Towards Rational Design of Selective Molecularly Imprinted Polymers (MIPs) for Proteins: Computational and Experimental Studies of Acrylamide-Based Polymers for Myoglobin

Mark V. Sullivan,[a] Sarah R. Dennison,[b] Georgios Archontis,[c] Subrayal M. Reddy,*[a] Joseph M. Hayes*[b]

[a] *School of Physical Sciences & Computing, Division of Chemistry, University of Central Lancashire, Preston PR1 2HE, United Kingdom*

[b] *School of Pharmacy & Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, United Kingdom*

[c] *Department of Physics, University of Cyprus, CY1678 Nicosia, Cyprus*

[*] Corresponding authors
tel: +441772893291 (SMR); fax: +441772894981; email: smreddy@uclan.ac.uk
tel: +441772894334 (JMH); fax: +441772894981; email: jhayes@uclan.ac.uk

Abstract: Molecularly imprinted polymers (MIPs) have potential as alternatives to antibodies in the diagnosis and treatment of disease. However, atomistic level knowledge of the prepolymerization process is limited that would facilitate rational design of more efficient MIPs. Accordingly, we have investigated using computation and experiment the protein-monomer binding interactions that may influence the desired specificity. Myoglobin was used as the target protein and five different acrylamide-based monomers considered. Protein binding sites were predicted using SiteMap and binding free energies of monomers at each site calculated using MM-GBSA. Statistical thermodynamic analysis and study of atomistic interactions facilitated rationalization of monomer performance in MIP rebinding studies (% rebind; imprinting factors). CD spectroscopy was used to determine monomer effects on myoglobin secondary structure, with all monomers except the smallest monomer (acrylamide) causing significant changes. A complex interplay between different protein-monomer binding effects and MIP efficacy was observed. Validation of hypotheses for key binding features was achieved by rational selection of two different co-monomer MIP combinations that produced experimental results in agreement with predictions. The co-monomer studies revealed that uniform, non-competitive binding of monomers around a target protein is favourable. This study represents a step towards future rational *in silico* design of MIPs for proteins.

1. Introduction

Molecular imprinting is fast becoming an established technique to produce polymers with molecular recognition properties.[1-4] These molecularly imprinted polymers (MIPs) are inspired by nature and are capable of binding various analytes with an affinity and selectivity similar to those of their biological equivalents. The advantages of MIPs over biological receptors includes durability, specificity and ease of mass production.[5-7] They have broad applications in areas such as chromatography, solid phase extractions, catalysis and (bio)sensors.[8-11] While protein MIPs have considerable potential to replace antibodies in the diagnosis and treatment of disease, they still have limitations. MIPs currently lack the selectivity required for practical applications partly due to a poor understanding of how to optimise protein-MIP interactions at cavity sites. The MIP cavities are borne out of the non-covalent interactions within the pre-polymerization mixture, leading to binding site heterogeneity especially within bulk MIP.[9] This means that a better understanding of the protein-monomer interactions that are favourable for producing selective MIPs is needed to progress them further as a viable alternative to their biological counterparts. There has been a particular increase in macromolecular imprinting over the past decade, mainly due to demand in the areas of biosensors, clinical diagnostics and therapeutics.[12, 13] Furthermore, towards an alternative to the expensive and degradable antibodies currently being used in analysis,[14-17] there is particular interest in selectively imprinting macromolecules such as proteins, viruses and cells, for their subsequent determination within a complex medium. Employing hydrogels for the imprinting of proteins has shown considerable promise towards this aim.[18] The hydrogels consist of significantly dilute cross-linked systems that form an insoluble network of hydrophilic cross-linked polymer chains.

MIP formation is achieved when a mixture consisting of functional monomers and cross-linker polymerises around a target molecule.[7, 19, 20] The target molecule in effect acts as a template for MIP formation. After polymerisation, the template molecule is removed, to leave binding sites on the newly formed MIP capable of selectively recognising the target molecule.[21, 22] Molecular imprinting is a consequence of monomer-template interactions, with non-covalent interactions such as hydrogen bonding, van der Waals and ionic interactions, together with hydrophobic effects used in the self-assembly process of the prepolymerization complex.[23-25] In order to achieve a successful imprinting procedure, there needs to be a strong monomer-template complex formed during prepolymerization, which is preserved through to the final polymerisation step, meaning the choice of monomer is of paramount importance.[4, 26] A more stable monomer-template complex with strong intermolecular interactions can potentially deliver a MIP with high selectivity and affinity.[22] In this regard, *in silico* predictive methods can provide a rationale for improved monomer-template complexation.

To date, however, computational methods have been mainly applied to the molecular imprinting of low molecular weight target (analyte) molecules. Molecular modelling software and searching algorithms, traditionally applied in drug design, have been adapted to find monomers with the ability to form strong complexes with small molecule templates.[27] Virtual libraries of polymerisable, functional monomers have been created and used to successfully predict *in silico* selective MIPs for small molecule analytes such as ephedrine.[27, 28] Computational methods allow monomers to be ranked in terms of predicted binding affinity. For example, the highest ranked monomers from LUDI docking[29] were used to produce a rationally designed selective MIP for microcystin-LR, a toxic analyte in water.[30] Molecular dynamics (MD) have applications in modelling multi-component systems and assessing a large number of interactions,[27] and have been used to simulate and evaluate MIP prepolymerization mixtures.[31] These studies allowed a closer analysis of how the functional monomers, cross-

linker and solvent interacted with the template and the preferred arrangement of interactions in the prepolymerization mixtures.[31] Quantum mechanics (QM) methods are also becoming more common place in MIP design, although limited by the size of the studied systems and associated computational expense.[27] QM in the form of density functional theory (DFT) calculations was successfully applied for the selection of monomers to design MIPs for 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), a well-known toxin.[32]

There have been few computational studies using proteins as templates for MIPs.[33] Due to the size of proteins, a high molar ratio of ligand to protein is required, with ratios of 2000:1 commonly seen. Molecular docking can predict where the ligands prefer to bind.[34] Docking studies of common monomers and cross-linkers to target proteins have highlighted complementary shape and demonstrated that there is considerable amount of hydrogen bonding, electrostatic and hydrophobic interactions between the target and monomers for specific recognition.[34, 35] However, in addition to favourable interactions, there are also monomer interactions with the target protein backbone that have the potential to cause secondary structure changes affecting MIP selectivity.[34] Boroznjak et al. used a docking approach to rationally select a functional monomer from m-phenylenediamine (mPD), dopamine and 3,4-ethylenedioxythiophene for building a MIP suitable for immunoglobulin G (IgG).[33] The three monomers were docked into the fragment antigen-binding (FAB), fragment crystallisable (Fc) and hinge regions of IgG. All three monomers studied revealed similar docking scores but mPD molecules were found to be more uniformly arranged around IgG, leading to a MIP with more complementary cavities, and thus, better experimental performance.

In this work, supported by experimental data, we have investigated at the atomic level using computation the features that govern the binding and relative efficacies of different monomers

for use in MIPs demonstrating selective target protein binding. The model target protein for these studies was myoglobin, while five acrylamide-based monomers (Figure 1) were chosen that would allow the analysis of different structural effects. Myoglobin is a protein released during damage to muscle tissue. Clinically, it can be a sensitive marker for muscle injury and has potential to be used as a cardiac marker for myocardial infarction. Importantly for the purposes of this study, because of the protein's small size (153 residues, ~17 kDa) and with fewer monomer binding sites, interpretation and analysis of the modelling results is better facilitated. It should be noted that the more readily available equine isoform of the protein was used for analysis, which has high homology with human form (88% identity),[36] so that results in terms of actual medical applications would need to be confirmed. The potential binding sites of the protein were predicted using the SiteMap program,[37-39] followed by prediction of the binding site-specific interactions of each monomer studied using Glide docking and post-docking molecular mechanics with generalized Born and surface area continuum solvation (MM-GBSA) binding free energy (BFE) calculations.[39] Docking has previously been applied for studies of this type but mainly for small molecule analytes (as highlighted above), while post-docking MM-GBSA should theoretically improve the accuracy of predictions. A comprehensive statistical thermodynamic analysis and study of protein-monomer interactions allowed us to apply a rational approach to predict co-monomer combinations that produce hydrogel MIPs with different efficacies compared to those MIPs of the individual monomers. Additionally, the effects of monomers on protein secondary structure features were analysed by far-UV circular dichroism (CD) spectroscopy, with the SELCON3 method, as this can provide an accurate estimation of protein secondary structure.[40, 41]
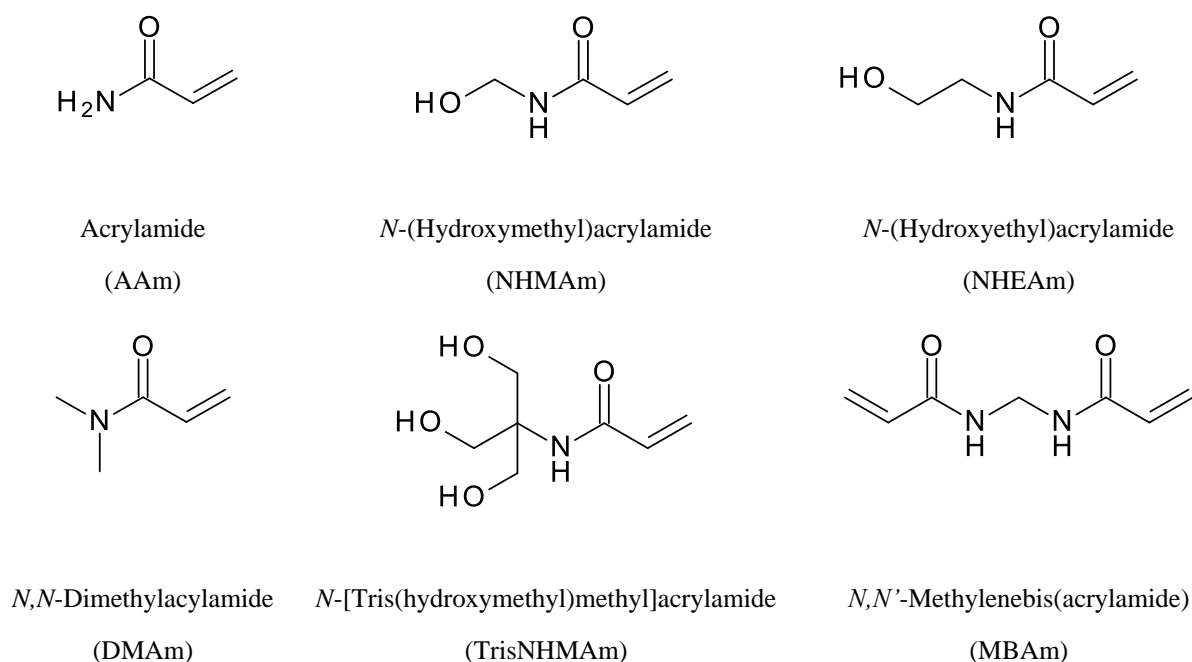
Figure 1: The five monomers (AAm, NHMAm, NHEAm, DMAm, TrisNHMAm) and cross-linker (MBAm) used in the polymerisation of the hydrogel MIPs.

2. Experimental Section

2.1 Computational details

2.1.1 Protein and Monomer Preparation

The initial setup of the myoglobin protein (PDB code: 5D5R) containing an iron centered heme group for calculations was performed using Schrödinger's "Protein Preparation Wizard".[39] Water molecules were deleted, bond orders were assigned and hydrogen atoms added, with protonation states for basic and acidic residues based on residue pK_a values at normal pH (7.0) calculated using PROPKA.[42] Subsequent optimization of hydroxyl groups, histidine protonation states and C/N atom flips, and side-chain O/N atom flips of Asn and Gln was based on optimizing hydrogen bonding patterns. Finally, an "Impref" minimization of the myoglobin was performed using the OPLS3 force field[43] to remove steric clashes and bad contacts, but with heavy atoms constrained to within 0.3 Å (RMSD) of their crystallographic positions. The

monomers shown in Figure 1 were prepared for calculations using LigPrep 3.5[39] with the

optimized potential for liquid simulations 3 (OPLS3) force field[43] and default settings.

2.1.2 Protein Binding Site Analysis

Potential monomer binding sites on myoglobin were predicted using SiteMap 3.9.[37, 39] The

settings included at least 5 site points per site, a 'less restrictive' definition of hydrophobicity,

along with a standard grid. Shallow binding sites were also detected. The sites generated were

cropped at a distance of 3 Å from the nearest site point, with a total of 14 predicted binding

sites identified. The surface of the protein was analysed using BioLuminate 16.2 with protein

surface analyzer.[39]

2.1.3 Docking Calculations

For each predicted myoglobin binding site, docking calculations on the five monomers were

performed using Glide 7.1 in standard-precision (SP) mode[39, 44, 45] with default OPLS3 atomic

charges and van der Waals scaling (0.8) for ligand nonpolar atoms to include modest 'induced

fit' effects. Docking grids were generated for each of the 14 different potential binding sites

from SiteMap, with the grid centre coordinates included in Table S1. Post-docking

minimization of poses with strain correction was employed. Up to 20 poses per input monomer

structure were saved for each docking run, in order to generate a large number of diverse poses.

2.1.4 Post-Docking MM-GBSA Calculations

Using the Glide-SP docking poses for each monomer at each site, MM-GBSA binding free

energies were calculated using Prime 4.4[39] and the following equation:

$$\Delta G_{PL} = \Delta E_{PL}^{MM} + \Delta G_{PL}^{solv} - T\Delta S_{PL}^{MM} \qquad (1)$$

where ΔE_{PL}^{MM} represents the molecular mechanics (MM) energy difference (internal,

electrostatic and van der Waals) between the protein (P) - monomer (L) bound and unbound

states, calculated using the OPLS3 forcefield.[43] DG_{PL}^{solv} is the corresponding solvation free energy change on binding calculated using a variable-dielectric generalized Born solvation model. While the protein was fixed for these calculations, ligand strain (reorganization) effects on binding were included. Estimates for solute entropy changes on binding ΔS_{PL}^{MM} were also included in terms of the change in vibrational, rotational and translational (VRT) entropy of ligands on binding. The latter was calculated using the Rigid Rotor Harmonic Oscillator (RRHO) approximation (default settings) with MacroModel v11.3[39] and the OPLS3 forcefield.[43]

2.1.5 Statistical Thermodynamics Analysis of Monomer Binding

In order to analyze the probabilities for monomers to bind at each of the predicted binding sites in the prepolymerization state, we performed statistical thermodynamic analysis of the DG_{PL} data. A full derivation of the employed formulas is detailed in the Supporting Information (SI). We consider a protein P with M independent, non-equivalent binding sites, indexed by the integers $i = \{1,...,M\}$. We assume that each protein site can be empty or bind one monomer. Binding of monomer L at site i is associated with a solution free energy change $DG_{PL}(i)$.

Single Monomer Studies:

We first consider the case of a solution with dissolved protein and *only one* monomer species L. The concentration of *free* monomer in solution, $x \circ [L]/C_0$, measured in units of a standard-state concentration C_0, is computed from the self-consistent solution of the following equation (Eq. (3) of SI material):

$$\frac{[L]}{C_0} = x = \frac{[L]_{tot}}{C_0} - \frac{[P]_{tot}}{C_0} \sum_{i=1}^{M} \frac{xe^{-bDG_{PL}(i)}}{1 + xe^{-bDG_{PL}(i)}} \tag{2}$$

where β equates to *1/kT*. $[L]_{tot}$ is the total ligand concentration (0.76 M) and

$[P]_{tot} = [L]_{tot}/1081 = 7.03 \times 10^{-4}$ M is the total protein concentration. A similar equation for

equivalent sites is derived in reference [46].

The average number of monomer L molecules bound to a protein molecule as a function of

the free ligand concentration x is [SI Eq. (2)]:

$$\overline{M}_L(x) = \sum_{i=1}^{M} \frac{x e^{-\beta \Delta G_{PL}(i)}}{1 + x e^{-\beta \Delta G_{PL}(i)}} \tag{3}$$

Co-monomer Studies.

We now consider co-monomer solutions, in which two monomers L_1 and L_2 are simultaneously

present in the solution. The free-monomer concentrations are evaluated by solving self-

consistently the following system of equations [SI Eq. (6)]:

$$
\begin{aligned}
x_1 &= \frac{[L_1]_{tot}}{C_0} - \frac{[P]_{tot}}{C_0} \sum_{i=1}^{M} \frac{x_1 e^{-\beta \Delta G_{PL_1}(i)}}{1 + x_1 e^{-\beta \Delta G_{PL_1}(i)} + x_2 e^{-\beta \Delta G_{PL_2}(i)}} \\
x_2 &= \frac{[L_2]_{tot}}{C_0} - \frac{[P]_{tot}}{C_0} \sum_{i=1}^{M} \frac{x_2 e^{-\beta \Delta G_{PL_2}(i)}}{1 + x_1 e^{-\beta \Delta G_{PL_1}(i)} + x_2 e^{-\beta \Delta G_{PL_2}(i)}}
\end{aligned}
\tag{4}
$$

The average numbers of monomers, L_1 or L_2, bound to a protein molecule are computed by the

relations [SI Eqs. (5)]:

$$\overline{M}_{L_1}(x_1) = \sum_{i=1}^{M} \frac{x_1 e^{-\beta \Delta G_{PL_1}(i)}}{1 + x_1 e^{-\beta \Delta G_{PL_1}(i)} + x_2 e^{-\beta \Delta G_{PL_2}(i)}} \tag{5a}$$

and

$$\overline{M}_{L_2}(x_2) = \sum_{i=1}^{M} \frac{x_2 e^{-\beta \Delta G_{PL_2}(i)}}{1 + x_1 e^{-\beta \Delta G_{PL_1}(i)} + x_2 e^{-\beta \Delta G_{PL_2}(i)}} \tag{5b}$$

Finally, the relative probabilities of monomers L_1 and L_2 to bind at the same protein site k (in the co-solvent solution) are calculated as [SI Eq. (14))]:

$$\frac{p(\text{site i is occupied by L}_1)}{p(\text{site i is occupied by L}_2)} = \frac{x_1 e^{-bDG_{PL_1}(i)}}{x_2 e^{-bDG_{PL_2}(i)}} @ \frac{e^{-bDG_{PL_1}(i)}}{e^{-bDG_{PL_2}(i)}} \tag{6}$$

The last equality is valid when the monomers have the same free concentration ($x_1 @ x_2$).

2.2 MIP Synthesis

2.2.1 Materials.

Acrylamide (AAm), ammonium persulphate (APS), α-lactalbumin, glacial acetic acid (AcOH), myoglobin (Mb) (from equine skeletal muscle) (Mb), *N*-(Hydroxymethyl)acrylamide (NHMAm), *N*-(Hydroxyethyl)acrylamide (NHEAm), *N,N*-Dimethylacylamide (DMAm), *N,N'*-methylenebisacrylamide (mBA), *N*-[Tris(hydroxymethyl)methyl]acrylamide (TrisNHMAm), sodium dodecyl sulphate (SDS), and tetramethylethyldiamide (TEMED), were all purchased and used without purification from Sigma-Aldrich, Poole, Dorset, UK.

2.2.2 Methods.

Solution preparation. A solution of 10% (w/v):10% (v/v) SDS:AcOH was prepared for use in the washing (protein elution) stages before the template reloading stage. SDS (10 g) and AcOH (10 mL) was dissolved in 990 mL of deionised (DI) water, to produce 1 L of the washing solution.

MIP preparation. Bulk MIP hydrogels were produced, using an optimised methodology,[18] where a 10 % cross-linking monomer/N,N'-methylenebisacrylamide hydrogel was found to produce the optimal imprint for Mb, in terms of specificity and rebinding efficiency of the MIP, compared with the non-imprinted polymer (NIP).[18]

The MIPs were produced with different monomers (AAm, NHMAm, NHEAm, DMAm, TrisNHMAm) and a 10% cross-linking density for the myoglobin as a template using the

following method. Into an Eppendorf tube, 12 mg of myoglobin template was dissolved in 970 µL of deionised water vortexed for 1 minute, followed by the addition of 7.6 x10^{-4} mol of functional monomer and mBA (cross-linker) at a ratio of 9:1 by weight, then vortexed for a further minute. Finally, 10 µL of a 5% TEMED (v/v) solution and 20 µL 5% APS (w/v) solution were added and the mixture was vortexed for 1 minute. Solutions were purged with nitrogen for 5 minutes and polymerisation occurred overnight at room temperature (~20 °C). Corresponding NIPs were produced using the same method, but in the absence of a protein template.

After polymerisation, the gels were granulated separately using a 75 µm sieve. The refined gels were washed with five 1 mL volumes of deionised water followed by five 1 mL volumes of 10% (w/v):10% (v/v) SDS:AcOH eluent; this allowed for the removal of the template protein from the MIP cavities. Following this, the gels were washed with five 1 mL volumes of deionised water to remove all residual 10% (w/v):10% (v/v) SDS:AcOH from the MIP gels. Each wash step was followed by centrifugation, whereby the gels were vortexed then centrifuged (using SLS Lab basics centrifuge) for 5 minutes at 15000 rpm (RCF: 15100 × g). Corresponding non-imprinted polymers (NIPs) were synthesized using the same procedure as the MIPs, but in the absence of the template molecule.

2.2.3 MIP Rebinding Studies

The subsequent rebinding effect of the conditioned and equilibrated MIPs and NIPs were characterized using the BioDrop µLITE UV/visible spectrometer. Hydrogels (200 mg) were then each treated with 400 µL of a 3 mg mL^{-1} myoglobin (template protein) solution. The polymer/protein solutions were mixed on a rotary vortex mixer and allowed to associate at room temperature (~20 °C) washed four times with 1 mL of deionized water. Each reload and wash step for hydrogels was followed by centrifugation for 5 minutes at 15000 rpm (RCF:

15100 × g). All supernatants were collected for analysis by spectrophotometry (at 405 nm for myoglobin).

2.3 CD Spectroscopy

The secondary structure of myoglobin was determined using a J-815 Jasco Circular Dichrosim spectropolarimeter at 20 °C. Essentially, protein samples (0.00075 mg/mL) were spiked with monomer solution to produce a molar ratio of 1:1081, then placed into a quartz cell (Starna Scientific UK) with a 10 mm path-length. Using CD spectrometer far-UV CD spectra were collected for the protein samples, where the instrument was set to standard, ten scans per sample were performed over a wavelength range of 260 to 180 nm at a band width interval of 1 nm, data pitch 0.5 nm, scanning mode set to continuous and a scan speed set at 100 nm min⁻¹. All spectra were baseline corrected and the percentage secondary structure was determined using the Selcon3 method (protein reference set 3) from the DichroWeb server.[47-49] These experiments were repeated 3 times and the secondary structure percentages were averaged.

3. Results & Discussion

3.1 Single Monomer Computational and Experimental Studies

3.1.1. Single Monomer MIPS Experimental Results

Each of the five monomers in Figure 1 (AAm, NHMAm, NHEAm, DMAm, TrisNHMAm) were used to prepare five different MIPs and their corresponding NIPs, where the NIPs were control polymers for MIP selectivity. The NIPs were produced in an identical manner to the MIPs, but in the absence of the protein template, and thus contained the same chemical properties as a MIP, but without the specific cavities. Rebinding studies were performed on the

MIPs and NIPs, and the percentage of target protein rebind and subsequent imprinting factors (IFs) were calculated using the following equation:

$$IF = \frac{\%\ protein\ rebind\ to\ MIP}{\%\ protein\ bind\ to\ NIP} \qquad (7)$$

Thus, IF is calculated as a ratio of protein rebind to the MIP versus protein that binds to the corresponding NIP. IF is commonly used to evaluate the imprinting effect and as a measure of the strength of interaction between the functional monomer and the template/target molecule. The higher the IF, the more selective the MIP is for the target molecule and an IF > 1.20 is generally considered favorable.[9] Whereas it is accepted that a selectivity factor (ratio of MIP binding to target versus a non-target protein) is a better determinant of selectivity, use of this factor is beyond the scope of this present study.

The single monomer experimental results (Table 1) revealed that NHMAm was the best monomer used to produce a MIP for myoglobin recognition, with an excellent 98.9% rebind and an imprinting factor of 1.90. The worst performing MIP was DMAm, with 72.0% rebind and an imprinting factor 1.48. With respect to overall MIP efficacy, the monomers could be ordered by performance NHMAm > AAm > NHEAm > TrisNHAm > DMAm. The monomer NHEAm was judged to be better than TrisNHAm on the basis of IF (1.77 versus 1.10, respectively), despite having a slightly lower % rebind (77.2% versus 79.9%, respectively).

Table 1: Percentage of the myoglobin target protein rebind to the five different acrylamide based MIPs and NIPs, and their corresponding impact factors. Results for co-monomer combinations A (TrisNHAm + DMAm) and B (NHEAm + DMAm) are also shown.

| Monomer/ Co-monomer	MIP Percentage of Protein Rebind (%)	NIP Percentage of Protein Bind (%)	Imprinting Factor (IF)
NHMAm	98.9 ± 0.2	51.8 ± 0.4	1.90
AAm	85.4 ± 1.0	47.5 ± 4.2	1.80
NHEAm	77.2 ± 3.0	43.6 ± 1.3	1.77
TrisNHAm	79.9 ± 4.8	72.3 ± 1.7	1.10
DMAm	72.0 ± 3.0	48.8 ± 0.9	1.48

TrisNHAm + DMAm (A)	85.5 ± 1.6	63.4 ± 1.1	1.30
NHEAm + DMAm (B)	63.2 ± 0.6	50.6 ± 2.2	1.30

The self-assembly method for MIP formation relies upon a degree of association between the monomer and the template, mainly based on hydrogen-bond interactions.[9] Monomers containing -OH functional groups (NHMAm, NHEAm and TrisNHAm), therefore, have the potential to bind stronger than the monomers AAm and DMAm. In accordance with this, TrisNHAm and NHMam NIPs have the highest protein binding values values. However, if one ranks the monomers by the MIP percentage protein rebind, then we have NHMAm > AAm > TrisNHAm > NHEAm > DMAm. Thus, in terms of MIP, the AAm monomer without any –OH groups is the second best performing monomer, superior to both NHEAm and TrisNHAm. In addition to this, TrisNHAm with three -OH functional groups does not perform particularly well with a percentage rebind of 79.9%, only performing slightly better than DMAm (72.0%). TrisNHAm also has an unsatisfactory IF value of 1.10, making it unsuitable for use as a myoglobin targeted MIP. This indicated that there are other factors, and not just the potential strength of protein-template interactions, to consider when choosing a suitable monomer for a target protein.

3.1.1 SiteMap Prediction of Myoglobin Binding Sites

Following preparation of the crystal structure of myoglobin for calculations, the potential binding sites were predicted using SiteMap, producing a map of 14 different binding sites. Visual inspection revealed that the 14 binding sites (**1-14**) were quite evenly distributed around the protein surface, as shown in Figure 2. This can be interpreted as favourable for MIP target protein binding, with monomers potentially more uniformly distributed around the template protein, compared to MIPs where the monomers are isolated to a specific area.[33]
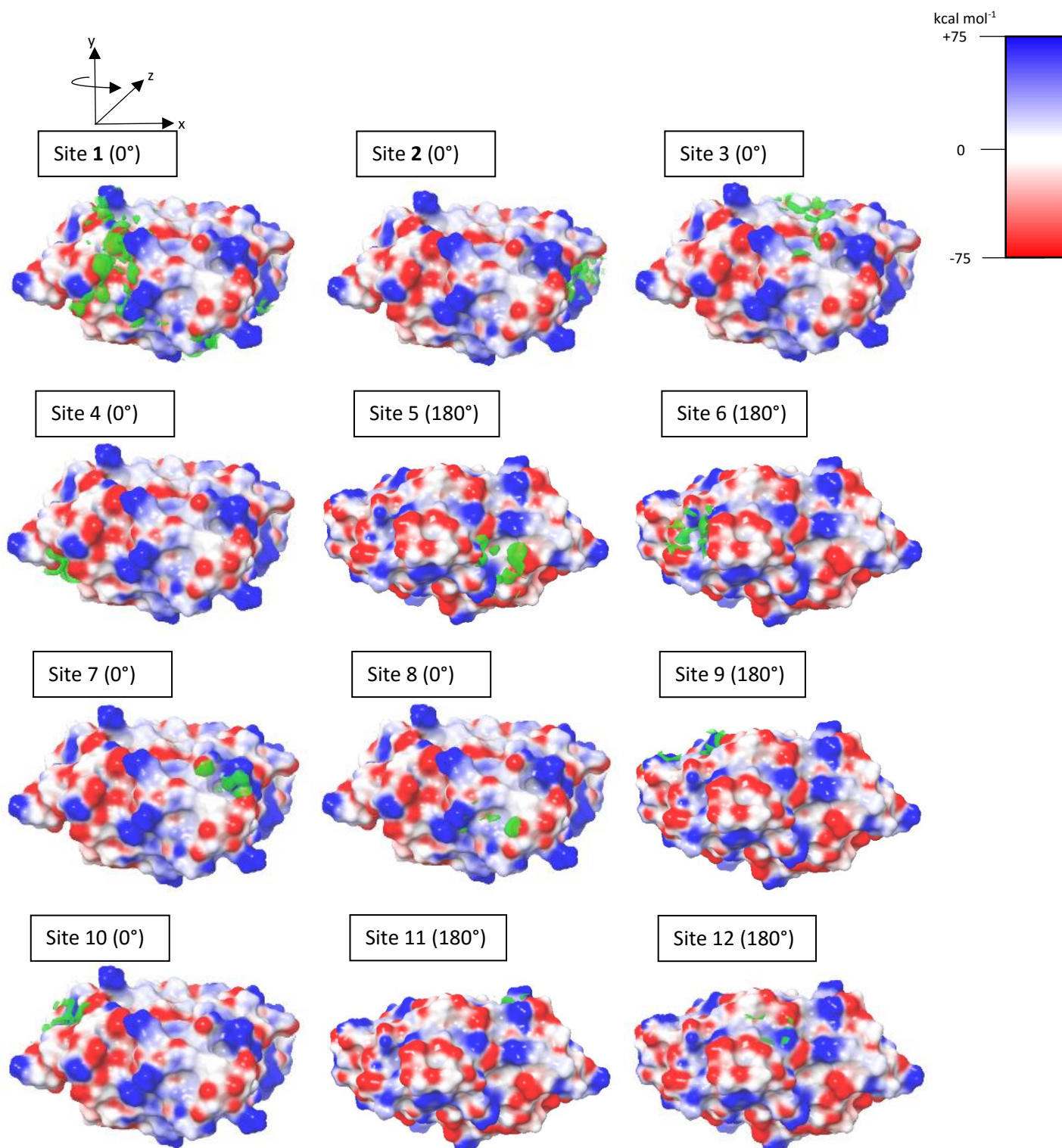
Figure 2: The potential monomer binding sites **1-12** (green) on myoglobin as predicted by SiteMap, together with the electrostatic potential (blue (positive) and red (negative)) mapped on to the protein molecular surface calculated using Maestro. Sites **13** and **14** are not displayed as they were buried and not on the protein surface; it transpires they were predicted as not favorable for monomer binding (Table 2).

3.1.2 Docking Results for Monomer Binding

Glide-SP docking of all five monomers into each of the 14 binding sites was performed, with the best GlideScores of monomers at each site presented in Table S2. The only monomer, which successfully docked to all 14 sites, is the smaller AAm monomer. NHMAm docks into 13 binding sites, TrisNHAm docks into 12 binding sites and the last two monomers (NHEAm and DMAm) dock into 11 binding sites. Hence, the best performing monomers experimentally as MIPs (NHMAm and AAm) docked into the most binding sites. TrisNHAm, however, had a number of unfavorable GlideScores at different sites. The docking calculations revealed key hydrogen bonding interactions between the monomers and the protein molecule. This involved either monomer carbonyl, amine or hydroxyl groups with similar functional groups on protein residues. Importantly, however, these hydrogen bond interactions were with both protein backbone and sidechain atoms, the degree of which with protein backbones can be significant [34]. Prior to detailed analysis, more accurate binding predictions were performed using the docking poses and the theoretically more rigorous post-docking MM-GBSA method,[50] whose connection with statistical thermodynamics has been established.[51]

3.1.3 MM-GBSA Results for Monomer Binding

For each site, post-docking MM-GBSA binding free energies (Eq. (1)) were calculated using the Glide-SP docking poses. The best values for each monomer and site are shown in Table 2. Breakdown of the contributions as per Eq. (1) are shown in Table S3. As explained in the methods, these values incorporate contributions due to protein-monomer interactions in the binding conformation (the MM term), solvent effects (the GBSA term), monomer reorganization and changes in monomer entropy upon binding, evaluated by the RRHO approximation.

Table 2: The ΔG_{PL} values (kcal mol^{-1}) of the five monomers (AAm, NHMAm, NHEAm, DMAm, TrisNHMAm) calculated using MM-GBSA at each of the 14 predicted binding sites of myoglobin. For co-monomer mixtures MIP A (TrisNHAm & DMAm) and MIP B (NHEAm & DMAm), the predicted ΔG_{PL} values given are for the monomer most probable to bind at the site, together with its probability value for binding (in parentheses) as calculated using Eq. (6).[a]

	Binding Site													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Monomer														
AAm	**2.5**	-9.0	-3.8	-2.2	-7.3	-2.8	**1.6**	-3.3	-1.2	-3.6	**0.1**	-1.7	**7.3**	**9.2**
NHMAm	-3.4	-10.8	-5.1	**16.2**	-9.8	-2.9	**0.7**	-5.3	-4.8	-9.9	-0.7	**0.5**	**36.8**	*n/a*
NHEAm	*n/a*	-11.3	-7.0	-5.9	-8.6	-3.2	-0.5	-6.6	-4.4	-12.9	**6.9**	-3.2	*n/a*	*n/a*
DMAm	*n/a*	-8.8	-9.1	-6.5	-7.3	-3.8	**0.7**	-5.4	-4.3	-10.7	-3.3	-2.3	*n/a* (-)	*n/a* (-)
TrisNHAm	-10.3	-5.2	-10.7	-10.7	-12.1	-7.4	-0.5	-11.3	-1.3	-24.9	-5.2	-5.2	*n/a*	*n/a*
Co-Monomer														
MIP A	-10.3 (1.000)	-8.8 (0.998)	-10.7 (0.953)	-10.7 (0.999)	-12.1 (1.000)	-7.4 (0.998)	-0.5 (1.00)	-11.3 (1.00)	-4.3 (0.994)	-24.9 (1.00)	-5.2 (0.960)	-5.2 (0.992)	*n/a* (-)	*n/a* (-)
MIP B	*n/a* (-)	-11.3 (0.985)	-9.1 (0.971)	-6.5 (0.732)	-8.6 (0.898)	-3.8 (0.732)	-0.5 (1.00)	-6.6 (0.542)	-4.4 (0.500)	-12.9 (0.979)	-3.3 (0.858)	-3.2 (0.793)	*n/a* (-)	*n/a* (-)

[a] Values that are positive (unfavourable binding) are highlighted in bold. *n/a* indicates that there were no predicted binding poses.

Binding of the various monomers at the different sites is associated with a range of free-energy values $\mathrm{D}G_{PL}$. Sites **2**, **3**, **5**, **8**, and **10** are associated with lower (more negative) values, ranging from ~ -4 to -12 kcal/mol (exception TrisNHMAm, *vide infra*), and generally appear to be key binding sites. Analysis of the $\mathrm{D}G_{PL}$ data from Table 2 revealed greater magnitude of values for monomers with more hydroxyl groups. TrisNHMAm (three hydroxyls) has the most negative $\mathrm{D}G_{PL}$ values, and demonstrates an especially strong preference for protein binding site **10** ($\mathrm{D}G_{PL}$ = -24.9 kcal/mol). NHEAm and NHMAm, each with one hydroxyl, consistently have better (more negative) $\mathrm{D}G_{PL}$ values compared to AAm (no hydroxyls); DMAm, also lacking hydroxyls, has slightly better $\mathrm{D}G_{PL}$ values than NHMam and slightly worse than NHEAm.

Figure 3 displays the average number of bound monomers (prior to polymerization) as a function of the free monomer concentration $x = [L] / C_0$, computed with Eq. (3). The largest value of the x-axis corresponds to the experimental total monomer concentration (0.76 M). Monomers with lower (more negative) binding free energies $\mathrm{D}G_{PL}$ are expected to bind with larger average numbers across the different binding sites. TrisNHMAm has the most monomers bound on the protein surface and AAm the least at all concentrations. The relative ranking depends somewhat on the monomer concentration, reflecting the fact that each monomer has a different distribution of binding free energies. For example, the monomer ranking at very small monomer concentrations (TrisNHMAm > NHEAm > NHMam ~ DMAm > AAm) is determined by the number of binding sites with $\mathrm{D}G_{PL} \leq$ -10 kcal/mol, (six, two, one, one and zero, respectively, as shown in Table 2). DMAm becomes a better binder than NHMam and NHEAm at higher concentrations, because it has a better distribution of binding free energies with less negative values (-10 < ΔG < 0 kcal/mol).
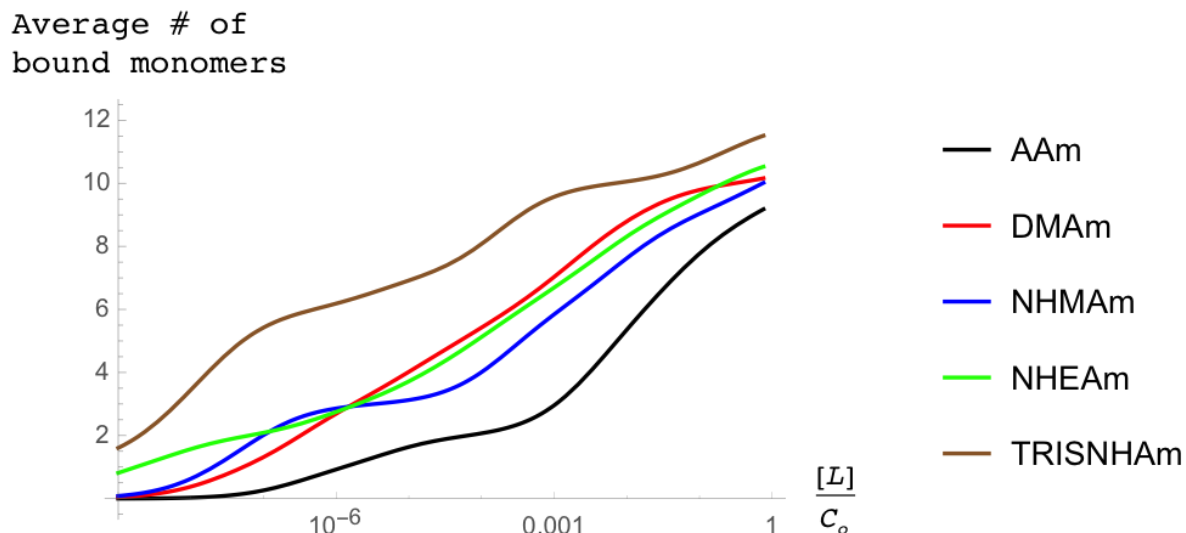
Figure 3: Average number of bound monomers as a function of the free monomer concentration $x = [L]/C_0$, computed with Eq. (3). The x-axis is shown in a logarithmic scale. The largest concentration corresponds to the total monomer concentration (0.76 M) at which point the average numbers are much closer.

Using Eq. (4) and the binding free energies reported in Table 2, the expected *free* monomer concentrations and the corresponding average number of bound monomers (prior to polymerization) were computed for each solution (Table 3). Due to the large monomer:protein ratio (1081:1), the free-monomer concentrations are similar to the total monomer concentration (0.76 M). Thus, the average number of bound monomers correspond to the values at the highest-concentration end of the curves in Figure 3. The relative ordering of our computed $\bar{M}_L(x)$ values, AAm < NHMAm < DMAm < NHEAm < TrisNHMAm, roughly reproduces the NIP efficiency of the various monomers (Table 1): NHEAm < AAm < DMAm < NHMAm < TrisNHMAm. This is reasonable, as the ability of the protein to bind prepolymerized individual monomers should be correlated with the non-imprinted cavity free polymer NIP.[46]

Table 3: Free monomer concentration and average number of monomers bound to the protein for each solution.

Monomer/Co-monomer Solution	Concentration of free monomer[a]	Average number of bound monomers[b]
	x (mol/L)	$\bar{M}_L(x)$
AAm	0.7536	9.17
NHMAm	0.7530	10.0
NHEMAm	0.7526	10.5
DMAm	0.7529	10.2
TrisNHAm	0.7519	11.5
DMAm/TrisNHAm	0.3785/0.3734[c]	2.14/9.34[c]
DMAm/NHEMAm	0.3766/0.3761[c]	4.85/5.62[c]

[a] Computed from Eq. (4). [b] Computed from Eq. (3), using the free monomer concentration of column 2. [c] The numbers correspond to the first and second listed monomer, respectively.

Even though the free and total monomer concentrations are approximately equal, the computed estimates of Table 3 imply that the large majority of protein binding sites are occupied by monomers. For example, TrisNHAm binds an average of 11.5 monomers in 12 sites with negative free energies; AA seems to have the worst performance, with 9.17 monomers in 14 sites (average binding numbers depend strongly on the $\mathrm{D}G_{PL}$ values). The reported $\mathrm{D}G_{PL}$ values are based on several approximations. Eqs. (3-4) assume that the various sites are independent from each other. It is likely that the MM-GBSA approximation yields too negative binding free energies, as also observed elsewhere.[52, 53] Furthermore, below we show through CD spectroscopy that with the exception of AA, binding of all other monomers modifies to a similar extent the protein secondary structure. In the MM-GBSA analysis, we neglect energy

contributions due to protein re-organization; their inclusion increases the error in the binding free energies particularly when different binding sites are considered. It has been shown that relative binding free energies for congeneric series of ligands binding at proteins can be better reproduced when protein reorganization is neglected (presumably cancellation of errors).[54] Due to these approximations, it is likely that the obtained average numbers are only indicative but reflect the important relative trends among the different monomers.
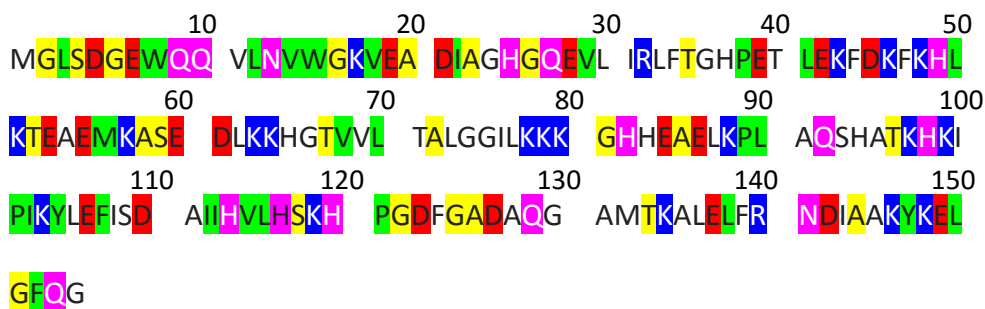
3.1.4 Protein Surface Analysis

Whilst monomers with more OH groups produce more favorable negative ΔG_{PL} values, this does not follow the experimental trend for monomer related MIP performance. It is clear that other factors such as a closer analysis of the binding interactions involved, as well as protein stability in the MIP prepolymerization mixture need consideration.

Protein-monomer prepolymerization complexes can potentially force proteins into conformations or aggregates which when imprinted produce MIPs that lack the desired target selectivity. The monomers interacting with the protein disrupt the secondary structure within the protein, particularly if these interactions are with protein backbone atoms.[34] On this basis, we have analyzed the surface of myoglobin using BioLuminate for the different amino acids involved in hydrogen bonding with the different monomers and whether these interactions are with backbones/sidechains of amino acid residues in helices, β-strands and loops. The surface of the protein is mainly of a positive nature (positive surface area of 4062 $Å^2$; negative surface area 2116 $Å^2$) but the overall formal charge on the protein was 0 (pI myoglobin = 6.8). In terms of potential hydrogen bonding, the donor surface area (1434 $Å^2$) exceeds the acceptor surface area (1035 $Å^2$), indicating that hydrogen bonding to acrylamide analogue carbonyl groups could be similarly favored. The residues on the protein surface that are positively and negatively charged, polar, hydrophobic and small non-polar are highlighted in the FASTA

primary structure sequence shown in Figure 4(A), further highlighting the charged/polar nature of the surface. Figure 4(B) reveals that positively charged Lys residues (11) are the most favored for protein-monomer hydrogen bonding but involvement of the acidic Glu (9) and Asp (6) residues are common as well. The hydrogen bonding with these is mainly sidechain, although some backbone interactions still occur. The rest of the residue types have ≤ 3 amino acids involved in hydrogen bonding interactions. Some non-polar residues are also on the surface but by nature of their structure, only have backbone hydrogen bond interactions. The contributions of hydrogen bonding energies to $\mathrm{D}G_{PL}$ values are shown in Table S4 and range from -0.3 – -2.5 kcal/mol; however, there is not much variance in their magnitudes between the different sites.

(A)

(B)
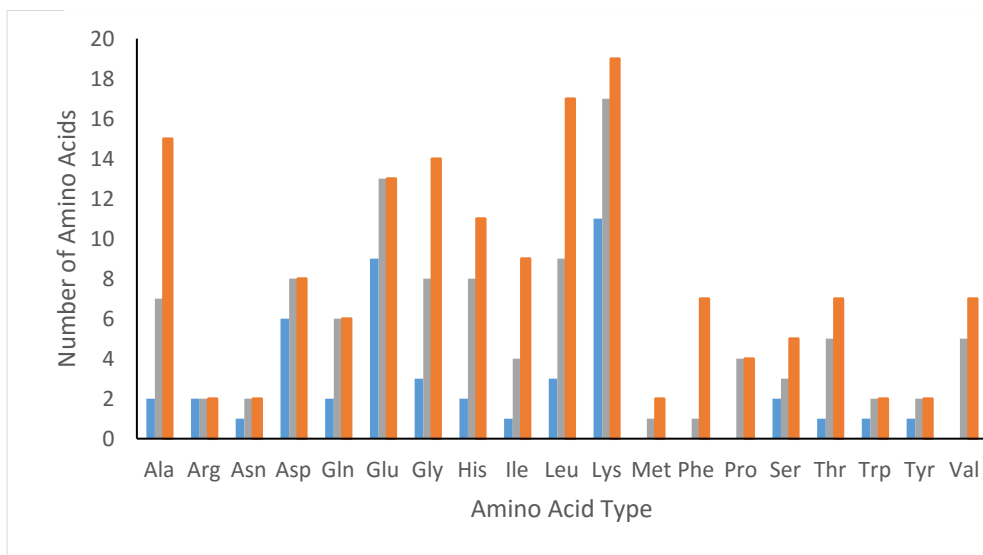
Figure 4: Myoglobin protein residue and surface analysis using Bioluminate. **(A)** FASTA primary structure sequence with the positive, negative, polar, hydrophobic and small non-polar surface residues shown in blue, red, magenta, green and yellow, respectively. **(B)** Total number of the different types of amino acids in the protein (orange) compared to the number found on the surface of the protein (grey) and the number of these involved in hydrogen bond interactions with any monomer (blue).

A more monomer-specific analysis of protein-monomer hydrogen bonding is shown in Table 4. The TrisNHMAm monomer has the highest number (42) of total hydrogen bond interactions with the protein and significantly, also the highest number (8) of helical backbone interactions. It is these interactions, in particular, that can affect protein secondary structure and have a negative impact on MIP selectivity.[34] So, while monomers with hydroxyl groups such as TrisNHMAm can form stronger interactions with protein molecules, these interactions do not necessarily lead to a better MIP due to potential changes in the secondary structure caused by the additional hydrogen bonds. Figure 5(A) and 5(B) show the binding of TrisNHMAm and AAm at binding site **5**, respectively. AAm forms only hydrogen bonds with sidechains, whereas TrisNHMAm is involved in a network of interactions, some of which are backbone. DMAm is predicted to have relatively few hydrogen bond interactions (14) but a high proportion are hydrogen bonds with the helical backbones (4). As DMAm does not have any hydrogen bond donors, these are all with DMAm as hydrogen bond acceptor and the hydrogen bonding energies of DMAm are generally of smaller magnitude (Table S4). Hence, unlike all other monomers, DMAm is unable to exploit hydrogen bond interactions with the many Glu/Asp sidechains on the surface of the protein (Figure 4), also in line with its poor overall MIP performance (Table 1). NHMAm and NHEAm have a low proportion of helical backbone interactions (3 and 1, respectively) in comparison to their relatively large number of interactions (27 and 22, respectively). It is notable also that the best performing monomer as MIP (NHMAm) has the most hydrogen bond contacts overall, TrisNHAm excepted.

Table 4: Summary of the predicted hydrogen bonding interactions for each of the five monomers (Figure 1) with myoglobin as calculated using MM-GBSA. For co-monomer MIPs A (TrisNHAm & DMAm) and B (NHEAm & DMAm), these interactions were estimated based on the dominant binding monomer at a particular site (Table 2).[a]

	Helical Interactions		Loop Interactions		Total
	Backbone	**Sidechain**	**Backbone**	**Sidechain**	
Monomer					
AAm	5	9	4	4	22
NHMAm	3	15	5	4	27
NHEAm	1	12	3	6	22
DMAm	4	5	2	3	14
TrisNHMAm	8	23	5	6	42
Co-Monomer[b]					
MIP A	5	18	3	8	34
MIP B	6	9	4	4	23

[a] Protein contained no β-sheet secondary structure hydrogen bond interactions.
[b] Interactions based on dominant binding monomer for each site (Table 2). In case of equal probabilities for monomers, monomer with most interactions was considered.
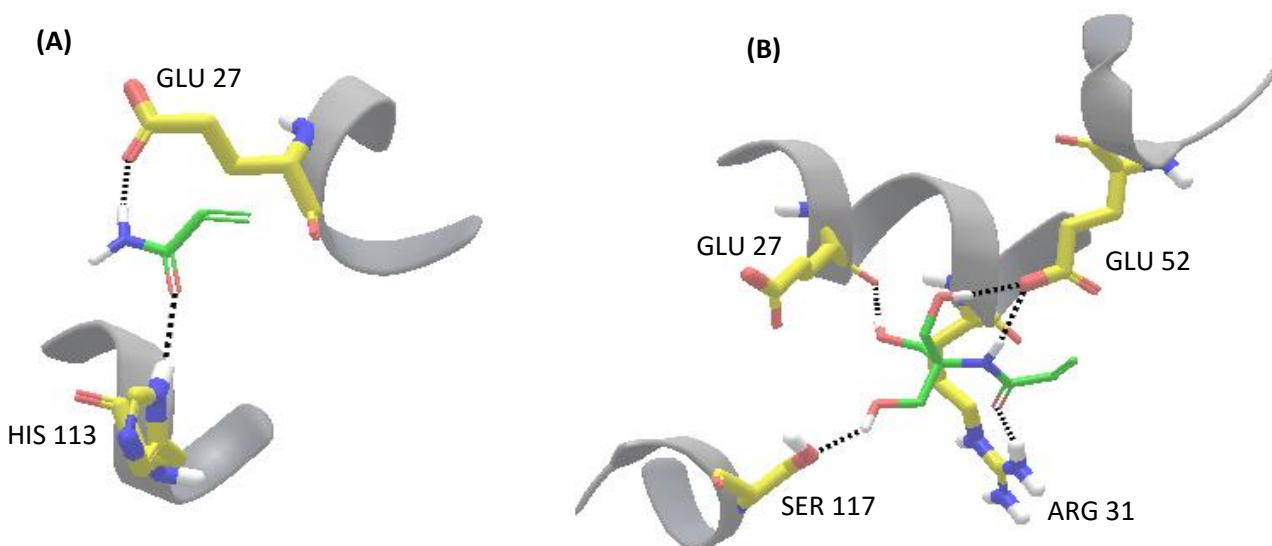
Figure 5: Hydrogen bond interactions between myoglobin residues at binding site **5** and the monomers **(A)** AAm and **(B)** TrisNHAm. Whereas AAm only has sidechain interactions, TrisNHMAm has a network of interactions with both backbone and sidechain.

3.1.5 CD Spectroscopy Result

The computational analysis revealed that the monomers potentially interact with protein helical backbone to varying degrees. As mentioned, this can cause changes in myoglobin secondary structure, leading to poor selectivity for target rebinding to the MIP. Accordingly, CD spectroscopy was used to explore how the different monomers can affect myoglobin secondary structure before polymerization, with the resulting percentage structural composition shown in Figure 6 and the corresponding spectra included in Figure S1.
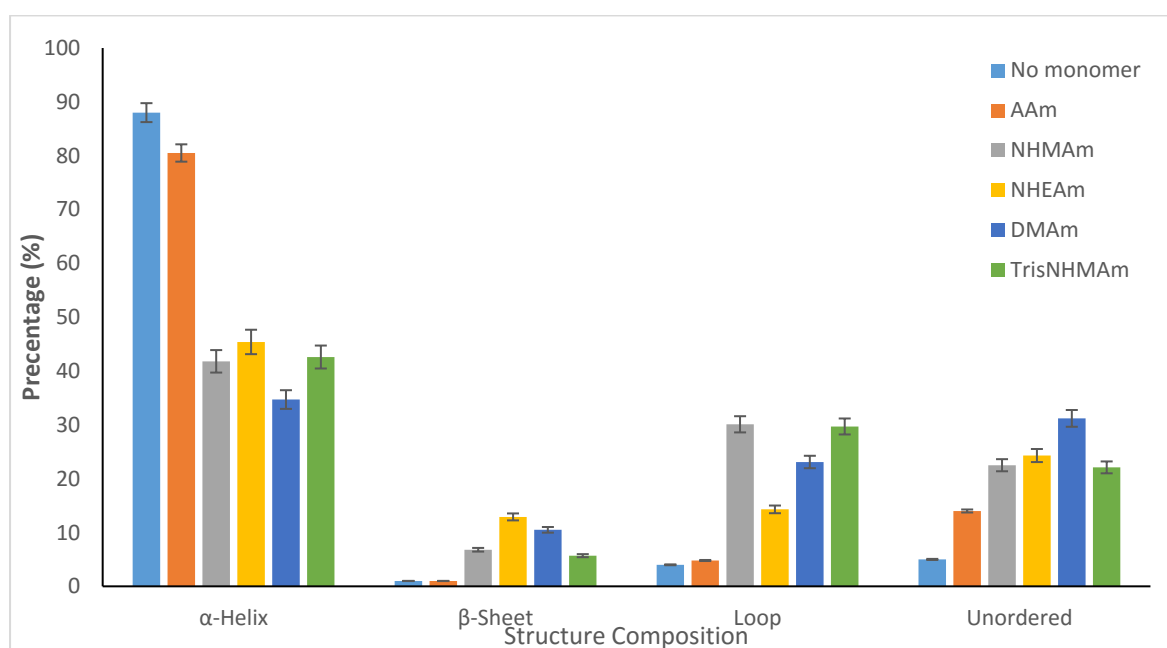
Figure 6: Secondary structure composition of myoglobin after being mixed with each of the 5 monomers (AAm, NHMAm, NHEAm, DMAm, TrisNHMAm) at a protein:monomer ratio of 1:1081, the same as the polymerisation solution used in hydrogel MIP formation.

The native structure of myoglobin was determined to be 88.0±2 % α-helix, 1.0±2 % β-sheet, 4.0±1 % turn, and 5.0±1 % unordered/random. Each monomer was added to a solution of the protein at a ratio of 1:1081, which is consistent with the experimental protein:monomer ratios in MIP prepolymerization mixtures. The monomer AAm produced a structural composition, which is similar to that of native myoglobin, with 80.5±2 % α-helix, 1.0±2 % β-sheet 4.8±1 % turn, and 14.0±1 % unordered/random. While there is a slight decrease in α-helix composition,

this result indicated a significant basis for AAm performing well in the target rebinding studies (Table 1). AAm did have some predicted helical backbone interactions (Table 4) but it is likely that its smaller size and lack of competing hydrogen bond contacts with protein compared to, for example, TrisNHMAm (Figure 5(B)) may be a crucial factor. In fact, all other monomers lead to significant secondary structure changes. The monomer DMAm produced a structural composition of only 34.7±2 % α-helix, 10.5±2 % β-sheet 23.1±1 % turn, and 31.2±1 % unordered/random. Hence, there is considerable change of the myoglobin native secondary structure features, suggestive of selectivity issues and another contributory factor for DMAm producing the worst performing monomer-based MIP (Table 1). DMAm is predicted to form relatively fewer hydrogen bonds and a large proportion were helical backbone. There is additionally potential steric effects due to the dimethyl group. The monomer TrisNHMAm also produced structural compositions with considerable changes in secondary structure, with 42.6±2 % α-helix, 5.7±1 % β-sheet 29.7±1 % turn and 22.1±1 % unordered/random observed which could be expected based on extensive hydrogen bonding interactions (3 monomer hydroxyls; 8 predicted helical backbone interactions). Somewhat surprisingly, however, the structurally similar NHMAm (3 helical backbone interactions) and NHEAm (1 helical backbone interaction) each with one hydroxyl, also cause secondary structure changes with values similar to those of TrisNHMAm. These monomers did not have many helical backbone interactions, yet secondary structure was affected. It does appear, therefore, that monomer size is a factor. Despite this, NHMAm and NHEAm performed well in the MIP target rebinding experiments, with the NHMAm based MIP in fact being the top-performer. NHMAm is smaller and less flexible than extended by -CH_2- analogue, NHEAm. In comparison, flexibility of cross-linker in MIP design is reported as favoured [55]. In conclusion, the CD data has indicated that secondary structure alteration (or lack of it) is likely a key factor in MIP performance

(AAm, TrisNMMAm. DMAm); however, NHMAm and NHEAm despite having a clear effect on secondary structure features still perform well in the MIP experiments.

Overall, analysis of the single monomer computational and experimental data unquestionably indicates a complex interplay of potential factors (including site binding free energies, nature of protein-monomer interactions and secondary structure changes) for a successful monomer used for producing protein-targetted MIPs. This led us to consider combining features of different monomers in carefully selected co-monomer experiments predicting outcomes in advance to validate our hypotheses.

3.2 Co-monomer MIPs Studies

3.2.1 Rational Design of Co-Monomer MIPs

Based on the single monomer studies, co-monomer MIPs were rationally designed, synthesized and further experimental rebinding studies performed. The criteria we set for good performance were that monomers bind strongly across the different predicted binding sites. Additionally, predicted binding of the co-monomer mixtures should have limited helical backbone interactions and have minimal competition with each other for each binding site. Two co-monomer MIPs were selected for synthesis. The first MIP (co-monomer mixture A) was designed to produce, in theory, a co-monomer MIP with better performance than either of the MIPs for the individual monomers. The second MIP (co-monomer mixture B) was predicted to perform worse than the individual co-monomers MIPs. Eq. (6) was employed to calculate the site-specific relative binding probabilities of monomers for the two co-monomer solutions (Table 2). The average number of each monomer $\bar{M}_L(x)$ bound for the MIPs are shown in Table 3.

Co-monomer mixture A contained a 50:50 mixture of the monomers TrisNHMAm and DMAm. This combination represented a good test of our hypotheses for improvement in MIP performance. The computed binding free energies suggested that DMAm binds strongly into the binding sites where TrisNHMAm does not (sites **2** and **9**). This suggests that for MIP A, there will be potentially strong uniform binding to 12 of the 14 predicted binding sites. The TrisNHAm:DMAm combination also takes into account the number of helical backbone interactions. Theoretically, a monomer is placed into a binding site resulting in less helical hydrogen bonding, and therefore potential for less structural changes in the protein compared to TrisNHMAm alone (Table 4). The average number of bound monomer $\bar{M}_L(x)$ is expected to be mainly TrisNHMam (9.34) compared to DMAm (2.14) due to the much more negative TrisNHMam $\mathrm{D}G_{PL}$ values (Table 3).

Co-monomer mixture B consisted of the monomers NHEAm and DMAm. These monomers were chosen based on predicted binding competition at sites **3**, **4**, **5**, **6**, **8**, **9**, **10**, **11** and **12**. The average number of bound monomers $\bar{M}_L(x)$ was expected to be similar for each monomer, NHEam with 5.62 compared to DMAm (4.85) (Table 3). This competition of monomers for binding sites reduces the selectivity of a MIP [34]. Also, for this combination, the sites where there is the strongest binding has considerable helical backbone binding (Table 4), and hence increasing the potential for secondary structure changes. Overall, therefore, we expected a less effective co-monomer MIP B compared to those prepared using the individual NHEAm and DMAm monomers.

3.2.2 Experimental Rebinding Co-monomer Results

The percentage rebind and imprinting factors (IF) results for the co-monomer MIPs A (TrisNHMAm and DMAm) and B (NHEAm and DMAm) and their corresponding NIPs are shown in Table 1.

MIP A with a 1:1 ratio of TrisNHAm and DMAm produced a MIP rebind of 85.5 % which was much better % rebind than when either of the individual monomers were used (79.9 % and 72.0 % for TrisNHAm and DMAm, respectively). The percentage rebind for the NIP is 63.4% leading to an IF value (1.3) greater than the established 1.2 threshold. In agreement with our predictions, therefore, this rationally designed combination of monomers to provide a more uniform coverage of the protein binding sites resulted in improved performance of the MIP. Meanwhile, MIP B produced a MIP rebind of 63.2%, considerably worse than each of the individual monomers (77.2% and 72.0% for NHEAm and DMAm, respectively). Again, in agreement with our predictions, this combination with predicted competition for binding sites lead to decreased performance of the MIP compared to individual monomers. Exploiting computational and rational design such as employed here, therefore, we have demonstrated the potential for prediction of the performance of a MIP in advance is possible.

3. Conclusions

We have exploited computational methods to gain a further understanding of the performance of hydrogel-based protein imprinted MIPs, based on predictions of protein-monomer interactions at the atomic level. Identification of 14 potential binding sites on the chosen target protein myoglobin was followed by molecular docking of 5 acrylamide-based test monomers (AAm, NHMAm, NHEAm, TrisNHMAm, DMAm). MM-GBSA post-docking calculations were performed to calculate the binding free energies and used to assess monomer performance. The results revealed that monomers with hydroxyl functional groups (NHMAm, NHEAm, TrisNHMAm) have generally stronger binding free energy (ΔG_{PL}) values than those

without the hydroxyl groups (AAm). While this showed some correlation with experimental data and with the NIP performance of the various monomers, it did not reflect, for example, the better MIP performance of the AAm monomer, with few protein-ligand hydrogen bonds (smaller ΔG_{bind} magnitudes) and the poorer performance of TrisNHMAm, with many hydrogen bonds (larger ΔG_{PL} absolute values). Analysis of protein-monomer interactions helped in part to explain these anomalies. In comparison to AAm, TrisNHMAm had a large amount of (inter-helical) protein backbone interactions, which could result in the secondary structure composition changes observed by CD spectroscopy, reducing the selectivity of the MIP. The monomers NHEAm and NHMAm despite affecting myoglobin secondary structure did produce MIPs with good selectivity, with the better MIP performance of NHMAm over NHEAm potentially related in part to less conformational flexibility. In fact, the smallest monomers AAm and NHMAm produced the best performing MIPs indicating that this simple parameter is an important factor. Analyzing predictions for binding of monomers at the individual sites (in terms of site specific ΔG_{PL} values and protein-monomer interactions), two co-monomer MIPs, MIP A and MIP B were rationally designed. For MIP A, monomers predicted to increase uniform binding around the surface of the protein were selected, while minimising helical backbone interactions and competition for binding sites. This resulted in a co-monomer MIP that performed better than the individual MIPs. For MIP B, rationally designed to perform less well than the individual monomers, two monomers that had competition for binding sites and that interacted strongly with the protein backbone and helical structure were selected, all of which should reduce MIP performance. Again, experimental validation was in line with the predictions. Overall, therefore, a better understanding of key atomistic interactions within a MIP-protein complex has been established in this study and represents a further step towards the rational design of MIPs with predictable characteristics in the future.

Associated Content

Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

Detailed derivation of the equations used for the thermodynamic analysis, the supplementary tables and figures are included as Supporting Information.

Acknowledgements

References

(1) Haupt, K. Biomaterials: Plastic Antibodies. *Nat. Mater.* **2010**, *9*, 612-614.

(2) Mosbach, K.; Ramstron, O. The Emerging Technique of Molecular Imprinting and its Future Impact on Biotechnology. *Nat. Biotechnol.* **1996**, *14*, 163-170.

(3) Ye, L.; Haupt, K. Molecularly Imprinted Polymers as Antibody and Receptor Mimics for Assays, Sensors and Drug Discovery. *Anal. Bioanal. Chem.* **2004**, *378*, 1887-1897.

(4) Zhang, H.; Ye, L.; Mosbach, K. Non-Covalent Molecularly Imprinting with Emphasis on its Application in Separation and Drug Development. *J. Mol. Recognit.* **2006**, *19*, 248-259.

(5) Reddy, S. M.; Sette, G.; Phan, Q. T. Electrochemical Probing of Sellective Haemoglobin Binding in Hydrogel-Based Molecularly Imprinted Polymers. *Electrochim. Acta* **2011**, *56*, 9203-9208.

(6) El-Sharif, H.; Hawkins, D. M.; Stevenson, D.; Reddy, S. M. Determination of Protein Binding Affinities within Hydrogel-Based Molecularly Imprinted Polymers (HydroMIPs). *Phys. Chem. Chem. Phys.* **2014**, *16*, 15483-15489.

(7) El-Sharif, H.; Yapati, H.; Kalluru, S.; Reddy, S. M. Highly Selective BSA Imprinted Polyacrylamide Hydrogels Facilitated by a Metal-Coding MIP Approach. *Acta Biomater.* **2015**, *28*, 121-127.

(8) Alexander, C.; Andersson, H. S.; Andersson, L. I.; Ansell, R. J.; Kirsch, N.; Nicholls, I. A.; O'Mahony, J.; Whitcombe, M. J. Molecularly Imprinting Science and Technology: A Survey of the Literature for the Years Up to and Including 2003. *J.Mol. Recognit.* **2006**, *19*, 106-180.

(9) Vasapollo, G.; Del Sole, R.; Mergola, L.; Lazzoi, M. R.; Scardino, A. Molecularly Imprinted Polymers: Present and Future Prospective. *Int. J.Mol. Sci.* **2011**, *12*, 5908-5945.

(10) Lieberzeit, P. A.; Samardzic, R.; Kotova, K.; Hussain, M. MIP Sensors on the Way to Biotech Application:Selectivity and Ruggedness. *Procedia Eng.* **2012**, *47*, 534-537.

(11) Urraca, J. L.; Hall, A. J.; Moreno-Bondi, M. C.; Sellergren, B. A Stoichiometric Molecularly Imprinted Polymer for the Class-Selective Recognition of Antibiotics in Aqueous Media. *Angew. Chem., Int. Ed.* **2006**, *45*, 5158-5161.

(12) Ayankojo, A. G.; Tretjakov, A.; Reut, J.; Boroznjak, R.; Opik, A.; Rappich, J.; Furchner, A.; Hinrichs, K.; Syritski, V. Molecularly Imprinted Polymer Integrated with a Surface Acoustic Wave Technique for Detection of Sulfamethizole. *Anal. Chem.* **2016**, *88*, 1476-1484.

(13) Kidakova, A.; Reut, J.; Rappich, J.; Opik, A.; Syritski, V. Preparation of a Surface-Grafted Protein-Selective Polymer Film by Combined use of Controlled/Living Radical Photopolymerization and Microcontact Imprinting. *React. Funct. Polym.* **2018**, *125*, 47-56.

(14) Deng, Q.; Wu, J.; Zhai, X.; Fang, G.; Wang, S. Highly Selective Fluorescent Sensing of Proteins Based on a Fluorescent Molecularly Imprinted Nanosensor. *Sens.* **2013**, *13*, 12994-13004.

(15) Gupta, S.; Saxena, M.; Saini, N.; Mahmooduzzafar; Kumar, R.; Kumar, A. An Effective Strategy for a Whole-Cell Biosensor Based on Putative Effector Interaction Site of the Regualtory DmpR Protein. *PLoS One* **2012**, *7*, e43527.

(16) Renkecz, T.; Mistlberger, G.; Pawlak, M.; Horvath, V.; Bakker, E. Molecularly Imprinted Polymer Microspheres Containing Photoswitchable Spiropyran-Based Binding Sites. *ACS Appl. Mater. Interfaces* **2013**, *5*, 8537-8545.

(17) Krstulija, A.; Lettieri, S.; Hall, A. J.; Roy, V.; Favetta, P.; Agrofoglio, L. A. Tailor-made Molecularly Imprinted Polymer for Selective Recognition of the Urinary Tumor Marker Pseudouridine. *Macromol. Biosci.* **2017**, *17*, 1700250.

(18) Hawkins, D. M.; Stevenson, D.; Reddy, S. M. Investigation of Protein Imprinting in Hydrogel-Based Molecularly Imprinted Polymers (Hydrogels). *Anal. Chim. Acta* **2005**, *542*, 61-65.

(19) Reddy, S. M.; Phan, Q. T.; El-Sharif, H.; Govada, L.; Stevenson, D.; Chayen, N. E. Protein Crystallization and Biosensor Applications of Hydrogel-Based Molecularly Imprinted Polymers. *Biomacromolecules* **2012**, *13*, 3959-3965.

(20) Poma, A.; Guerreiro, A.; Whitcombe, M. J.; Piletska, E. V.; Turner, A. P. F.; Piletsky, S. A. Solid-Phase Synthesis of Molecular Imprinted Polymer Nanoarticles with a Reusable Template - "Plastic Antibodies". *Adv. Funct. Mater.* **2013**, *23*, 2821-2817.

(21) Ge, Y.; Turner, P. F. Too Large to Fit? Recent Developments in Macromolecular Imprinting. *Trends Biotechnol.* **2008**, *26*, 218-224.

(22) Tse Sum Bui, B.; Haupt, K. Molecularly Imprinted Polymers: Synthetic Receptors in Bioanalysis. *Anal. Bioanal. Chem.* **2010**, *398*, 2481-2492.

(23) Alizadeh, T.; Zare, M.; Ganjali, M. R.; Norouzi, P.; Tavana, B. A New Molecularly Imprinted Polymer (MIP)-Based Electrochemical Sensor for Monitoring 2,4,6-Trinitrotoluene (TNT) in Waste Waters and Soil Samples. *Biosens. Bioelectron.* **2010**, *25*, 1166-1172.

(24) Piletska, E. V.; Guerreiro, A.; Whitcombe, M. J.; Piletsky, S. A. Influence of the Polymerization Conditions on the Performance of Molecularly Imprinted Polymers. *Macromolecules* **2009**, *42*, 4921-4928.

(25) Piletsky, S. A.; Piletska, E. V.; Karim, K.; Freebairn, K. W.; Legge, C. H.; Turner, A. P. F. Polymer Cookery: Influence of Polymerization Conditions on the Performance of Molecularly Imprinted Polymers. *Macromolecules* **2002**, *35*, 7499-7504.

(26) Karim, K.; Breton, F.; Rouillon, R.; Piletska, E. V.; Guerreiro, A.; Chianella, I.; Piletsky, S. A. How to Find Effective Functional Monomers for Effective Molecularly Imprinted Polymers? *Adv. Drug Delivery Rev.* **2005**, *57*, 1795-1808.

(27) Cowen, T.; Karim, K.; Piletsky, S. A. Computational Approaches in the Design of Synthetic Receptors- A Review. *Anal. Chim. Acta* **2016**, *936*, 62-74.

(28) Piletsky, S. A.; Karim, K.; Piletska, E. V.; Day, C. J.; Freebairn, K. W.; Legge, C. H.; Turner, A. P. F. Recognition of Ephedrine Enantiomers by Molecularly Imprinted Polymers Designed using a Computational Approach. *Analyst* **2001**, *126*, 1826-1830.

(29) Bohm, H. J. The Computer Program LUDI: A New Method for the De Novo Design of Enzyme Inhibitors. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 61-78.

(30) Chianella, I.; Lotierzo, M.; Piletsky, S. A.; Tothill, I. E.; Chen, B.; Karim, K.; Turner, A. P. F. Rational Design of a Polymer Specific for Microcystin-LR using a Computational Approach. *Anal. Chem.* **2002**, *74*, 1288-1293.

(31) Olsson, G. D.; Niedergall, K.; Bach, M.; Karlsson, B. C. G.; Tovar, G.; Nicholls, I. A. Simulation of Imprinted Emulsion Prepolymerization Mixtures. *Polym. J.* **2015**, *47*, 827-830.

(32) Khan, M. S.; Pal, S.; Krupadam, R. J. Computational Strategies for Understanding the Nature of Interation in Dioxin Imprinted Nanoporous Trappers. *J. Mol. Recognit.* **2015**, *28*, 427-437.

(33) Boroznjak, R.; Reut, J.; Tretjakov, A.; Lomaka, A.; Opik, A.; Syritski, V. A Computational Approach to Study Functional Monomer-Protein Molecular Interactions to Optimize Protein Molecular Imprinting. *J. Mol. Recognit.* **2017**, *30*, e2635.

(34) Kryscio, D. R.; Shi, Y.; Pengyu, R.; Peppas, N. A. Molecular Docking Simulations for Macromolecularly Imprinted Polymers. *Ind. Eng. Chem. Res.* **2011**, *50*, 13877-13884.

(35) Pan, J.; Xue, X.; Wang, J.; Xie, H.; Wu, Z. Recognition Property and Preparation of *Staphylococcus Aureus* Protein A-Imprinted Polyacylamide Polymers by Inverse-Phase Suspension and Bulk Polymerization. *Polym.* **2009**, *50*, 2365-2372.

(36) Morton, D. J.; Van Wagoner, T. M.; Seale, T. W.; Whitby, P. W.; Stull, T. L. Utilization of Myoglobin as a Heme Source by Haemophilus Influenzae Requires Binding of Myoglobin to Haptoglobin. *FEMS Microbiol. Lett.* **2006**, *258*, 235-240.

(37) Halgren, T. A. New Method for Fast and Accurate Binding-Site Identification and Analysis. *Chem. Biol.Drug Des.* **2007**, *69*, 146-148.

(38) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377-389.

(39) Schrodinger In *Schrodinger Release 2016-3;* LLC: New York, NY, 2016.

(40) Tetin, S. Y.; Prendergast, F. G.; Venyaminov, S. Y. Accuracy of Protein Secondary Structure Determination from Circular Dichroism Spectra Based on Immunoglobulin Examples. *Anal. Biochem.* **2003**, *321*, 183-187.

(41) Sreerama, N.; Woody, R. W. Esitmation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Anal. Biochem.* **2000**, *287*, 252-260.

(42) Sondergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Imroved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK(a) Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284-2295.

(43) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L., et al. OPLS3: A Force Field Providing Broad Coverage of Drug-Like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281-296.

(44) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49*, 6177-6196.

(45) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K., et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739-1749.

(46) Curk, T.; Dobnikar, J.; Frenkel, D. Rational Design of Molecularly Imprinted Polymers. *Soft Matter* **2016**, *12*, 35-44.

(47) Whitmore, L.; Woollett, B.; Miles, A. J.; Janes, R. W.; Wallace, B. A. The Protein Circular Dichroism Data Banck, a Web-Based Site for Access to Circular Dichroism Spectroscopic Data. *Struct.* **2010**, *18*, 1267-1269.

(48) Whitmore, L.; Wallace, B. A. Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers* **2008**, *89*, 392-400.

(49) Whitmore, L.; Wallace, B. A. DICHROWEB, an Online Server for Protein Secondary Structure Analyses from Circular Dichroism Spectroscopic Data. *Nucleic Acids Res.* **2004**, *32*, W668-W673.

(50) Hayes, J. M.; Archontis, G. In *MM-GB(PB)SA Calculations of Protein-Ligand Binding Free Energies;* Wang, L., Ed.; Molecular Dynamics - Studies of Synthetic and Biological Macromolecules; InTech: 2012; Vol. 8, 171-190.

(51) Swanson, J. M. J.; Henchman, R. H.; McCammon, J. A. Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. *Biophys. J.* **2004**, *86*, 67-74.

(52) Woo, H. J.; Roux, B. Calculation of Absolute Protein-Ligand Binding Free Energy from Computer Simulations. *Proc. Natl. Acad. Sci. USA.* **2005**, *102*, 6825-6830.

(53) Hayes, J. M.; Skamnaki, V. T.; Archontis, G.; Lamprakis, C.; Sarru, J.; Bischler, N.; Skaltsounis, A. L.; Zographos, S. E.; Oikonomakos, N. G. Kinetic, *In Silico* Docking, Molecular Dynamics, and MM-GBSA Binding Studies on Prototype Indirubins, KT5720, and Staurosporine as Phosphorylase Kinase ATP-Binding Site Inhibitors: The Role of Water Molecules Examined. *Proteins: Struct. Funct. Bioinf.* **2011**, *79*, 703-719.

(54) Rapp, C. S.; Kalyanaramen, C.; Schiffmiller, A.; Schoenbrun, E. L.; Jacobson, M. P. A Molecular Mechanics Approach to Modeling Protein-Ligand Interactions: Relative Binding Affinities in Congeneric Series. *J. Chem. Inf. Model.* **2011**, *51*, 2082-2089.

(55) Sibrian-Vazquez, M.; Spivak, D. A. Improving the Strategy and Performance of Molecularly Imprinted Polymers using Cross-Linking Functional Monomers. *J. Org. Chem.* **2003**, *68*, 9604-9611.