SUPPLEMENTARY INFORMATION:

Computational design of cyclic peptides for the customized oriented immobilization of globular proteins

Miguel A. Soler[1], Alex Rodriguez[2], Anna Russo[1], Abimbola Feyisara Adedeji[1,3], Cedrix J. Dongmo Foumthuim[1], Cristina Cantarutti[1], Elena Ambrosetti[3,4], Loredana Casalis[4], Alessandra Corazza[1], Giacinto Scoles[1], Daniela Marasco[5], Alessandro Laio[2], Sara Fortuna[1,#,]*

[1]Department of Medical and Biological Sciences, University of Udine, Piazzale Kolbe, 4 - 33100 Udine, Italy

[2]SISSA, Via Bonomea 265, I-34136 Trieste, ITALY

[3] PhD School of Nanotechnology, Department of Physics, University of Trieste, Via Valerio 2, 34127 Trieste, Italy

[4]Nanoinnovation Lab - Elettra-Sincrotone S.C.p.A., ss 14 km 163,5 in AREA Science Park I-34149, Basovizza-Trieste, Italy

[5]Department of Pharmacy, CIRPEB: Centro Interuniversitario di Ricerca sui Peptidi Bioattivi- University of Naples "Federico II", DFM-Scarl, 80134, Naples, Italy.

[#] now at Center for biomedical sciences and engineering, University of Nova Gorica, Dvorec Lanthieri, Glavni Trg 8, 5271 Vipava, Slovenia

*sara.fortuna@ung.si

TABLE S1. Scoring Energies (E, in kcal/mol) and sequences of the soluble peptides with the lowest Scoring E. The peptides are the output of two runs of the algorithm. Each run consists of three parallel replica exchange Monte Carlo optimisation run at the pseudotemperature T =0.3, 0.6, 0.9. The solubility has been predicted with Innovagen's peptide calculator webtool [PepCalc.com. http://pepcalc.com/. Innovagen AB, 2012.]

	Run 01			Run 02		
Index	Step (*T*)	Scoring E	Sequence	Step (*T*)	Scoring E	Sequence
1	**357** (0.6)	-25.4	CKWREQDRLPEC	374 (0.3)	-24.6	CYIGENDWRFHC
2	339 (0.3)	-25.3	CKWREQDRLSEC	352 (0.3)	-24.4	CYIGESDWRFHC
3	366 (0.6)	-25.0	CKWREFDRLPEC	468 (0.3)	-24.3	CYMTKSDRRFQC
4	329 (0.3)	-24.9	CVRYEWQHYRHC	**395** (0.6)	-24.3	CYRGQNDWRFHC
5	375 (0.3)	-24.7	CQRYSHQHYRHC	396 (0.6)	-24.2	CYRGQTDWRFHC
6	328 (0.3)	-24.6	CVRYEMQHYRHC	**461** (0.6)	-24.1	CYFTKTDRRFQC
7	432 (0.3)	-24.6	CVRYEEQWYRHC	460 (0.3)	-24.1	CYFTKSDRRFQC
8	**296** (0.3)	-24.5	CTRYFEQFYRHC	**444** (0.3)	-24.0	CYVIKADRRFQC
9	316 (0.3)	-24.5	CVRYERQHYRHC	402 (0.3)	-24.0	CYDGENDWRFHC
10	**381** (0.3)	-24.4	CRRYSHQHYRHC	481 (0.3)	-24.0	CYKFLDDWRFTC
11	406 (0.3)	-24.4	CYRYSEQHYRHC	484 (0.3)	-23.9	CYMTKVDIRFQC
12	**396** (0.3)	-24.3	CRRYSSQHYRHC	**476** (0.3)	-23.9	CYMTKSDIRFQC
13	387 (0.3)	-24.2	CRRYSGQHYRHC	404 (0.3)	-23.9	CYDGEADWRFHC
14	284 (0.3)	-24.2	CTRYILQKYRHC	460 (0.3)	-23.8	CYFGKSDRRFQC
15	**416** (0.3)	-24.1	CYRYSEQRYRHC	442 (0.6)	-23.8	CYTIKADRRFQC
16	236 (0.3)	-24.1	CTRYDWQHTRHC	406 (0.3)	-23.8	CYDGEADWRFSC
17	413 (0.3)	-24.0	CERYSEQHYRHC	380 (0.3)	-23.8	CYQGENDWRFHC
18	419 (0.3)	-24.0	CVRYSEQHYRHC	**449** (0.6)	-23.7	CYVGKADRRFQC
19	417 (0.3)	-23.9	CVRYSEQRYRHC	411 (0.3)	-23.7	CYRDKTDWRFHC
20	**415** (0.3)	-23.8	CERYSEQRYRHC	403 (0.3)	-23.6	CYRGKTDWRFHC

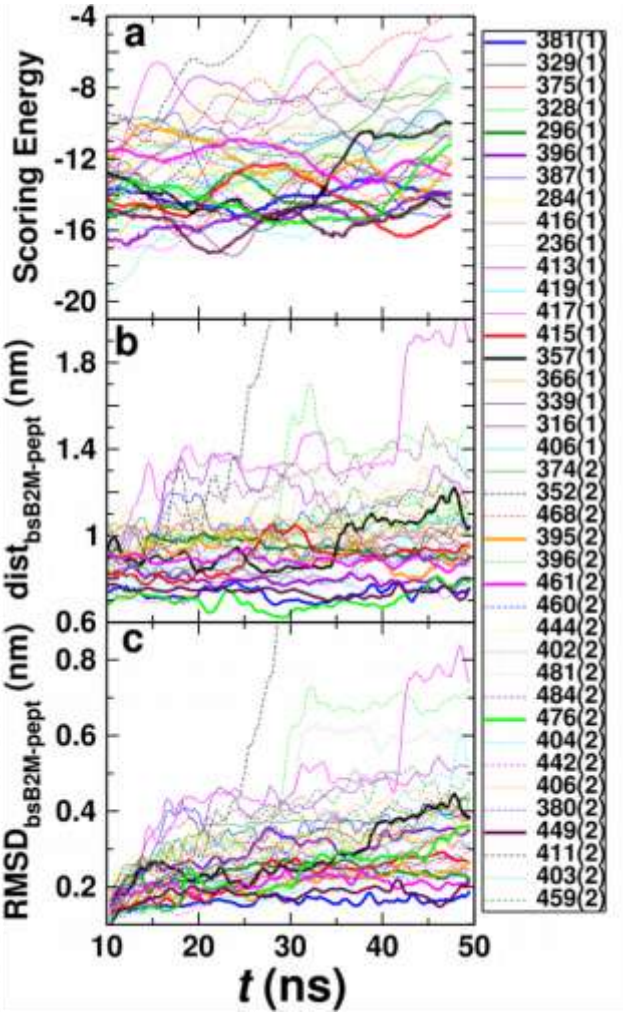

Figure S1. Primary screening: MD simulation results. Selected peptides are highlighted by thicker lines. (a) Average scoring energy, (b) Separation distance between β2m binding site (BS-B2M) and peptide, and (c) RMSD of BS-B2M+peptide along the simulation time.

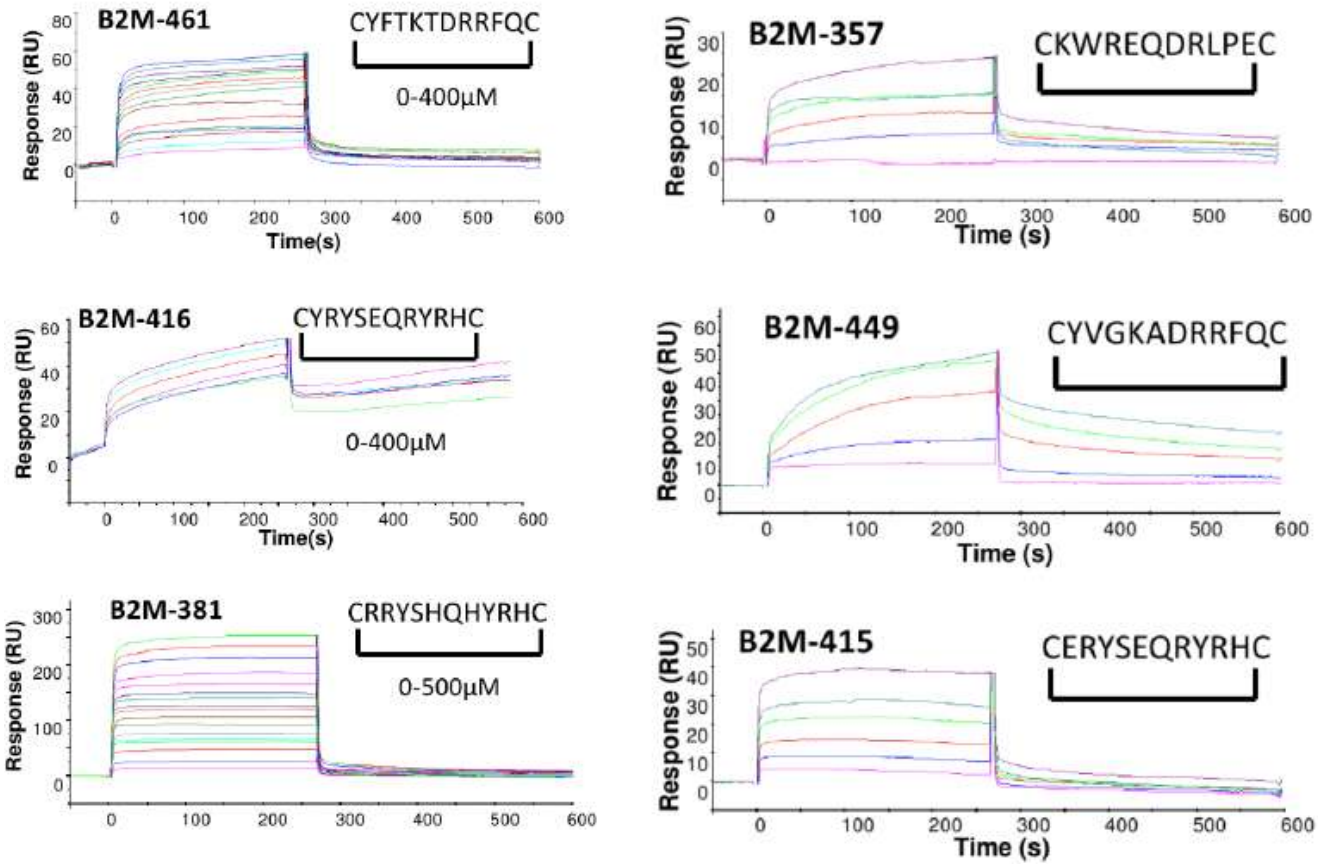

Figure S2. SPR results.

Table S2. Peptide sequences elected through SPR secondary screening, pI (isoelectric point) and dissociation constants toward β2m are also reported (K_d).

Id	sequence	pI	K_d (µM)
357	[C]KWREQDRLPE[C]	6.17	30 ± 9
381	[C]RRYSHQHYRH[C]	9.50	38 ± 9
415	[C]ERYSEQRYRH[C]	8.06	57 ± 9
461	[C]YFTKTDRRFQ[C]	8.90	40 ± 9
449	[C]YVGKADRRFQ[C]	8.90	36 ± 8
416	[C]ERYSEQRYRH[C]	8.86	70 ± 18

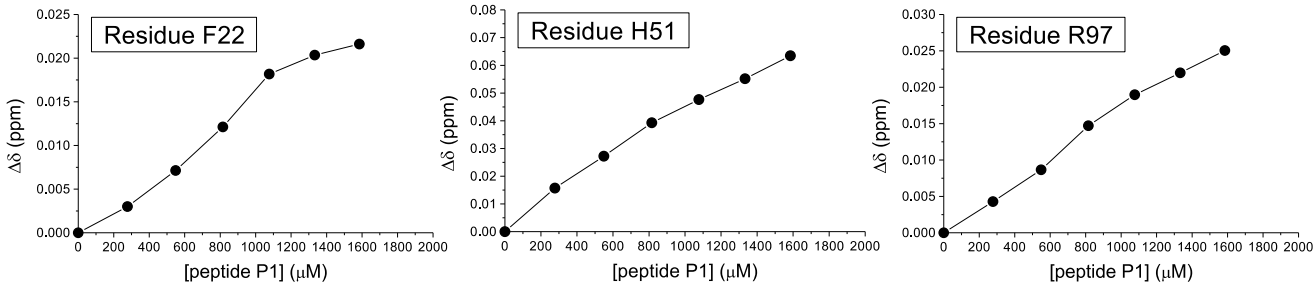

Figure S3. Representative chemical shift deviations measured from 2D [^{1}H, ^{15}N] HSQC experiments. Lines are only intended to guide the eye.

Determination of the Dissociation Constant by Atomic Force Microscopy (AFM)

Supplementary Figure S4 shows three cyclic peptide-DNA assemblages confined within protein-repellant ethylene terminated alkylthiol monolayer on an ultra-flat gold substrate, with the same surface coverage (S/A) of 2.56, generated for each independent experimental section using the synergetic approach described in the materials and methods section that were tested for β2m recognition at different concentrations. Also, different surface coverages (S/A = 1 - 10) of cyclic peptide-DNA assemblage were tested for β2m recognition at different concentration and we were able to tune the surface coverage of cyclic peptide-DNA assemblage by varying nanografting parameter known as S/A using the procedure established in the previous works of our group[1-3]. However, only cyclic peptide-DNA assemblages with surface coverage S/A = 2.56 were considered for the determination of dissociation constant, because this surface coverage yields a consistent and reproducible binding signal (differential height) for all concentration of β2m.

In Figure S4a the cyclic peptide-DNA assemblage was tested towards the recognition of β2m proteins in a 50 µM solution of β2m, and the side by side topographic height profile measurement in (Figure S4a (iii)) shows no change in the height of the assemblage after β2m recognition assay. However, in Figure S4b which is another assemblage shows a small increase in height after its interaction with 80 µM solution of β2m as shown in Figure S4b(iii). Lastly, the assemblage in Figure S4c shows a higher incremental height after its exposure to 100 µM solution of β2m as shown in Figure S4c(iii). In addition to the quantitative measurements, we also observed a qualitative result in term of the changes in the brightness and the contrast of assemblages (as show in the assemblages in Figure S4b and S4c) after successful β2m recognition by the cyclic peptide-DNA assemblages. The scan size of the micrograph is 3 µm X 3 µm while the size of cyclic peptide-DNA assemblage is 1 µm X 1 µm. This is applicable to all AFM micrographs in Figure S4 and S5.

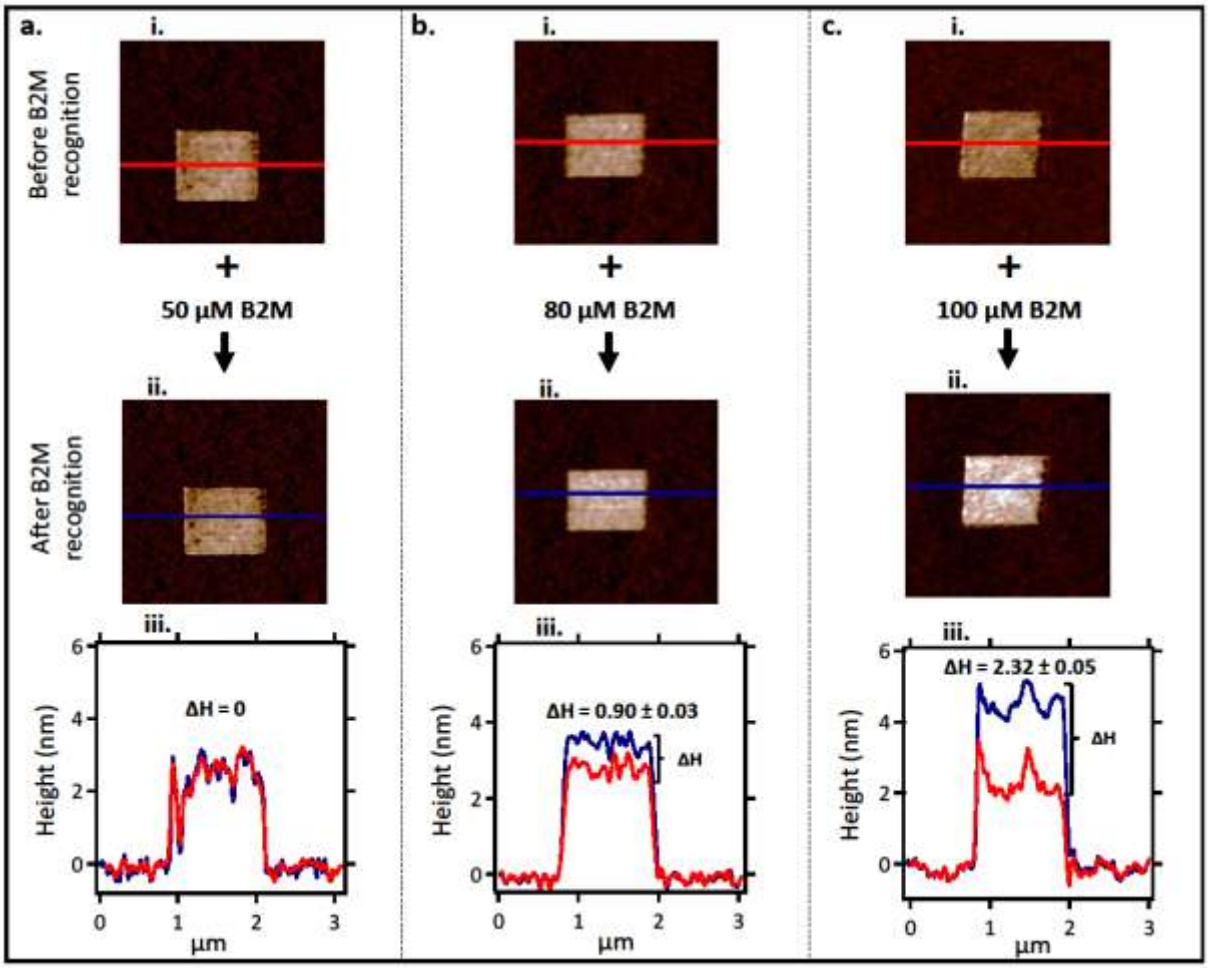

Figure S4: AFM-based quantitative approach for analyzing surface-based cyclic peptide-DNA assemblages before and after β2m recognition at different target (β2m) concentration. Three cyclic peptide-DNA nano-assemblages generated in different experimental sections were tested for β2m recognition at different concentration of proteins, 50 µM, 80 µM, and 100 µM as shown in (a), (b) and (c) respectively. AFM topographic height measurement was employed in each case, to determine the differential height, which is the binding signal for the detection of β2m in each case, this is shown in (a (iii)), (b (iii)) and (c (iii)) respectively.

To determine the dissociation constant of cyclic peptide –DNA conjugate assemblages towards β2m protein, the differential height (ΔH), that is the change in height between the height of the cyclic peptide-DNA assemblage before β2m recognition and the height of the same assemblage after β2m recognition assay, was measured as shown in Figure S5**.** The

procedure described in Figure S5 was subsequently done for each cyclic peptide-DNA assemblage tested for β2m recognition at different β2m concentration. For each concentration, the average differential height $\boldsymbol{\Delta H}_i$ for the i-th experiment was calculated as follows;

$$\boldsymbol{\Delta H}_i = \frac{\boldsymbol{\Delta H_1 + \Delta H_2 + \Delta H_3 + \Delta H_4}}{\boldsymbol{4}} \quad \textbf{(S1)}$$

Where ΔH_1, ΔH_2, ΔH_3, and ΔH_4 are the differential heights for each plot in Figure S5 obtained from peak-to-peak subtraction of the profiles.

Lastly, the overall average differential height ($\boldsymbol{\Delta H}$) obtained for at least three independent experiments for a given β2m concentration was taken into account.

Mathematically, we calculated overall average differential height as follows;

$$\boldsymbol{\Delta H} = \frac{1}{n} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\Delta H}_i \quad \textbf{(S2)}$$

Where $\boldsymbol{\Delta H}_i$ is the average differential height for experiment i (@ 100 μM of β2m), and n is the number of experiments performed.

These calculations were done for all the cyclic peptide-DNA assemblages tested for β2m recognition at different concentration and the overall ΔH(s) were plotted as a function of β2m concentration in Figure 6i. The experimental data were fitted with Hill's equation using IGOR PRO 6.3A analysis software.

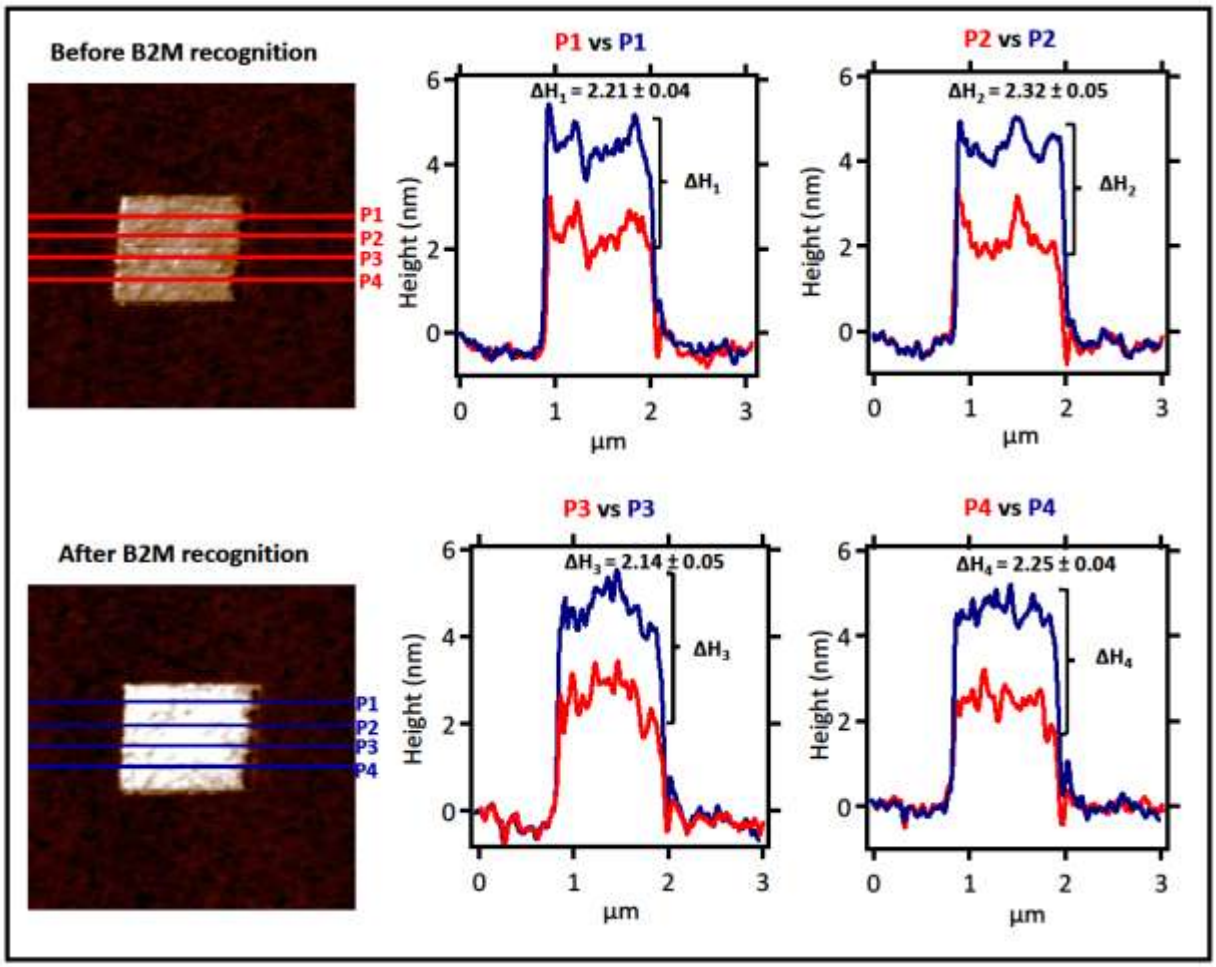

Figure S5: An example of a subsequent height measurement procedure employed in determining the average height difference across the topographic AFM micrograph of cyclic peptide-DNA assemblage before and after β2m recognition (100 μM concentration). To determine the differential height value for each plot, we employed peak-to-peak subtraction of the height waves for each plot using IGOR-waves analysis. Followed by statistical analysis of the output wave from the peak-to-peak wave subtraction. The same analysis was employed for all the four plots. This yields the average value for ΔH_1, ΔH_2, ΔH_3, and ΔH_4 with their corresponding standard deviation. Lastly, the mean ΔH ($\boldsymbol{\Delta H}$) which is 2.23 ±0.05 nm in this case, becomes the binding signal in this particular experiment.

References

1. Mirmomtaz, E. et al. Quantitative study of the effect of coverage on the hybridization efficiency of surface-bound DNA nanostructures. *Nano letters* **8**, 4134-4139 (2008).
2. Castronovo, M. et al. Control of steric hindrance on restriction enzyme reactions with surface-bound DNA nanostructures. *Nano letters* **8**, 4140-4145 (2008).
3. Bano, F. et al. Toward multiprotein nanoarrays using nanografting and DNA directed immobilization of proteins. *Nano letters* **9**, 2614-2618 (2009).