Determination of the Structures of Molecularly Imprinted Polymers and Xerogels Using an Automated Stochastic Approach

ARTICLE *in* ANALYTICAL CHEMISTRY · AUGUST 2013

Impact Factor: 5.64 · DOI: 10.1021/ac402004z · Source: PubMed

8 AUTHORS, INCLUDING:

Elmer-Rico Mojica
Pace University
58 PUBLICATIONS **186** CITATIONS
SEE PROFILE

Diana S Aga
University at Buffalo, The State University of N…
125 PUBLICATIONS **3,643** CITATIONS
SEE PROFILE

analytical chemistry

Article

pubs.acs.org/ac

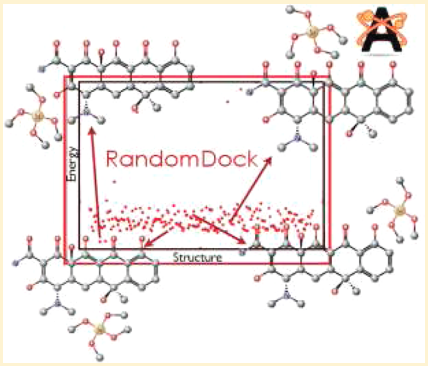
Determination of the Structures of Molecularly Imprinted Polymers and Xerogels Using an Automated Stochastic Approach

Amanda Wach, Jiechen Chen, Zackary Falls, David Lonie, Elmer-Rico Mojica,† Diana Aga, Jochen Autschbach,* and Eva Zurek*

Department of Chemistry, University at Buffalo, State University of New York, Buffalo, New York 14260-3000, United States

Supporting Information

ABSTRACT: An automated stochastic docking program with a graphical user interface, RANDOMDOCK (RD), has been developed to aid the development of molecularly imprinted polymers and xerogels. RD supports computations with ab initio and semiempirical quantum chemistry programs. The RD algorithms have been tested by searching for the most stable geometries of a varying number of methacrylic acid molecules interacting with nicotinamide. The optimal structures found are either as stable or more stable than those previously proposed for this molecularly imprinted polymer, illustrating that RD is capable of identifying the lowest-energy structures out of a potentially vast number of possible configurations. RD was subsequently applied to determine the most favorable binding sites between silane molecules and tetracycline (TC) as well as TC analogues. Hydrogen bonding between the templates and a silane is an important determinant of stability. Dispersion interactions are also sizable, sometimes dominant, especially between the largest silane and TC analogues not possessing a site readily available for hydrogen bonding. We highlight the importance of exploring the full intermolecular potential energy landscape when studying systems which may not afford highly specific interactions.



The dramatic rise in the number of papers focusing on molecular imprinting of polymers (MIPs) in the last two decades highlights the importance of this rapidly burgeoning field. The interest in MIPs[1−3] is derived from their broad range of applications, including chemical sensing,[4−6] efficient chromatographic separations, and solid-phase extractions, among many others.[7,8] Molecular-imprinted xerogels (MIXs) are based on sol−gels as opposed to organic acrylate or acrylic type polymers. MIXs may have superior properties to MIPs, such as enhanced selectivity and improved mass transport.[9−11] Finding the best monomer or combinations of monomers for the imprinted material is crucial. Many factors affect the binding strength and affinity of the MIP or MIX (MIPX) to the template.[12−14] Functional groups on the template and monomer can be more important than the shape of the template.[15] Experimental trial−and−error techniques to identify the most promising candidate among a large number of monomers may be time-consuming and inefficient.[16] Combinatorial chemistry has previously been used to find MIPX combinations.[17]

Computational methods may augment or replace experimental trial and error searches. The past two decades have witnessed a spectacular development of hardware as well as computational chemistry algorithms. It is now possible to test a library of promising MIPX constituents with arbitrary templates in silico. A review has been provided by Nicholls et al.[18] Different computational methods such as semiempirical[19] or ab initio calculations[20] have been used. Fundamental computational studies or development of models on the process of MIPX formation have been reported.[21−23] Molecular dynamics[24−26] and Monte Carlo simulations[21,23] have been carried out to gain a better understanding of the molecular imprinting process and microsctructure of a MIPX.

Many studies employ a scoring function based on the interaction energies of one or more monomers and the template molecule in the prepolymerization step.[21,27−30] A large binding energy between the monomer and template may be indicative of specific interactions between the two, resulting in the formation of an effective MIPX. In-silico modeling of kinetic quantities or structural changes of the monomer−template complex upon polymerization would require a significantly larger computational effort.

Even when only the interaction energy is of interest, there is potentially a very large conformational space to explore. In representative studies,[31−33] a knowledge-based guess was made as to where, spatially, the monomer would bind to the template. Maximizing the number of strong interactions such as hydrogen bonding can be a useful guiding principle. However, as the number of monomers in the computations grows it is less obvious which structures are going to be the most stable. Furthermore, in cases where the binding between the template and monomer is less specific[8] many different conformations

Received: April 12, 2013
Accepted: August 8, 2013

   dx.doi.org/10.1021/ac402004z | *Anal. Chem.* XXXX, XXX, XXX−XXX

may have very similar binding energies. Consider a recent experimental−computational study by Mojica et al.[32,34] where Hartree−Fock calculations and density functional theory (DFT) with small basis sets were utilized for the ranking of MIX monomers using tetracycline (TC) and TC analogues for templating. The computational component employed a single monomer−template complex for each monomer type and a few structural candidates. The study was also plagued by weak and unspecific interactions between TC and the silanes. As a result, the utility of the computations was limited.

Herein, we demonstrate a stochastic search method, RANDOMDOCK (RD), which broadly explores the potential energy landscape of a given set of molecules. Previous experimental findings of Mojica et al.,[32,34] hitherto unexplained, are rationalized by the present calculations. Further, our results illustrate that good MIPXs strongly benefit from specific interactions such as hydrogen bonding. Chemical intuition breaks down when nonspecific intermolecular interactions such as dispersion dominate or compete with specific interactions, or when the conformational search space is large. In these cases it is highly beneficial to utilize an automated approach.

The RD program has been interfaced with a number of molecular quantum chemistry packages in order to allow the MIPX search to employ a variety of semiempirical and ab initio quantum chemical methods as well as molecular mechanics for structural optimizations and energy evaluations. In the present work, dispersion-corrected DFT is used for the structure searches. The graphical user interface (GUI) renders RD very easy and efficient to use. The program is open source and can be downloaded for free from the Internet. The monomers used in this study are frequently used functional monomers for creating molecularly imprinted polymers and xerogels. It is reasonable to expect that the monomers resulting with the lowest binding energy based in RD will facilitate the selection of a particularly promising combination of functional monomer and other templates that will display a high imprinting factor.

COMPUTATIONAL DETAILS

The computational approach proposed herein is particularly useful for systems that exhibit nonspecific interactions, or many competing interactions. Full details are provided in the Supporting Information (SI). This manuscript and the SI are intended to serve as a short manual for the program and to help researchers in developing suitable computational protocols. The AVOGADRO[35] builder was used to construct and preoptimize the molecules considered herein, as well as to run the RANDOMDOCK (RD) extension. RD generates a user−specified number of geometries which are sent out for optimization by an external quantum chemical program. The workflow is typically as follows: The substrate (template) and matrix (monomer) molecules are drawn and pre−optimized by AVOGADRO or imported from a file using one of the common chemical structure formats. The number of matrix and substrate molecules are specified in the GUI. Either isolated molecules or complexes (dimers, trimers, etc.) may be employed. Conformers of the substrate and/or matrix molecules are then generated and pre−optimized using one of the force field implementations in the OPENBABEL library.[36,37] Either an exhaustive search can be performed, or a specified number of conformers may be created. The latter option does not guarantee that the most stable conformers will be found, since the conformer-generating algorithm stops once the specified number has been reached.

Conformers of the matrix and/or substrate molecules are then chosen randomly with a probablilistic criterion reflecting the stability based on the force-field energy. The matrix molecules are placed randomly about the substrate subject to user−specified distance constraints. A range for the interatomic distances (IADs) between the matrix and substrate atoms may be set. The IADs are measured between all atoms in the matrix/template complex. This is done in a manner that ensures that all atoms between any template-matrix pair are separated by at least the minimum specified distance, but at least one pair (per matrix molecule) is closer than the maximum distance. An option to force the formation of hydrogen bonds ensures that each configuration has at least one short X−H−X contact between the matrix and the substrate, where X is F, O, or N. Additional options and the settings used for the conformer generation are described in the SI. Currently, ADF,[38] GAMESS,[39] and MOPAC[40] are supported as the back-end quantum chemistry engines, providing a broad range of semiempirical, DFT, and correlated wave function options.

We used the "Build Cluster" and "Force hydrogen bonds" options. The MMFF94 force−field was employed for pre−optimizing geometries and calculating the conformer energies. Two hundred structures were created in each RD run. Full DFT geometry optimizations with good-quality basis sets were not performed on these initial structures, due in part to the computational cost involved and also because a great number of the randomly generated configurations are not particularly stable. Thus, the initial geometries were loosely optimized with ADF,[41] using single-ζ Slater-type basis sets (SZ) with 1s frozen for N, C, and O, as well as 1s−2p for Si. The revPBE density functional[42] was employed, along with Grimme's semiempirical correction for dispersion[43] (DFT-D3). The twenty most stable structures obtained from each RD run were further optimized to obtain a more accurate energy ordering using a valence triple-ζ Slater-type basis set with polarization functions for all atoms (TZP). The integration accuracy and convergence criteria were increased to values stricter than the ADF defaults. Additional structures were also generated manually by taking the most stable configurations from the high-quality optimizations and enforcing symmetry constraints, or by exploring the nearby configuration space via slight geometrical adjustments. The interaction energy, ΔE, was calculated from

$$\Delta E = E^{\text{complex}} - E^{\text{template}} - \sum_i E_i^{\text{monomer}} \quad (1)$$

using the energies of the most stable TC and silane conformers. The basis set superposition error (BSSE) corrected interaction energy, ΔE^{cp}, was calculated using the Counterpoise method. The energies and Cartesian coordinates of the most stable structures are provided in the SI.

RESULTS AND DISCUSSION

Benchmarks: Nicotinamide−Methacrylic Acid Complexes. In order to test the ability of RANDOMDOCK to find the global energy minimum of a MIP and to develop a suitable computational protocol the methacrylic acid/nicotinamide (MAA/NAM) system was chosen. NAM is a vitamin and an important food additive, and it has many other uses, for example in cancer therapy and in health related applications.[44] MAA binds to NAM via specific hydrogen−bonding interactions. The intermolecular bonding between these molecules was recently investigated with B3LYP/TZVP//RI-BLYP/SV(P) computations.[45] The interaction energy of the

dx.doi.org/10.1021/ac402004z | Anal. Chem. XXXX, XXX, XXX−XXX

most stable structure, illustrated in Figure 1a, was calculated as −14.7 kcal/mol. BSSE corrections only slightly decreased the magnitude to −14.1 kcal/mol. Structures were proposed for complexes containing up to four MAA molecules, as shown in Figure 1.

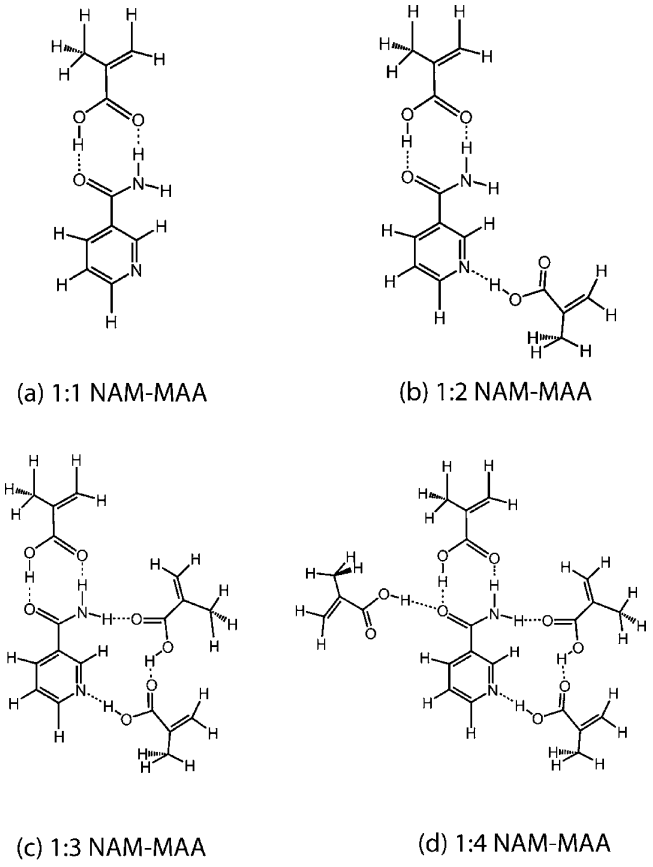


Figure 1. Most stable nicotinamide (NAM) and methacrylic acid (MAA) prepolymerization complexes, for varying ratios of NAM:-MAA, as found in ref 45.

Because of the potential importance of dispersion forces, we carried out geometry optimizations of the complexes previously proposed using DFT-D3, as described in the Computational Details. The interaction energies, see Table 1, are 10−23% larger than those calculated in ref 45, likely a result of the attractive intermolecular dispersion forces which are not captured by the B3LYP functional. Indeed calculations carried out with a correlated wave function-based method (RI-MP2/TZVPP) suggested that the binding energies should be about 2 kcal/mol stronger per hydrogen bond than the B3LYP results, yielding an interaction energy of −15.9 and −29.0 kcal/mol for the 1:1 and 1:2 species.[45] The MP2 energies are in very good agreement with the present DFT-D3 results.

Table 1. BSSE Uncorrected Interaction Energies (kcal/mol) for the Complexes in Figure 1

system	B3LYP[a]	DFT-D3[b]
1:1 NAM-MAA	−14.7	−16.2
1:2 NAM-MAA	−25.9	−30.0
1:3 NAM-MAA	−38.5	−45.9
1:4 NAM-MAA	−46.2	−56.8

[a]B3LYP/TZVP//RI-BLYP/SV(P) data from ref 45. [b]Present work, dispersion-corrected revPBE functional. The interaction energies have been calculated based on the NAM and MAA conformers illustrated in Figure 4a and c in the SI, in order to aid comparison with ref 45.

Twenty structural candidates from an RD search were further optimized using a larger basis set and higher integration accuracy as described in the Computational Details. Small deviations from planar arrangements due to finite convergence criteria of the optimizer were corrected by symmetrization. The lowest-energy structures emerging from these calculations are shown in Figure 2. The most stable 1:1, 1:2, and 1:3 complexes located by RD afforded a double hydrogen bond between the amide and carboxyl groups of NAM and the hydroxyl and carboxyl groups of one of the MAA molecules, and are similar but not identical to those previously suggested in ref 45. With the exception of NAM−$(MAA)_4$, the structures illustrated in Figure 1 were indeed found by the RD algorithm, but they were not the most stable. The structure in Figure 1d was likely not located among the twenty lowest in energy since its energy is rather high. The interaction energies for the systems illustrated in Figure 2 are collected in Table 2, where we also compare their energies to those calculated for the structures in Figure 1.

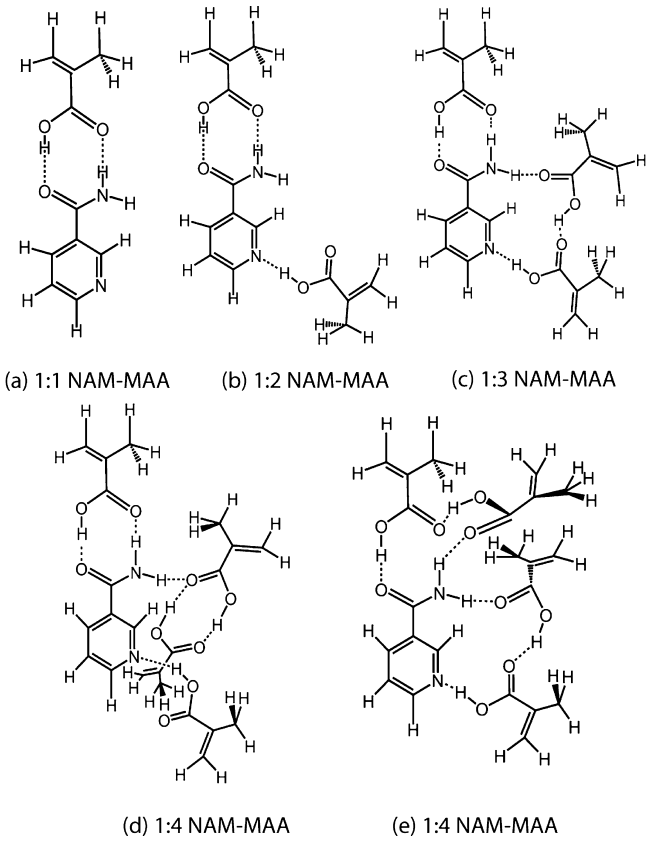


Figure 2. Most stable NAM−MAA MIP systems found with RANDOMDOCK and DFT-D3.

The main difference between the complexes presented here and the previous set of structures is that RD finds the MAA molecule is in a more stable *entgegen* (*E*) conformation illustrated in Figure 1b of the SI (note that we use an *E*/*Z* nomenclature here to label planar rotamers with respect to rotation about the C—C single bond connecting the C=O and C=C bonds, and not to indicate the configuration around the C=C double bond). In NAM−$(MAA)_2$, the hydroxyl group of the second MAA binds to the nitrogen atom in the NAM ring,

dx.doi.org/10.1021/ac402004z | *Anal. Chem.* XXXX, XXX, XXX−XXX

Table 2. Relative Energies of the Most Stable Structures Found Using RANDOMDOCK (Figure 2) and Those Presented Previously in Reference 45 (Figure 1)[a]

MIP	$E_{\text{Figure 2}} - E_{\text{Figure 1}}$	ΔE	ΔE^{cp}
1:1 NAM-MAA	−0.2	−16.1	−15.3
1:2 NAM-MAA	−0.2	−30.1	−28.6
1:3 NAM-MAA	−1.4	−47.1	−44.5
1:4 NAM-MAA	−5.3	−60.6	−56.8
1:4 NAM-MAA	−3.0	−58.6	−54.9

[a]DFT-D3 interaction energies without (ΔE) and with BSSE correction (ΔE^{cp}) for the NAM−MAA complexes illustrated in Figure 2. The interaction energies are given with respect to the NAM and MAA conformers illustrated in Figure 4b and c of the SI. All of the energies are given in units of kilocalories per mole.

but this MAA possesses the *Z* conformation, which is somewhat higher in energy. Rotating about the C−C bond so that the *E* conformation is present instead decreases the energy of the system by a further 0.5 kcal/mol. In order to find this geometry via RD, it would have been necessary to increase the number of structures which are optimized by the first−principles program. In the 1:3 complex, the third MAA molecule hydrogen-bonds via a hydroxyl to the carboxyl group of the second MAA, and to the amine via a carboxyl. All three of the MAA molecules adopt the *E* conformation. The second column in Table 2 shows that the systems located by RD were all lower in energy than those previously found in ref 45.

In our initial exploration of the potential energy landscape of NAM−$(MAA)_4$, most of the optimized structures contained hydrogen−bonded dimers of MAA which remained separated from the rest of the cluster. Another run was carried out using the most stable NAM−$(MAA)_3$ from Figure 2 as the template and a single MAA as the monomer. The energies of the two most favorable structures found in this fashion were substantially lower than the ones obtained in the previous run, and they were at least 3 kcal/mol more stable than the structures proposed in ref 45. Due to the larger configuration space, it would have been necessary to create significantly more structures within an unbiased RD run consisting of one template and four monomers in order to find these more stable geometries. By biasing the search toward structures which contained the preferred motifs from the NAM−$(MAA)_3$ cluster, the procedure locates low-energy geometries rather rapidly. The most stable configuration, shown in Figure 2d, resembles the 1:2 cluster, except that the MAA molecule which is bonded to the nitrogen atom in the NAM ring is rotated 180° about the hydrogen bond. This rotation makes space for an MAA dimer, which hydrogen bonds to the amide group of NAM, forming a structure which is no longer planar. Interestingly, one of the MAA molecules in this dimer is not bound to NAM. This cluster is more than 5 kcal/mol lower in energy than the one in Figure 1d, illustrating that energy gained from a hydrogen bond between two MAA molecules is more important than the second hydrogen bond between the oxygen atom in NAM and a hydroxyl in MAA. This suggests that the monomers may bind more strongly to each other than they do to the template. The second most stable structure in Figure 2e resembles the 1:3 configuration, except that the fourth MAA hydrogen bonds to the amide group in NAM via a carboxyl, as well as to another MAA. The two NAM−$(MAA)_4$ clusters found in this work, as well as the one proposed in ref 45, all contained six hydrogen bonds.

Our initial tests illustrated that the stochastic approach proposed herein is able to locate clusters which are either as stable, or are more stable, than ones which have been determined via chemical intuition. We therefore proceeded to apply RD to systems where it is difficult to determine the most favorable intermolecular interactions a priori.

Xerogels Imprinted by Tetracycline and Its Analogues. *Tetracycline Templates and Silane Monomers.* Because tetracyclines (TCs) are commonly employed in antibiotics and are additives in animal feeds, residues are often found in the environment, rendering TCs important templates in MIXs.[46] Tetracycline itself is an unstable molecule and readily transforms to its anhydro and epimer forms. At a pH < 2, TC becomes anhydrotetracycline (ATC),[47] and at pH 2−6, TC metamorphoses to 4-epitetracycline (ETC) and 4-epianhydrotetracycline (EATC).[48] Due to the acidic conditions experienced by TC during the polymerization of the xerogel, it has been proposed that imprinted xerogels will not be exclusively selective toward TC, but may also be selective toward ATC, EATC, and ETC as well.[32] For this reason we considered each TC analog, the most stable conformers of which are shown in Figure 3a−d. Since TC and its analogues

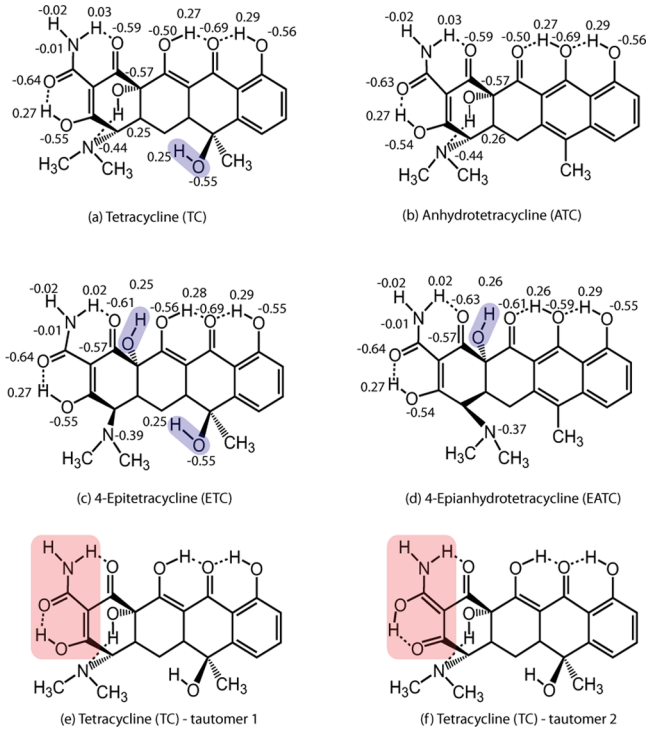
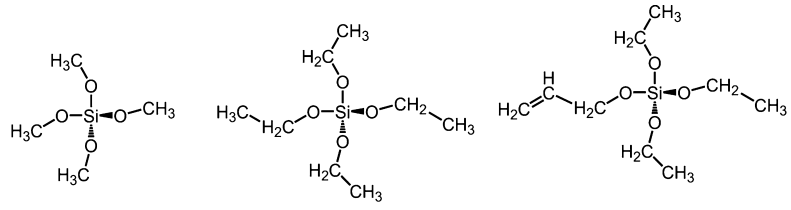


Figure 3. (a−d) Most stable conformers of the TC analogs have the largest number of intramolecular hydrogen bonds. The hydroxyl units highlighted in blue do not form intramolecular hydrogen bonds. Mulliken charges are provided for select atoms. (e and f) Two TC tautomers considered in this work. The area within the molecule which undergoes tautomerization is highlighted.

have the ability to undergo tautomerization, we included both tautomers in our analysis as well. The second tautomer considered, see Figure 3e and f, was the most stable one for TC and ATC, but only by ∼0.3 kcal/mol. For ETC the two tautomers were isoenergetic, and in the case of EATC the first tautomer was 0.3 kcal/mol lower in energy than the second. Silanes are widely used for MIXs due to their ability to form sol−gels.[49] Sol−gels are optically transparent, chemically inert,

dx.doi.org/10.1021/ac402004z | Anal. Chem. XXXX, XXX, XXX−XXX



(a) Tetramethoxysilane (TMOS) (b) Tetraethoxysilane (TEOS) (c) Allyltriethoxysilane (ATriEOS)

Figure 4. Silane structures.

rigid, and porous, making them ideal for molecular imprinting.[10,50] The silanes considered in this work, tetramethoxysilane (TMOS), tetraethoxysilane (TEOS), and allyltriethoxysilane-(AtriEOS), are illustrated in Figure 4. AtriEOS was chosen because of possible $\pi-\pi$ and hydrophobic interactions with the target template, and a hydrogen bond that may form between the hydrolyzable moeity ($-OCH_3$) of the silanes and the target template. The unhydrolyzed silanes were used in order to examine the prepolymerization interactions, which could provide insight on the possible performance of the resulting MIX materials. Furthermore, this allows to verify or refute previous computational work on the same systems and provide an explanation of the experimentally determined MIX performances.[32,34]

Interaction Energies. After carrying out the computational protocol described in the Computational Details, we carefully analyzed the structures which were within 5 kcal/mol of the most stable geometry. Duplicate structures, ie. those displaying similar bonding motifs and total energies, were identified and eliminated from further analysis. A number of duplicates with slightly differing geometries were found in each run, indicative of shallow intermolecular potential energy surfaces. The conformation of the silane had a minimal effect on the overall energy of the system. Thus, if two structures displayed similar bonding motifs but one had a different conformation of the silane that raised the overall energy slightly, the latter was not subjected to further study. We then proceeded to analyze the geometries, as well as the intra- and intermolecular bonding present in the most stable unique structures. An RD run was also carried out on the second TC tautomer (TC2) with TMOS, and the most stable structures displayed the same bonding motifs as those found for TC1-TMOS. We therefore assumed that the preferred binding locations of the monomer are comparable for both sets of tautomers in each TC analogue. Thus, the initial geometries for the second set of tautomers were constructed manually using the most stable structures found for the first set. This approach avoided a full new set of stochastic searches with the larger silanes TEOS and ATriEOS.

Tables 3–5 provide the interaction energies for the unique TC1–silane systems which were within 5 kcal/mol of the structure with the lowest energy. The corresponding interaction energies for the TC2–silane structures are also given, as well as the relative energies of the systems containing the two different tautomers. The BSSE corrections are much smaller than those obtained in the previous works on MIXs with TC analogues,[32,34] due to the good quality basis sets employed herein. For TMOS, two unique structures were found to lie within 5 kcal/mol of each other for each TC analogue, and in some cases, the energy difference between them was within the error bars of the calculations (i.e. in ATC1-TMOS and EATC2-TMOS). The difference in energy between the two tautomers

Table 3. Interaction Energies without (ΔE) and with BSSE Corrections (ΔE^{cp}) of the Most Stable MIX Systems Consisting of a Given TC Analogue and One TMOS Molecule[a]

	tautomer 1		tautomer 2		
MIX	ΔE	ΔE^{cp}	ΔE	ΔE^{cp}	E_{rel}
TC–TMOS	−15.5	−13.6	−15.3	−13.6	0.1
	−12.4	−10.7	−13.5	−11.8	1.4
ATC–TMOS	−11.8	−10.3	−11.2	−9.8	−0.3
	−11.8	−10.2	−14.4	−12.5	3.0
ETC–TMOS	−13.6	−11.8	−13.4	−11.7	0.0
	−11.9	−10.3	−11.5	−10.0	−1.5
EATC–TMOS	−11.8	−10.3	−14.4	−12.4	2.2
	−10.2	−8.3	−14.4	−12.4	3.9

[a] E_{rel} is the relative energy difference (BSSE corrected) between the systems containing two different TC tautomers. All energies are given in units of kilocalories per mole.

Table 4. Same as in Table 3, but for the Silane TEOS

	tautomer 1		tautomer 2		
MIX	ΔE	ΔE^{cp}	ΔE	ΔE^{cp}	E_{rel}
TC–TEOS	−16.4	−14.7	−16.4	−14.6	0.2
	−12.7	−11.1	−12.4	−10.7	−0.1
ATC–TEOS	−8.4	−7.2	−8.5	−7.3	0.4
	−8.3	−7.0	−10.7	−9.1	2.7
ETC–TEOS	−18.8	−16.7	−19.4	−17.2	0.3
	−17.0	−14.8	−19.3	−17.1	1.5
	−14.2	−12.4	−17.4	−15.2	4.5
EATC–TEOS	−10.4	−9.1	−10.3	−9.0	−0.4
	−9.7	−8.3	−9.7	−8.3	−0.3
	−8.7	−7.4	−8.6	−7.4	−0.4

(i.e. TC-TMOS, ATC-TMOS, ETC-TMOS) was found to be small. Our results indicate that even for this relatively simple silane, a MIX system may contain both sets of tautomers as well as multiple binding motifs with similar interaction energies, implying that the bonding may be unspecific. For TEOS the number of structures within 5 kcal/mol of the global minimum was larger in the case of ETC and EATC. For ATriEOS, even more low-energy structures were located. For these two more complex silanes, a number of unique but nearly isoenergetic structures were found including: two configurations of ATC1-TEOS, ETC2-TEOS, TC2-ATriEOS, ETC1-ATriEOS, ETC2-ATriEOS, as well as both sets of tautomers of each of the TC analogues and TEOS. In general the interaction energies of the silanes with the second set of tautomers were smaller in magnitude than with the first, but not by much.

Binding Sites. Careful analysis of the most stable geometries found in the structure searches (see the SI), combined with

dx.doi.org/10.1021/ac402004z | Anal. Chem. XXXX, XXX, XXX–XXX

Table 5. Same as in Table 3, but for the Silane ATriTEOS

	tautomer 1		tautomer 2		
MIX	ΔE	ΔE^{cp}	ΔE	ΔE^{cp}	E_{rel}
TC−ATriEOS	−13.3	−11.1	−10.6	−8.6	−2.4
	−10.6	−8.6	−10.0	−8.0	−0.3
	−10.1	−8.3	−9.1	−7.3	−0.7
	−8.6	−6.7	−7.6	−6.2	−0.7
ATC−ATriEOS	−15.9	−13.2	−14.9	−12.2	−0.7
	−13.4	−11.7	−12.9	−11.2	−0.2
	−13.4	−11.5	−12.5	−10.7	−0.5
ETC−ATriEOS	−15.9	−13.9	−16.7	−14.7	0.7
	−15.9	−13.5	−16.3	−13.9	0.4
	−11.1	−10.0	−11.6	−10.4	0.5
EATC−ATriEOS	−16.3	−13.7	−15.4	−13.0	−1.3
	−13.2	−11.6	−12.5	−10.9	−1.0
	−13.1	−11.8	−12.6	−11.3	−0.9
	−12.8	−11.1	−12.3	−10.6	−0.8
	−11.9	−10.9	−12.6	−11.5	0.3
	−11.7	−10.2	−11.0	−10.3	−1.0

additional calculations designed to provide some insights for the reasons why a specific configuration was found, revealed a number of trends. A full description of the analysis and our findings is provided in the SI. In a nutshell:

1. Any possible *intramolecular* hydrogen bonding within TC or a TC analogue is maximized.
2. If possible an intermolecular hydrogen bond between the silane and template may also be formed. However, see point 3. Some of the intermolecular hydrogen bonds classify as "dispersion-assisted" per the findings noted in the paragraph following point 4.
3. The strength of the dispersion interactions is another key player. There is an intricate balance between intra- and intermolecular hydrogen bonding, as well as dispersion, often resulting in a number of possible configurations that are within a few kilocalories per mole of each other.
4. Since there are multiple potential bonding locations in these MIX systems, the formation of complexes between the template and more than one silane are likely to arise when an excess of the monomer is available.

In order to quantify the important influence of dispersion on the geometries and the resulting binding energies, the most stable structures were reoptimized without dispersion. In every instance, the intermolecular distances increased significantly and the binding energies approached zero. Even in situations where, based upon distance and charge criteria, there appeared to be a hydrogen bond between the silane and the TC analogue, the optimization without dispersion resulted in a drastic lengthening of this contact. Thus, dispersion interactions are needed to support any intermolecular hydrogen bonding. Similar results have been obtained in our studies of 4-fluorostyrene.[51]

Most Favorable MIX Systems. Assuming that the imprinting factor, IF, is correlated with the strength of the binding interaction between the TC analogue and the silane, one can use the calculations to determine which combination of template and silane will yield the highest IF. We have taken an average of the BSSE corrected interaction energies calculated for the two tautomers; see Table 6. The magnitude of the interaction energy is the largest for ETC-TEOS. As pointed out in the binding site analysis (see the SI), some

Table 6. BSSE Corrected Interaction Energies (ΔE^{cp}; kcal/mol) of the Most Stable Template−Silane Structures, Averaged for the Two Tautomers of Each System[a]

	TMOS	TEOS	ATriEOS
TC	−13.6 (H)	−14.6 (H)	−9.9 (H)
ATC	−11.4 (D)	−7.3 (D)	−12.7 (D)
ETC	−11.7 (H)	−17.0 (H)	−14.3 (H,D)
EATC	−11.4 (D)	−9.7 (H)	−13.3 (D)

[a]The dominant interaction type between the template and the silane for the most stable structures, where H and D denote hydrogen bonding and dispersion, respectively, is also provided. In order to determine the "dominant interaction type", we have considered the bonding motif present in the most stable structures for each tautomer. If two structures were nearly isoenergetic, then both of their tautomers were considered (i.e., we looked at the two most stable ETC1−ATriEOS and ETC2−ATriEOS structure; see Table5).

dispersion-assisted hydrogen bonding is involved in the structure. Except for the largest silane ATriEOS, systems which exhibit intermolecular hydrogen bonding have more favorable interaction energies. TC and ETC give the strongest binding with TEOS as a result of hydrogen bonding. Conversely, ATC and EATC bind most strongly to ATriEOS because of the greater dispersion interactions which are possible with the bigger silane. Even though TMOS does not exhibit the strongest binding to any of the templates, the interaction energy is relatively consistent, ranging from −11.3 to −13.6 kcal/mol. This may be desired in situations when the TC is expected to convert between analogs.

In a previous study by Mojica et al.[32] designed to rationalize experimental trends from ref 34 the magnitude of the calculated interaction energy was generally found to be ETC > EATC > ATC > TC, and for ATriEOS EATC had a larger interaction energy than ETC. We find that ETC indeed often gives the strongest binding, with the exception of TMOS which prefers to bind to TC. We must note again that Mojica et al. used structures that were optimized with DFT, but not including dispersion and with rather small basis sets. More generally, however, the conclusions from ref 32 remain valid regarding the following points: Both experiment and computations show that a combination of TC, ETC, ATC, and EATC gets imprinted and that different silanes interact differently with the four analogues. The likelihood of epimerization of TC to ETC followed by degradation in the experiments, and the overall rather unspecific nature of the template−silane interactions, rationalizes why neither calculations nor experiments were able to identify a strong rank order for the IFs for various types of silanes.[34] As an exception, the observed good performance of ATriEOS and TEOS noted in the experiments[34] may in fact be an indication that ETC was preferentially imprinted. Indeed, ion chromatography data suggested that ETC was more abundant than TC during the imprinting process.[32] Incidentally, for the TEOS and ATriEOS xerogels, the calculated IF is largest with ETC. Furthermore, the interactions for ATC and EATC are much less specific (dispersion dominated).

SUMMARY AND CONCLUSIONS

An open-source program has been developed which can be employed for stochastic searches of the most stable configuration of an ensemble of molecules. Constraints limiting the search space to reasonable structures, such as imposing intermolecular distance cutoffs, forcing hydrogen bonds, or

dx.doi.org/10.1021/ac402004z | Anal. Chem. XXXX, XXX, XXX−XXX

forming "clusters" can be enforced. The search may be limited to a single molecular configuration. Optionally, all conformers of a molecule may be generated. The conformations used to create the structures which are sent out for further structural optimization by a quantum chemical program are chosen probabilistically, based on their molecular mechanics energies. This stochastic structure search program and its accompanying GUI, collectively named RANDOMDOCK (RD), currently supports ADF, GAMESS, and MOPAC as external optimizers.

RD has been employed to search for the most stable configurations of complexes containing one nicotinamide and up to four methacrylic acid molecules to model interactions leading to a MIP. These systems afford very specific hydrogen bonding interactions between the template and the monomers. For the 1:1, 1:2, and 1:3 systems, the most stable structures had similar binding motifs to those proposed in previous literature. For the 1:4 ratio, however, the automated structure search located configurations which were more than 3 kcal/mol lower in energy.

RD was then used to explore the potential energy landscape of a system containing TC or a TC analogue and one xerogel silane molecule. This situation is more complex and potentially contains chemically nonintuitive binding sites because of the relatively weak hydrogen bonding between the template and the monomer, and also because these two molecules can interact via dispersion forces. In general we found that intramolecular hydrogen bonding within the tetracycline analogue was the strongest interaction present in these systems. Intermolecular hydrogen bonding between a sterically unhindered hydroxyl group in the template and an oxygen atom in the silane was somewhat less strong, followed by dispersion interactions. However, for ATC and EATC, which did not contain sites that allowed the molecule to readily form intermolecular hydrogen bonds, the best interaction was with the bulkiest silane considered (ATriEOS), due to the uniformly attractive nature of the dispersion interaction.

The results shed light onto the findings from two previous studies[32,34] which hinted at an abundance of ETC present in imprinting studies of TC, due to epimerization. The silanes ATriEOS and TEOS yielded the best performing MIXs in these studies. The present computational results show that for these xerogels there are hydrogen bonding interactions present in dimers of ATriEOS and ETC or TC, with ETC being significantly more strongly interacting than TC for ATriEOS and somewhat stronger for TEOS. The interactions for the other TC analogs are less strong than for ETC. Moreover, in general they are dominated by unspecific dispersion interactions which therefore is less likely to result in an MIX that performs well. Therefore, the computational data from the stochastic search fully support the experimental results of ATriEOS and TEOS having suitable imprinting factor for "tetracycline", as long as one also takes epimerization of TC to ETC into consideration. Further optimization of xerogel formulations would likely need to establish more specific interactions between the template and the MIX monomer components.

■ ASSOCIATED CONTENT

Supporting Information

Illustrations, Cartesian coordinates, and energies of the optimized structures. Further computational details. Screenshots and a brief overview of the RANDOMDOCK extension. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: jochena@buffalo.edu (J.A.); ezurek@buffalo.edu (E.Z.).

Present Address

[†]Department of Chemistry and Physical Sciences, Pace University, New York, NY 10038

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the NSF [DMR-1005413 (E.Z.) and CHE-0750321 (J.A., D.A.)] for financial support, and the Center for Computational Research (CCR) at SUNY Buffalo for computational support. We thank the NSF-REU program (CHE-0851700) for the summer fellowship support for Z.F., and the NSF-CSUMS program (grant number 0802994) for supporting A.W. and J.C. E.Z. thanks the Alfred P. Sloan Foundation for a Research Fellowship (2013-2015).

■ REFERENCES

(1) Alexander, C.; Andersson, H.; Andersson, L.; Ansell, R.; Kirsch, N.; Nicholls, I.; O'Mahony, J.; Whitcombe, M. *J. Mol. Recognit.* **2006**, *19*, 106−180.

(2) Haupt, K. *Anal. Chem.* **2003**, *75*, 376A−383A.

(3) Wulff, G. *Angew. Chem., Int. Ed.* **1995**, *34*, 1812−1832.

(4) Haupt, K.; Mosbach, K. *Chem. Rev.* **2000**, *100*, 2495−2504.

(5) Holthoff, E.; Bright, F. *Acc. Chem. Res.* **2007**, *40*, 756−767.

(6) Kriz, D.; Ramstrom, O.; Mosbach, K. *Anal. Chem.* **1997**, *69*, 345A−349A.

(7) Zhang, H.; Ye, L.; Mosbach, K. *J. Mol. Recognit.* **2006**, *19*, 248−259.

(8) Whitcombe, M.; Vulfson, E. *Adv. Mater.* **2001**, *13*, 467−478.

(9) Marx, S.; Liron, Z. *Chem. Mater.* **2001**, *13*, 3624−3630.

(10) Cummins, W.; Duggan, P.; McLoughlin, P. *Anal. Chim. Acta* **2005**, *542*, 52−60.

(11) Silva, R. G. C.; Augusto, F. *J. Chromatogr. A* **2006**, *1114*, 216−223.

(12) Farrington, K.; Regan, F. *Talanta* **2009**, *78*, 653−659.

(13) Piletsky, S. A.; Andersson, H. S.; Nicholls, I. A. *Macromolecules* **1999**, *32*, 633−636.

(14) Alexander, C.; Andersson, H.; Andersson, L.; Ansell, R.; Kirsch, N.; Nicholls, I.; O'Mahony, J.; Whitcombe, M. *Anal. Chim. Acta* **2001**, *435*, 9−18.

(15) Ren, K.; Zare, R. *ACS Nano* **2012**, *6*, 4314−4318.

(16) Diñeiro, Y.; Menéndez, M.; Blanco-López, M.; Lobo-Castañón, M.; Miranda-Ordieres, A.; Tuñón-Blanco, P. *Biosens. Bioelectron.* **2006**, *22*, 364−371.

(17) Takeuchi, T.; Fukuma, D.; Matsui, J. *Anal. Chem.* **1999**, *71*, 285−290.

(18) Nicholls, I. A.; Andersson, H. S.; Charlton, C.; Henschel, H.; Karlsson, B. C.; Karlsson, J. G.; O'Mahony, J.; Rosengren, A. M.; Rosengren, K. J.; Wikman, S. *Biosens. Bioelectron.* **2009**, *25*, 543−552.

(19) Meng, Z.; Yamazaki, T.; Sode, K. *Biosens. Bioelectron.* **2004**, *20*, 1068−1075.

(20) Dineiro, Y.; Menendez, M. I.; Blanco-Lopez, M. C.; Lobo-Castanon, M. J.; Miranda-Ordieres, A. J.; Tunon-Blanco, P. *Anal. Chem.* **2005**, *77*, 6741−6746.

(21) Herdes, C.; Sarkisov, L. *Langmuir* **2009**, *25*, 5352−5359.

(22) Sarkisov, L.; Van Tassel, P. R. *J. Phys. Chem. C* **2007**, *111*, 15726−15735.

(23) Dourado, E. M. A.; Sarkisov, L. *J. Chem. Phys.* **2009**, *130*, 214701 (1−9).

(24) Azenha, M.; Szefczyk, B.; Loureiro, D.; Kathirvel, P.; Cordeiro, M. N. D.; Fernando-Silva, A. *Langmuir* **2011**, *27*, 5062−5070.
(25) Chianella, I.; Lotierzo, M.; Piletsky, S.; Tothill, I.; Chen, B.; Karim, K.; Turner, A. *Anal. Chem.* **2002**, *74*, 1288−1293.
(26) Li, Y.; Li, X.; Li, Y.; Dong, C.; Jin, P.; Qi, J. *Biomaterials* **2009**, *30*, 3205−3211.
(27) Liu, Y.; Wang, F.; Tan, T.; Lei, M. *Anal. Chim. Acta* **2007**, *581*, 137−146.
(28) Wu, L.; Zhu, K.; Zhao, M.; Li, Y. *Anal. Chim. Acta* **2005**, *549*, 39−44.
(29) Wu, L.; Li, Y. *J. Mol. Recognit.* **2004**, *17*, 567−574.
(30) Wu, L.; Sun, B.; Li, Y.; Chang, W. *Analyst* **2003**, *128*, 944−949.
(31) Diñeiro, Y.; Menéndez, M.; Blanco-López, M.; Lobo-Castañón, M.; Miranda-Ordieres, A.; Tuñón-Blanco, P. *Anal. Chem.* **2005**, *77*, 6741−6746.
(32) Mojica, E.; Autschbach, J.; Bright, F.; Aga, D. *Analyst* **2011**, *136*, 749−755.
(33) Azenha, M.; Kathirvel, P.; Nogueira, P.; Fernando-Silva, A. *Biosens. Bioelectron.* **2008**, *23*, 1843−1849.
(34) Mojica, E.; Autschbach, J.; Bright, F.; Aga, D. *Anal. Chim. Acta* **2011**, *684*, 63−71.
(35) Hanwell, M.; Curtis, D.; Lonie, D.; Vandermeersch, T.; Zurek, E.; Hutchison, G. *J. Cheminformatics* **2012**, *4*, 1−17.
(36) *The Open Babel Package*, development version. http://openbabel.org (accessed Feb 2013).
(37) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchson, G. R. *J. Cheminformatics* **2011**, *3*, 1−14.
(38) Baerends, E. J. et al. *ADF2010.01*; SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, 2010; http://www.scm.com.
(39) Schmidt, M.; Baldridge, K.; Boatz, J.; Elbert, S.; Gordon, M.; Jensen, J.; Koseki, S.; Matsunaga, N.; Nguyen, K.; Su, S.; Windus, T.; Dupuis, M.; Montgomery, J. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(40) Stewart, J. J. P. *Molecular Orbital PACkage*; http://openmopac.net (accessed Feb 2013.)
(41) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931−967.
(42) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1998**, *80*, 891.
(43) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(44) Del Sole, R.; Scardino, A.; Lazzooi, M. R.; Vasapollo, G. *J. Appl. Polym. Sci.* **2011**, *120*, 1634−1641.
(45) Sole, R. D.; Lazzoi, M.; Arnone, M.; Sala, F.; Cannoletta, D.; Vasapollo, G. *Molecules* **2009**, *14*, 2632−2649.
(46) O'Connor, S.; Aga, D. *Trac-Trend. Anal. Chem.* **2007**, *26*, 456−465.
(47) Durckheimer, W. *Angew. Chem., Int. Ed.* **1975**, *14*, 721−734.
(48) McCormick, J.; Fox, S.; Smith, L.; Bitler, B.; Reichenthal, J.; Origoni, V.; Muller, W.; Winterbottom, R.; Doerschuk, A. *J. Am. Chem. Soc.* **1957**, *79*, 2849−2858.
(49) Collinson, M. *Crit. Rev. Anal. Chem.* **1999**, *29*, 289−311.
(50) Lee, S.; Lin, H.; Chen, H. *J. Appl. Polym. Sci.* **2009**, *114*, 3994−3999.
(51) Jewell, A.; Simpson, S.; Enders, A.; Zurek, E.; Sykes, C. *J. Phys. Chem. Lett.* **2012**, *3*, 2069−2075.