CHEMISTRY
A EUROPEAN JOURNAL

Supporting Information

Anchoring of histidine-tagged proteins to molecular printboards: self-assembly, thermodynamic modeling and patterning

Manon J. W. Ludden,[a] Alart Mulder,[b] Katrin Schulze,[b] Vinod Subramaniam,[c] Robert Tampé*,[b] Jurriaan Huskens*[a]

Dedicated to Professor David N. Reinhoudt on the occasion of his 65th birthday

[a] M. J. W. Ludden, and Prof. Dr. J. Huskens, Molecular Nanofabrication group, MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, Fax: (+31) 53-489-4645.

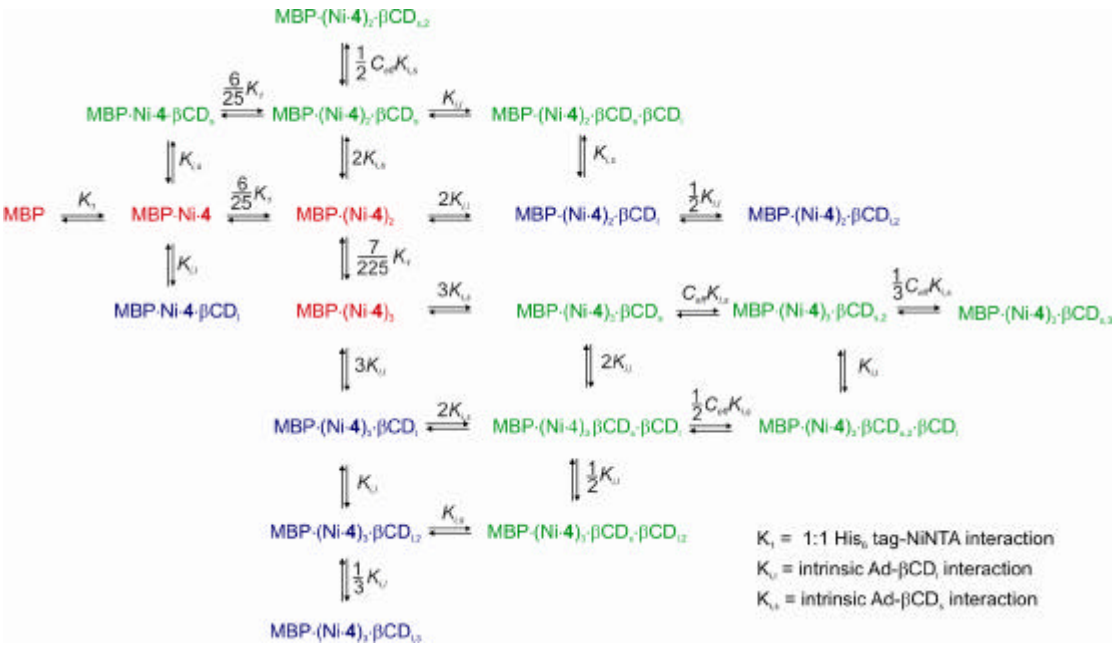
[b] K. Schultze, Dr. A. Mulder, Prof. Dr. R. Tampé, Institut für Biochemie, Biozentrum Frankfurt, Johann-Wolfgang-Goethe-Universität, Marie-Curie-Strasse 9, 60439 Frankfurt am Main, Germany, Fax: (+49) 69-798-29495.

[c] Prof. Dr. V. Subramaniam, Biophysical Engineering group, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, Fax: (+31) 53-489-1105.

** Corresponding authors: j.huskens@utwente.nl, tampe@em.uni-frankfurt.de*

Modeling of the binding of His_6-MBP to the molecular printboard

The binding of His_6-MBP to *b*CD SAMs via Ni•**4** can be monovalent, divalent, or trivalent. In Scheme 2 all possible equilibria are presented, assuming the Ni•**4** complex forms completely, as discussed in the main text. At *b*CD SAMs, all His_6-MBP units complexed to one Ni•**4** will behave as monovalent guests, binding to surface-confined *b*CD (*b*CD_s) in a similar fashion as to *b*CD in solution (*b*CD_l). For His_6-MBP units that are bound via two or three Ni•**4** complexes, the binding to *b*CD SAMs is governed by an effective concentration term (C_{eff}), which is the driving force for the formation of multivalent complexes at *b*CD SAMs. It has to be noted, that the effective concentration represents the probability that an unused guest site finds a complementary host site and thus incorporates all entropic multivalency factors, including e.g. entropy changes due to conformational losses between the guest sites. For the system at hand, the effective concentration concept provides a rigorous quantitative description of the equilibria involved.[1,2]

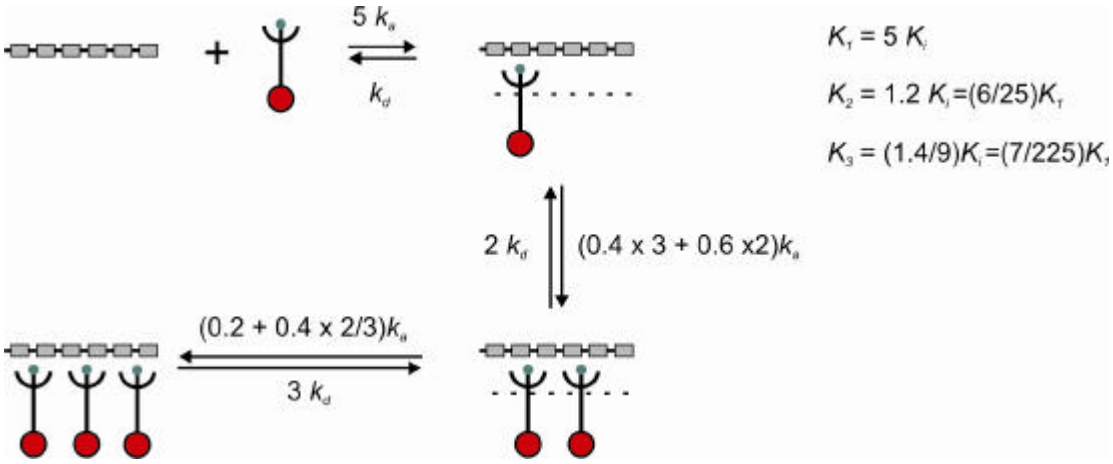


Scheme 2. Equilibria for all species (solution and surface) for the attachment of His_6-MBP at the molecular printboard (charges are omitted for clarity). Subsequent complexation steps of Ni•**4** to MBP are shown in red, and all surface species are given in green.

A general description is given for the multivalent binding of His_6-MBP•(Ni•**4**)$_x$ (x=1-3) to the molecular printboard. The stepwise adsorption of e.g. His_6-MBP•(Ni•**4**)$_3$ to the surface involves an intermolecular adsorption step and two intramolecular binding steps, the latter of which are both governed by C_{eff}. All solution and surface species of MBP are shown in Scheme 1. All intrinsic stability constants for *b*CD_l and *b*CD_s are assumed equal for all steps given in Scheme 2.[1]

Similar to the binding studies described before,[1,2] SPR titrations performed for the binding of His_6-MBP to *b*CD SAMs in the presence of Ni•**4** are fitted here, yielding K_1 values for the His_6 tag-Ni•**4** interaction, while the intrinsic binding constant of an adamantyl guest to *b*CD in solution and the intrinsic binding constant of an adamantyl guest to a surface-confined *b*CD cavity, $K_{i,l}$ and $K_{i,s}$, respectively, are fixed to the values determined by ITC and SPR.

The statistical factors relating K_2 and K_3 for additional Ni•**4** were determined by noting that: (i) Ni•**4** binds to two neighboring histidines, (ii) binding Ni•**4** to His_6-MBP is 5 times as likely as binding to a His_2 unit (which is the intrinsic interaction motif in this case), (iii) there are 2 or 3 His_2 sites free for interaction of a second Ni•**4** to MBP•Ni•**4** in 60% and 40% of the MBP•Ni•**4** complexes, respectively, and (iv) only for 46.7% of the MBP•(Ni•**4**)$_2$ complexes there is an additional free His_2 site available for a third Ni•**4** unit (See Scheme 3). This leads to the prefactors of $K_2 = \frac{6}{25} K_1$ and $K_3 = \frac{7}{225} K_1$ as given in Scheme 2.



Scheme 3. The statistical factors relating K_1, K_2 and K_3 to K_i.

Since all measurements have been performed at pH = 7.5, Ni•**4** is always formed completely (see main text). The mass balances that can be constructed based on Scheme 2 are the following (charges are omitted for clarity):

$$
\begin{aligned}
[\mathrm{MBP}]_{tot} = {} & [\mathrm{MBP}]_{free} + [\mathrm{MBP{\bullet}Ni{\bullet}}\mathbf{4}] + [\mathrm{MBP{\bullet}Ni{\bullet}}\mathbf{4}{\bullet}\beta\mathrm{CD_l}] + [\mathrm{MBP{\bullet}Ni{\bullet}}\mathbf{4}{\bullet}b\mathrm{CD_s}] + \\
& [\mathrm{MBP{\bullet}\,(Ni{\bullet}}\mathbf{4})_2] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_2{\bullet}b\mathrm{CD_s}] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_2{\bullet}(b\mathrm{CD_s})_2] + \\
& [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_2{\bullet}b\mathrm{CD_l}] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_2{\bullet}(b\mathrm{CD_l})_2] + \\
& [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_2{\bullet}b\mathrm{CD_s}\beta\mathrm{CD_l}] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}(b\mathrm{CD_s})_2] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3] + \\
& [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}b\mathrm{CD_l}] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}(b\mathrm{CD_l})_2] + \\
& [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}(b\mathrm{CD_l})_3] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}b\mathrm{CD_s}] + \\
& [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}b\mathrm{CD_s}\beta\mathrm{CD_l}] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}b\mathrm{CD_s}(b\mathrm{CD_l})_2] + \\
& [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}(b\mathrm{CD_s})_2 b\mathrm{CD_l}] + [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}(b\mathrm{CD_s})_2] + \\
& [\mathrm{MBP{\bullet}(Ni{\bullet}}\mathbf{4})_3{\bullet}(b\mathrm{CD_s})_3] \qquad (1)
\end{aligned}
$$

$$[\mathbf{4}]_{tot} = [Ni\bullet\mathbf{4}]_{free} + [Ni\bullet\mathbf{4}\bullet b\mathrm{CD_l}] + [Ni\bullet\mathbf{4}\bullet b\mathrm{CD_s}] + [MBP\bullet Ni\bullet\mathbf{4}] + [MBP\bullet Ni\bullet\mathbf{4}\bullet b\mathrm{CD_s}] + [MBP\bullet Ni\bullet\mathbf{4}\bullet b\mathrm{CD_l}] + 2([MBP\bullet(Ni\bullet\mathbf{4})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_2 b\mathrm{CD_s}] + [MBP\bullet(Ni\bullet\mathbf{4})_2\bullet(b\mathrm{CD_s})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_2\bullet b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_2\bullet(b\mathrm{CD_l})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_2\bullet b\mathrm{CD_l}\bullet b\mathrm{CD_s}]) + 3([MBP\bullet(Ni\bullet\mathbf{4})_3] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_s}] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_s})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_s})_3] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_l})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_l})_3] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_s}\bullet b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_3(b\mathrm{CD_s})_2 b\mathrm{CD_l}] + [MBP(Ni\bullet\mathbf{4})_3 b\mathrm{CD_s}(b\mathrm{CD_l})_2]) \quad (2)$$

$$[\mathbf{3}]_{tot} = [\mathbf{3}]_{free} + [\mathbf{3}\bullet b\mathrm{CD_l}] + [\mathbf{3}\bullet b\mathrm{CD_s}] \quad (3)$$

$$[b\mathrm{CD_s}]_{tot} = [b\mathrm{CD_s}]_{free} + [MBP\bullet Ni\bullet\mathbf{4}\bullet b\mathrm{CD_s}] + [MBP\bullet(Ni\bullet\mathbf{4})_2\bullet b\mathrm{CD_s}] + [MBP\bullet(Ni\bullet\mathbf{4})_2\bullet b\mathrm{CD_s}\bullet b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_s}] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_s}\bullet b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_s}\bullet(b\mathrm{CD_l})_2] + 2([MBP\bullet(Ni\bullet\mathbf{4})_2\bullet(b\mathrm{CD_s})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_s})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_s})_2 b\mathrm{CD_l}]) + 3[MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_s})_3] + [\mathbf{3}\bullet b\mathrm{CD_s}] + [Ni\bullet\mathbf{4}\bullet b\mathrm{CD_s}] \quad (4)$$

$$[b\mathrm{CD_l}]_{tot} = [b\mathrm{CD_l}]_{free} + [MBP\bullet(Ni\bullet\mathbf{4})_2\bullet b\mathrm{CD_s} b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_2\bullet b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_s} b\mathrm{CD_l}] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_s})_2 b\mathrm{CD_l}] + 2([MBP\bullet(Ni\bullet\mathbf{4})_2\bullet(b\mathrm{CD_l})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_l})_2] + [MBP\bullet(Ni\bullet\mathbf{4})_3\bullet b\mathrm{CD_s}(b\mathrm{CD_l})_2]) + 3[MBP\bullet(Ni\bullet\mathbf{4})_3\bullet(b\mathrm{CD_l})_3] + [\mathbf{3}\bullet b\mathrm{CD_l}] + [Ni\bullet\mathbf{4}\bullet b\mathrm{CD_l}] + [MBP\bullet Ni\bullet\mathbf{4}\bullet b\mathrm{CD_l}] \quad (5)$$

Species involving $b\mathrm{CD_s}$ are expressed in volume concentrations.[1] The binding of the divalent MBP•(Ni•**4**)$_2$ and trivalent MBP•(Ni•**4**)$_3$ to $b\mathrm{CD_l}$ involves statistical factors (Scheme 2) arising from the probabilities for binding relative to the monovalent species, in this case according to a normal 1:3 complexation sequence.

The binding constants for first intermolecular binding events of the divalent and trivalent species at the surface are:

$$K = \frac{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_2 \cdot b\mathrm{CD_s}]}{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_2][b\mathrm{CD_s}]} = 2K_{i,s} \tag{6}$$

$$K = \frac{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_3 \cdot b\mathrm{CD_s}]}{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_3][b\mathrm{CD_s}]} = 3K_{i,s} \tag{7}$$

The second, intramolecular, binding event for the di- and trivalent species, and third, for the trivalent species (equations 9, 10, and 11) are governed by an effective concentration term, which is defined as given in equation 8.[1,2] The effective concentration is given by multiplying the maximum effective concentration, $C_{eff,max}$, which is the number of accessible host sites in the probing volume, with the fraction of free host sites at the surface.

$$C_{eff} = C_{eff,\max} \frac{[b\mathrm{CD_s}]}{[b\mathrm{CD_s}]_{\mathrm{tot}}} \tag{8}$$

$$K = \frac{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_2 \cdot (b\mathrm{CD_s})_2]}{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_2 \cdot b\mathrm{CD_s}][b\mathrm{CD_s}]} = \tfrac{1}{2} C_{eff} K_{i,s} \tag{9}$$

$$K = \frac{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_3 \cdot (b\mathrm{CD_s})_2]}{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_3 \cdot b\mathrm{CD_s}][b\mathrm{CD_s}]} = C_{eff} K_{i,s} \tag{10}$$

$$K = \frac{[\mathrm{MPB} \cdot (\mathrm{Ni} \cdot \mathbf{4})_3 \cdot (b\mathrm{CD_s})_3]}{[\mathrm{MBP} \cdot (\mathrm{Ni} \cdot \mathbf{4})_3 \cdot (b\mathrm{CD_s})_2][b\mathrm{CD_s}]} = \tfrac{1}{3} C_{eff} K_{i,s} \tag{11}$$

Since the SPR experiments were performed in a flow system, all solutions species concentrations can be calculated from simplified forms of equations 1-3 and 5. After numerical optimization of these equations, the values obtained for the solution species concentrations were used in the full equations 1-5 for calculations of the surface species.

Substitution of the equilibrium constant definitions into the mass balances for $[MBP]_{tot}$, $[b\text{CD}_s]_{tot}$, $[b\text{CD}_l]_{tot}$, $[\mathbf{3}]_{tot}$, and $[\mathbf{4}]_{tot}$ (equations 1-5) provides a set of numerically solvable equations with [MBP], $[b\text{CD}_s]$, $[b\text{CD}_l]$, [**3**], and [**4**] as the variables.

Starting from an initial estimate for K_1 (defined as: $K_1 = \frac{[\text{MBP} \cdot (\text{Ni} \cdot \mathbf{4})]}{[\text{MBP}][\text{Ni} \cdot \mathbf{4}]}$) using fixed values for $C_{eff,max}$ (0.1 M) and the other stability constants, this set of equations is solved numerically using a Simplex algorithm in a spreadsheet approach.[3] When fitting SPR data, K_1 is optimized in a least-squares optimization routine, assuming that the SPR response (intensity) is linearly dependent on the coverages of MBP, **3**, and **4** adsorbed to the *b*CD SAM, regardless of the type of species. The maximum intensity (I_{max} of MBP) is then optimized as an independent fitting parameter as well while those of **3** and **4** were determined by independent SPR measurements.

Based on Scheme 1, the overall stability constant for $\text{His}_6\text{-MBP}\bullet(\text{Ni}\bullet\mathbf{4})_3\bullet(b\text{CD}_s)_3$ can be given by equation 12:

$$K = (K_{i,s})^3 K_1 K_2 K_3 C_{eff}^2 \cdot [\text{Ni} \cdot \mathbf{4}]^3 \qquad (12)$$

Assuming $K_{i,s}$ = 1.2 × 10^4 M^{-1}, K_1 = 7.8 × 10^3 M^{-1} ($K_2 = \frac{6}{25} K_1$ and $K_3 = \frac{7}{225} K_1$), $C_{eff} = C_{eff,max}$ = 0.1 M (at relatively low coverages), and [Ni•**4**] = 1 μM, an apparent conditional binding constant of ~ 10^5 M^{-1} can be estimated.

References

1. J. Huskens, A. Mulder, T. Auletta, C. A. Nijhuis, M. J. W. Ludden, D. N. Reinhoudt, *J. Am. Chem. Soc.* **2004**, *126*, 6784-6797.
2. A. Mulder, T. Auletta, A. Sartori, S. Del Ciotto, A. Casnati, R. Ungaro, J. Huskens, D. N. Reinhoudt, *J. Am. Chem. Soc.* **2004**, *126*, 6627-6636.
3. J. Huskens, H. van Bekkum, J. A. Peters, *Comp. Chem.* **1995**, *19*, 409-416.