Selective Enrichment of Hypermethylated DNA by a Multivalent Binding Platform

Ruben W. Kolkman, Loes I. Segerink,* and Jurriaan Huskens*

The preselection of hypermethylated DNA (hmDNA) from liquid biopsy samples is key to enable early-stage cancer diagnostics. Due to limited selectivity of the existing preselection approaches, however, wide integration in the clinic is currently prohibited. Here, it is argued that an affinity method on a surface, such as used in affinity chromatography, can be significantly improved by employing the principles of multivalency and superselectivity. In the here proposed method, a methyl binding domain (MBD) protein immobilized at a surface is used as a receptor for (hyper)methylated DNA. By the organization of multiple MBDS on a surface, a multivalent binding platform is achieved. The MBD surface receptor density on that platform is key to increase the enrichment selectivity of hmDNA as a single MBD protein binds both methylated and non-methylated DNA with a small difference in affinity. When the receptor density is varied, multivalent analyte binding typically responds in a non-linear fashion, which phenomenon is called superselectivity. By careful tuning of the MBD density, it is envisaged that the selectivity for methylated over non-methylated DNA can be optimized. Strong applicability is foreseen in a medical setting by implementing such an enrichment step in an analytical process or a lab-on-a-chip device.

1. Introduction

Cancer is a growing health concern worldwide.[1] Despite the increase in cancer survival rates over the last decades, cancer remains one of the deadliest diseases.[2,3] One of the key aspects of surviving cancer is to detect the disease as early as possible, thereby enlarging the window of opportunity of curing it.[4,5] A recent trend for early cancer diagnostics is the pre-symptomatic detection of cancer biomarkers in liquid biopsy samples and, therefore, improving the survival rate.[6,7]

Hypermethylated DNA (hmDNA) is one of the typical biomarkers found in liquid biopsy samples of cancer patients.[8,9] The presence of hmDNA in blood and urine is correlated to multiple types of cancer, including gastric, lung and ovarian cancer.[10–13] From a genetic point of view, DNA methylation takes place predominantly at promoter regions with a relative high amount of cytosine bases that are followed directly by a guanine (CpGs) in the 5′-to-3′ direction, the so-called CpG-rich regions.[11] The methylation of a CpG is an epigenetic alternation in which a methyl group is covalently bonded to the cytosine base at the fifth carbon. By CpG methylation, gene expressions are controlled in cells. As a consequence, methylation can play a role in tumor development when tumor suppressor genes are methylated, thereby repressing the transcription and thus silencing these genes.[14] Hypermethylation refers to the situation that methylation of a promoter region occurs, which in a healthy situation does not occur.

Early disease detection as well as disease progress and the effectiveness of a therapy can be monitored by measuring the hmDNA concentration over time.[15] Yet, the concentration of hmDNA, especially in early-stage cancers, can be as low as a few DNA copies per liquid biopsy sample.[16–18] The current approach to measure hmDNA employs DNA isolation and bisulfite conversion, making it time-consuming and labor intensive. As a result, the method is not widely applicable in the clinic. Typically, hmDNA is detected in bisulfite-treated DNA samples by quantitative polymerase chain reaction (qPCR, of the region of interest; can be cancer dependent). Alternative hmDNA detection approaches that may be suitable for use in point-of-care applications, include electrochemistry,[19] CRISPR,[20] and isothermal amplification.[21]

The detection of specific hmDNA biomarkers by sequencing or biosensing approaches involves the ability to distinguish between hmDNA and non-methylated DNA.[22–24] Commonly a preselection step is applied, and three different approaches exist in order to differentiate between hmDNA and non-methylated
“background” DNA.\cite{25-27} The first approach uses methyl-sensitive restriction enzymes.\cite{28-30} These enzymes are able to cleave DNA at a specific recognition sequence only when the CpG is (non-)methylated, depending on the enzyme used. For example, after digesting the DNA sample with a specific enzyme, all the non-methylated DNA is cleaved, while the methylated DNA remains intact. The second approach to differentiate between the two types of DNA is bisulfite conversion.\cite{31,32} Using this chemical modification method, all non-methylated cytosines are converted into the base uracil, while all methylated cytosines remain unconverted. As a result, the base sequence of the non-methylated DNA changes, allowing an easier differentiation using PCR. The third method uses affinity chromatography purification with an anti-5-methyl-cytosine antibody\cite{33} or a methyl binding domain (MBD)\cite{34,35} protein attached to the surface of beads employed in the method. This approach is based on enriching a DNA sample with hmDNA due to binding of the methylated DNA to the biorecognition element-modified surface. The used biorecognition element, or receptor, that is immobilized on the surface possesses a higher affinity for the hmDNA compared to non-methylated DNA thus enabling the separation. Each of the aforementioned preselection methods has its limitations regarding the selectivity, thus resulting in false positives and affecting negatively the ease of application in early-stage cancer diagnostics.\cite{25-27,36-38} Specifically, the use of restriction enzymes can cause incomplete digestion;\cite{25,27,36-38} bisulfite conversion results in DNA degradation;\cite{25,27-29,36-38}; and affinity purification suffers from the unwanted enrichment of non-methylated DNA.\cite{25,37,38} For these reasons, there is a need for improving the preselection of hmDNA before implementa-
tion in the clinic can be considered.

Here, we will focus on a concept that provides the enrichment of hmDNA using an MBD protein bound to a capture surface by applying the principles of multivalent binding, and the superselective binding of hmDNA that results from it. Multivalency is a principle of binding that provides affinity, selectivity, and organization.\cite{39-42} The enhancement in binding strength of multivalent targets in comparison to monovalent ones is known as the multivalency effect or avidity.\cite{43} When varying the receptor density on a surface, superselectivity can occur.\cite{44} Superselectivity is the on/off binding behavior of multivalent targets that results from a non-linear dependence of the target coverage on the receptor density. Below a specific threshold receptor density, no binding takes place, while above this density the target binds with close to maximum surface coverage. Despite many of the fundamental aspects of multivalency and superselectivity have been addressed, few practical applications have resulted from it. Here, in the specific case of DNA binding to an MBD-covered surface, we envisage that the threshold density at which the transition from non-binding to binding occurs will depend on the number and density of CpG sites on the DNA and on whether these sites are methylated or not. Thus, one can rationalize conceptually that a careful tuning of the MBD density can result in hmDNA enrichment of DNA isolated from a liquid biopsy while non-methylated or less-methylated DNA is not enriched, thus, enabling selective hmDNA-based cancer diagnostics (Figure 1). Thus we aim to outline the strategy to develop a diagnostic tool that employs the principles of multivalency and superselectivity in hmDNA enrichment.

In the concept proposed here, the enrichment of hmDNA from a solution can be achieved by immobilizing an MBD protein on a surface, possibly that of a column material or micro- and nanoparticles.\cite{34,37,45} The MBD proteins on the surface function as receptors that can interact with the methylated CpGs (CpG); the ligands from the DNA in solution. The binding site of MBD proteins interacts both with non-methylated CpG (CpG) and CpG, but the affinity for CpG is higher.\cite{46-52} The multiple CpG sites in hmDNA allow the formation of multiple non-covalent bonds to surface-immobilized MBD, resulting in an overall higher binding strength.

Here, we discuss how the general strategy of hmDNA enrichment using a capture surface coated with an MBD protein can be designed optimally by exploiting the principles of superselectivity and multivalent binding. We first discuss the binding performance and selectivities of the protein subtypes.

![Figure 1. Conceptual illustration of the envisaged hmDNA enrichment process. DNA is isolated from a blood liquid biopsy containing both hmDNA and non-methylated DNA. CpG and CpG are illustrated here by red and green hexagons, respectively, attached to the DNA at the top and bottom strand. Next, superselective enrichment of hmDNA on a capture surface with an optimal MBD surface receptor density. Only DNA with multiple CpGs are able to bind to the surface due to a sufficient avidity. The hmDNA-enriched DNA sample is analyzed with qPCR to detect a cancer-specific hmDNA sequence.](image-url)
present within the MBD protein family toward methylated and non-methylated DNA. Thereafter we explain the principles of the enrichment strategy using an MBD protein, and we discuss the prerequisites to achieve highly selective multivalent binding of targets to a receptor-modified surface in various applications. Finally, we will conclude with an outlook on how the enrichment of hmDNA with superselectivity properties can be developed.

2. MBD Protein As a Receptor for hmDNA

The family of MBD proteins contains five members that have a higher binding affinity for methylated DNA compared to non-methylated DNA. These are methyl-CpG binding protein 2 (MeCP2), MBD1, MBD2, MBD3, and MBD4. Despite all are selective for a C*pG, the absolute binding affinities of the individual proteins within the MBD family for C*pG and CpG differ. The binding affinities of an MBD protein that is interacting with C*pG or CpG is important for the overall affinity with (non-)methylated DNA and for the selectivity of the enrichment of hmDNA.

The dissociation constants ($K_d$) of the C*pG-selective MBD proteins for both C*pG or CpG are compared in Table 1. The data reported in Table 1 shows that in particular MeCP2, MBD1, MBD2, MBD3, and MBD4 are comparable, affinities for CpG. However, the direct comparison of the $K_d$ values between different studies is complicated by the use of different DNA lengths and CpG positions, of slightly different reaction buffer compositions, and of different techniques to measure $K_d$. In the study by Fraga et al. it was concluded that the $K_d$ values were strongly dependent on the used DNA sequence.

In particular the ratio of the $K_d$ values for CpG and C*pG, which is a measure for the selectivity, can be a proper measure to identify the most optimal receptor for the enrichment. It can be expected that the factors complicating the comparison of $K_d$ values between different proteins is largely eliminated by taking ratios, so by looking at the selectivities. As seen from Table 1, high selectivities ($K_d$ of CpG/C*pG) are achieved by MeCP2, MBD1 and MBD2, with values ranging from 50–295. These of proteins, MBD1 seems to have the best selectivity.

From this analysis, decent individual selectivities are obtained for MeCP2, MBD1 and MBD2, which are much better than those of MBD3 and MBD4. Most likely, any of the first three can be used in achieving enrichment. Yet, as will become clear from the multivalency and superselectivity discussion below, the intrinsic selectivity (i.e., the ratio of $K_d$ values for individual CpG/C*pG sites) is not enough to predict the best selectivity in a multivalent binding event for (methylated) DNA that contains multiple C*pG sites. Qualitatively, the overall selectivity for binding hmDNA will be better when a good intrinsic selectivity is coupled to a relatively weak binding constant (high $K_d$), in particular for CpG, and high valency (number of interactions) leading to a “multiplication” of several individual selectivities. In that way, a large difference in overall $K_d$ values for hmDNA and non-methylated DNA can be achieved in the biologically relevant concentration range. The next sections focus only on the use of the MBD2 protein, as this is the most used MBD protein nowadays in the literature for the enrichment of hmDNA.

3. Performance of Existing hmDNA Enrichment Methods Using MBD2

The enrichment of hmDNA is commonly performed by a surface modified with the MBD2 protein. The general approach of enrichment consists of three consecutive steps. In the first step, MBD2 is immobilized onto the surface, such as a column material or nano/micro-size particles.

Table 1. Reported $K_d$ (nm) values of the MBD protein family upon interaction with DNA containing a C(*p)G in various studies. In the table the used selectivity ($K_d$/C*pG) DNA length in base pairs (bp), number of C(*p)G per sequences, location of C(*p)G in bp position from 5′ to 3′ direction, characterization technique, and the studies are displayed. Data not determined or displayed in a study is shown as “ND.”

| MBD type | $K_{d,CpG}$ (nm) | $K_{d,C*pG}$ (nm) | Selectivity | DNA length (bp) | Number of C(*p)Gs | Location of C(*p)G (bp) | Characterization technique | Reference |
|----------|------------------|-------------------|-------------|----------------|------------------|----------------------|--------------------------|-----------|
| MeCP2    | 10               | 500               | 50          | 12             | 1                | 5                    | Fluorescence polarization | Hashimoto et al. [46] |
|          | 272              | 458               | 3           | 42             | 1                | 10                   | Capillary electrophoresis | Fraga et al. [47] |
|          | 8                | ND                | ND          | 24             | 1                | 12                   | Electroforetic mobility shift assay | Buchmuller et al. [48] |
| MBD1     | 5                | 1400              | 280         | 12             | 1                | 5                    | Fluorescence anisotropy | Liu et al. [49] |
|          | 244              | 72 000            | 295         | 22             | 1                | 10                   | Fluorescence polarization | Hashimoto et al. [50] |
| MBD2     | 60               | 6500              | 108         | 12             | 1                | 5                    | Fluorescence polarization | Hashimoto et al. [51] |
|          | 3                | 188               | 63          | 42             | 1                | 20                   | Capillary electrophoresis | Fraga et al. [52] |
|          | 6                | ND                | ND          | 22             | 1                | 10                   | Fluorescence spectroscopy | Heimer et al. [53] |
|          | 12               | ND                | ND          | ND             | ND               | ND                   | Fluorescence anisotropy | Liu et al. [54] |
|          | 29               | ND                | ND          | 17             | 1                | 8                    | Fluorescence polarization | Pan et al. [55] |
| MBD3     | 1300             | 6600              | 5           | 12             | 1                | 5                    | Fluorescence polarization | Hashimoto et al. [56] |
|          | 580              | 684               | 1           | 42             | 1                | 20                   | Capillary electrophoresis | Fraga et al. [57] |
| MBD4     | 220              | 1070              | 5           | 12             | 1                | 5                    | Fluorescence polarization | Hashimoto et al. [58] |
In general, proteins can be immobilized by a wide variety of approaches, but a detailed description of the protocols is outside the scope of this article. The second step is the actual binding of hmDNA to the MBD2-modified surface. The hmDNA binds to multiple MBD2 proteins when the spacing between C\(^5\)pGs is between 3–20 base pairs. This minimal distance is required to prevent steric hindrance. In the last step, the bound DNA is removed from the surface with an elution buffer to enable subsequent characterization. Table 2 provides an overview of various MBD2-based enrichment methods including DNA processing steps applied pre/post hmDNA enrichment.

A performance study of MBD2 used for hmDNA enrichment was performed by Nair et al. Here, biotinylated MBD2 was immobilized on streptavidin-coated magnetic beads. Three different genes were used: a partly methylated DNA, fully methylated DNA, and fully non-methylated DNA originating from the cancerous LNCaP cell lines. As a control, the PrEC cell line was used, where all different genes were non-methylated. All non-methylated DNA was not significantly interacting with the MBD2-modified beads, while the partly or fully methylated DNA genes were. The authors were able to discriminate between the partly and fully methylated DNA using a salt wash with a NaCl concentration ranging between 0.2 and 2 M, as was validated by quantitative PCR (qPCR).

In a study by Warton et al., a method was developed to improve the methylated DNA enrichment by decreasing the amount of MBD2-functionalized magnetic beads upon enrichment of DNA. They found a positive effect on the enrichment factor (the enriched amount of methylated DNA divided by the enriched amount of non-methylated DNA in a separate experiment) that was increased from a factor 7 to 30 by decreasing the amount of magnetic beads. An alternative is the incorporation of sperm DNA as blocking agent upon hmDNA enrichment. The usage of the blocking agent improved the enrichment factor from 2.1 to 14.2.

The selectivity of hmDNA enrichment was studied in more detail by Yegnasubramanian et al. by using fragmented genomic DNA from white blood cells at input amounts ranging between 32 pg and 100 ng. The genomic DNA was methylated using the M.SssI enzyme to create hmDNA. Here, magnetic particles covered with anti-histidine antibodies were modified with His-tagged MBD2 in the presence of 200 ng non-methylated plasmid DNA and used for the enrichment. After a 1 h incubation step with DNA, the magnetic beads were isolated with a magnetic rack. Subsequently, the beads were washed and followed by the elution of bound DNA using a temperature of 95 °C. Enrichment factors of 500-700-fold were found by qPCR at 4 and 100 ng hmDNA and non-methylated DNA input concentrations. Nevertheless, the authors showed that the selectivity of an MBD2-based enrichment method is strongly dependent on the DNA input concentration (Figure 2). For example, enrichment factors of only 1 and 5 were achieved when using 11 ng non-methylated DNA and 0.03 or 0.1 ng of hmDNA, respectively. To improve the limited MBD2 enrichment selectivity, especially at low hmDNA and high non-methylated DNA input concentrations, they used methyl-sensitive restriction enzymes prior to the enrichment. This resulted in the detection of 32 pg hmDNA in a sample containing more DNA.

**Table 2.** Applications of DNA samples enriched in hmDNA using MBD2-modified surfaces. Application, processing pre/post hmDNA enrichment, DNA origin, and references are displayed. Circulating free DNA and recombinase polymerase amplification are abbreviated as cfDNA and RPA, respectively.

| Application | DNA processing steps (pre/post hmDNA enrichment) | DNA origin | References |
|-------------|---------------------------------------------------|------------|------------|
| CpG methylation profiling | Sonication of genomic DNA (pre and whole genome amplification (post)) | Prostate normal and cancer cells | Nair et al. [65] |
| Next-generation sequencing | - | cfDNA from plasma/serum | Warton et al. [38] |
| Calorimetric and/or electrochemical detection of genome methylation and methylation-specific genes | Restriction enzyme digest of | Cancer/control cell lines and urine of prostate cancer patients | Wee et al. [66] |
| hmDNA-based prostate cancer detection | Methyl-sensitive restriction enzyme digest (pre) | Prostate cancer cell line and tissue | Yegnasubramanian et al. [37] |

![Figure 2](image_url) DNA input concentrations versus the amount of output DNA after hmDNA enrichment as quantified by qPCR. Fragmented genomic DNA from white blood cells was used in its non-methylated form (WBC, white filled circles) and in the hmDNA form (WBC M.SssI, black filled circles). hmDNA was formed by methylation of the non-methylated genomic DNA with the M.SssI enzyme. Reproduced with permission. [37] Copyright 2006, Oxford University Press.
than a 3000-fold higher concentration of non-methylated genomic DNA.

All the standalone MBD2-based enrichment systems reported and discussed here did only report the fold of enrichment based on the quantification after enriching a solution with (hyper)methylated DNA or non-methylated DNA in separate experiments. Unfortunately, the selectivity of enriching DNA solutions containing both (hyper)methylated DNA and non-methylated DNA are currently unknown despite its importance for clinical implications. The enrichment selectivity determination of DNA mixtures is crucial as the concentration of hmDNA biomarkers in liquid biopsies can be orders of magnitude lower compared to that of non-methylated “background” DNA.[16–18] Furthermore, detailed understanding of the enrichment selectivity of DNA mixtures would provide information in which clinically relevant hmDNA concentration can be differentiated from the non-methylated DNA. The MBD2-based enrichment systems were also not quantitatively tested with various DNA input concentrations in most of the studies.[38,65,66] Furthermore, the highly selective enrichment system developed by Yegnasubramanian et al.[37] is unfortunately rather complex, lengthy to perform, and needs methyl-sensitive restriction enzymes which are restricted in use due to the limited amount of available recognition sequences. Overall, the MBD2-based enrichment systems are only able to differentiate well between hmDNA and non-methylated DNA at high input concentrations but not at combined low and high input concentrations of hmDNA and non-methylated DNA, respectively. For that reason, the current hmDNA enrichment selectivity is too limited to function as a simple standalone system without any post-modification steps to enable wide integration in the clinic in early cancer diagnostics. To enable early-stage cancer diagnostics, the hmDNA enrichment selectivity should be increased. Additionally, the effect on the hmDNA enrichment selectivity using DNA mixtures with various hmDNA/non-methylated DNA ratios should be determined. Finally, the impact of different DNA input concentrations on the hmDNA enrichment should be verified.

4. Multivalency As a Tool for Superselective Biosensing

hmDNA contains multiple C*pGs, each of which functions as a ligand that can bind to the MBD2 receptors on the surface during enrichment. At increasing ligand and receptor densities, more non-covalent interaction pairs can be formed, which is a typical example of multivalent binding (Figure 3A).[43,67,68] Interaction pairs are formed when the geometrical locations between the ligand and receptor match.[66] Multivalent binding results into strong but reversible complex formation for which the overall binding strength, defined by the avidity constant ($K_a$), is significantly higher compared to a monovalent binding event.[69] Additionally, the $K_a$ is increasing faster than the sum of the individual binding interactions. The avidity is strongly dependent on both the ligand and receptor densities, and more stable complexes are formed at higher densities.[70] An example is the binding of Affitin receptors toward the pathogen Staphylococcus aureus.[71] The binding affinity was increased by the formation of a dendrimer containing multiple Affitin receptors. As a result, an approximately 600-fold improvement in the $K_a$ value down to $0.2\text{ nm}$ was achieved for the dendrimer-based system in comparison to the monovalent binding of Affitin to Staphylococcus aureus. Another example is the trapping of viruses by multivalent interactions with DNA-based shells functionalized with antibodies.[72] The functionalized shell binds as sufficiently strong to the virus owing to multivalent interactions resulting in the suppression of virus binding to the cell surface.

Multivalent binding is composed of inter- and intramolecular interactions.[73] At interfaces, all ligands of an individual target bind to a receptor immobilized on a surface, for example, of a micro/nanoparticle, a cell, or a self-assembled monolayer. The advantages of surface binding is that a higher avidity constant and thus more stable complexes can be formed as a result of a higher effective molarity ($EM$).[74,75] $EM$ describes the probability for intramolecular ring closure.[76] The overall binding or avidity constant, $K_{av}$, is dependent on the monovalent ligand-receptor binding strength ($K_a$), $EM$ and the number of ligand-receptor interaction pairs ($n$), as described by Equation (I).[43]

![Figure 3](https://www.advancedsciencenews.com/)

**Figure 3.** A) Illustration of the monovalent and multivalent interaction of a target to a receptor-modified surface. B) Overview of the degree of target binding as function of the receptor density for monovalent targets with low (green) and high (blue) ligand affinity for the receptor modified surface and the multivalent target with low ligand affinity for the receptor (red).[46] $n_R$ displays the surface receptor density and $\theta$ the target surface coverage. Reproduced with permission.[46] Copyright 2011, National Academy of Sciences.
Superselectivity is used among others as a tool for targeting cells based on variations in their receptor densities to improve cancer cell-specific drug delivery. Different ErbB2 receptor densities on cell surfaces were used to determine the effect on targeting by EC1 ligand-modified nanoparticle micelles. Three regimes of receptor densities were created with low, medium and high number of ErbB2 receptors per cell representing healthy, early and late stage cancer cells, respectively. A decrease of the $K_\infty$ from 12.8 to 0.4 µM using a micelle modified with 40 ligands compared to a monovalent interaction was observed, as was validated using surface plasmon resonance spectroscopy. Additionally, cellular uptake of the ligand-modified micelles depended on both the ligand and receptor density, as observed by flow cytometry. At healthy receptor densities, limited endocytosis took place for an 8-valent nanoparticle which was approximately twofold lower than for the 40-valent one. The endocytosis remained limited for the 8-valent nanoparticles in early and high cancer stage cells but was significant for the 40-valent one, which is attributed to a larger $EM$ and $n$ which resulted in a $K_\infty$ of the micelle that is sufficient for binding to the cancer cells.

A multivalent system is not directly a superselective system. Superselectivity is dependent on the multivalent enhancement factor $K_{EM}$. In fact, if $K_{EM}$ is too high, less or no superselectivity is observed. In such a case, the dependence of target binding at increasing receptor density is less strong. The preference of superselectivity for a low $K_{EM}$ is explained by noting that in such a “weak multivalent” system multiple ligand–receptor interaction pairs are mandatory to enable binding of the target to the surface. The location of ligands on a polymer that binds multivalently at a receptor-modified surface is crucial as well for the degree of superselectivity, as was described in a theoretical study by Tito et al. The superselectivity was more pronounced when a polymer with locally assembled ligands was used at reduced ligand densities. The increase in superselectivity was attributed to reduced competition at the sensor surface upon binding, lesser blocking of the binding sites, and increase in cooperativity.

The importance of superselectivity in discriminating between different targets on a receptor-modified surface was reported in experimental work by Dubacheva et al. Their system was based on the interaction between a polymer modified with cyclodextrin ligands and a monolayer-modified surface containing a tunable receptor density of either ferrocene or adamantane. The two different receptors were used to study the effect of different binding strength, while the $K_\infty$ values of ferrocene and adamantane are 200 and 10 µM, respectively. The higher affinity between the adamantane receptor and the target resulted in superselective binding at an approximately 16-fold lower receptor density to achieve maximum surface coverage compared to the usage of ferrocene receptors. Furthermore, maximal surface coverage of the target on the adamantane-coated surface could be achieved while binding of the target to the ferrocene-coated surfaces remained fully absent. This work has clearly shown that superselectivity enables to discriminate between the target and non-target by optimizing the receptor density.

This section has shown that it is critical to control the multivalent enhancement factor $K_{EM}$ to develop a superselective platform for the enrichment of (hyper)methylated DNA. The effect of the selective MBDs (McCP2, MBD1 and MBD2) on the superselective enrichment of hmDNA should be investigated to decide which of these proteins provides the best superselectivity. The selectivity will, therefore, be a combination of both a good intrinsic selectivity and a low $K_{EM}$. The possible $n$ for both (hyper)methylated and non-methylated DNA depends on the DNA length. In liquid biopsies, such as blood and urine, the DNA is fragmented in comparison with cellular DNA, which can result in a decrease of the valency $n$ for both types of DNA. Due to lower $n$ values of DNA from liquid biopsies, the $K_\infty$ will be lower and, therefore, surface binding of DNA that lacks methylation will be reduced as well.

Conceptually the superselective device should be able to discriminate between hmDNA and non-methylated DNA binding (Figure 4). At low MBD densities, the number of MBD interactions per DNA is too low for both hmDNA and non-methylated DNA to reach a $K_\infty$ sufficient for surface binding. In the medium MBD density regime, multiple MBD interactions are formed with methylated DNA resulting in a sufficiently high $K_\infty$ for surface binding. However, the number of MBD interactions is too little to enable surface binding of non-methylated DNA because of the lower $K_\infty$. On the other hand, if the MBD density becomes too large, too many MBD interactions can be formed with both hmDNA and non-methylated DNA. Consequently, $K_\infty$ is sufficiently high to give surface binding for both types of DNA, resulting in a decrease of the selectivity of hmDNA enrichment.

Equation (1) can be read as follows. Monovalent binding ($n = 1$) leads to a binding constant of $K_i$. For each additional interaction pair in a multivalent complex, the affinity increases with a factor $K_{EM}$. This factor is therefore called the multivalent enhancement factor.

An example of multivalency in nature is the binding of a virus particle on a cell surface. Efficient binding, possibly followed by endocytosis, of the virus particle is only possible in situations when the virus binds multivalently at the cell surface with sufficiently high $EM$ and $n$. Therefore, virus binding can be prevented by the use of monovalent drug molecules that bind at the virus protein sites to decrease $EM$ and $n$ such that $K_\infty$ becomes too weak for surface binding.

Superselectivity is dependent on the multivalent enhancement factor $K_{EM}$. For each additional interaction pair in a multivalent complex, the affinity increases with a factor $K_{EM}$. This factor is therefore called the multivalent enhancement factor.

The importance of superselectivity in discriminating between different targets on a receptor-modified surface was reported in experimental work by Dubacheva et al. Their system was based on the interaction between a polymer modified with cyclodextrin ligands and a monolayer-modified surface containing a tunable receptor density of either ferrocene or adamantane. The two different receptors were used to study the effect of different binding strength, while the $K_\infty$ values of ferrocene and adamantane are 200 and 10 µM, respectively. The higher affinity between the adamantane receptor and the target resulted in superselective binding at an approximately 16-fold lower receptor density to achieve maximum surface coverage compared to the usage of ferrocene receptors. Furthermore, maximal surface coverage of the target on the adamantane-coated surface could be achieved while binding of the target to the ferrocene-coated surfaces remained fully absent. This work has clearly shown that superselectivity enables to discriminate between the target and non-target by optimizing the receptor density.
5. Conclusion and Perspectives

In summary, we have presented arguments that the limited selectivity of current hmDNA enrichment methods based on affinity chromatography can be improved by applying the principles of multivalency and superselectivity. Nowadays hmDNA enrichment methods predominantly use surface-bound MBD2 in which the MBD2 protein functions as a receptor. The limitation of MBD2 is that it possesses a decently, but not exhaustively, higher (63- to 108-fold) binding affinity for C*pG compared to CpG, while similar binding affinities are observed for MeCP2 and MBD1 of the MBD protein family. Because of these relatively small differences in binding affinities, the hmDNA enrichment selectivities are limited. Reduced enrichment selectivities are especially profound upon the enrichment of DNA mixtures when the hmDNA concentration is (much) lower than that of non-methylated DNA. Liquid biopsies also contain hmDNA and non-methylated DNA, therefore, explaining the limited use of hmDNA enrichment in the clinics nowadays.

The here proposed method for increasing the selectivity of hmDNA enrichment employs the MBD surface receptor density as a control parameter, and it is affected by several crucial parameters. First, the density affects both the multivalency enhancement factor and the intrinsic selectivities of the receptors toward a C*pG and CpG. A high hmDNA selectivity will be favored by a low multivalency enhancement factor and a high intrinsic selectivity of the receptors. It should be determined which of the MBDs MeCP2, MBD1, and or MBD2 can fulfill the first parameter the best. Second, the number of CpGs within a sequence and its degree of methylation play an important role. To illustrate, the number of C*pGs in a DNA fragment can be as high as 1 C*pG per 10 base pairs, but this is sequence dependent. Increasing CpG methylation results into a higher $K_\text{av}$ with the MBD-modified surface. Therefore, lower MBD surface receptor densities can be used at higher degrees of CpG methylation. Future work should clarify what the role of increasing CpG methylation is on the binding at various MBD surface receptor densities. Third is the role of the DNA length in the enrichment of hmDNA. The effect of the DNA length is an important aspect of study, since various lengths of DNA are present in liquid biopsies. For example, the main DNA fragment sizes in blood and urine are $\approx$160 bp and $\approx$90 bp, respectively. The longer the DNA, the more MBD-DNA interaction pairs can be formed, thus increasing $K_\text{av}$. It will be crucial to understand what the difference in the MBD surface receptor density will be for surface binding of, for example, 100 bp-long hmDNA compared to 200 bp-long non-methylated DNA. Furthermore, it should be investigated whether hmDNA can be enriched efficiently from a DNA sample. Likely this is especially important to enable hmDNA-based, early-stage cancer diagnostics where the hmDNA concentrations are typically very low (few copies per milliliters of sample). Overall, this study indicates that the enrichment of hmDNA can be improved significantly by optimizing the MBD surface receptor densities according to the principles of multivalency and superselectivity. The development of the superselective hmDNA enrichment system will be of interest for the clinics after the integration of the proposed method in a lab-on-a-chip device, ensuring fast and simplistic usage.

Acknowledgements

The authors thank the Weijerhorst Foundation for funding and Nienke van Dongen for help with the schematic illustrations.

Conflict of Interest

The authors declare no conflict of interest.
Ruben W. Kolkman obtained his Bachelor’s degree in Chemistry at the Saxion University of Applied Sciences and Master’s degree in Chemical Engineering at the University of Twente. Since August 2018 he is a Ph.D. candidate in two research groups at the University of Twente under supervision of Prof. Jurriaan Huskens and Prof. Loes Segerink. His research interests are multivalency, supramolecular chemistry, surface chemistry, microfluidics, and cancer detection.

Loes I. Segerink (1984) obtained her Ph.D. in 2011 at the University of Twente and her impact of her work has been recognized by the Simon Stevin Leerling award (2011), Simon Stevin Gezel award (2012). She started as Postdoc in the same group (Albert van den Berg) and worked also for several months in the group of Helene Andersson Svahn (KTH, Sweden). In 2013 she received her Veni grant. In 2014 she started her tenure track position (University of Twente), and became Full Professor in July 2021. Her work focuses on biomedical microdevices for diagnostic and organ-on-chip applications.

Jurriaan Huskens (1968) obtained his Ph.D. (1994) at the Delft University of Technology with Herman van Bekkum. After postdoctoral stays with Dean Sherry and Manfred Reetz, he became Assistant Professor with David Reinhoudt at the University of Twente in 1998, and Full Professor of the Molecular NanoFabrication group in 2005. He received the Unilever Research Award (1990), a Marie Curie fellowship (1997), the Gold Medal 2007 of the Royal Netherlands Chemical Society, and a Fellowship from the Institute of Advanced Study, Durham University, UK (2019). Present research interests encompass: supramolecular chemistry at interfaces, supramolecular materials, multivalency, (bio)sensing, and sustainable chemistry.