Introduction

While computational methods have long been used in the design of synthetic receptors, a comprehensive approach has largely been unavailable for use as a standard protocol in the production of high affinity biomimetic materials. Molecularly imprinted polymer (MIP) synthetic receptors or 'plastic antibodies' are prepared by the formation of a cross-linked polymer in the presence of a molecular template. The self-assembly of functional monomers with complementary functional groups to those of the template results in formation of a pre-polymerization complex, which is stabilized by cross-linking during polymer formation. Appropriate selection of functional monomers with strong interaction energies with the template will favour the associated state, resulting in the formation of MIPs with high affinity and specificity for the target. In the absence of a rational approach to the design of MIPs, monomer selection is often made on the basis of previous experience or chemical intuition; in many cases methacrylic acid is used as the sole functional monomer due to its importance in the history of molecular imprinting. Laboratory-based approaches to the optimization of monomer compositions centre on combinatorial synthesis and screening [1,2], an approach limited by the large number of different polymers required to account for the many potentially suitable monomers.

With thousands of functional monomers commercially available or readily synthesized, a more rational approach to monomer selection was obviously required for further advancement of the field. The literature provides numerous examples where researchers have investigated the strength of monomer–template interactions through the use of molecular mechanics (MM) [3–6], molecular dynamics (MD) [7–9], and quantum mechanics (QM) [10,11], based molecular modelling techniques, but these are still limited by the sequential screening of each monomer manually.

We have developed a protocol utilizing a molecular mechanics/molecular dynamics approach to the automated screening of a library of candidate functional monomers for their interaction with a chosen template. For this purpose, we have used a commercial suite of software (SYBYL® from Tripos Inc.) that uses a small library of commercially available...
monomers. The validation of the described computational protocol as a means of rapid and reliable MIP design is provided by reference to many published examples of high-affinity MIPs for a diverse range of targets prepared according to this design strategy.

Incentives

The foundational principle of computational MIP design is that the stability of the template–monomer complex is directly related to the quality of imprinted sites created in the polymer after cross-linking. Much of the contemporary use of molecular modelling in the design of MIPs therefore is centred on the equation:

\[ \Delta E = E_c - (E_i + \Sigma E_{M}) \]  

(1)

Where \( E_c \), \( E_i \) and \( E_{M} \) are the lowest calculable energies of the template–monomer complex, template and monomer respectively. Comparison of the different values of \( \Delta E \) gives an indication of the relative stability of different components in the system, and thus provides an appropriate guide to the selection of an appropriate monomer [12–14], to find the most suitable template–monomer ratio [15–17], or both [18–20]. This method has become popular along with the use of QM-based (predominantly DFT) techniques in MIP design as a result of advances in hardware making greater computer power available. However, this approach is still time consuming and computationally demanding, and QM is for this reason associated with the screening of a relatively small number of monomers (typically five or fewer). There are rare exceptions to this in which 20 or more functional monomers, along with a number of cross-linkers, have been ranked against a particular template [21–23], but the time presumably invested in this does not lend to this being an appropriate general model of MIP design.

While QM has advantages in accuracy over the alternatives which makes it desirable in the comparison of different polymerization constituents, either directly via the above equation, or by frontier orbital analysis [24–26], which can be used as an indicator of kinetic stability [27], MM and MD have the power to perform hundreds of tasks simultaneously, such as the simulation of pre-polymerization systems consisting of thousands of molecules, and provide an analysis of interactions occurring between molecules over many nanoseconds [28]. These models are also continually being revisited, often with the adoption of techniques that are new to MIP design, such as analysis by determination of the radial distribution functions (RDFs) of atoms likely to be involved in hydrogen bond formation. RDF methods provide the distances between atoms and allows selection of appropriate chemical components in the polymerization system (monomers, solvents, etc.), when used as a tool in predicting the likelihood of successful template complexation and polymer synthesis [29–31].

The protocol described herein began development in 2000 [32], and an early form of the procedure was employed for the first time to design MIPs for creatinine [33], ephedrine [34], and microcystin–LR [35]. Dozens of papers have since been published describing the use of this protocol for a broad range of templates, with the technique being continually modified to provide a reliable method of designing high affinity imprinted polymers. Here can be seen the incremental advancements describing how molecular modelling techniques can be used to rapidly screen large databases of functional monomers in order to identify suitable candidate monomers for MIP preparation. The computational time and resources required for performing these MM and MD calculations of monomer–template interactions are modest and can produce results within a few hours. The method represents a generic procedure for the selection of monomer mixtures for the imprinting of virtually any template.

Experimental

All calculations and procedures were carried out on a desktop PC running RHEL 3.0 or later (Linux platform), executing the software package SYBLY 7.3 (Tripos Inc.). The protocol described was developed using the SYBLY software but can be adapted for application in other programs. Standard procedures are followed regarding preparation of the selected template, which may be either the whole molecule (as is typical in smaller structures) or an appropriate epitope may be used to represent the binding region of a biochemical macromolecule. These structures are often obtained from online sources such as PubChem [36], ZINC [37,38], or RCSB PDB [39], when possible to ensure the correct appropriate template geometries are presented in screening.

Automatic monomer screening

Templates constructed manually may be minimized and processed by simulated annealing using any available force field, but for greater compatibility with the LeapFrog protocol the Tripos force field and Gasteiger–Hückel charges are preferred. All structures must be available in a mol2 file format.

The monomer library can be constructed by a number of approaches. Using the SYBLY software a large number of molecules can be saved under one file name or retained in one folder easily, facilitating the writing of a script which sequentially loads a monomer, records the total internal energy of the monomer and template in isolation, forms a complex by energy minimization, and records the energy of the new arrangement before restarting with a new monomer. This process can be easily automated using a simple algorithm written in SYBLY Programming Language (SPL), and can be easily adapted for use in other software. Here however we emphasize the benefits of adapting LeapFrog for use in the screening process; Leapfrog includes a function to add an observed structure to a database (‘add fragments’), or a large number of monomer can be automatically added with simple SPL algorithms (An example script is given in Appendix 1).

Once the library is established the screening can be initiated by launching the Leapfrog program. Using the ‘dream’ mode allows greater freedom to modify parameters and ensuring the ‘calculate’ option is enabled and set to ‘all atoms’ allows observation of the whole template as opposed to the binding cavity of a macromolecule. In the ‘tradeoff’ between quality and variety the former must be maximized in the ‘tradeoff’
dialogue, including removal of ‘protected’ atoms and exclusion of desolvation and chemical synthesis effects on the binding energy. Inclusion of hydrogen bond energy must also be requested via the tailored ‘energy startup’ options.

For effective adaption of Leapfrog the ‘relative move frequencies’ must also be modified from their default settings. This largely involves simplifying the relatively sophisticated program, which builds idealized receptor binding compounds from libraries of simple synthon equivalents. By removing the option to form bonds between these simple structures in the cavity (i.e. around the template) an effective screening protocol can be developed. The frequencies of all actions are therefore set as zero, except for ‘new’, ‘twist’, ‘save’ and ‘weed’, which are typically set at 10, 5, 5 and 1 respectively.

Stoichiometric refinement by molecular dynamics

On completion of the Leapfrog run a table is produced listing each of the monomers by their binding energy to the template. The highest ranking monomer or monomers can be observed in their highest affinity positions around the template and selected appropriately for stoichiometric refinement by molecular dynamics simulation. SYBYL provides a number of methods of solvating the template in the chosen monomer(s), and a number of approaches may be taken. Here we apply the XFIT algorithm, which sequentially adds solvents around the template in a close packing arrangement. This continues until the edges of the simulation environment, a small cube of dimensions automatically determined by the dimensions of the template and monomer. Molecular dynamics simulations are performed with an NTP ensemble at 300 K for 1 ns with a dielectric constant of 1. The Tripos force field is typically used with Gasteiger–Hückel charges and a non-bonding interaction limit of 8 Å is applied. The initial velocity is set from the relevant Maxwell–Boltzmann distribution. On completion the system is minimized and the interactions between the template and the solvent monomer are observed, the complex present being indicative of the appropriate ratios for the highest affinity ratio of monomers for the polymerization mixture.

Results and Discussion

Application of the method

The protocol will be of interest to researchers involved in the design and synthesis of MIPs in any format (e.g. micro- and nano- particles, films or monoliths), and suitable for the design of high affinity MIPs for diverse templates including clinical targets (drugs), environmental/food targets (e.g. toxins) and for MIPs to be used in extreme environments. This protocol is particularly suitable for use with low molar mass templates and where the development of high affinity MIPs is required: such as (i) in the separation and purification of high–value products; (ii) analytical sample pre-treatment and solid–phase extraction; (iii) drug or fragrance release matrices; (iv) adsorbents for clinical or environmental applications; (v) sensors and assays for environmental analysis, food analysis and clinical diagnostics. The benefits for end–users of this technology have been identified within clinical analysis, in diagnostics, in pharmaceutical manufacturing and by environmental agencies.

The protocol has many advantages over other modelling techniques for monomer selection and MIP design. A library of 20 polymerizable monomers (27 accounting for charged and neutral forms) can be rapidly screened with a template in around 30 minutes, or approximately 60 minutes with a database of 100 monomers. The tasks described here can be accomplished using an unmodified desktop PC in reasonable run times, obviating the need for supercomputing facilities. In the case of rationally–designed polymers (RDPs, non–imprinted polymers bearing functional monomers selected on the basis of their interaction with the intended target), the same screening protocol can be used without the need for MD analysis, further reducing the time required for effective design. For some applications (such as in environmental and food analysis) RDPs may be preferred over MIPs as they possess a high binding capacity and reduced cost, while retaining good selectivity and affinity for the target.

Database design and automated screening

The monomer library in figure 1 is the original selection used in the automated MIP design procedure, and contains a range of acidic, basic and neutral monomers that may be capable of interacting with the template through non–covalent interactions. In this article the screening of members of the virtual monomer library for their interactions with the template is carried out using the LeapFrog algorithm within SYBYL. Existing tools for automatically ranking the greatest interaction of each of a library of compounds with another compound are not known to the authors, but the design or adaption of existing programs is possible. LeapFrog is a generic algorithm that provides a means for docking ligand precursor molecules with receptor binding sites in order to determine the optimum ligand structure. Interaction points within the receptor site are identified by sampling the environment within the binding site and determining the electrostatic, steric, and lipophilic characteristics, giving an indication of appropriate geometric and electronic properties of new drugs for the receptor. In choosing the correct parameters for the docking, the ‘ligands’ may be the library of functional monomers, and the ‘receptor sites’ regions of high or low electron density immediately surrounding the template.

Using the LeapFrog approach, each monomer is placed in close proximity to the interaction points identified on the template surface and a binding energy calculated. The monomers are then rotated by a set small distance around the interaction point or moved to a different site, and the binding energy is again recorded. Figure 2 shows the interaction points (represented here as red, blue and yellow spheres) around a creatinine template with various monomers being analyzed. Upon completion the monomers are presented with each of their strongest binding interactions and ranked accordingly. The positions of each of the monomers, both in the position of greatest affinity and lesser arrangements, can also be visualized with the template to provide a sense of regionality. Table 1 shows a typical binding energy table ranking monomers with highest binding with N–3–oxo–dodecanoyl–L–homoserine lactone template.

Monomers giving the highest binding scores will be those that form the strongest complexes with the template and represent...

Figure 1: Virtual library of functional monomers first used in the automated screening protocol. The library has since been expanded, but the monomers shown are still regularly used in synthesis.

Figure 2: General (top left) and 3D ball-and-stick (top right) representations of creatinine, and monomer interaction with creatinine template (bottom). The colored balls in red, blue and yellow donate the interaction sites used to run the LeapFrog algorithm. These sites represent points of electron density maxima and minima around the template creatinine.
the best candidates for polymer preparation. By visualizing the interactions for several high affinity monomers, and each of these monomers in positions of slightly lower affinity, an indication is given of possible combinations of monomers which could be used to develop higher affinity MIPs.

**Refinement using MD simulations**

The appropriate ratio of polymerization mixture components is determined by performing MD simulations. This may involve the use of a single monomer species, or if the screening shows that two different monomers interact with different regions of the template, then these monomers may act synergistically in the imprinting of that template and both will then be used.

A pre-computed box of fixed dimensions is prepared by saturating the space around the template with the monomer selected from the results obtained during the screening. Figure 3 shows a graphical representation of a pre-computed box with the template at the center of the box (shown in purple) surrounded by itaconic acid. Upon equilibration the system is energetically minimized to clearly show the interactions found between the template and the monomers. Analysis of this complex provides a guide to the appropriate ratios of monomer to template in the pre-polymerization mixture. While the screening process yields good predictions of appropriate functional monomer selection alone, this process of solvating the template in the highest affinity monomer/monomers adds an additional refinement step which in not accounted for by any other approach to design.

**Protocol development**

The procedure was first demonstrated in a simple form some time ago for the design of a MIP for creatinine [34]. Since that time dozens of papers have been published describing the successful design of MIPs for a broad range of templates [40,41], some of which are listed in table 2 [33–35, 42–52]. The protocol has been advanced through its initial use by refinement of the parameters and the introduction of further functional monomers. A number of these new additions are more specialist compounds that must first be synthesized and cannot be readily obtained commercially, but are useful in controlling target affinity and selection for certain polymer properties.

In the case of RDPs the choice of monomer selected via the screening process has been shown to be sufficient for the synthesis of high affinity materials [43,44]. For the synthesis of these materials further refinement is not required, and so the whole design procedure can be completed in under an hour. Typically however a stoichiometric ratio must be determined for effective complexation in the pre-polymerization mixture, for an imprinted polymer, and thus the MD protocol must be followed. The examples in table 2 range from (i) good imprinting factors [33,35]; (ii) high recovery of template from using solid phase extraction (SPE) [42,43,50,51]; (iii) controlled release [46,47]; (iv) dissociation constant in nM [35]; and (v) industrial applications [44,48]. This procedure therefore can be observed to produce excellent results with minimal time requirements, making this whole process highly efficient in comparison with alternative approaches.

**Acknowledgments**

This research used the SPECTRE High Performance Computing Facility at the University of Leicester. We would like to thank Dr Andrew Myers for his IT support.

**Author Contributions**

K. K., E. P. and S. P. developed the protocols. A. G. and M. W. contributed the data. K. K., T. C. and S. P. wrote the paper. All authors have discussed the results and approved the final manuscript.
Table 2: Examples of MIPs synthesized using this protocol. Several examples have been collected which can be used to emphasize the high affinity and selectivity which can be achieved with this approach.

<table>
<thead>
<tr>
<th>Template</th>
<th>Polymer Highlights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ephedrine [33]</td>
<td>Imprinting factor of 1.42–2.09</td>
</tr>
<tr>
<td>Creatinine [34]</td>
<td>Superior selectivity in comparison to the polymer prepared using traditional approach with a detection limit of 25 M</td>
</tr>
<tr>
<td>Microcystin-LR [35]</td>
<td>$K_i$ of MIP is 0.3 nM (polyclonal antibody 0.5 nM)</td>
</tr>
<tr>
<td>Ochratoxin-A [42,43]</td>
<td>75% Recovery of ochratoxin using MIP SPE. High affinity of NIP SPE with detection limit of 1 ng of mycotoxins</td>
</tr>
<tr>
<td>Afflatoxin B1 [43]</td>
<td></td>
</tr>
<tr>
<td>Abacavir [44]</td>
<td>High binding capacity, up to 157 mg of drug/g of absorbent, suitable for industrial applications</td>
</tr>
<tr>
<td>Cocaine [45], Deoxyephedrine [45], Methadone [45], Morphone [45]</td>
<td>Imprinted factors of 1.2 (cocaine), 2.5 (deoxyephedrine), 3.5 (methadone) and 3 (morphone)</td>
</tr>
<tr>
<td>Triazines [46,47]</td>
<td>Correlation between monomer-template binding energy with experimental binding using NIP SPE. The high affinity polymer released ~2% and low affinity polymer released ~27% of the template over 25 days.</td>
</tr>
<tr>
<td>Tylosin [48]</td>
<td>The polymer capacity for tylosin was estimated as 6.4 mg/g for MIP, which was suitable for practical applications</td>
</tr>
<tr>
<td>Curcumin [49]</td>
<td>Comparison of the batch analysis of the MIP- and NIP-grafted nanoparticles shows superior MIP binding to curcumin (μg per g particles)</td>
</tr>
<tr>
<td>Kukoamine A [50]</td>
<td>Kukamine can be purified (90%) from potato extract using MIP</td>
</tr>
<tr>
<td>Artemisinin [50]</td>
<td>Quantitative recovery of artemisinin (87%)</td>
</tr>
<tr>
<td>N-acyl-homoserine lactones [52]</td>
<td>Computationally designed polymers could sequester a signal molecule of V. fischeri bacteria</td>
</tr>
</tbody>
</table>

References
