JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Ring Closure To Form Metal Chelates in 3D Fragment-Based de Novo Design

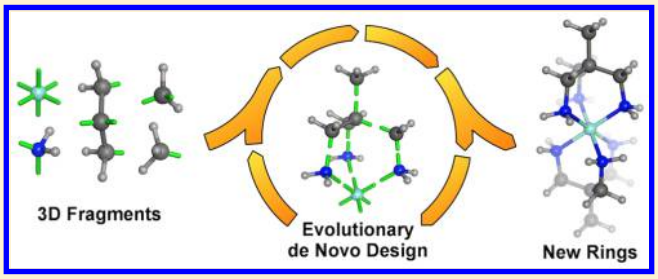
Marco Foscato,[†] Benjamin J. Houghton,[‡] Giovanni Occhipinti,[†] Robert J. Deeth,[‡] and Vidar R. Jensen*[,†]

[†]Department of Chemistry, University of Bergen, Allégaten 41, N-5007 Bergen, Norway

[‡]Inorganic Computational Chemistry Group, Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, Great Britain

Supporting Information

ABSTRACT: We describe a method for the design of multicyclic compounds from three-dimensional (3D) molecular fragments. The 3D building blocks are assembled in a controlled fashion, and closable chains of such fragments are identified. Next, the ring-closing conformations of such formally closable chains are identified, and the 3D model of a cyclic or multicyclic molecule is built. Embedding this method in an evolutionary algorithm results in a de novo design tool capable of altering the number and nature of cycles in species such as transition metal compounds with multidentate ligands in terms of, for example, ligand denticity, type and length of bridges, identity of bridgehead terms, and substitution pattern. An application of the method to the design of multidentate nitrogen-based ligands for Fe(II) spin-crossover (SCO) compounds is presented. The best candidates display multidentate skeletons new to the field of Fe(II) SCO yet resembling ligands deployed in other fields of chemistry, demonstrating the capability of the approach to explore structural variation and to suggest unexpected and realistic molecules, including structures with cycles not found in the building blocks.



INTRODUCTION

Cyclic arrangements of connected atoms (i.e., rings) are fundamental structural features in chemical compounds and have a special role in transition metal chemistry. A single metal atom may be bonded to two or more separate binding sites on the same ligand, thus forming a chelate.[1] Such complexes between metals and multidentate ligands are usually more stable than their acyclic counterparts involving monodentate ligands. This effect, known as the "chelate effect",[2−4] explains why multidentate ligands are so common in transition metal chemistry.

The properties of such chelate complexes depend on the structure of the multidentate ligands.[5] Rational design of these ligands may thus contribute to the development of stable compounds with desirable properties and is a worthwhile goal.[6,7] The designer of multidentate ligands may tune the nature of the binding sites, the ligand denticity, the nature and length of the bridges connecting the binding sites, and the functionalization of both bridges and pendant groups (i.e., nonbridge substituents on binding moieties).[8,9] These variations result in a vast number of possible multidentate ligands. The potential number of metal complexes bearing these ligands is even larger, as there may be other ancillary ligands on the same metal center. The picture is further complicated by the possibility that monodentate ligands self-assemble to give multidentate coordination.[10−12]

All of this variability means that the designer needs powerful methods to search for good ligands and complexes. To this end, some workers have projected portions of the vast structural space onto a smaller and chemically meaningful descriptor space, where ligands are represented by a set of comparable descriptors characterizing the whole ligand, or the ligand−metal pair, in terms of backbone flexibility, bite angle, steric bulk, and electronic contributions.[8,13] This approach has been used to map the properties of mono- and bidentate ligands,[14−18] enabling rational selection of existing ligands[19] and prediction of the properties of novel ligands.[13,20,21] Moreover, in combination with predictive models built from empirical knowledge collected by high-throughput experimental techniques, the descriptor space has been used to design new catalysts with bidentate ligands that were automatically generated and evaluated in silico.[22−24]

A kind of in silico method that, at least in principle, may evolve to become a particularly powerful design tool involves automated generation of new virtual candidates and is termed "de novo design" when the candidates are built from smaller molecular fragments. Such "from-scratch" methods have been developed for the design of organic drug molecules.[25] Some of these drug design methods allow for creation of new rings and thus could, in principle, also design new multidentate ligands for transition metals.[26−35] However, transition metal chemistry requires molecular modeling tools adapted for, e.g., unusual atom and bond types, which is beyond what methods used in

Received: July 4, 2015
Published: September 1, 2015

ACS Publications © 2015 American Chemical Society

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

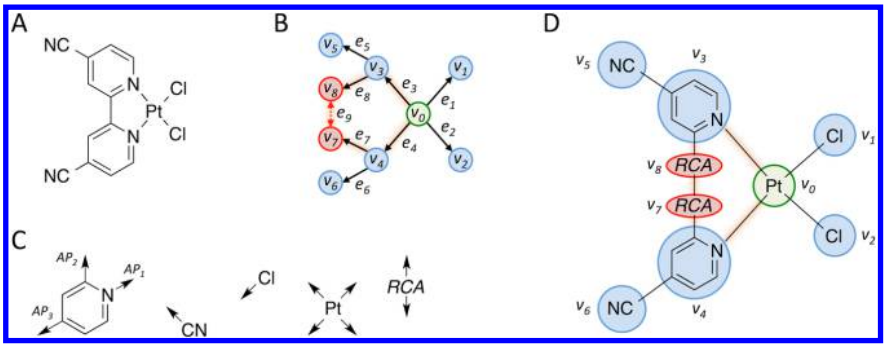


Figure 1. Graph representation of chemical species containing rings. A molecule (A) is represented by a graph (B) collecting fragments (C) in a spanning tree (vertices v_0 to v_8 and edges e_1 to e_8 (black arrows) in B) and the fundamental cycle (the highlighted ring of edges in B) defined by the chord v_7–v_8 (e_9, red arrow in B). The resulting overlap of the outer-layer graph with the molecular fragments (inner-layer graphs) is shown in (D). The root fragment is rendered in green, and the ring-closing vertices, each of which hosts a dummy ring-closing atom (RCA), are shown in red.

the design of organic drugs can handle.[36] This deficiency largely explains why drug design methods have thus far found little use in transition metal and organometallic chemistry.[37,38]

This is now changing, as recent years have seen the development of de novo design tools specifically for transition metal chemistry.[8,36,39] An example is HostDesigner, which was developed to design multidentate binding sites (i.e., hosts) for metal ions[39] and has been extended to support general host–guest design.[40−44] Binding sites preorganized for a particular guest entity are connected in an energetically favorable manner thanks to the method's identification of suitable linker units.[45] HostDesigner has been applied to the de novo design of multidentate ligands for metal complexes,[43,46−49] supramolecular coordination compounds,[50] and receptors for inorganic anions[51] and neutral organic molecules.[52] Nevertheless, HostDesigner focuses only on the bridges between given binding sites and does not perform modulation of coordinating groups and substituents, which can significantly affect functional properties other than the geometrical preorganization of the binding sites, e.g., electron donation. Moreover, the indiscriminate assembly of linkers and complex fragments leads to numerous synthetically inaccessible candidates.[52]

Another de novo method currently under development is aimed at the design of functional transition metal compounds in general.[36] In this method, a given criterion representing the property of interest is used as a fitness or scoring function to optimize (i.e., evolve) a population of molecular candidates according to the principles of natural evolution. Artificial evolution of the population is made possible by the use of a "genomelike" chemical representation, namely, a graph consisting of an ordered collection of molecular fragments. Each such fragment may be deleted, replaced, or decorated (i.e., further substituted), thereby providing the means for automated creation and modification of chemical species via genetic operators such as mutation and crossover.

This artificial evolution approach has thus far been applied to predict improved ruthenium catalysts for olefin metathesis[36] and organic dye molecules for solar cells.[53] Recent improvements of the underlying method include handles for controlling the synthetic accessibility of the automatically generated molecules[54] and the quality of the three-dimensional (3D) molecular models, even for transient species and transition states.[55] However, to date it has only been possible to treat rings as complete entities that must already be present from the start of the design process, i.e., they must be present in the fragments (the molecular building blocks), and no new rings can be formed or opened during the process. This has limited the method to decoration and combination of existing ring systems only and has hampered the identification of new multidentate ligands and functional metal chelates.

To overcome this limitation and to handle also closure, rupture, modification, and inheritance of rings during de novo design experiments, we have extended our 3D fragment-based method. The methodological extensions are described in Methods, while validation and applications of the new method can be found in the Results and Discussion.

METHODS

Graph Representation. In the following we describe extensions of the recently developed method for automated building of realistic 3D molecular models.[36,54,55] Molecular entities are represented by a graph $G(V, E)$, where $V = \{v_1, ..., v_n\}$ is the set of vertices, each of which contains a molecular fragment, and $E = \{e_1, ..., e_m\}$ is the set of edges connecting the vertices, which represent the bonding interactions between molecular fragments (Figure 1). The graph G is conveniently decomposed into a spanning tree $T(V, E')$, which spans all of the vertices of G and contains no cycle of vertices ($E' \subseteq E$), and the set of fundamental cycles $Fc = \{C_1, ..., C_k\}$ of G (naming and notation are as in ref 56).

Each *vertex* v_i contains a 3D molecular fragment of any size that carries one or more attachment points (APs), which define possible end points of incident edges. Each AP defines the position of a potential connection to another fragment (as the atom, or dummy atom, directly involved in the connection, i.e., the source of the AP), the type of the attachment point (also termed the "class", i.e., the AP class),[54] and the spatial direction of the possible connection (the AP vector).[55]

Each *edge* e_i connects two vertices and identifies a pair of end points, where each end point defines the identity of a vertex and one of its APs. All edges belonging to E' are directed, i.e., they define a source vertex and a target vertex. The edge e_j that connects vertices v_i and v_j is the only connection between these two vertices (multiple edges are not allowed in G), and moreover, $v_i \neq v_j$ (i.e., G has no self-loops). From a chemical point of view, e_j represents a formal bond between two molecular fragments to give a united entity. Therefore, an edge does not tell whether the bond is covalent, ionic, or of the coordination type, and bond order is treated as an optional property of e_j.[55]

Since each vertex of G is a molecular fragment that is represented by a graph, there are de facto two layers of graphs:

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

an outer layer, G, of which the vertices are molecular fragments of any size, and an inner layer of graphs, each representing the structure of a single fragment, in which the vertices represent atoms. Whereas each inner-layer graph is treated as an unmodifiable unit, the outer-layer graph (G) perceives each such inner-layer graph as a single vertex. Nevertheless, the inner-layer graph is clearly required to convert G into a molecular model, and the APs form the bridge between the two layers, as they define the relationship between connections of vertices (outer layer) and connections of molecular fragments (inner layer) (see Figure 1).

The *spanning tree* $T(V, E')$ is conventionally rooted on the starting (root) vertex of G and is a subgraph of G that spans all of the vertices and enough edges to make each vertex reachable from the root by a directed path (black arrows in Figure 1B). As a tree, T has no cycle of vertices. Therefore, if G has no cycle, then $G = T$. Although a spanning tree does not contain cycles of vertices, it can indeed describe cyclic molecules if rings are embedded within single vertices. This approach allows handling of immutable ring systems within a treelike graph.[36,54]

Each of the fundamental cycles of vertices collected in the set $Fc = \{C_1, ..., C_k\}$ corresponds to a chord of T that is an edge found in E but not in E' (i.e., chords belong to G but not to T).[56] Chords are made only between special-purpose vertices termed *ring-closing vertices* (RCVs), which are chemically "empty" fragments containing only a dummy atom (a *ring-closing atom* (RCA); see Figure 1).

Molecular Design. *Graph Creation and Modification.* New molecular entities are obtained by either (a) construction from scratch or (b) modification of existing graphs. In the framework of the evolutionary algorithm within which the present ring-closing procedure is embedded,[36] both (a) and (b) have been described previously for the particular case of tree-shaped molecules, for which the graph G is identical to the spanning tree T.[36] To summarize, the construction of a graph involves a three-step process: (i) definition of the root vertex of T, (ii) growth of the system by probability-controlled recursive attachment of properly chosen vertices to free attachment points, and (iii) saturation of attachment points that cannot remain unused (open valences, etc.) by specific capping vertices (typically, but not exclusively, H atoms and methyl groups). This process is governed by parameters that control the substitution probability and limit the size of the generated graph. In addition, each connection of vertices must respect a set of AP class compatibility rules to achieve synthetic accessibility; these rules are collected in a so-called compatibility matrix (see ref 54). On the other hand, modification of an existing graph occurs on T via either mutation or crossover.[36] The former consists of replacing a vertex v_i and all vertices reachable from v_i by a directed path in T by a different vertex or by no vertex (deletion). The crossover operation swaps selected branches between the tree graphs of two parent graphs. Both construction and modification of graphs are controlled by the AP class compatibility rules.[54] In the present method, these protocols are integrated with new functionalities to allow identification and handling of the set of fundamental cycles Fc. The molecular graph G is built from T using the definition of Fc.

Modification of a graph affects the underlying tree, and whenever a modification involves a vertex taking part in a fundamental cycle, the fundamental cycle and the corresponding chord are removed. The method evaluates the possibility of introducing one or more new chords to define one or more new fundamental cycles. To this end, during the construction and modification of T, specific vertices are used as placeholders for possible chord end points. These vertices, which are called *ring-closing vertices*, are regular vertices that contain special fragments including a single, two-AP dummy atom, termed a *ring-closing atom* (see Figure 1). Only one of the two APs is available during construction and modification of trees, while the other is reserved for potential chord formation. As for any other vertex, the location of RCVs in T depends on the definition of the AP class compatibility, the available fragment space (i.e., the combined ensemble of fragments and connection rules), and the defined probability of appending substituents at each stage of the building process (see Computational Details in the Supporting Information). A tree T might hold a list of RCVs, each of which identifies the position and type of one of the two end points of a candidate chord. Since the creation of a chord corresponds to the connection of two such end points, the combinations of pairs of RCVs define the set of possible chords of a given tree. Even a few RCVs give a multitude of candidate chords, only a fraction of which are sensible from a chemical point of view. A chord between two RCVs in the outer-layer graph G (v_7 and v_8 in Figure 1) corresponds to a chemical bond passing through the two RCAs that are contained in the RCVs, connecting the two molecular fragments of the vertices to which the RCVs are attached (e.g., v_3 and v_4 in Figure 1). Therefore, compatibility criteria (see below) are used to select only chemically reasonable pairs of RCVs. A given tree may still give rise to several combinations of compatible RCV pairs, each of which defines a different molecular entity (excluding entities related by symmetry). Thus, in order to obtain a graph that represents a single entity, one of the possible combinations is randomly chosen for each generated graph independently and used to define the set of chords and the fundamental cycles Fc. The graph resulting from the combination of T and Fc is then finalized by a capping procedure that switches unused RCAs into minimal molecular fragments (i.e., H atoms or methyl groups) to saturate free valences.[54]

The compatibility criteria for RCVs are (i) the AP class compatibility of the AP holding the RCVs (e.g., AP_2 of v_3 and AP_2 of v_4 in Figure 1), (ii) the RCA type compatibility, and (iii) the geometrical chain closability criteria. While the first two criteria are given by user-provided rules similar to those used for acyclic vertex−vertex connections,[54] chain closability requires the evaluation of the geometrical possibility of forming a bond between the two extremities of the atom chains holding the two RCAs, corresponding to the path connecting the RCVs in T. It should be noted that since fragments can contain rings, the atom path may not be unique, and thus, the shortest atom chain is chosen for the evaluation of closability[57] (see Computational Details).

The ring closability is a function of the 3D features of the system that requires the evaluation of conformations allowing sufficient overlap of the pair of atom-to-RCA bond vectors (see Computational Details for the threshold values used for head-to-tail interatomic distances and the cosine of the angle between the vectors). Our current implementation does not include adaptation of the bond angles along the evaluated chain of atoms, which is justified by the fact that the bond angles of the 3D fragments (i.e., the building blocks) already are expected to be reasonable. In contrast, in loop modeling for protein structure prediction, methods have been developed to handle ring closure that account for some degree of bond angle

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

and bond length deformation,[58−62] and similar techniques will be integrated in future versions of the present method.

Finally, as a further condition for path closability, two or more fundamental cycles may share one or more bonds and are therefore conformationally dependent. The simultaneous closability of such interdependent fundamental cycles is verified only if at least one conformation of the shared bonds allows simultaneous closure of all of the mutually dependent cycles.

Graph to 3D Model. The construction of a 3D molecular model from the graph representation exploits the rich geometrical information provided by the 3D fragments.[55] Connection of these molecular fragments is performed sequentially following the spanning tree T, thus at first ignoring the chords. Fragments are spatially arranged by using the geometrical information reported in each 3D fragment and its attachment points. This information suffices to define fragment−fragment bond lengths and angles but not the corresponding interfragment torsions. The result is a treelike molecule that does not comprise the rings defined by *Fc* and that contains RCAs in lieu of the chord end points.

Closure of all of the rings defined by the fundamental cycles is achieved by forming the fragment−fragment bonds corresponding to the chords of T. This is a challenging conformational problem that, in addition to each chain of bonds corresponding to the molecular projection of a fundamental cycle of vertices C_i, also involves the conformation of each decorating substituent on these chains, the conformation of every group in spatial proximity to these substituents, and all of the interdependent fundamental cycles of C_i. The solution to this problem corresponds to finding the best conformation that allows closure of all of the rings while avoiding unrealistic and too-strained geometries. Under the assumption that the rings can be closed with negligible bond angle deformation, this conformation is identified using a global optimization technique in the rotational space, implemented in the PSSROT program of the Tinker[63] package,[64−67] coupled with a *ring-closing potential* (RCP) specifically developed for this purpose, which is described in the next section. Running the RCP−PSSROT procedure on a treelike molecule bearing RCA pairs gives two possible outcomes: (i) a partially folded conformation that does not allow closure of all of the rings because of steric hindrance or (ii) a fully folded conformation in which the pairs of atoms holding compatible RCAs are sufficiently close and obey the threshold values of their bonding geometries (see Computational Details). In the latter case, the 3D molecular model is finalized by removing the RCAs and modifying the connectivity by introducing all of the new ring-closing bonds, i.e., those between the atoms hosting each of the two RCAs connected by a chord.

Ring-Closing Potential. The RCP is an empirical potential energy function designed to perceive ring closures as favorable events while at the same time including nonbonded interactions from the van der Waals (vdW) term of the universal force field (UFF).[68] The RCP is also designed to be compatible with the potential smoothing protocol implemented in Tinker.[66] The RCP does not include bond stretching and bending terms, as it is currently meant only for calculations in the rotational space defined by the list of rotatable bonds and fragment−fragment connections.[55] Moreover, since 1−4 repulsion is embedded in the vdW potential and the kinds of rotatable bonds considered here are mainly defined by 1−4 repulsive interactions (i.e., single, nonresonating bonds),[69] all of the bond torsion terms are also ignored in a first approximation.

The RCP is calculated as the sum of two energy terms (eq 1):

$$E_{RCP}(t) = E_{vdW}(t) + E_{rct}(t) \quad (1)$$

where $E_{vdW}(t)$ is a modified vdW term, $E_{rct}(t)$ is the ring-closing term, and t is the parameter that controls the potential smoothing procedure, with $t = 0$ for the unsmoothed potential.[66] $E_{vdW}(t)$ reflects the standard Tinker implementation of the vdW term for calculations with smoothed potentials, the only difference being the atom proximity reduction factors. Since ring closure occurs between atoms that are not connected (e.g., the source atoms, SA_A and SA_B, in Figure 2), the scaling

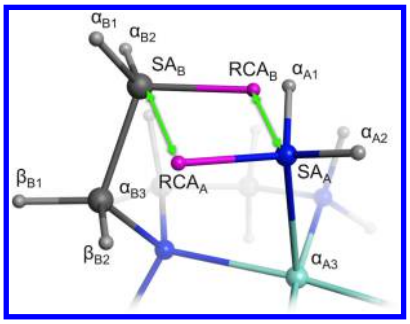


Figure 2. Ring-closing terms (green arrows) defined between the two ends of closable chains. Magenta spheres represent the dummy ring-closing atoms (RCAs).

factors intended for atoms in 1−2, 1−3, 1−4, and 1−5 relationships have to be applied as if the source atoms holding the RCAs were connected. For instance, in Figure 2, the 1−2 scaling factor should be used for the pairs SA_A−SA_B, RCA_A−α_{Bn}, and RCA_B−α_{An}, the 1−3 factors for SA_A−α_{Bn}, SA_B−α_{An}, and SA_A−α_{B1}, the 1−4 factors for SA_A−β_{Bn}, and so forth. In addition, a formal 1−1 relation (i.e., atom overlap) is used for each RCA-to-atom relation between the two extremes of the closing chains (RCA_A−SA_B and RCA_B−SA_A in Figure 2). These modifications remove some of the repulsive components of the vdW potential and enable the two ends of the closing chains to approach each other. These chain ends also experience an attractive force from the ring-closing term, $E_{rct}(t)$, involving only pairs of RCAs and SAs at the two closing chain extremities, as exemplified by RCA_A−SA_B and RCA_B−SA_A in Figure 2. RCA and SA are images of the same atom, and the role of the ring-closing term is to achieve overlap of these two images. Such overlap requires a geometrical arrangement inviting bond formation between SA_A and SA_B. To this end, the ring-closing term has been given a long-range attractive tail so that RCA_i and SA_j experience an attractive force even when placed at the extremes of stretched chains as well as a finite and negative value at zero distance. The ring-closing term consists of two contributions per fundamental cycle:

$$E_{rct}(t) = \sum_{C \in Fc} [f_{rct}(RCA_{A,C}, SA_{B,C}, t) + f_{rct}(RCA_{B,C}, SA_{A,C}, t)] \quad (2)$$

Each contribution consists of a combination of three Gaussian functions (identified by the index $k = 1, 2, 3$):

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

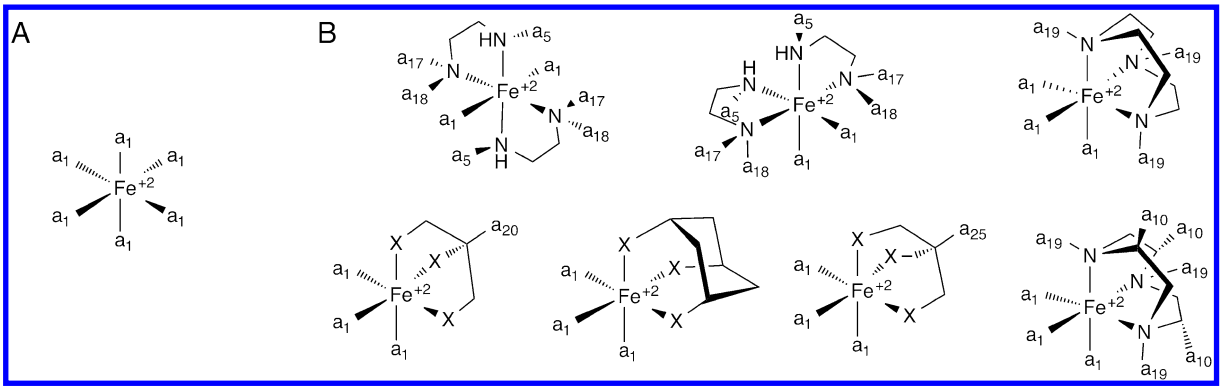


Figure 3. Library of 3D fragments (scaffolds) used to initiate the construction of Fe(II) complexes in experiments of types **A** and **B**. Attachment points are represented with dummy atoms using the attachment point class (a_n) as labels. For clarity, "X" is used in lieu of an sp^3 nitrogen atom with two attachment points, both of class "a_{19}".

$$f_{\mathrm{rct}}(\mathrm{RCA_i}, \mathrm{SA_j}, t) = \sum_{k=1}^{3} \frac{a_{\mathrm{ij},k}}{(1 + 4Dt)^{3/2}} \exp\left(-\frac{b_{\mathrm{ij},k} r_{\mathrm{ij}}^{\,2}}{1 + 4Dt}\right) \tag{3}$$

in which D is the constant regulating the rate of smoothing of the potential energy surface of the ring-closing term (see ref 66), r_{ij} is the distance between RCA_i and SA_j (green arrows in Figure 2), and

$$a_{ij,k} = a_k \alpha_{\mathrm{ij}} \tag{4}$$

$$b_{ij,k} = \frac{b_k}{\beta_{\mathrm{ij}}^{\,2}} \tag{5}$$

where a_k and b_k are constants defining each of the three Gaussian functions combined linearly in eq 3 and α_{ij} and β_{ij} are force field parameters defined by the types of RCA_i and RCA_j (or the geometric average of different types). The values of the parameters used in this study are given in Computational Details.

RESULTS AND DISCUSSION

The machinery described in the previous sections was applied to the design of mononuclear iron(II) complexes with multidentate nitrogen-based ligands. Such complexes are among the most extensively studied spin-crossover (SCO) compounds[70−72] and may, in the condensed phase, with cooperation between SCO centers,[73] provide materials with peculiar properties that can be exploited in display, sensing, and memory devices.[74−77] Depending on the ligand field surrounding the metal center, the complexes may exist in either the low-spin (LS, $S = 0$, strong ligand field) or high-spin (HS, $S = 2$, weak ligand field) state. With intermediate ligand field strength, a spin transition can be induced by an external perturbation (light, heat, pressure), thus opening SCO complexes for technological applications. The selection of a ligand set capable of producing the proper ligand field is therefore fundamental for the very nature of the SCO phenomenon.[78] The combination of Fe(II) with nitrogen-based ligands spans the SCO divide and leads to a significant metal−ligand bond length change during the spin transition. In the solid state, such bond length changes promote cooperativity between the metal centers[73] that can generate abrupt spin transitions and hysteresis,[79] both of which are tailored properties for technological applications.[80] Therefore, hundreds of SCO candidates have been synthesized combining amines, imines, azo- and *N*-heteroaromatic ligands with denticities from one (i.e., monodentate) to six or more.[72]

On the computational side, Deeth and co-workers have developed a force field in the context of the ligand field molecular mechanics (LFMM)[81] method that allows inexpensive modeling of both spin states, thus facilitating virtual screening and de novo design techniques.[82,83] In fact, although LFMM parameters for modeling of SCO compounds are currently available only for amine ligands, the ligand field generated by the six amine-type ligands can, to some extent, be tuned by altering the structure of the molecular skeleton that holds the coordinating atoms. Therefore, in the de novo design experiments presented below, ligand sets are explored combining (i) different ligand denticities (from bidentate to hexadentate, with at most two different multidentate ligands) and variations of (ii) bridge lengths, (iii) bridge types (i.e., aromatic, aliphatic cyclic or acyclic), and (iv) the identity of the bridgehead atoms, plus (v) modulation of pendant groups on coordinating atoms and on bridges. With the exception of the latter, these structural modifications involve alteration of multiple metal−organic rings, highlighting the crucial role of the ring generation and modification method described above.

In the context of evolutionary-algorithm-based global optimization,[36] each candidate SCO complex needs to be evaluated on the fly by a numerical scoring function representing its fitness for the role of SCO compound. In this first application, we have chosen to construct an empirical fitness function that favors (i) a small energy difference between the minimum-energy geometries of the two spin states, (ii) an LS rather than HS ground state, and (iii) a small value of the estimated enthalpy contribution to the energy barrier associated with the spin transitions in both directions (LS → HS, HS → LS). The formulation of the fitness function is reported in eq 6:

$$\mathrm{fitness} = -(\Delta E_{\mathrm{rH-rL}} + p)^2 - w\Delta E_{\mathrm{vHL}}^{\,2} - w\Delta E_{\mathrm{vLH}}^{\,2} \tag{6}$$

where ΔE_{rH-rL} is the LFMM energy difference between the most stable conformations of the two spin states of a given Fe(II) complex (i.e., relaxed spin transition),[84] p is the bias favoring the LS ground state over the HS ground state ($p = -3.0$ kcal/mol), $w > 0$ is an empirical weighting constant ($w = 0.05$) to reduce the contributions from ΔE_{vHL} and ΔE_{vLH}, which are the values of the vertical spin transitions[84] between the LFMM potential energy surfaces at the lowest-energy geometries of the HS and LS states, respectively. A high

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

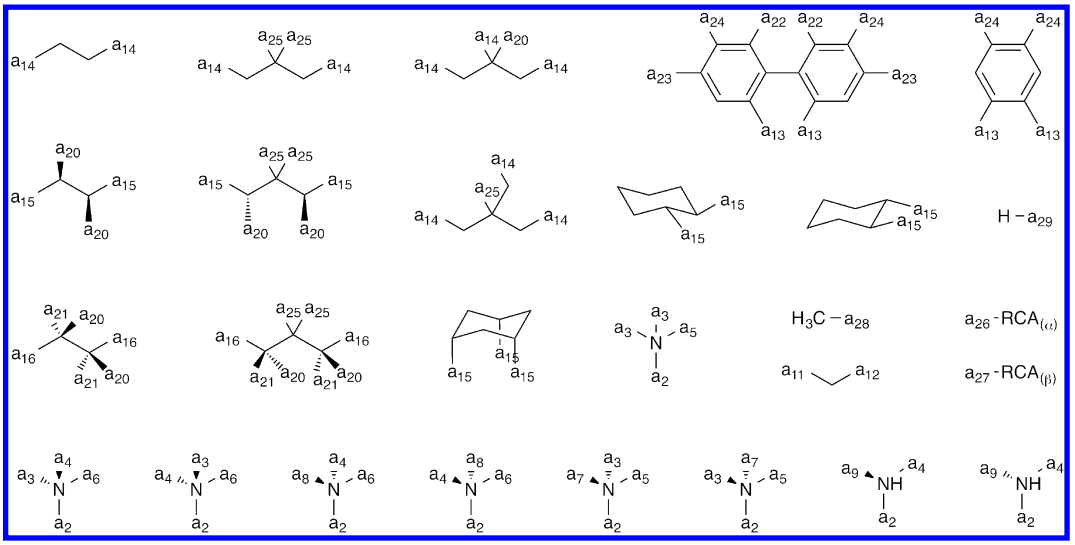


Figure 4. Library of 3D molecular fragments. Attachment points (APs) are represented with dummy atoms labeled according to their AP class (a_n). The type of RCA is indicated by the subscripts α and β.

contribution from the last two terms is here taken as a rough indication of a high enthalpic component to the energy barriers associated with the spin transitions, i.e., the minimum-energy crossing point.

Two types of de novo design experiments are presented: type **A**, in which Fe(II) complexes were built starting from single-atom fragments (i.e., a naked Fe(II) atom), and type **B**, in which portions of the chelating system were already introduced in the root fragment (also called the scaffold; Figure 3). While it is clear that experiments of type **A** can cover a wider structural variability than those of type **B** and are therefore less biased by preconcepts that typically affect intuition-based design, type-**B** experiments are presented to show how preliminary knowledge of the chemistry under study can be exploited to focus the experiment onto a specific set of predefined skeletons, which are kept fixed and possibly incorporated into wider coordinating systems with higher denticity. Moreover, approaches of type **B** significantly reduce the complexity of the search space as well as that of the ring-closing problem by lowering the number of ring closures needed to obtain a valid structure. The two experiments discussed below, though starting from different initial scaffolds (Figure 3), made use of the same connecting rules (see Computational Details) and library of fragments (Figure 4). The latter was designed to include the most frequent linkers connecting amine ligands in iron complexes. All of these building blocks were prepared as 3D fragments specifically meant for this application (see Computational Details). In particular, fragments with bridging capability were generated with bond angles and bond lengths suitable for the formation of rings with five, six, or more members. Since the current implementation of the ring closability conditions does not modify bond angles but instead makes use of loose allowance values (see Methods), changes in ring size are permitted only within the range defined by such allowances. In particular, most rings could change size from five to six members and vice versa, but not from five to four or three members.

Artificial evolution experiments started with the generation of an initial random population of 50 complexes that was subsequently evolved for 100 generations, producing five new candidates in each generation (see Computational Details). The outcomes of two artificial evolution runs are summarized in Figure 5 (**A**) and Figure 6 (**B**), where the fitness of the instantaneous population (i.e., the best 50 candidates produced up to a certain generation) is displayed through the generations. In both cases, the overall performance of the population improves during the experiment, though mostly by removal of bad candidates and with only a little increase of the highest fitness. Such asymptotic behavior, however, is an

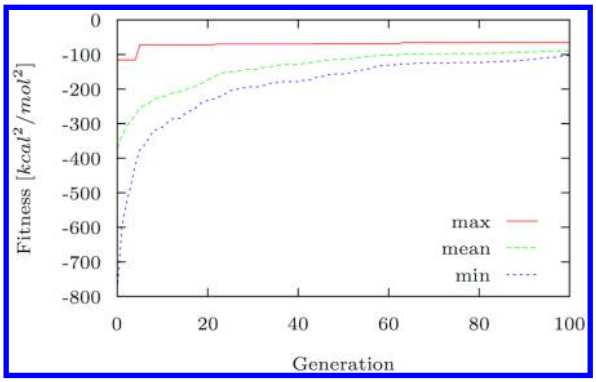


Figure 5. Evolution of the maximum (red), mean (green), and minimum (blue) fitness value through an experiment of type **A**.

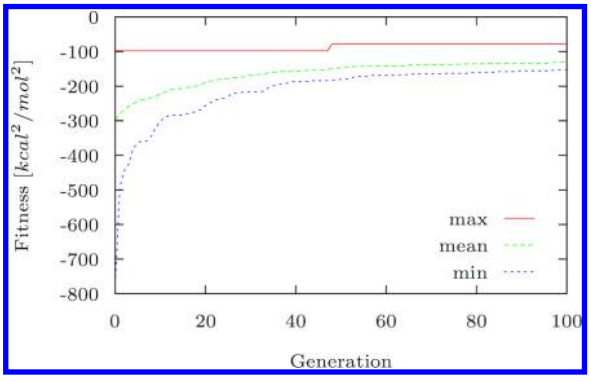


Figure 6. Evolution of the maximum (red), mean (green), and minimum (blue) fitness value through an experiment of type **B**.

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

expected result of the definition of the fitness function in eq 6, which cannot give values higher than zero.

The role of the ring-closing machinery can be appreciated by observing the characterization of ligand denticity during the experiments. Figures 7 and 8 display the changes in the occurrence of Fe(II) surrounded by one of the permitted combinations of ligand denticities: three bidentate ligands (κ^2, κ^2, κ^2), two tridentate ligands (κ^3, κ^3), one tetradentate ligand and one bidentate ligand (κ^4, κ^2), or one hexadentate ligand (κ^6). In both experiments, fluctuations of the relative amounts were registered, thus showing that the denticity sets compete mainly on the basis of the performance of the generated compounds. In fact, if well-performing candidates of one such set are identified, then their occurrence in the population will tend to increase. This is valid under the assumption that the possibility exists to generate similarly well-performing analogues. While both experiments evolved toward populations enriched in complexes with two tridentate ligands, the occurrence of hexadentate ligands seemed to be affected by the ease of accessing such structures. In fact, the very small percentage of hexadentate ligands displayed in Figure 7 for experiment **A** is in contrast to that of experiment **B** (Figure 8), where such ligands occupy a consistent portion of the population during the whole experiment. The generation of a κ^6 candidate requires the simultaneous closure of at least six metal−organic rings in case of experiments of type **A**, but this number is reduced to only three for most of the scaffolds deployed in experiments of type **B** (Figure 3). The different accessibility of hexadentate ligands seems to affect the production of molecules by lowering the effective probability of generating molecules with high cyclicity, which are more challenging to generate but, in the particular case of multidentate ligands, not necessarily more challenging from the chemical synthesis point of view.

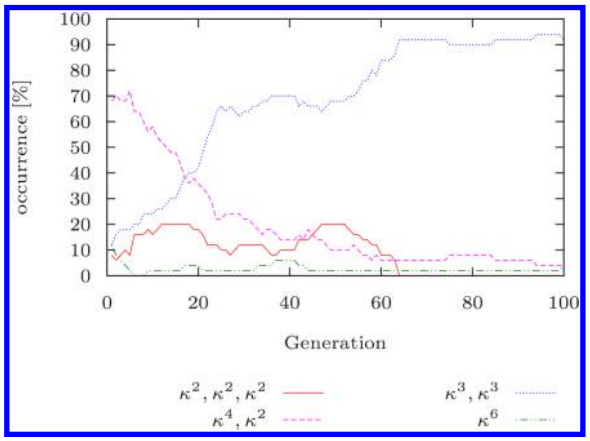


Figure 7. Characterization of the Fe(II) coordinating environment in the instantaneous population throughout the experiment of type **A**. Each κ^n represents a single *n*-dentate ligand.

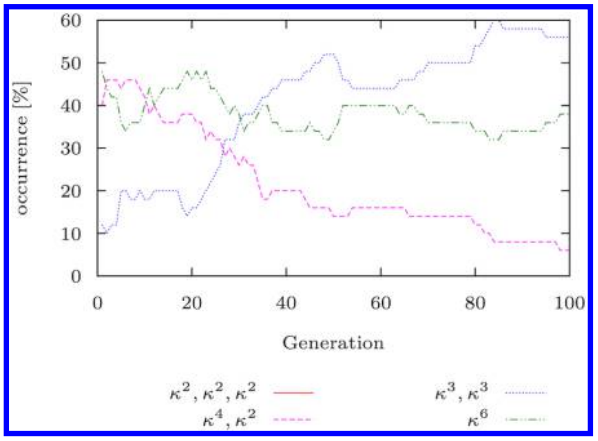


Figure 8. Characterization of the Fe(II) coordinating environment in the instantaneous population throughout the experiment of type **B**. Each κ^n represents a single *n*-dentate ligand.

Both experiments are characterized by a trend that prefers complexes with two tridentate ligands (κ^3, κ^3), but while analysis of the final populations confirmed that the best candidates identified by **A** have two tridentate ligands (**1** and **2** in Figure 9 and Table 1), the fittest molecule from **B** has a hexadentate ligand (**4** in Figure 9 and Table 1). Nevertheless, about 36% of the final population of **B** includes the tridentate skeleton of **2** (tacph in Figure 10), and that of **1** (ampda in Figure 10) is also found in about the same amount of molecules, also including **4** and **6**. Thus, while the different library of scaffolds used in **B** (Figure 3) precluded any possibility of building **2**, the two experiments agreed on the fact that ligands based on tacph and ampda lead to high fitness.

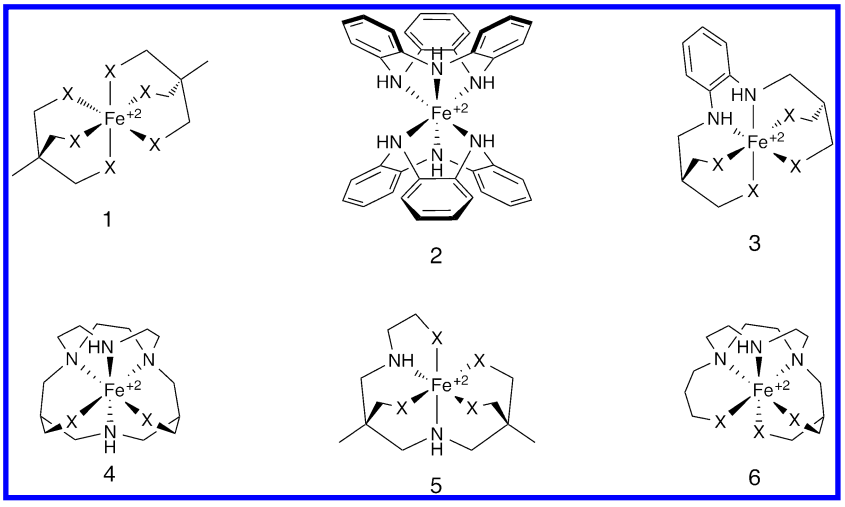


Figure 9. Sketches of representative molecules with high fitness discussed in the text (X = NH_2).

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

Table 1. Energies (in kcal/mol) of the Relaxed (ΔE_{rH-rL}) and Vertical (ΔE_{vHL} and ΔE_{vLH}) Spin Transitions for Selected SCO Candidates and Existing Molecules

	ΔE_{rH-rL}		ΔE_{vLH}		ΔE_{vHL}	
mol. ID	LFMM[e]	DFT	LFMM[e]	DFT	LFMM[e]	DFT
1[a]	0.42	3.05	29.06	32.68	20.15	18.60
2[a]	0.84	3.36	29.38	33.51	20.86	30.35
3[a]	1.30	4.51	30.07	34.30	21.62	20.47
4[b]	4.93	5.47	33.29	32.13	19.35	18.78
5[a]	0.94	4.55	31.46	33.21	24.38	20.74
6[b]	−1.96	2.31	27.93	30.72	26.06	19.32
$[Fe(tacn)_2]^{2+}$ [c]	1.40	5.29	42.14	35.88	25.22	18.63
$[Fe(NH_3)_6]^{2+}$ [d]	−6.27	−7.36	31.99	22.82	28.56	27.45
mean absolute deviation		2.72		4.14		3.85

[a]New candidate generated by an experiment of type **A**. [b]New candidate generated by an experiment of type **B**. [c]Previously existing compound; undergoes thermal SCO in solution.[91,92] [d]Previously existing compound; HS ground state, no SCO behavior. [e]From LFMM as implemented in Tinker.[93]

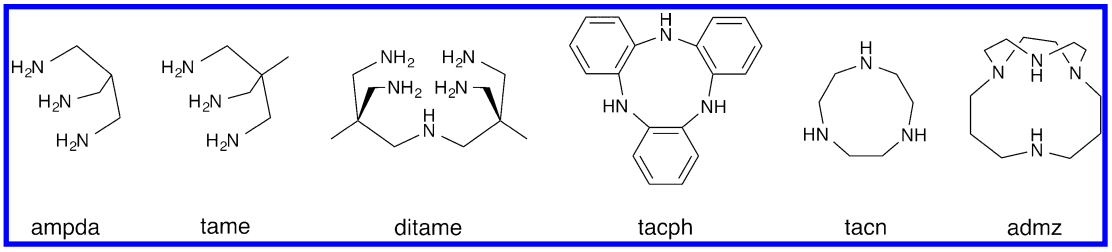


Figure 10. Sketches of ligand skeletons mentioned in the text.

The use of larger scaffolds in **B** also has another effect. Despite the diffusion of the ampda building block, the root fragment used to generate most of the final population of experiment **B** is that containing the tacn skeleton (Figure 10) rather than that containing ampda. This is explained by the fact that the best candidates in **B**, namely, **4**, **6**, and similar molecules, are all built from the tacn scaffold, and thus, it is likely for new molecules generated in **B** to inherit the tacn core by means of the genetic operators crossover and mutation.

To better evaluate the quality of the selected structures, the relaxed and vertical spin-state transitions were calculated with density functional theory (DFT) for the most representative molecules (Table 1). While the DFT method adopted here differs from that used as a reference by Deeth and co-workers in the parametrization of the LFMM force field (see Computational Details),[82,83] the results obtained for $[Fe(NH_3)_6]^{2+}$ are in good agreement with high-accuracy computations reported in the literature.[84,85] Moreover, the trend expected for Fe(II) compounds used to parametrize the LFMM force field is mostly recovered with an accuracy compatible to that expected for DFT calculations of spin-state transitions (mean absolute deviation = 2.8 kcal/mol; see the Supporting Information).[84,86−89] It is known that the performance of DFT in this task might be very sensitive to the functional and basis set[88] and may also be system-specific.[89] Therefore, in the present context the DFT calculations serve mostly to exclude gross errors in the empirical LFMM estimate and to evaluate whether the best candidates selected by LFMM have relaxed spin-state energetics that are also within a 5−6 kcal/mol window at the DFT level, which has been suggested as indicative of potential SCO behavior.[89,90]

The data in Table 1 show better agreement between LFMM and DFT on the relaxed spin-state transitions, which were used to parametrize the force field,[83] than on the vertical transitions. While these results confirm that the best candidates satisfy the empirical criteria encoded in the fitness function, it is clear that further and more accurate analysis needs to be performed, also including an evaluation of the minimum-energy crossing point (MECP) that characterizes the spin transition.

The overall best candidate is **1** $[Fe(tame)_2]^{2+}$ (tame = 1,1,1-tris(aminomethyl)ethane; see Figures 9 and 10). Bis(tame) complexes have been experimentally characterized for Ni(II), Co(III),[94,95] and Pt(IV),[96] but to date no attempt to synthesize **1** has been reported in the literature. Notably, tame has a central role as a tripodal building block for hexadentate ligands and has also been exploited in such a way for many candidate Fe(II) SCO compounds.[97] The energy of the relaxed transition calculated for **1** at the DFT level (Table 1) is sufficiently close to the SCO divide and suggests an LS ground state, which is in agreement with the LFMM prediction. An additional promising result is that the vertical transition calculated with either LFMM or DFT is slightly smaller than that of $[Fe(tacn)_2]^{2+}$, which undergoes SCO in solution.[91,92] The DFT-optimized structures of the two spin states show that the elongation of the Fe−N bonds resulting from the LS → HS transition is projected nearly exclusively onto the direction of the C_3 symmetry axis, causing an elongation of the molecule of 0.28 Å between the methyl groups. A C_3 symmetry axis also characterizes **2**, which is the second-best molecule according to the fitness calculated by LFMM. Nevertheless, the value of ΔE_{vHL} is significantly underestimated by LFMM (by almost 10 kcal/mol), and thus, the fitness of **2** is overestimated. The LFMM force field is probably at fault here since aniline-type N donors were not part of the training set.[82,83] The NH_2 group in aniline is conjugated with the benzene ring, leading to a planar geometry around nitrogen and the assignment of an sp^2 NC=C atom type in MMFF. The sp^3 N atom type used here thus overestimates the degree of pyramidalization when coordinated

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

to a metal center, and this effect varies with spin state and the concomitant change in M−N bond strength. Nevertheless, while the tacph ligand in **2** is certainly peculiar, syntheses of the *N*,*N*′-dimethyl analogues have been reported and also supported with crystallographic structures.[98] The same authors suggested that the crown conformation shown in Figure 9 could coordinate metal atoms and provide a bowl-shaped pocket suitable for supramolecular interactions, which in **2** could also be exploited to alter the properties of the potential SCO materials.[99] However, such a conformation is less stable by 12.9 kcal/mol (OPBE/cc-pVTZ PCM(water); see Computational Details) than the one observed in the crystallographic structure of the *N*,*N*′-dimethyl analogues, which does not allow tridentate coordination. Besides, comparison of the DFT geometries of the HS and LS states (see the Supporting Information) shows how the rigidity of the tacph skeleton imposes a substantial and asymmetric displacement of the ligand (about 1 Å displacement of one nitrogen) that could be too great to be accommodated in a solid lattice, thus preventing spin crossover.[80] Moreover, the high value of ΔE_{vHL} from DFT might indicate a higher MECP. Despite promising indications, namely, the high fitness, the values of the relaxed spin transition both from LFMM and DFT, and the appealing possibility of combining the SCO center with a supramolecular interaction pocket,[98] the complex could be too strained, and the ligand lacks preorganization. From the methods point of view, this result highlights the role of the allowance in the evaluation of ring-closing conditions (see Computational Details), which might have to be reduced if borderline structures are to be avoided. Nevertheless, it should be noted that the calculation of the fitness does not include any evaluation of the conformational preorganization of the ligand. Thus, the affinity for the metal is not evaluated, and ligands displaying low affinity may be generated. To avoid this effect, besides a reduction of the allowances, a comparison of the coordinating conformation with the free-ligand conformation could be introduced in the evaluation of the fitness as well as that of the metal−ligand binding energy, but the computational cost of each candidate would rise significantly.

As anticipated, what emerges from the analysis of the other molecules displayed in Figure 9 is the dominant role of ampda, the tripodal building block constituting the fundamental organic core of the ligand in **1**. One or two ampda fragments are present in all of the fittest structures except for **2** and other molecules based on tacph. For instance, the addition of the phenyl bridge in **3** connects the two tripodal building blocks, giving a hexadentate ligand. The higher denticity should favor coordination of the metal while retaining the proper spin-state energy difference as demonstrated by the DFT calculation. Instead, the monodirectional elongation observed for **1** cannot take place, and the LS → HS transition is accompanied by a more complex expansion of the coordination sphere that for the most part involves the four primary amines.

The tetradentate bicyclic skeleton of **4** (admz in Figure 10)[100] has been used with Co(II), Ni(II), Cu(II), and Zn(II),[100] and a related compound with a bowl of smaller size (i.e., 1,4,7,10-tetraazabicyclo[5.5.2]tetradecane) has been coordinated to Fe(II).[101,102] The DFT calculation agrees on the LS ground state of **4**, but the presence of the metal-coordinating primary amines forces the propyl bridges belonging to the 11-membered ring to assume a more strained conformation than that of the tetradentate analogue. Thus, the stability of the hexadentate complex might be compromised. On the contrary, the last two skeletons, **5** and **6**, are much more flexible and should allow for more relaxed structures. In particular, **6** is a modified version of **4** obtained by removal of a methylene bridge, with the consequent elimination of the cyclic constraint and the release of the conformational strain mentioned above. According to LFMM, this structural change is accompanied by a change in the ground state and consequential reduction of the fitness. On the contrary, DFT predicts an LS ground state, and the values for the relaxed and vertical transitions **6** are still promising. Finally, the hexadentate ligand of **5** is characterized by the pentadentate ditame core (see Figures 9 and 10), which has been synthesized and used to prepare complexes of late transition metals[103] but not Fe(II) complexes.

It should be noted that the acyclic ligand in **5** is the result of the closure of five metal−organic rings. In fact, the strategy followed to build the molecules in both **A**- and **B**-type experiments has been that of growing molecular branches on any of the attachment points of the initial scaffold, which corresponds to building a spanning tree with many primary branches, and then closing the metal−organic rings, which might possibly lead to the simultaneous formation of organic rings, thus increasing the cyclicity of the free organic ligand. For example, the molecular representation of the spanning tree of molecule **2**, prior to the ring closure step, had six equal branches surrounding the metal atom, and each was formed by the same five fragments: N, H, Ph, $RCA_{(\alpha)}$, $RCA_{(\beta)}$. Therefore, the two tacph rings in **2** are the result of the closure of three metal−organic cycles each. It is clear that this approach, although general and, as shown above, capable of exploring vast structural diversity, is perpendicular to the preparation of transition metal complexes based on the synthesis and isolation of the organic ligands (closure of all organic rings) followed by formation of the complex (closure of all metal−organic rings). In fact, herein the construction of complexes is to some extent similar to synthetic approaches where the metal ion acts as a template (i.e., template synthesis).[104−106] Therefore, for the strategy applied here, the definition of the connectivity rules may be difficult when the aim is precise control of the chemical traits of the organic ligand. An alternative building strategy would be to grow spanning trees with as many primary branches on the metal atoms (i.e., the root fragments) as the number of organic ligand molecules desired (optionally with symmetric reproduction of only one ligand)[54] and to prohibit cross-connection between the branches. This approach is less general and is de facto a special case of the previous and more general approach applied here. Nevertheless, it allows more precise control of the synthetic accessibility of the sole organic ligand and could be used effectively when ligands with variable denticity are generated and a nonspecified number of solvent molecules are meant to fill all of the vacant coordination sites on the metal.

Finally, although the results obtained demonstrate the capabilities of the method, its implementation can be improved. Low productivity of the molecular generation, as the number of graphs leading to a possible SCO candidate over the number of generated graphs, has emerged as a bottleneck. While the test case imposed strict requirements, i.e., minimum and maximum numbers of rings to close and rejection of monodentate ligands, three main methodological issues contribute to the low productivity: (i) randomly generated graphs rarely satisfy the graphs' filtering criteria; (ii) lack of early-stage detection of duplicate and isomorphic graphs; and (iii) primitive imple-

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

mentation of atom path closability evaluation. As explained in Computational Details, the graphs were filtered to retain only those representing molecules with given chemical traits, for instance, avoiding monodentate ligands and controlling the number of possible rings. In fact, the vast majority of the randomly generated graphs were rejected because of restrictions on the number of possible ring closures: either too few, which gives one or more monodentate ligands (61.9% for type **A** and 64.3% for type **B**), or too many, which may lead to complex cages[107] (9.7% for type **A** and 11.6% for type **B**). Moreover, rejection rates of 16.8% and 8.7% for **A** and **B**, respectively, were caused by insufficient functionalization of the metal, which is due to the choice of favoring low substitution probability (see Computational Details). Then, most of the graphs that were finally converted into molecules were also rejected because of identification of duplicates (5.3% and 9.0%) and incomplete ring closure (6.2% and 5.3%). Duplicate graphs can easily be generated because, though the tree is always rooted on a given vertex, multiple different trees can be built and combined with different sets of fundamental cycles to give the exact same graph. This source of redundancy is further expanded with that intrinsically derived from the fragment library, namely, the fact that arrangements of different fragments, though defining different graphs, may lead to the same molecular structure. Therefore, early detection of graph isomorphism is needed, and the currently implemented approach based on removal of duplicate InChI keys,[108,109] although capable of accounting for the stereochemistry of the organic ligands,[110] seems deficient and is applied at a late stage (i.e., after construction of the molecular model). Instead, while incomplete ring closure may certainly derive from steric hindrance between substituents decorating an otherwise closable chain of atoms, part of the reason why some molecules did not lead to complete ring closure is due to the lack of any evaluation of the simultaneous path closability condition. In fact, attempts to include such evaluation, by detection of at least one common entry in the lists of closable conformations for all interdependent fundamental cycles, have turned out to be too computationally costly for systematic application. In conclusion, 0.05% and 0.18% of the graphs for experiments of type **A** and **B**, respectively, were submitted to the fitness calculation. These values also indicate the extent to which the larger scaffold used in **B** reduced the complexity of the whole experiment and increased the productivity, but at the cost of generality loss.

Further developments will therefore be focused on improving the productivity of the graph generation process and the evaluation of the path closability conditions, for which methods based on inverse kinematics have demonstrated performance.[58−62] Moreover, the randomized sampling that initiates the population of artificial evolution experiments risks not being sufficiently representative of the diversity contained in the fragment space. Thus, further developments must also include diversity-oriented sampling of the search space.[19]

CONCLUSIONS

The capability to perform ring closures has been integrated into 3D fragment-based de novo design to provide a tool capable of creating and modifying multicyclic compounds in the context of an evolutionary-algorithm-based design framework. The overall method has been proven capable of building and modifying systems with various cyclicities, providing the means to identify new candidate Fe(II) spin-crossover compounds based on multidentate amine ligands. Candidates were automatically generated by two different strategies exploring structural features that, in addition to simple substitution patterns on both coordinating atoms and bridges, also include ring-related features such as ligand denticity, nature of the bridges, and identity of the metal-coordinating terms connected by the bridges. The resulting best candidates displayed skeletons that have not yet been explored for SCO behavior but are similar to ligands deployed for other metals or to molecules designed for different aims. The results demonstrate the capability of the presented method to suggest unexpected yet still reasonable ligands. Moreover, both LFMM and DFT calculations suggested that some of the best candidates identified by the automated tool are close to the SCO divide, and some deserve further investigation, which should include an evaluation of the stability of the complexes and the minimum-energy crossing point. However, the analysis of the results has pointed out an efficiency bottleneck, and the need for further developments of the method has been highlighted and discussed.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00424.

Computational details, benchmark for the DFT method, Cartesian coordinates, and sample Gaussian input files (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: vidar.jensen@kj.uib.no.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The Research Council of Norway (RCN) is acknowledged for financial support via the eVITA (Grant 205273/V30) and GASSMAKS (208335/E30) Programs and for CPU and storage resources granted through the NOTUR (NN2506K) and NORSTORE (NS2506K) Supercomputing Programs. B.J.H. acknowledges the support (Doctoral Training Account) by the Engineering and Physical Sciences Research Council (EPSRC), and R.J.D. acknowledges Chemical Computing Group for their generous software support. Network support by COST Action CM1305 (eCOSTBIO) is also gratefully acknowledged. The authors thank Vishwesh Venkatraman and Bjørn K. Alsberg for discussions. ChemAxon (http://www.chemaxon.com) is thanked for free academic use of the Marvin package.

REFERENCES

(1) McNaught, A. D.; Wilkinson, A. *IUPAC Compendium of Chemical Terminology (The "Gold Book")*, 2nd ed.; Blackwell Scientific Publications: Oxford, U.K., 1997.

(2) Schwarzenbach, G. Der Chelateffekt. *Helv. Chim. Acta* **1952**, *35*, 2344−2359.

(3) Davydova, E. I.; Sevastianova, T. N.; Timoshkin, A. Y.; Suvorov, A. V.; Frenking, G. Chelate Effect: The Importance of Reorganization Energy. *Int. J. Quantum Chem.* **2004**, *100*, 419−425.

(4) Vallet, V.; Wahlgren, U.; Grenthe, I. Chelate Effect and Thermodynamics of Metal Complex Formation in Solution: A Quantum Chemical Study. *J. Am. Chem. Soc.* **2003**, *125*, 14941−14950.

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

(5) Comba, P.; Kerscher, M. Computation of Structures and Properties of Transition Metal Compounds. *Coord. Chem. Rev.* **2009**, *253*, 564−574.
(6) Houk, K. N.; Cheong, P. H.-Y. Computational Prediction of Small-Molecule Catalysts. *Nature* **2008**, *455*, 309−313.
(7) Thiel, W. Computational Catalysis—Past, Present, and Future. *Angew. Chem., Int. Ed.* **2014**, *53*, 8605−8613.
(8) Hageman, J. A.; Westerhuis, J. A.; Frühauf, H.-W.; Rothenberg, G. Design and Assembly of Virtual Homogeneous Catalyst Libraries − Towards *in Silico* Catalyst Optimisation. *Adv. Synth. Catal.* **2006**, *348*, 361−369.
(9) Gillespie, J. A.; Dodds, D. L.; Kamer, P. C. J. Rational Design of Diphosphorus Ligands − a Route to Superior Catalysts. *Dalton Trans.* **2010**, *39*, 2751−2764.
(10) Jiang, X.-B.; Lefort, L.; Goudriaan, P. E.; de Vries, A. H. M.; van Leeuwen, P. W. N. M.; de Vries, J. G.; Reek, J. N. H. Screening of a Supramolecular Catalyst Library in the Search for Selective Catalysts for the Asymmetric Hydrogenation of a Difficult Enamide Substrate. *Angew. Chem., Int. Ed.* **2006**, *45*, 1223−1227.
(11) Goudriaan, P. E.; van Leeuwen, P. W. N. M.; Birkholz, M.-N.; Reek, J. N. H. Libraries of Bidentate Phosphorus Ligands; Synthesis Strategies and Application in Catalysis. *Eur. J. Inorg. Chem.* **2008**, *2008*, 2939−2958.
(12) Raynal, M.; Ballester, P.; Vidal-Ferran, A.; van Leeuwen, P. W. N. M. Supramolecular Catalysis. Part 1: Non-Covalent Interactions as a Tool for Building and Modifying Homogeneous Catalysts. *Chem. Soc. Rev.* **2014**, *43*, 1660−1733.
(13) Fey, N. The Contribution of Computational Studies to Organometallic Catalysis: Descriptors, Mechanisms and Models. *Dalton Trans.* **2010**, *39*, 296−310.
(14) Fey, N.; Harvey, J. N.; Lloyd-Jones, G. C.; Murray, P.; Orpen, A. G.; Osborne, R.; Purdie, M. Computational Descriptors for Chelating P,P- and P,N-Donor Ligands1. *Organometallics* **2008**, *27*, 1372−1383.
(15) Jover, J.; Fey, N.; Harvey, J. N.; Lloyd-Jones, G. C.; Orpen, A. G.; Owen-Smith, G. J. J.; Murray, P.; Hose, D. R. J.; Osborne, R.; Purdie, M. Expansion of the Ligand Knowledge Base for Chelating P,P-Donor Ligands (LKB-PP). *Organometallics* **2012**, *31*, 5302−5306.
(16) Jover, J.; Fey, N. Screening Substituent and Backbone Effects on the Properties of Bidentate P,P-Donor Ligands (LKB-PPscreen). *Dalton Trans.* **2013**, *42*, 172−181.
(17) Burello, E.; Rothenberg, G. Topological Mapping of Bidentate Ligands: A Fast Approach for Screening Homogeneous Catalysts. *Adv. Synth. Catal.* **2005**, *347*, 1969−1977.
(18) Burello, E.; Marion, P.; Galland, J.-C.; Chamard, A.; Rothenberg, G. Ligand Descriptor Analysis in Nickel-Catalysed Hydrocyanation: A Combined Experimental and Theoretical Study. *Adv. Synth. Catal.* **2005**, *347*, 803−810.
(19) Maldonado, A. G.; Hageman, J. A.; Mastroianni, S.; Rothenberg, G. Backbone Diversity Analysis in Catalyst Design. *Adv. Synth. Catal.* **2009**, *351*, 387−396.
(20) Burello, E.; Rothenberg, G. *In Silico* Design in Homogeneous Catalysis Using Descriptor Modelling. *Int. J. Mol. Sci.* **2006**, *7*, 375−404.
(21) Fey, N.; Garland, M.; Hopewell, J. P.; McMullin, C. L.; Mastroianni, S.; Orpen, A. G.; Pringle, P. G. Stable Fluorophosphines: Predicted and Realized Ligands for Catalysis. *Angew. Chem., Int. Ed.* **2012**, *51*, 118−122.
(22) Urbano-Cuadrado, M.; Carbó, J. J.; Maldonado, A. G.; Bo, C. New Quantum Mechanics-Based Three-Dimensional Molecular Descriptors for Use in QSSR Approaches: Application to Asymmetric Catalysis. *J. Chem. Inf. Model.* **2007**, *47*, 2228−2234.
(23) Strassberger, Z.; Mooijman, M.; Ruijter, E.; Alberts, A. H.; Maldonado, A. G.; Orru, R. V. A.; Rothenberg, G. Finding Furfural Hydrogenation Catalysts via Predictive Modelling. *Adv. Synth. Catal.* **2010**, *352*, 2201−2210.
(24) Maldonado, A. G.; Rothenberg, G. Predictive Modeling in Homogeneous Catalysis: A Tutorial. *Chem. Soc. Rev.* **2010**, *39*, 1891−1902.
(25) Kutchukian, P. S.; Shakhnovich, E. I. De Novo Design: Balancing Novelty and Confined Chemical Space. *Expert Opin. Drug Discovery* **2010**, *5*, 789−812.
(26) Gillet, V. J.; Newell, W.; Mata, P.; Myatt, G.; Sike, S.; Zsoldos, Z.; Johnson, A. P. SPROUT: Recent Developments in the de Novo Design of Molecules. *J. Chem. Inf. Model.* **1994**, *34*, 207−217.
(27) Glen, R. C.; Payne, A. W. R. A Genetic Algorithm for the Automated Generation of Molecules within Constraints. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 181−202.
(28) Globus, A.; Lawton, J.; Wipke, T. Automatic Molecular Design Using Evolutionary Techniques. *Nanotechnology* **1999**, *10*, 290.
(29) Nachbar, R. B. Molecular Evolution: Automated Manipulation of Hierarchical Chemical Topology and Its Application to Average Molecular Structures. *Genet. Program. Evolvable Mach* **2000**, *1*, 57−94.
(30) Douguet, D.; Thoreau, E.; Grassy, G. A Genetic Algorithm for the Automated Generation of Small Organic Molecules: Drug Design Using an Evolutionary Algorithm. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 449−466.
(31) Barone, R.; Barone, R.; Arbelot, M.; Chanon, M. Gasp: A Computer Program to Generate Automatically Polycyclic Structures. *Tetrahedron* **2001**, *57*, 6035−6042.
(32) Brown, N.; McKay, B.; Gilardoni, F.; Gasteiger, J. A Graph-Based Genetic Algorithm and Its Application to the Multiobjective Evolution of Median Molecules. *J. Chem. Inf. Model.* **2004**, *44*, 1079−1087.
(33) Lima, L.; Barreiro, E. Bioisosterism: A Useful Strategy for Molecular Modification and Drug Design. *Curr. Med. Chem.* **2005**, *12*, 23−49.
(34) Wong, S. S. Y.; Luo, W.; Chan, K. C. C. EvoMD: An Algorithm for Evolutionary Molecular Design. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2011**, *8*, 987−1003.
(35) Mishima, K.; Kaneko, H.; Funatsu, K. Development of a New *De Novo* Design Algorithm for Exploring Chemical Space. *Mol. Inf.* **2014**, *33*, 779−789.
(36) Chu, Y.; Heyndrickx, W.; Occhipinti, G.; Jensen, V. R.; Alsberg, B. K. An Evolutionary Algorithm for *de Novo* Optimization of Functional Transition Metal Compounds. *J. Am. Chem. Soc.* **2012**, *134*, 8885−8895.
(37) Sadowski, J.; Gasteiger, J.; Klebe, G. Comparison of Automatic Three-Dimensional Model Builders Using 639 X-Ray Structures. *J. Chem. Inf. Model.* **1994**, *34*, 1000−1008.
(38) Gasteiger, J.; Sadowski, J.; Schuur, J.; Selzer, P.; Steinhauer, L.; Steinhauer, V. Chemical Information in 3D Space. *J. Chem. Inf. Model.* **1996**, *36*, 1030−1037.
(39) Hay, B. P.; Firman, T. K. HostDesigner: A Program for the *de Novo* Structure-Based Design of Molecular Receptors with Binding Sites That Complement Metal Ion Guests. *Inorg. Chem.* **2002**, *41*, 5502−5512.
(40) Hay, B. P.; Oliferenko, A. A.; Uddin, J.; Zhang, C.; Firman, T. K. Firman, T. K. Search for Improved Host Architectures: Application of de Novo Structure-Based Design and High-Throughput Screening Methods To Identify Optimal Building Blocks for Multidentate Ethers. *J. Am. Chem. Soc.* **2005**, *127*, 17043−17053.
(41) Hay, B. P.; Firman, T. K.; Moyer, B. A. Structural Design Criteria for Anion Hosts: Strategies for Achieving Anion Shape Recognition through the Complementary Placement of Urea Donor Groups. *J. Am. Chem. Soc.* **2005**, *127*, 1810−1819.
(42) Bryantsev, V. S.; Hay, B. P. *De Novo* Structure-Based Design of Bisurea Hosts for Tetrahedral Oxoanion Guests. *J. Am. Chem. Soc.* **2006**, *128*, 2035−2042.
(43) Young, N. J.; Hay, B. P. Structural Design Principles for Self-Assembled Coordination Polygons and Polyhedra. *Chem. Commun.* **2013**, *49*, 1354−1379.
(44) Hay, B. P. *De Novo* Structure-Based Design of Anion Receptors. *Chem. Soc. Rev.* **2010**, *39*, 3700−3708.
(45) Hay, B. P.; Hancock, R. D. The Role of Donor Group Orientation as a Factor in Metal Ion Recognition by Ligands. *Coord. Chem. Rev.* **2001**, *212*, 61−78.

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

(46) Vukovic, S.; Hay, B. P. *De Novo* Structure-Based Design of Bis-Amidoxime Uranophiles. *Inorg. Chem.* **2013**, *52*, 7805−7810.
(47) Bryan, J. C.; Hay, B. P.; Sachleben, R. A.; Eagle, C. T.; Zhang, C.; Bonnesen, P. V. Design, Synthesis, and Structure of Novel Cesium Receptors. *J. Chem. Crystallogr.* **2003**, *33*, 349−355.
(48) Hay, B. P.; Uddin, J.; Firman, T. K. Eight-Coordinate Stereochemistries of U(IV) Catecholate and Aquo Complexes. *Polyhedron* **2004**, *23*, 145−154.
(49) Hay, B. P.; Firman, T. K.; Lumetta, G. J.; Rapko, B. M.; Garza, P. A.; Sinkov, S. I.; Hutchison, J. E.; Parks, B. W.; Gilbertson, R. D.; Weakley, T. J. R. Toward the Computer-Aided Design of Metal Ion Sequestering Agents. *J. Alloys Compd.* **2004**, *374*, 416−419.
(50) Jia, C.; Hay, B. P.; Custelcean, R. De Novo Structure-Based Design of Ion-Pair Triple-Stranded Helicates. *Inorg. Chem.* **2014**, *53*, 3893−3898.
(51) Moyer, B. A.; Custelcean, R.; Hay, B. P.; Sessler, J. L.; Bowman-James, K.; Day, V. W.; Kang, S.-O. A Case for Molecular Recognition in Nuclear Separations: Sulfate Separation from Nuclear Wastes. *Inorg. Chem.* **2013**, *52*, 3473−3490.
(52) Hay, B. P.; Jia, C.; Nadas, J. Computer-Aided Design of Host Molecules for Recognition of Organic Guests. *Comput. Theor. Chem.* **2014**, *1028*, 72−80.
(53) Venkatraman, V.; Foscato, M.; Jensen, V. R.; Alsberg, B. K. Evolutionary *de Novo* Design of Phenothiazine Derivatives for Dye-Sensitized Solar Cells. *J. Mater. Chem. A* **2015**, *3*, 9851−9860.
(54) Foscato, M.; Occhipinti, G.; Venkatraman, V.; Alsberg, B. K.; Jensen, V. R. Automated Design of Realistic Organometallic Molecules from Fragments. *J. Chem. Inf. Model.* **2014**, *54*, 767−780.
(55) Foscato, M.; Venkatraman, V.; Occhipinti, G.; Alsberg, B. K.; Jensen, V. R. Automated Building of Organometallic Complexes from 3D Fragments. *J. Chem. Inf. Model.* **2014**, *54*, 1919−1931.
(56) *Handbook of Graph Theory*, 2nd ed.; Gross, J. L., Yellen, J., Zhang, P., Eds.; Discrete Mathematics and Its Applications; CRC Press: Boca Raton, FL, 2013.
(57) In this method only acyclic bonds are considered rotatable. Therefore, the choice between equally long atom paths in a cyclic system has no effect on the evaluation of the closability of the path.
(58) Bruccoleri, R. E.; Karplus, M. Chain Closure with Bond Angle Variations. *Macromolecules* **1985**, *18*, 2767−2773.
(59) Coutsias, E. A.; Seok, C.; Jacobson, M. P.; Dill, K. A. A Kinematic View of Loop Closure. *J. Comput. Chem.* **2004**, *25*, 510−528.
(60) Coutsias, E. A.; Seok, C.; Wester, M. J.; Dill, K. A. Resultants and Loop Closure. *Int. J. Quantum Chem.* **2006**, *106*, 176−189.
(61) Chys, P.; Chacón, P. Random Coordinate Descent with Spinor-Matrices and Geometric Filters for Efficient Loop Closure. *J. Chem. Theory Comput.* **2013**, *9*, 1821−1829.
(62) Zamuner, S.; Rodriguez, A.; Seno, F.; Trovato, A. An Efficient Algorithm to Perform Local Concerted Movements of a Chain Molecule. *PLoS One* **2015**, *10*, e0118342.
(63) Ponder, J. W. *TINKER: Software Tools for Molecular Design*, version 6.2; Washington University School of Medicine: Saint Louis, MO, 2013.
(64) Kostrowicki, J.; Scheraga, H. A. Application of the Diffusion Equation Method for Global Optimization to Oligopeptides. *J. Phys. Chem.* **1992**, *96*, 7442−7449.
(65) Nakamura, S.; Hirose, H.; Ikeguchi, M.; Doi, J. Conformational Energy Minimization Using a Two-Stage Method. *J. Phys. Chem.* **1995**, *99*, 8374−8378.
(66) Pappu, R. V.; Hart, R. K.; Ponder, J. W. Analysis and Application of Potential Energy Smoothing and Search Methods for Global Optimization. *J. Phys. Chem. B* **1998**, *102*, 9725−9742.
(67) Hart, R. K.; Pappu, R. V.; Ponder, J. W. Exploring the Similarities between Potential Smoothing and Simulated Annealing. *J. Comput. Chem.* **2000**, *21*, 531−552.
(68) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(69) The method identified rotatable bonds according to user-defined SMARTS queries.
(70) Gütlich, P. Spin Crossover − Quo Vadis? *Eur. J. Inorg. Chem.* **2013**, *2013*, 581−591.
(71) Gütlich, P.; Garcia, Y.; Goodwin, H. A. Spin Crossover Phenomena in Fe(II) Complexes. *Chem. Soc. Rev.* **2000**, *29*, 419−427.
(72) Halcrow, M. A. The Spin-States and Spin-Transitions of Mononuclear Iron(II) Complexes of Nitrogen-Donor Ligands. *Polyhedron* **2007**, *26*, 3523−3576.
(73) Hauser, A.; Jeftić, J.; Romstedt, H.; Hinek, R.; Spiering, H. Cooperative Phenomena and Light-Induced Bistability in Iron(II) Spin-Crossover Compounds. *Coord. Chem. Rev.* **1999**, *190−192*, 471−491.
(74) Bousseksou, A.; Molnár, G.; Demont, P.; Menegotto, J. Observation of a Thermal Hysteresis Loop in the Dielectric Constant of Spin Crossover Complexes: Towards Molecular Memory Devices. *J. Mater. Chem.* **2003**, *13*, 2069−2071.
(75) Kahn, O.; Martinez, C. J. Spin-Transition Polymers: From Molecular Materials Toward Memory Devices. *Science* **1998**, *279*, 44−48.
(76) Bousseksou, A.; Molnár, G.; Salmon, L.; Nicolazzi, W. Molecular Spin Crossover Phenomenon: Recent Achievements and Prospects. *Chem. Soc. Rev.* **2011**, *40*, 3313−3335.
(77) Gaspar, A. B.; Seredyuk, M. Spin Crossover in Soft Matter. *Coord. Chem. Rev.* **2014**, *268*, 41−58.
(78) Atmani, C.; El Hajj, F.; Benmansour, S.; Marchivie, M.; Triki, S.; Conan, F.; Patinec, V.; Handel, H.; Dupouy, G.; Gómez-García, C. J. Guidelines to Design New Spin Crossover Materials. *Coord. Chem. Rev.* **2010**, *254*, 1559−1569.
(79) Halcrow, M. A. Spin-Crossover Compounds with Wide Thermal Hysteresis. *Chem. Lett.* **2014**, *43*, 1178−1188.
(80) Halcrow, M. A. Structure:function Relationships in Molecular Spin-Crossover Complexes. *Chem. Soc. Rev.* **2011**, *40*, 4119−4142.
(81) Deeth, R. J. The Ligand Field Molecular Mechanics Model and the Stereoelectronic Effects of d and s Electrons. *Coord. Chem. Rev.* **2001**, *212*, 11−34.
(82) Deeth, R. J.; Anastasi, A. E.; Wilcockson, M. J. An *In Silico* Design Tool for Fe(II) Spin Crossover and Light-Induced Excited Spin State-Trapped Complexes. *J. Am. Chem. Soc.* **2010**, *132*, 6876−6877.
(83) Handley, C. M.; Deeth, R. J. A Multi-Objective Approach to Force Field Optimization: Structures and Spin State Energetics of d^6 Fe(II) Complexes. *J. Chem. Theory Comput.* **2012**, *8*, 194−202.
(84) Swart, M. Accurate Spin-State Energies for Iron Complexes. *J. Chem. Theory Comput.* **2008**, *4*, 2057−2066.
(85) Pierloot, K.; Vancoillie, S. Relative Energy of the High-($^5T_{2g}$) and Low-($^1A_{1g}$) Spin States of $[Fe(H_2O)_6]^{2+}$, $[Fe(NH_3)_6]^{2+}$, and $[Fe(bpy)_3]^{2+}$: CASPT2 versus Density Functional Theory. *J. Chem. Phys.* **2006**, *125*, 124303.
(86) Deeth, R. J.; Fey, N. The Performance of Nonhybrid Density Functionals for Calculating the Structures and Spin States of Fe(II) and Fe(III) Complexes. *J. Comput. Chem.* **2004**, *25*, 1840−1848.
(87) Harvey, J. N. On the Accuracy of Density Functional Theory in Transition Metal Chemistry. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2006**, *102*, 203−226.
(88) Swart, M. Spin States of (bio)inorganic Systems: Successes and Pitfalls. *Int. J. Quantum Chem.* **2013**, *113*, 2−7.
(89) Houghton, B. J.; Deeth, R. J. Spin-State Energetics of Fe^{II} Complexes − The Continuing Voyage Through the Density Functional Minefield. *Eur. J. Inorg. Chem.* **2014**, *2014*, 4573−4580.
(90) Ye, S.; Neese, F. Accurate Modeling of Spin-State Energetics in Spin-Crossover Systems with Modern Density Functional Theory. *Inorg. Chem.* **2010**, *49*, 772−774.
(91) Turner, J. W.; Schultz, F. A. Intramolecular and Environmental Contributions to Electrode Half-Reaction Entropies of $M(tacn)_2^{3+/2+}$ (M = Fe, Co, Ni, Ru; Tacn = 1,4,7-Triazacyclononane) Redox Couples. *Inorg. Chem.* **1999**, *38*, 358−364.

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856

(92) Turner, J. W.; Schultz, F. A. Solution Characterization of the Iron(II) Bis(1,4,7-Triazacyclononane) Spin-Equilibrium Reaction. *Inorg. Chem.* **2001**, *40*, 5296−5298.
(93) Foscato, M.; Deeth, R. J.; Jensen, V. R. Integration of Ligand Field Molecular Mechanics in Tinker. *J. Chem. Inf. Model.* **2015**, *55*, 1282−1290.
(94) Dalrymple, S. A.; Shimizu, G. K. H. Second-Sphere Coordination Networks: "Tame-Ing" (Tame = 1,1,1-Tris-(aminomethyl)ethane) the Hydrogen Bond. *J. Mol. Struct.* **2006**, *796*, 95−106.
(95) Dalrymple, S. A.; Shimizu, G. K. H. Crystal Engineering of a Permanently Porous Network Sustained Exclusively by Charge-Assisted Hydrogen Bonds. *J. Am. Chem. Soc.* **2007**, *129*, 12114−12116.
(96) Brown, K. N.; Hockless, D. C. R.; Sargeson, A. M. Synthesis and Electrochemistry of $[Pt(tame)_2]^{4+}$: Crystallographic Analysis of bis[1,1,1-Tris(aminomethyl)ethane-N,N′]platinum(II) Bis-(tetrachlorozincate) Dihydrate. *J. Chem. Soc., Dalton Trans.* **1999**, 2171−2176.
(97) Seredyuk, M.; Gaspar, A. B.; Kusz, J.; Gütlich, P. Mononuclear Complexes of Iron(II) Based on Symmetrical Tripodand Ligands: Novel Parent Systems for the Development of New Spin Crossover Metallomesogens. *Z. Anorg. Allg. Chem.* **2011**, *637*, 965−976.
(98) Panagopoulos, A. M.; Zeller, M.; Becker, D. P. Synthesis of an Ortho-Triazacyclophane: *N,N′,N″*-Trimethyltribenzo-1,4,7-Triazacyclononatriene. *J. Org. Chem.* **2010**, *75*, 7887−7892.
(99) Yan, Z.; Li, J.-Y.; Liu, T.; Ni, Z.-P.; Chen, Y.-C.; Guo, F.-S.; Tong, M.-L. Enhanced Spin-Crossover Behavior Mediated by Supramolecular Cooperative Interactions. *Inorg. Chem.* **2014**, *53*, 8129−8135.
(100) Broge, L.; Pretzmann, U.; Jensen, N.; Søtofte, I.; Olsen, C. E.; Springborg, J. Cobalt(II), Nickel(II), Copper(II), and Zinc(II) Complexes with [35]Adamanzane, 1,5,9,13-Tetraazabicyclo[7.7.3]-nonadecane, and [(2.3)2.21]Adamanzane, 1,5,9,12-Tetraazabicyclo[7.5.2]hexadecane. *Inorg. Chem.* **2001**, *40*, 2323−2334.
(101) Hubin, T. J.; McCormick, J. M.; Collinson, S. R.; Buchalova, M.; Perkins, C. M.; Alcock, N. W.; Kahol, P. K.; Raghunathan, A.; Busch, D. H. New Iron(II) and Manganese(II) Complexes of Two Ultra-Rigid, Cross-Bridged Tetraazamacrocycles for Catalysis and Biomimicry. *J. Am. Chem. Soc.* **2000**, *122*, 2512−2522.
(102) Hubin, T. J.; McCormick, J. M.; Collinson, S. R.; Alcock, N. W.; Clase, H. J.; Busch, D. H. Synthesis and X-Ray Crystal Structures of iron(II) and manganese(II) Complexes of Unsubstituted and Benzyl Substituted Cross-Bridged Tetraazamacrocycles. *Inorg. Chim. Acta* **2003**, *346*, 76−86.
(103) Hegetschweiler, K.; Maas, O.; Zimmer, A.; Geue, R. J.; Sargeson, A. M.; Harmer, J.; Schweiger, A.; Buder, I.; Schwitzgebel, G.; Reiland, V.; Frank, W. The Coordination Chemistry of the Pentadentate 2,2,6,6-Tetrakis(aminomethyl)-4-Azaheptane (ditame). *Eur. J. Inorg. Chem.* **2003**, *2003*, 1340−1354.
(104) Thompson, M. C.; Busch, D. H. Reactions of Coordinated Ligands. IX. Utilization of the Template Hypothesis to Synthesize Macrocyclic Ligands in Situ. *J. Am. Chem. Soc.* **1964**, *86*, 3651−3656.
(105) McMurry, T. J.; Raymond, K. N.; Smith, P. H. Molecular Recognition and Metal Ion Template Synthesis. *Science* **1989**, *244*, 938−943.
(106) Fabbrizzi, L.; Licchelli, M.; Mosca, L.; Poggi, A. Template Synthesis of Azacyclam Metal Complexes Using Primary Amides as Locking Fragments. *Coord. Chem. Rev.* **2010**, *254*, 1628−1636.
(107) With proper settings, the presented method can generate cagelike ligands both by construction from scratch and by alteration of previously generated structures.
(108) Heller, S. R.; McNaught, A. D. The IUPAC International Chemical Identifier (InChI). *Chem. Int.* **2009**, *31* (1), 7−9.
(109) Heller, S.; McNaught, A.; Stein, S.; Tchekhovskoi, D.; Pletnev, I. InChI - the Worldwide Chemical Structure Identifier Standard. *J. Cheminf.* **2013**, *5*, 7.
(110) The current generator of InChI keys (from CDK Library 1.4.19) supports only limited sources of stereochemistry. For example, while the cis and trans isomers of $[Fe(diammac)]^{2+}$ (diammac = 6,13-diamino-6,13-dimethyl-1,4,8,11-tetraazacyclotetradecane) return different InChI keys and can be distinguished automatically, the same InChI key is returned for both Δ- and Λ-$[Fe(en)_3]^{2+}$, preventing automated discrimination of these enantiomers.

DOI: 10.1021/acs.jcim.5b00424
J. Chem. Inf. Model. 2015, 55, 1844−1856