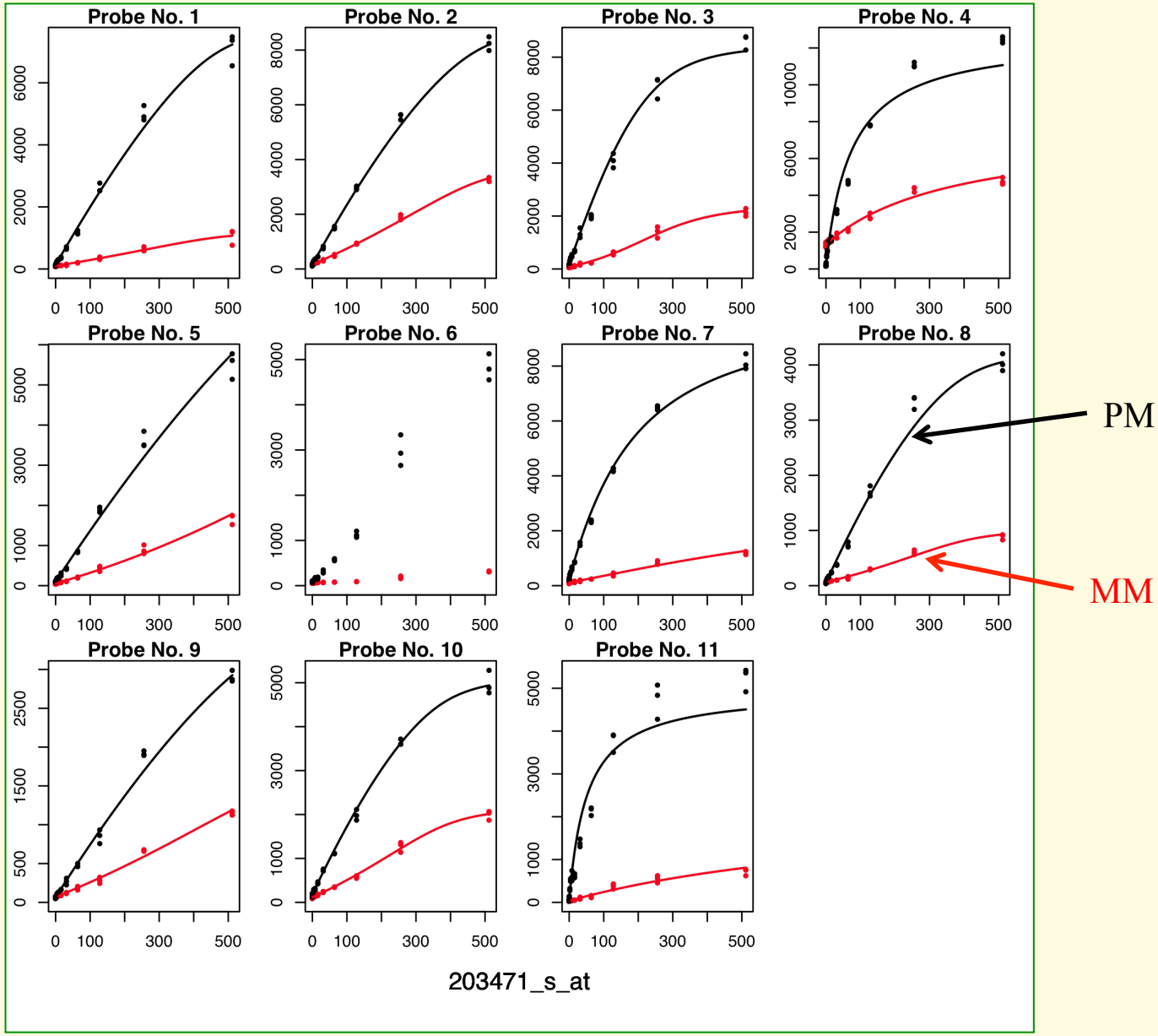
where θ_{sum} is the solution to

$$\theta_{\text{sum}} = \sum_{\text{features } f} \frac{K^f(x_S - p\theta_{\text{sum}})}{1 + K^f(x_S - p\theta_{\text{sum}})}$$

with the sum taken over all features contributing to the depletion of specific target species S.

The 4 fitting parameters A, B, K and p can be fitted well to spike-in data ...

Example of fits to Affy U133a spikes:



$$I(x_S) = A + B\frac{K(x_S - p\theta_{\text{sum}})}{1 + K(x_S - p\theta_{\text{sum}})}$$

The fitting parameters

A, B, K and p

are complicated functions of physical input parameters

x_N, K_r ($r = PS, PN, SN, \ldots$), p, a, b, w_S, w_N.

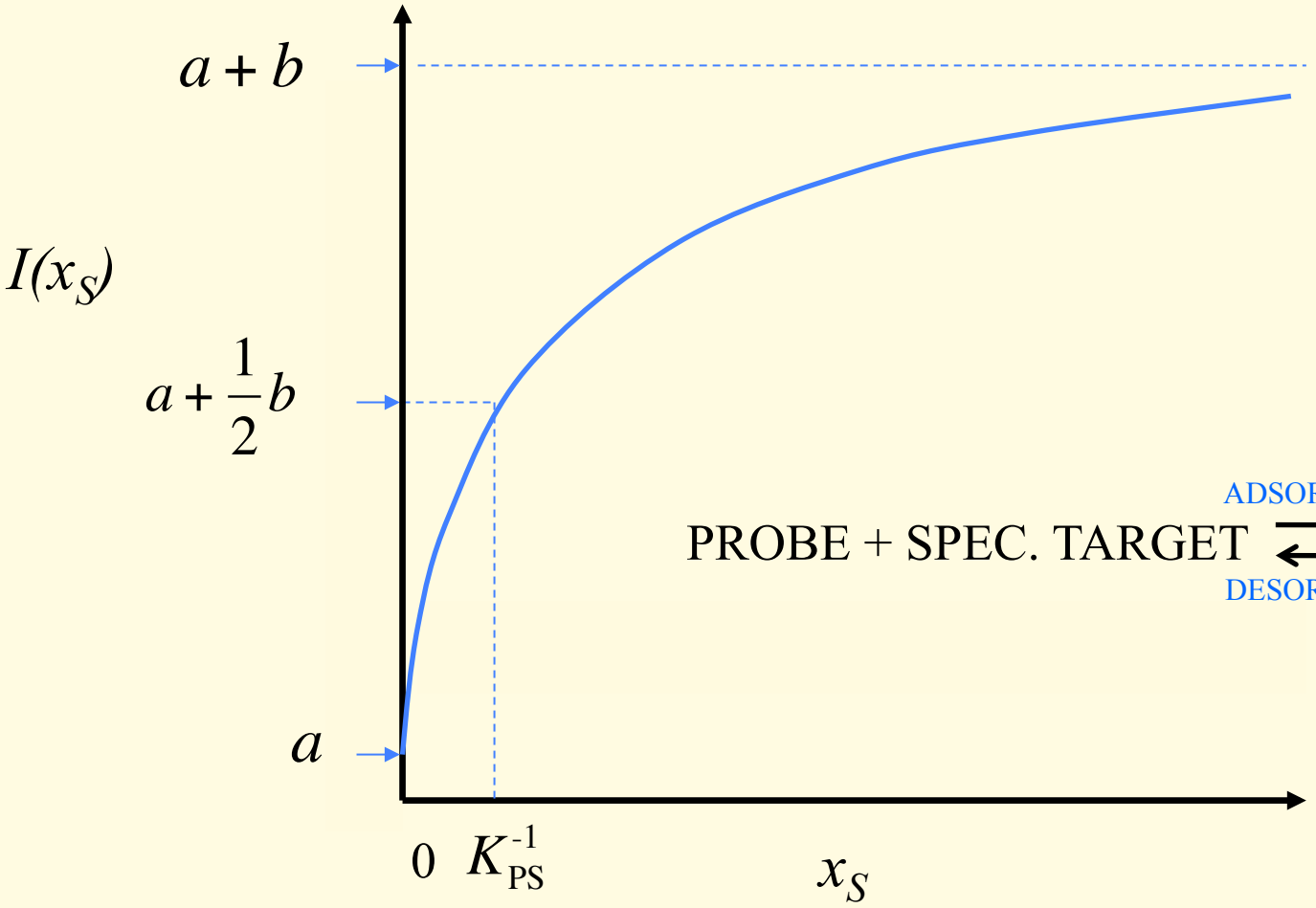
Is it practical to 'predict' these parameters from the available data (probe sequences, distribution of intensities across microarray, ...)?

Maybe/Maybe not, but the model does give a physical understanding which helps with building algorithms (Hook curve, Inverse Langmuir , ...)

Understanding the model bit by bit:

1. With no non-spec. hybridisaton, target depletion or washing, 'Langmuir isotherm' is:

$$I(x_S) = a + b\frac{K_{PS}x_S}{1 + K_{PS}x_S}$$



ADSORPTION

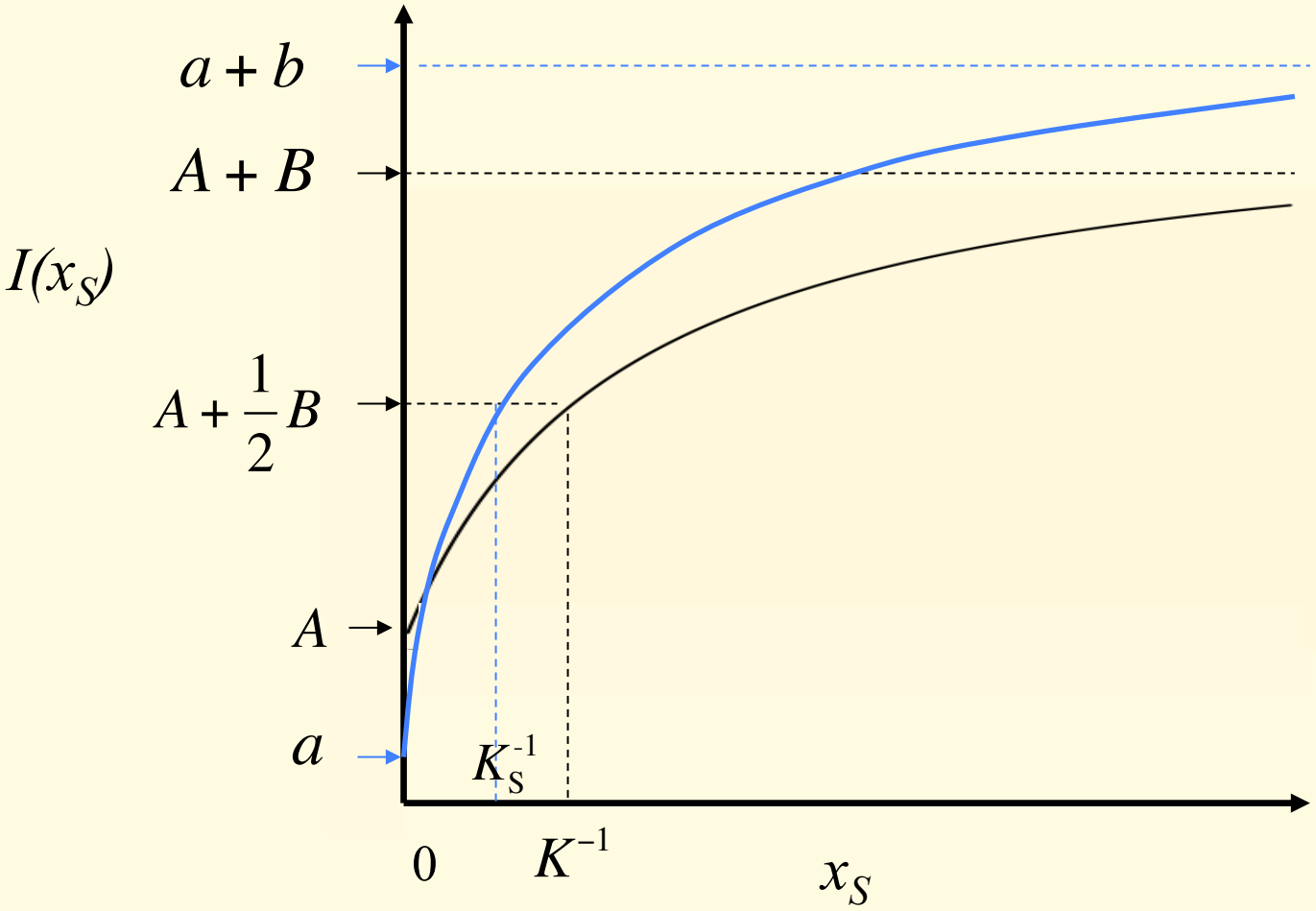
PROBE + SPEC. TARGET ⇌ DUPLEX

DESORPTION

Understanding the model bit by bit:

2. Switch on non-specific hybridisation and washing

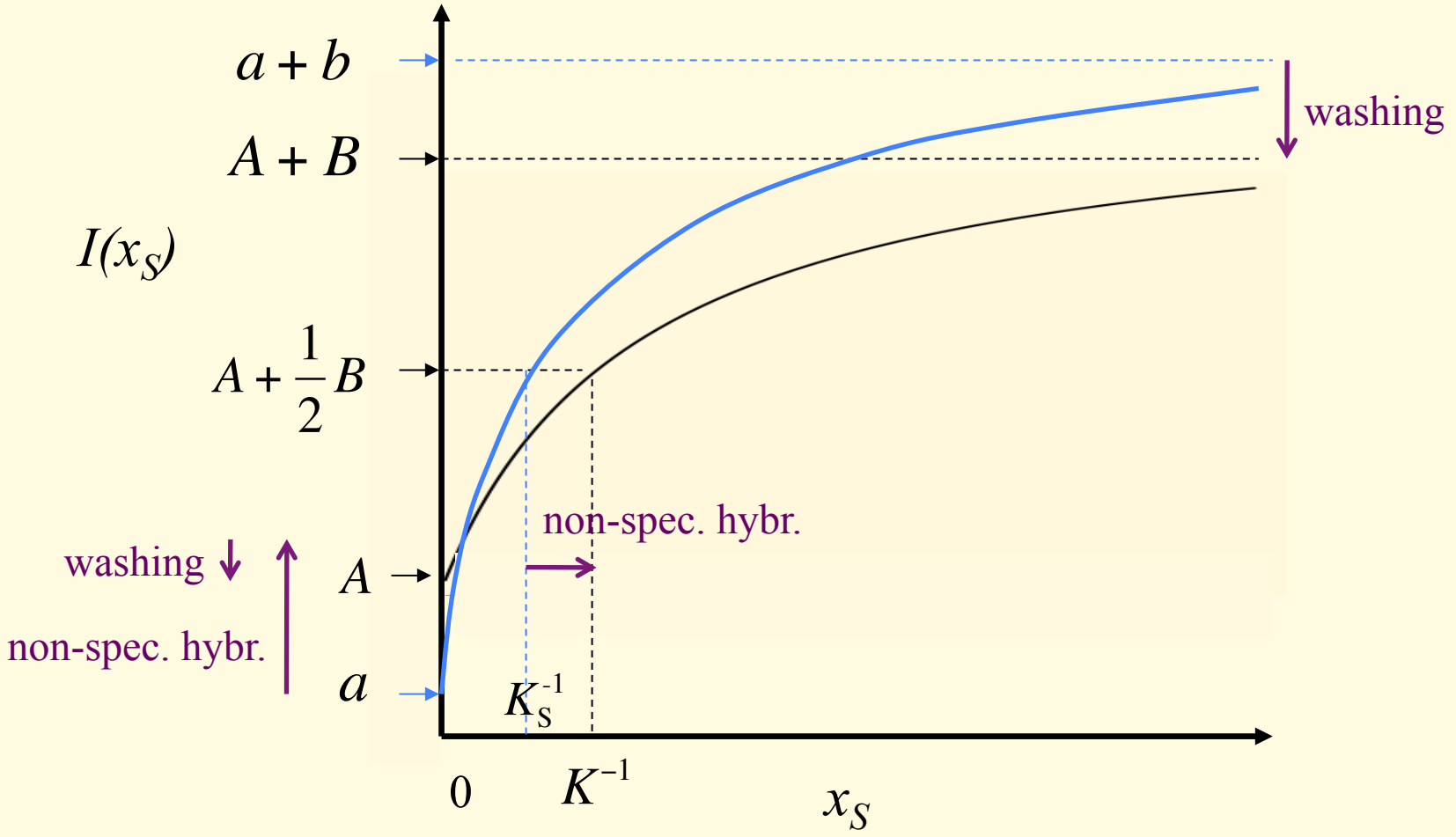
$$I(x_S) = A + B\frac{Kx_S}{1+Kx_S}$$



Understanding the model bit by bit:

2. Switch on non-specific hybridisation and washing

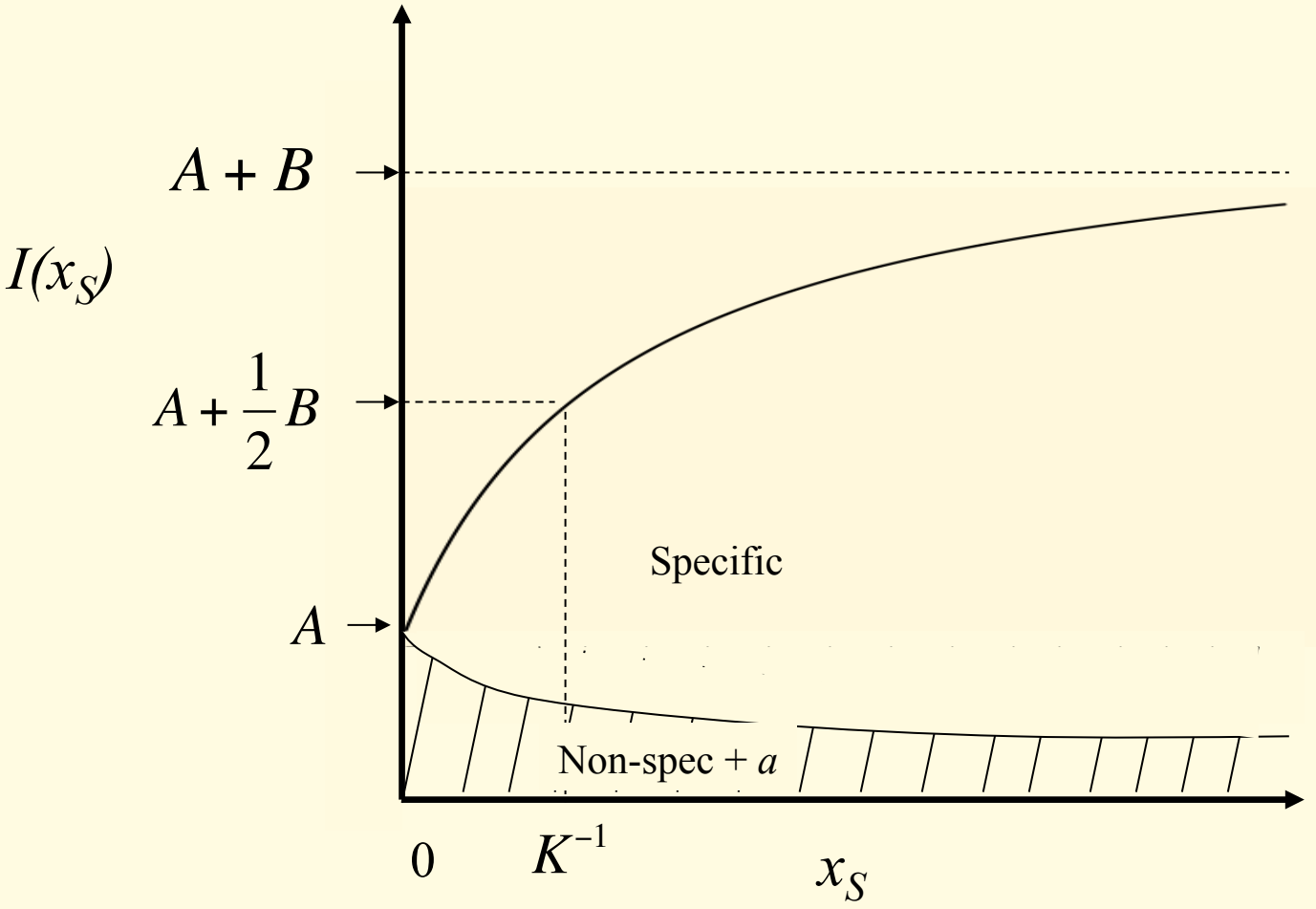
$$I(x_S) = A + B\frac{Kx_S}{1+Kx_S}$$



Understanding the model bit by bit:

2. Switch on non-specific hybridisation and washing

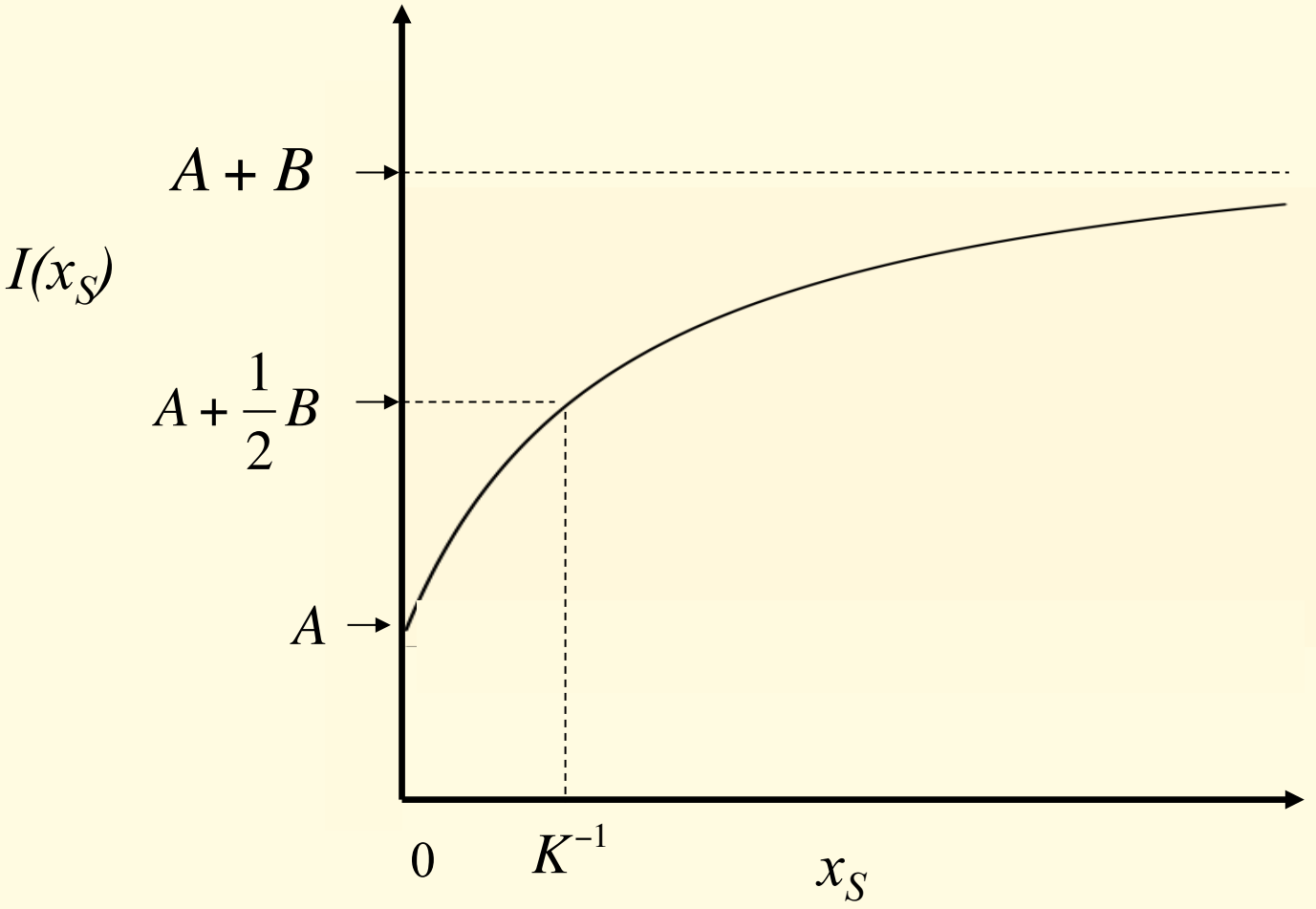
$$I(x_S) = A + B\frac{Kx_S}{1+Kx_S}$$



Understanding the model bit by bit:

2. Switch on non-specific hybridisation and washing

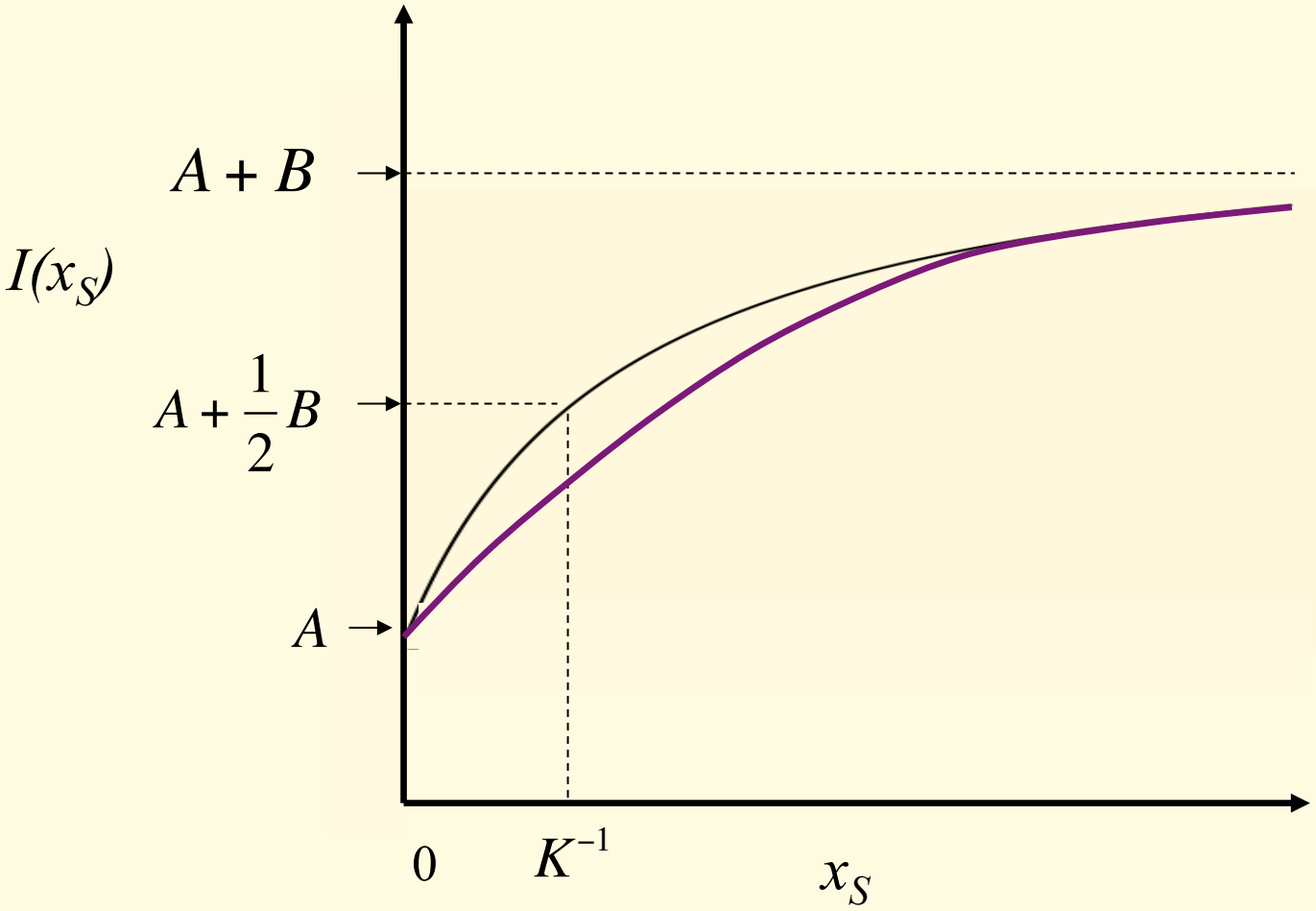
$$I(x_S) = A + B\frac{Kx_S}{1+Kx_S}$$



Understanding the model bit by bit:

3. Switch on target depletion

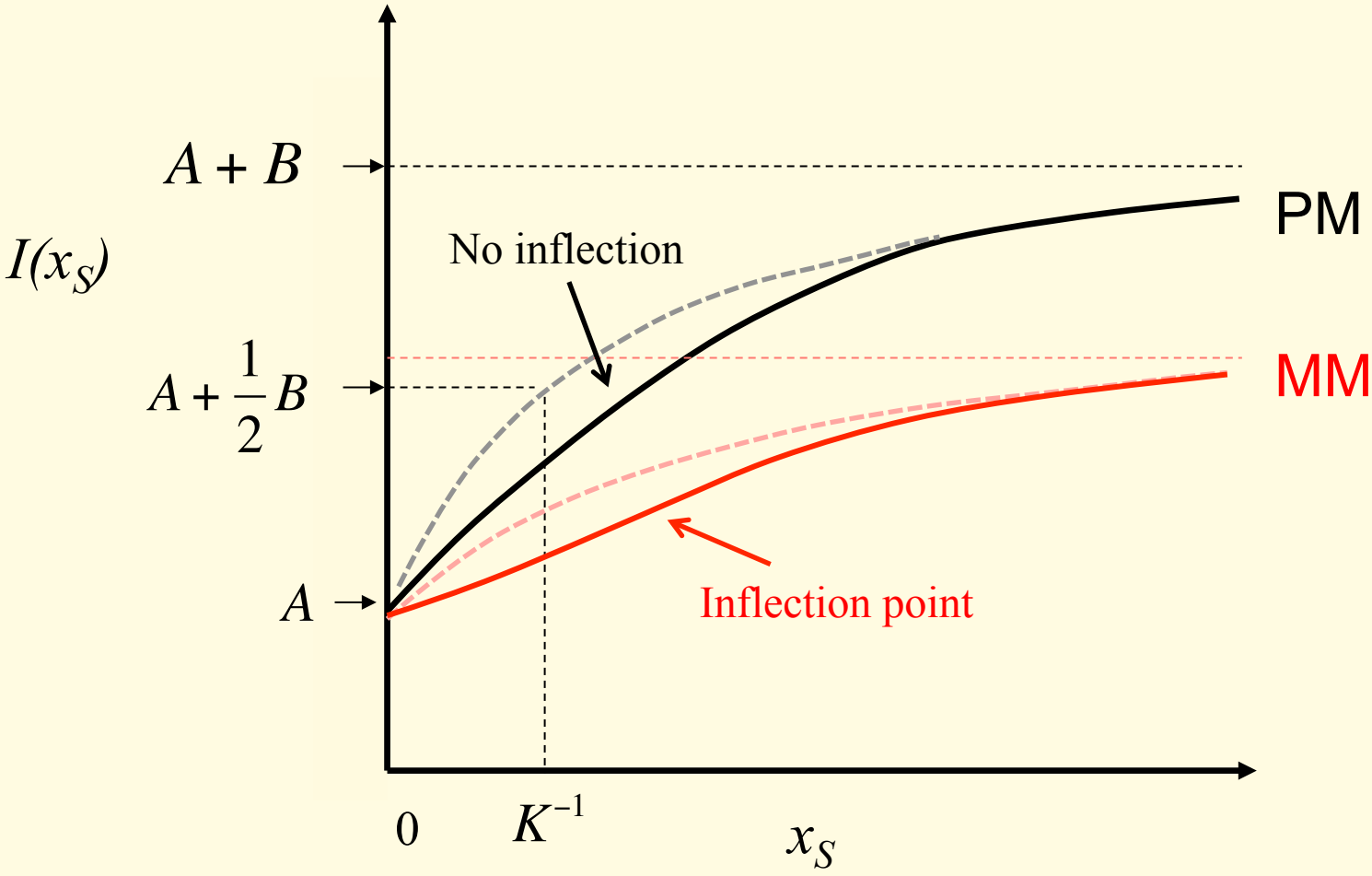
$$I(x_S) = A + B\frac{K(x_S - p\theta_{\text{sum}})}{1 + K(x_S - p\theta_{\text{sum}})}$$



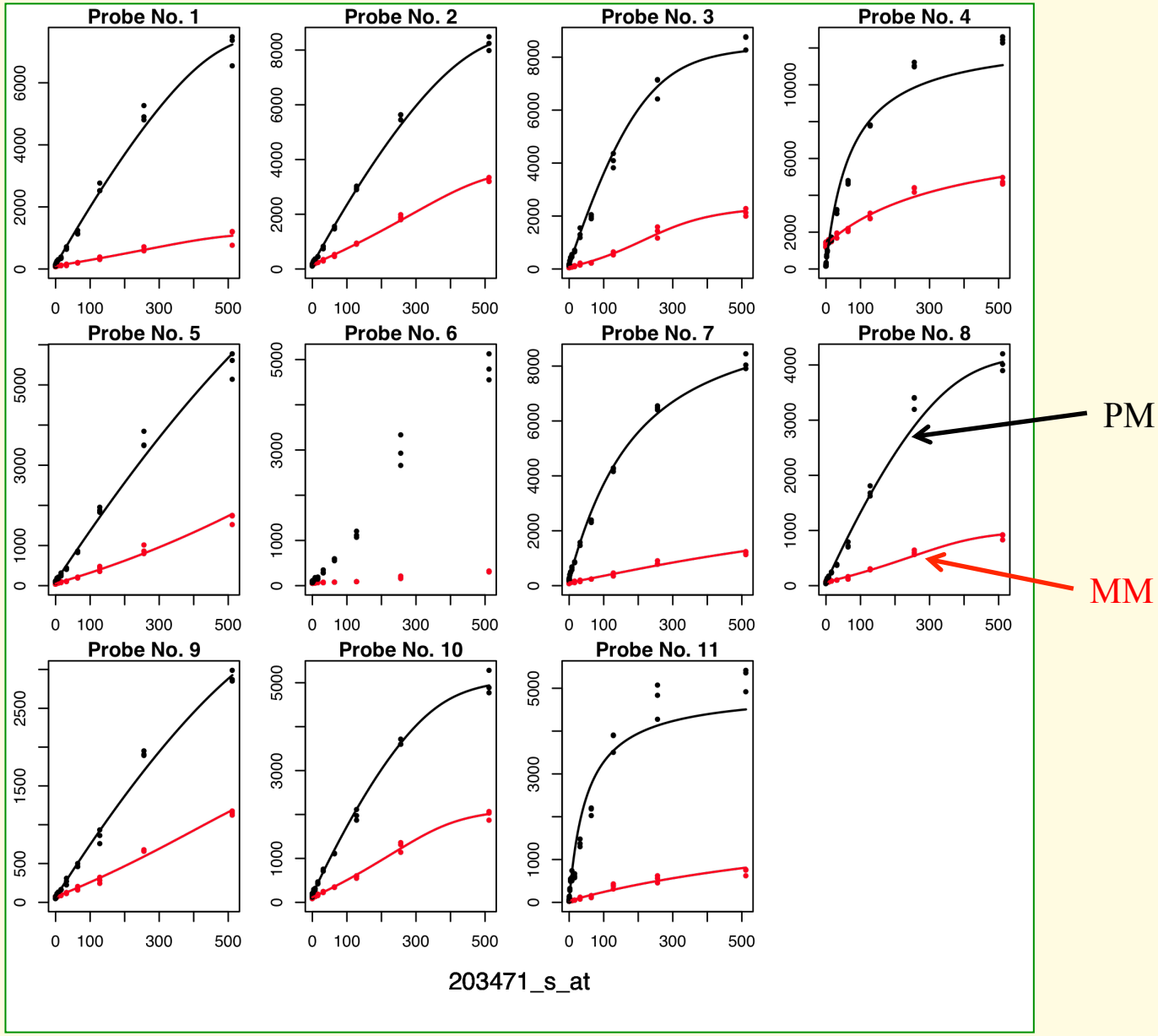
Understanding the model bit by bit:

4. Switch on target depletion and consider PM/MM pair

$$I(x_S) = A + B\frac{K(x_S - p\theta_{\text{sum}})}{1 + K(x_S - p\theta_{\text{sum}})}$$



Example of fits to Affy U133a spikes:



Example of fits to Affy U133a spikes: