**Supplementary Information**

Whole-cell biosensor with tuneable limit of detection enables low-cost agglutination assays for medical diagnostic applications

Nicolas Kylilisa,b, Pinpunya Riangrungroja,c, Hung-En Laia,d, Valencio Salemae, Luis Ángel Fernándeze, Guy-Bart V Stana,b, Paul S Freemonta,d, Karen M Polizzia,f,\*

aImperial College Centre for Synthetic Biology, Imperial College London, London SW7 2AZ, United Kingdom;

bDepartment of Bioengineering, Imperial College London, London SW7 2AZ, United Kingdom

cDepartment of Life Sciences, Imperial College London, London SW7 2AZ, United Kingdom

dSection of Structural Biology, Department of Medicine, Imperial College London, London SW7 2AZ, United Kingdom

eDepartment of Microbial Biotechnology. Centro Nacional de Biotecnología, CNB-CSIC. Darwin 3, Campus UAM, Cantoblanco, Madrid, Spain

f Department of Chemical Engineering, Imperial College London, London SW7 2AZ, United Kingdom

\*To whom correspondence should be addressed: [k.polizzi@imperial.ac.uk](mailto:k.polizzi@imperial.ac.uk); p.freemont@imperial.ac.uk

# 

Table of Contents

[SUPPLEMENTARY NOTE 1 3](#_Toc533144773)

[Anti-GFP antibody display and nanobody quantification 3](#_Toc533144774)

[SUPPLEMENTARY NOTE 2 5](#_Toc533144775)

[Development of the model analyte protein: tdGFP 5](#_Toc533144776)

[SUPPLEMENTARY NOTE 3 8](#_Toc533144777)

[Background of the cell agglutination model 8](#_Toc533144778)

[Normalized cell agglutination dynamics 8](#_Toc533144779)

[Free analyte concentration dynamics 9](#_Toc533144780)

[Bound nanobody density dynamics (on the cell surface) 9](#_Toc533144781)

[Mathematical model derivation 10](#_Toc533144782)

[Equation (1): Time-dependent dynamics of normalized cell agglutination, 10](#_Toc533144783)

[Equation (2): Time-dependent dynamics of bound nanobody density 12](#_Toc533144784)

[Sensitivity analysis 15](#_Toc533144785)

[SUPPLEMENTARY NOTE 4 17](#_Toc533144786)

[Diagnostic workflow 17](#_Toc533144787)

[SUPPLEMENTARY NOTE 5 18](#_Toc533144788)

[Anti-fibrinogen nanobody cell-surface display 18](#_Toc533144789)

[PAGE analysis of human fibrinogen 19](#_Toc533144790)

[SUPPLEMENTARY REFERENCES 21](#_Toc533144791)

# SUPPLEMENTARY NOTE 1

## Anti-GFP antibody display and nanobody quantification

An experiment was designed to investigate the localization of the anti-GFP nanobody and its functionality as a sensor element when displayed on the surface of the host cell. *E. coli* DH10B strain cells harboring the pNVgfp construct were induced with IPTG for NVgfp protein expression and subsequently labeled with saturating amounts of purified GFP as described in the Materials and Methods. After washing the cells to remove excess GFP, cells were analyzed by fluorescence microscopy (Supplementary Figure 1A) and flow cytometry (Supplementary Figure 1B) to confirm that the anti-GFP nanobody was expressed on the surface *E. coli* cells and was able to bind to its ligand.

Then, an experiment was designed to quantify the number of displayed nanobodies per cell. It was reasoned that the number of anti-GFP nanobodies displayed could be calculated by titrating known amounts of purified GFP concentrations against a constant number of *E. coli* cells. The number of displayed nanobodies could be derived from the point where increasing GFP concentrations do not result in increased cell fluorescence due to saturation of nanobody binding sites, similar to the approach reported previously for the quantification of scFv displayed on *E. coli* cells.1 The *E. coli* DH10B strain cells transformed with the pNVgfp construct were induced for NVgfp protein expression as described in the Materials and Methods. Following protein expression, the cells were centrifuged and re-suspended in PBS to an O.D.600 value of 1. Then, 200 μl of cell suspensions were incubated with varying amounts of purified GFP, excess unbound GFP was washed away and cell fluorescence was measured by flow cytometry.

The fluorescence signal of labeled cells saturated at concentrations between 2 and 5 picomoles of GFP for 200 μl cell suspension at O.D.600 = 1 (Supplementary Figure 1C). Thus, given a total number of nanobodies in the reaction (2-5 picomoles divided by Avogadro’s number) and the total number of cells in the reaction (0.2 x 109)2 we were able to calculate an average number of nanobodies per cell of ~6,000-15,000.



Supplementary Figure 1: Anti-GFP nanobody display by *E. coli* cells

Expression of anti-GFP nanobody displayed on the surface of *E. coli* DH10B cells analyzed by fluorescence microscopy (**A**) and flow cytometry (**B**). **C** Quantification of the number of anti-GFP nanobody expressed on the cell surface by flow cytometry analysis. Histograms of cell fluorescence intensity at various amounts of monomeric GFP (left) and the correlation of calculated geometric means at these GFP amounts (right). The minimum GFP amount initially reached to plateau region (red dash line) was considerably represented as the saturated amount occupied all nanobody molecules on the cell surface.

# SUPPLEMENTARY NOTE 2

## Development of the model analyte protein: tdGFP

The agglutination reaction requires that both the sensor element and the analyte molecule be multivalent. Analyte molecule multivalency could result from the presence of 2 epitope sites on the analyte molecule recognized by two different sensor elements or from the presence of the epitope on repeated domains in the analyte structure, for example in homodimeric proteins. With this latter strategy in mind, protein engineering was used to develop a fusion protein of two GFP monomers that would provide 2 epitope sites for the anti-GFP nanobody.

A tandem dimeric GFP (tdGFP) was constructed by cloning two GFP DNA sequences in tandem separated by a linker domain. The linker peptide sequence used was twelve amino acids (G H G T G S T G S G S S), and was chosen as it has been used previously by others to engineer an EGFP dimer fusion protein.3 Standard molecular biology procedures were used to construct the plasmid DNA for the tdGFP design.

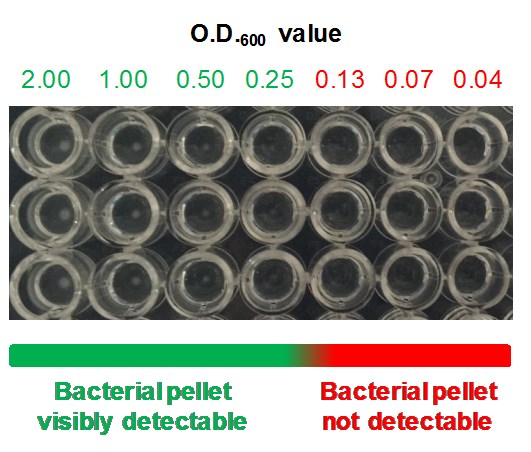
The folding state of the engineered tdGFP protein was verified by fluorescence spectroscopy. It was reasoned that correctly folded tdGFP protein would exhibit twice the fluorescence signal of a monomeric GFP protein at an equimolar concentration. Therefore, monomeric GFP and tdGFP were recombinantly expressed *E. coli* and purified to high purity as determined by SDS-PAGE electrophoresis (Supplementary Figure 1A) and protein concentration was quantified by a spectrophotometer. Purified proteins were diluted to 1.8 μM concentration and a fluorescence plate reader was used to measure the fluorescence emission of the samples at 510nm. Background fluorescence was accounted for by the use of an equal volume solution of the buffer system.

The fluorescent intensity of the signal from the tdGFP samples was almost double the intensity of the signal from the GFP samples (Supplementary Figure 1B). Fluorescence signal for tdGFP protein was measured to 121, 022 au, while the fluorescent signal for GFP protein was measured to 66,733 au. This represents a change of 1.81-fold increased fluorescent signal for tdGFP as compared to GFP monomer for equimolar protein concentrations. This result suggested that a large percentage of the tdGFP fusion protein was folded.



Supplementary Figure 2: GFP fluorescence versus tdGFP fluorescence.

A SDS-PAGE analysis of eluted fractions from size exclusion chromatography following nickel-affinity chromatography purification. B Bar chart of fluorescence signal emitted by monomeric GFP and tdGFP fusion protein, showing properly folded protein exhibiting twice the fluorescence signal from monomeric GFP. Proteins were purified, protein concentrations were made equimolar at 1.8 μM. Measurements were carried out on a microplate reader and error bars represent standard deviation of 3 technical repeats.



Supplementary Figure 3: Bacterial cell pellet as a visual output for bacterial agglutination reactions.

Image of a round bottom well microplate in which 200μl *E. coli* cell suspensions at various O.D.600 values were left to settle to cell pellets overnight.

# SUPPLEMENTARY NOTE 3

## Background of the cell agglutination model

In principle, the binding of two cell-surface nanobodies to the two epitopes on one tandem dimer GFP (tdGFP) enables tethering of two cells via a tdGFP bridge. As the amount of tdGFP increases, more inter-cell bridges can be formed, allowing the formation of larger cell aggregates. We have modeled the agglutination phenomenon by considering a system consisting of a finite number of cells with cluster size ranging from 1 (individual cells) to a large number (cluster containing the largest number of agglutinated cells), freely diffusing analyte and bound analyte on the cell surface. In this system, cells with free and bound nanobody can diffuse freely and collide with one another to form larger clusters via bivalent analytes.

To model our nanobody-induced cell agglutination system, we have adapted a mathematical model first developed by Dolgosheina *et al.*, which describes the agglutination of *Bacillus subtilis* cells in the presence of bivalent antibodies.4 The model framework can be described by two equations, the first describing the cell agglutination dynamics and the second describing bound nanobody density dynamics.

Time-dependent dynamics of normalized cell agglutination, , is given by

defining .

Equation (1)

where is the initial concentration of cells, is the normalized agglutination rate constant and (t) is the time-dependent bound nanobody density.

Time-dependent dynamics of bound nanobody density, , is given by

where and .

Equation (2)

where is the nanobody binding rate constant, is the initial concentration of analyte and

Derivation of Equations (1) and (2) are shown in the following section.

## Normalized cell agglutination dynamics

The time-dependent dynamics of the concentration of *k*-mer clusters, , where and is a non-zero integer, can be described by Equation (3) below.

Equation (3)

At the initial timepoint, it is assumed that there is no binding interaction between nanobody and analyte, so all clusters have one cell each (k=1). The concentration of clusters is, therefore, also the same as the initial concentration of cells, . This can be represented as , whereas for all . The bound nanobody density refers to the fraction of nanobody that is bound per cell, and ranges between 0 and 1 (where 1 refers to the case when all nanobody receptors are bound by analytes). Since at the initial timepoint, there is no binding interaction, .

Since the interaction between the bound analyte and free nanobody is a prerequisite of cell agglutination, the dynamics of cell agglutination is dependent on the agglutination rate constant , bound nanobody density, free nanobody density , and the total number of ways involved in constructing and deconstructing *k*-mer clusters, represented by the two summations. A halving factor is applied to the first summation to avoid double-counting the ways of constructing *k*-mer clusters, as the sum has already considered all possible ways to construct the *k*-mer cluster twice5 given that the order of coalescing clusters is not important.

## Free analyte concentration dynamics

Equation (4)

## Bound nanobody density dynamics (on the cell surface)

Equation (5)

Equation (4) describes the dynamics of free analyte concentration () in the system, where the concentration decreases when free analytes bind to nanobody but increases when nanobody-bound analytes dissociate to become free analytes. The term represents the concentration of free nanobody, whereas the term represents the concentration of bound nanobody. is the number of nanobody per cell, whereas is the initial concentration of cell, and the product of these two constants, , is the concentration of nanobody. is the analyte-nanobody binding rate constant whereas is the analyte-nanobody dissociation rate constant.

Equation (5) describes dynamics of bound nanobody density, where the density increases when free analyte () bind to free nanobody but decreases when analyte-nanobody dissociates to produce free analyte and free nanobody.

## Mathematical model derivation

### Equation (1): Time-dependent dynamics of normalized cell agglutination,

From Equation (3):

Let be the total concentration of agglutinating clusters and cells such that

Equation (6)

Summing both sides of Equation (3) from k=1 to ∞,

Equation (7)

On the LHS, summing the differential terms gives

On the RHS, looking at the double summation, since , the summation over to (omitting *t*), noting that for is as below.

Combining LHS and RHS of Equation (7) and applying Equation (6),

Equation (8)

Let be the normalized cell agglutination representing the fraction of cells involved in agglutination such that

where . It is important to note that as cell agglutination progresses, the total number of clusters of cells decreases (as more cells/clusters coalesce to form larger clusters), so there is a net decrease of over time from 1 to 0. This can be seen clearly by noting that the sign of in Equation (8) is always negative.

Substituting into Equation (8) yields

Equation (9)

Let , substituting into Equation (9) yields Equation (1).

Equation (1)

### Equation (2): Time-dependent dynamics of bound nanobody density

From Equation (4),

Extracting the common factor , and substituting Equation (5) to Equation (4),

Equation (10)

Let , where is the ratio of concentrations of nanobody to analyte, and is the total number of analytes,

Equation (11)

To calculate the constant , let . Since initially, the concentration of free analyte is equal to initial concentration of analyte , and there is no analyte bound so ,

Equation (12)

From Equation (5),

Substituting Equation (12) into Equation (5), and let ,

Equation (2)

A list of parameters and their corresponding descriptions are shown below in Supplementary Table 1. Typical values of certain parameters are also included where appropriate, with references indicated in square brackets.

Supplementary Table 1: List of parameters used in model simulations.

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Description** | **Value (units)** |
|  | Time-dependent concentration of k-mer cluster, where j + l = k | N/A |
|  | Time-dependent concentration of j-mer cluster, where j + l = k | N/A |
|  | Time-dependent concentration of l-mer cluster, where j + l = k | N/A |
|  | Normalized agglutination rate constant; kinetic constant related to agglutination interaction between cell clusters | 1 (h-1)**a** |
|  | Binding rate constant; kinetic constant related to binding interaction between free analyte and nanobody   * anti-GFP nanobody * anti-fib nanobody (hFibNb1) * anti-fib nanobody (hFibNb2) | 3600 x 7.68E+056  3600 x 6.62E+057  3600 x 1.02E+057  (M-1 h-1) |
|  | Dissociation rate constant; kinetic constant related to dissociation of bound analyte from nanobody   * anti-GFP nanobody * anti-fib nanobody (hFibNb1) * anti-fib nanobody (hFibNb2) | 3600 x 1.74E-046  3600 x 2.1E-037  3600 x 2.52E-027  (h-1) |
|  | Time-dependent bound nanobody density, where | N/A |
|  | Time-dependent free nanobody density | N/A |
|  | Time-dependent concentration of free analyte in the system | N/A |
|  | Number of nanobody per cell   * anti-GFP cells * anti-fib cells | 1E+04**b**  3.914E+04**b** |
|  | Initial concentration of cells   * anti-GFP cells * anti-fib cells | 1.38E-12 **c**  1.51E-12 **c**  (M) |
|  | Initial concentration of analyte | 1E-10 to 1E-07(M) **d** |
|  | Total concentration of agglutinating clusters and cells | N/A |
|  | Normalized cell agglutination, where | N/A |

Notes:

|  |  |
| --- | --- |
| **a** | is set to 1 h-1 as the default value, as there is insufficient experimental characterization data for estimating the kinetic constant of agglutination reaction. |
| **b** | Determined from nanobody quantification assays performed in this project. |
| **c** | This range is selected to correspond to the number of cells that allows visualization of cell pellets in round-bottom wells in agglutination assays |
| **d** | This range is selected to correspond to the conditions tested in agglutination assays |

## Sensitivity analysis

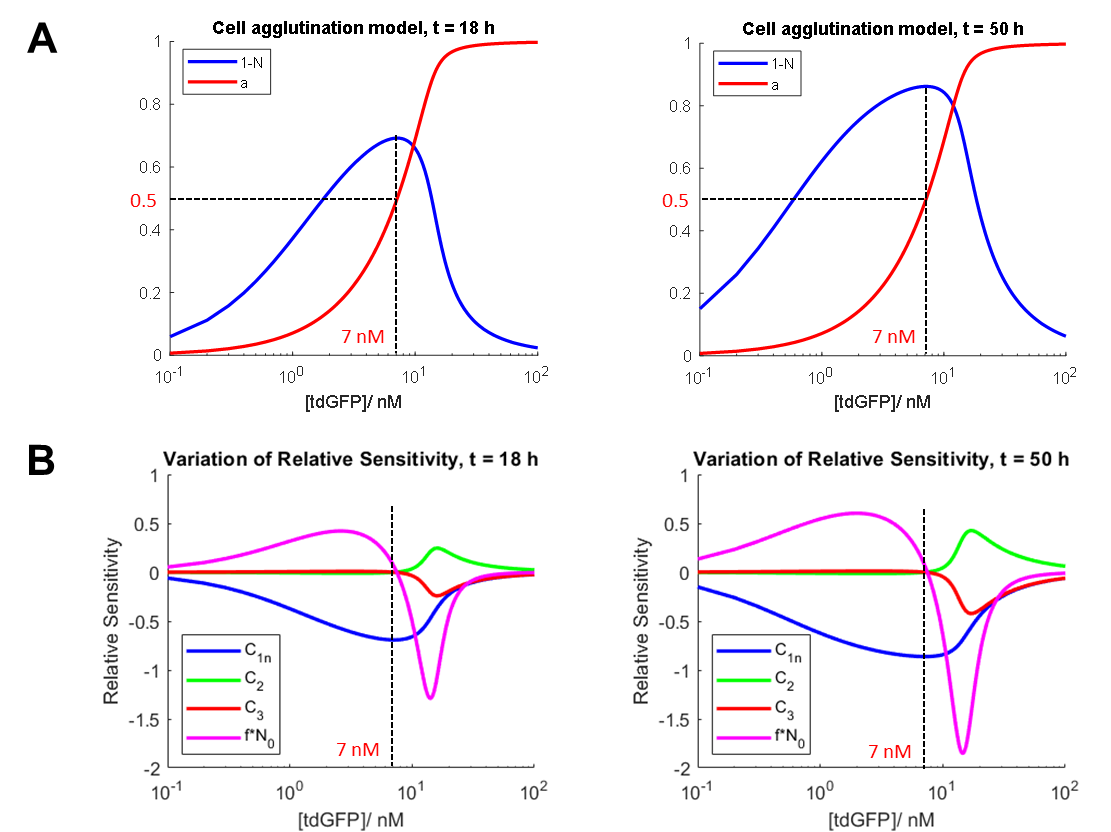
To calculate the relative sensitivity of parameters in the model, we adopted the SENS\_SYS Matlab script (scripts and accompanying documentation are available on [MathWorks forum](https://uk.mathworks.com/matlabcentral/fileexchange/1480-sensitivity-analysis-for-odes-and-daes))8 and plugged in our model parameters. Briefly, the absolute sensitivity of a parameter, *p*, in a function *y* is the partial differentiation of *y* with respect to *p*, whereas the relative sensitivity of a parameter is the product of its absolute sensitivity and the ratio of *p* over *y*. In our model, we are determining the sensitivity of the parameters , , and in the function . Due to the wide range of numerical values between the different parameters (up to 19 orders of magnitude), the relative sensitivity calculation would be more suitable in this model as it takes the magnitude of parameters into consideration.

The initial values of the parameters used in the sensitivity analysis can be found in Supplementary Table 2.

Supplementary Table 2: Initial parameter values for sensitivity analysis

|  |  |  |
| --- | --- | --- |
| Parameter | Initial value of parameter | Unit |
| Normalized agglutination rate constant, | 1 | h-1 |
| Binding rate constant, | 3600 x 7.68 x 105 | M-1 h-1 |
| Dissociation rate constant, | 3600 x 1.74 x 10-4 | h-1 |
| Total concentration of nanobody, | 1.38\*10-8 | M |

We have chosen to set the time at which the normalized cell agglutination, , is calculated to be 18 hours, as this is a typical period for an overnight incubation of the cell agglutination assay. We have probed the changes of model when this value of time is changed from 18 h to 50 h, and we found that while the shape of the cell agglutination profile with respect to [tdGFP] changes slightly, it does not change the position of the peak which remains at [tdGFP]max at 7 nM (Supplementary Figure 4A). The sensitivity analysis graphs also shows similar peak profiles (Supplementary Figure 4B), suggesting that 18 h is a representative timepoint for sensitivity analysis.



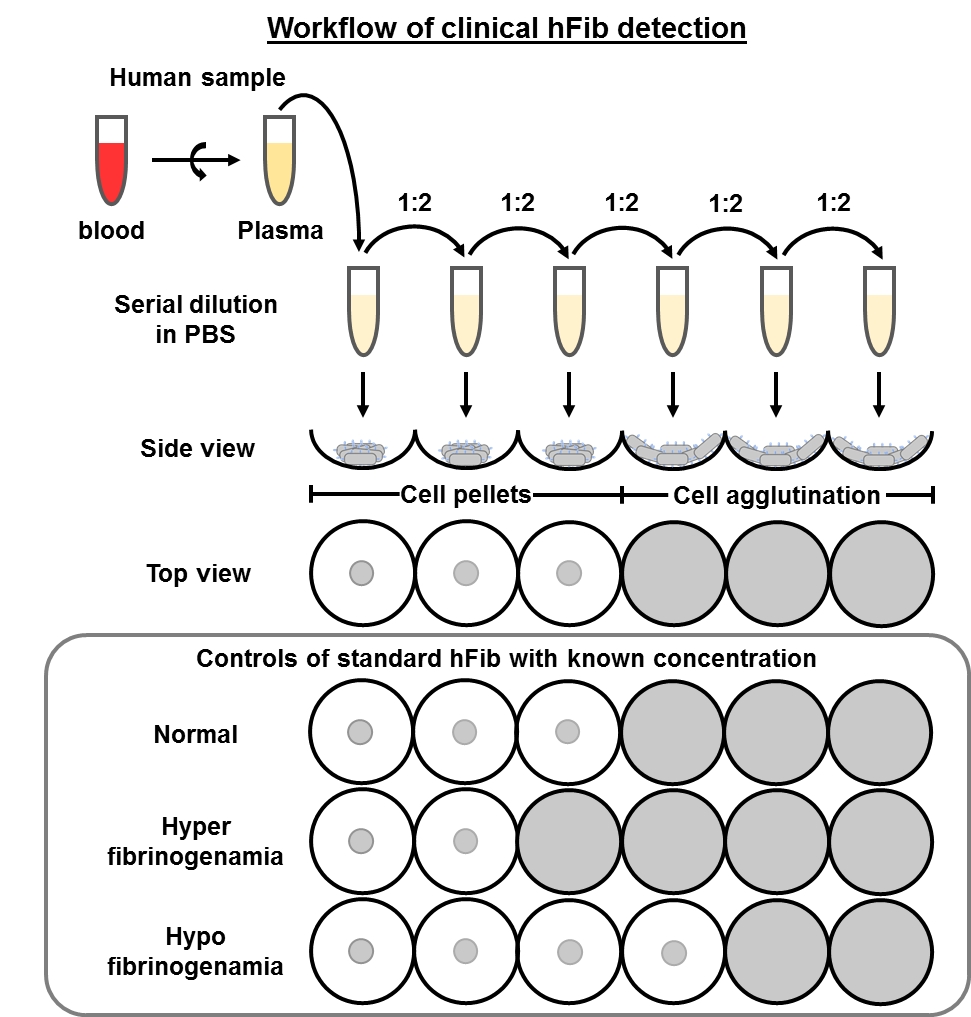
Supplementary Figure 4: Simulation of agglutination reaction and sensitivity analysis.

A Cell agglutination model simulated at 18 h and 50 h timepoints. The dotted line shows [tdGFP]max, which corresponds to max, and . B Relative sensitivity of model simulated at 18 h and 50 h time endpoints. Parameters are defined as in Supplementary Table 2. Dotted line indicates maximum agglutination at 7 nM tdGFP as the reference point.

# SUPPLEMENTARY NOTE 4

## Diagnostic workflow

The proposed diagnostic workflow mirrors what occurs in a hemagglutination assay (Supplementary Figure 5). Blood samples from a patient would be centrifuged to remove the red blood cells and plasma fraction used in the assay. Serial dilutions of the patient sample would be mixed with the whole-cell biosensor and incubated as described in the materials and methods. Depending on the concentration of hFib in the patient sample, the equivalence zone will change, e.g. samples containing a higher concentration will show agglutination with less serial dilution than those with a lower concentration. One or more standards of known concentration could be included to provide a concentration range estimate for the sample (semi-quantitative assay), which can be back-calculated from the degree of dilution required for agglutination to occur.



Supplementary Figure 5: Schematic of the workflow of clinical hFib detection

# SUPPLEMENTARY NOTE 5

Anti-fibrinogen nanobody cell-surface display

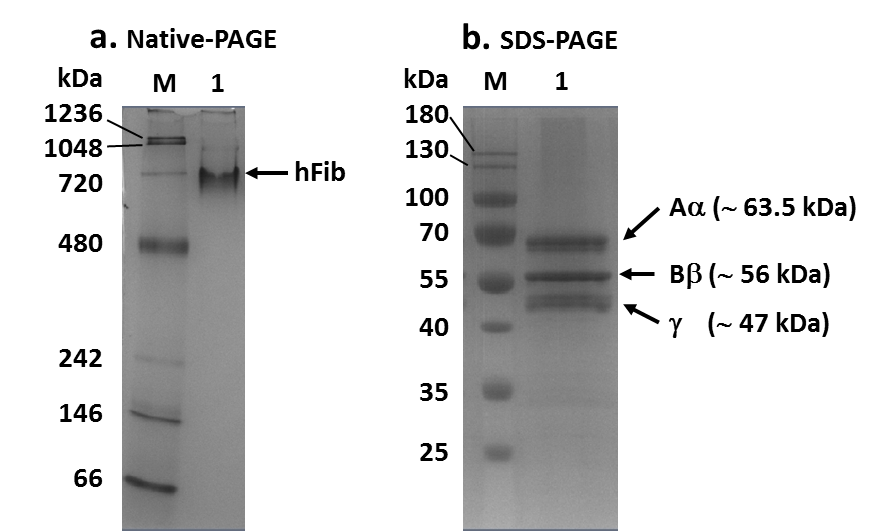
An experiment was designed to investigate the localization of the anti-fibrinogen nanobodies (NbFib1-2)) and their functionality as sensor elements when displayed on the surface of the host cell. *E. coli* DH10B strain cells harboring the pNVfib1 (or pNVfib2) were labeled with anti-myc antibody fluorescent conjugate as described in the Materials and Methods. After washing the cells to remove excess labelling antibody, cells were analyzed by flow cytometry (Supplementary Figure 6A) and fluorescence microscopy (Supplementary Figure 6B) to confirm cell surface display.

****

Supplementary Figure 6: Anti-hFib nanobody display

Cell surface expression of anti-hFib nanobodies (NbFib1-2) labeled with Myc-Tag (9B11) Mouse mAb (Alexa Fluor® 488 Conjugate) determined by flow cytometry (A) and fluorescence microscopy (B). Nanobodies expression was fully induced by the use of IPTG. Microscopic images were taken at 100 ms exposure time using a magnification of 60x.

PAGE analysis of human fibrinogen

****

**Supplementary Figure 7: Aggregation of hFib in solution** Purified human fibrinogen (hFib) was analysed on 4-16% Native-PAGE (a) and 12.5% SDS-PAGE (b). Lane M: NativeMark unstained protein standard (for Native gel) or Pageruler prestained protein ladder (for SDS gel). Lane 1: ~6.7 ng hFib. Arrows represent the size of hFib on native gel and subunits on SDS gel.

# SUPPLEMENTARY REFERENCES

(1) Chen, G.; Cloud, J.; Georgiou, G.; Iverson, B. L. A Quantitative Immunoassay Utilizing Escherichiacoli Cells Possessing Surface-Expressed Single Chain Fv Molecules. *Biotechnology Progress* **1996,** *12* (4), 572-574.

(2) Moran, U.; Phillips, R.; Milo, R. SnapShot: Key Numbers in Biology. *Cell* **2010,** *141* (7), 1262.

(3) Chen, Y.; Wei, L.-N.; Müller, J. D. Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. *Proceedings of the National Academy of Sciences* **2003,** *100* (26), 15492-15497.

(4) Dolgosheina, E. B.; Karulin, A.; Bobylev, A. V. A kinetic model of the agglutination process. *Mathematical Biosciences* **1992,** *109* (1), 1-10.

(5) Wattis, J. A. D. An introduction to mathematical models of coagulation–fragmentation processes: A discrete deterministic mean-field approach. *Physica D: Nonlinear Phenomena* **2006,** *222* (1), 1-20.

(6) Rothbauer, U.; Zolghadr, K.; Tillib, S.; Nowak, D.; Schermelleh, L.; Gahl, A.; Backmann, N.; Conrath, K.; Muyldermans, S.; Cardoso, M. C.et al. Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nature Methods* **2006,** *3*, 887-889.

(7) Salema, V.; Lopez-Guajardo, A.; Gutierrez, C.; Mencia, M.; Fernandez, L. A. Characterization of nanobodies binding human fibrinogen selected by E. coli display. *Journal of Biotechnology* **2016**, 58-65.

(8) Garcia Molla, V. M.; Gomez Padilla, R. Sensitivity analysis for ODEs and DAEs. **2002,** (02 June 2018). https://uk.mathworks.com/matlabcentral/fileexchange/1480-sensitivity-analysis-for-odes-and-daes