


OPEN

A Variable Neighbourhood Descent Heuristic for Conformational Search Using a Quantum Annealer

D. J. J. Marchand¹, M. Noori¹ , A. Roberts¹, G. Rosenberg¹, B. Woods¹, U. Yildiz¹, M. Coons², D. Devore² & P. Margl²

Discovering the low-energy conformations of a molecule is of great interest to computational chemists, with applications in in silico materials design and drug discovery. In this paper, we propose a variable neighbourhood search heuristic for the conformational search problem. Using the structure of a molecule, neighbourhoods are chosen to allow for the efficient use of a binary quadratic optimizer for conformational search. The method is flexible with respect to the choice of molecular force field and the number of discretization levels in the search space, and can be further generalized to take advantage of higher-order binary polynomial optimizers. It is well-suited for the use of devices such as quantum annealers. After carefully defining neighbourhoods, the method easily adapts to the size and topology of these devices, allowing for seamless scaling alongside their future improvements.

The study of molecular structures is foundational to attaining an understanding of chemical processes. Chemical behaviour is determined in large part by the arrangement of atoms within participating molecules as a chemical process unfolds. A widely used approach for studying aspects of molecular structure is to limit the molecular degrees of freedom to torsions only, considering bond lengths and bond angles to be fixed at some values. By doing so, a given molecule with a specific connectivity of its constituent atoms may take on a variety of three-dimensional spatial arrangements, known as *conformational isomers*, or simply *conformations*. Whereas the internal bonds of distinct conformations are the same, the rotation angles around rotatable bonds of the molecule are what distinguish one conformation from another.

In medicinal chemistry, conformational analysis often involves identifying bioactive conformations of ligand molecules^{1,2}. Protein folding is one illustration of function determined by conformation that is both very important to the fields of medicine and biochemistry and extremely difficult to compute, making it a “grand challenge” for science³. In a broader context, conformational analysis is a topic of utmost importance in chemical and materials research. For instance, the conformational behaviour of polymers is key to determining crystallinity, shape, and entanglement of individual chains⁴ that in turn affect macroscopic materials’ properties such as elasticity, strength, or toughness.

Importantly, the geometrical differences between conformations result in different values for the molecular potential energy^{5–7}, which is a key factor for molecular stability and reactivity. To this end, the problem of finding the conformations associated with the local minima of the potential energy surface (PES) of the molecule, referred to as the *conformational search problem*, has been of long-standing interest. The local minima of the PES are often called *conformers*⁸, and the conformer with the lowest energy is referred to as the *global minimum-energy conformer*.

Over time, several practical approaches to the conformational search problem have been developed and critically reviewed^{8–11}. For small molecules, one may be able to deterministically solve the conformational search problem via techniques such as branch and bound^{12–14}. However, this is impractical for even moderately sized molecules, as the size of the conformational search space grows exponentially with the number of rotatable bonds¹⁵. Such increasing difficulty makes the conformational search problem computationally intractable for many of the molecules that have real-world applications and an attractive target for exploration with novel computational technologies and techniques.

¹1QBit Information Technologies (1QBit), 458-550 Burrard Street, Vancouver, BC, V6C 2B5, Canada. ²The Dow Chemical Company, Core R&D, 1776 Building, Midland, MI, 48674, United States. Correspondence and requests for materials should be addressed to M.N. (email: moslem.noori@1qbit.com)

To address the computational complexity presented by larger molecules, many metaheuristic approaches have been studied. Examples of such approaches include genetic algorithms^{16,17}, conformational space annealing^{18,19}, tabu search^{20,21}, molecular dynamics (MD)^{22,23}, and basin/funnel hopping²⁴. Variations of the Monte Carlo (MC) method have also been widely used^{25,26} as a less computationally expensive alternative to MD⁸. In addition, parallel tempering (PT), also known as replica exchange^{27,28}, can be applied to both MC and MD to further improve their sampling performance of the conformational search space.

A comparatively recent approach to address the growing computational complexity of optimization problems relies on the putative future advantage of specialized hardware like quantum annealers to solve binary quadratic optimization problems (see Supplementary Information for more details). One challenge lies in reformulating the optimization problems, a task that often requires approximations or simplifications. Our motivation was to develop such a formulation for the conformational search problem that avoids drastic compromises while producing good conformers.

For this purpose, we propose an iterative heuristic method for the conformational search problem based on *variable neighbourhood descent* (VND). In each iteration of the method, we use the molecular structure to choose specialized conformational neighbourhoods that can be minimized efficiently. More specifically, using the structural graph of a given molecule, subsets of rotatable bonds are selected at each iteration. Fixing the values of other torsion angles, the problem of minimizing the molecular energy with respect to the selected torsion angles is then formulated as a binary program with an objective function that is a polynomial of a chosen degree. This allows the method to be adapted to the specifics of the optimizer by limiting the degree of the binary program. The values of the selected torsion angles are then set to the solution of this binary program before starting each subsequent iteration during which a new subset of rotatable bonds are optimized. The process continues until some stopping criteria are met.

Although the method can be readily extended to any chosen degree, we assume in this paper that a binary quadratic program is desirable as it is well-suited for optimization using quantum annealing^{29,30}. Furthermore, by changing the parameters of the neighbourhood selection procedure, the method can be easily adapted to the size of the conformational search problem in terms of the number of rotatable bonds, as well as the size and connectivity of available quantum annealers. The flexibility of solving any conformational search problem using currently available quantum annealers, without imposing restrictions on the granularity of the conformational space, differentiates our work from a previous study on protein folding using quantum annealing³¹.

We evaluate the performance of our proposed algorithm over three families of molecules relevant to industry, using an algorithm that returns an exact optimal solution and the D-Wave 2000Q quantum annealer^{32–34}. The latter provides an assessment using the latest available hardware at the time of writing of this work, whereas the former can be seen as a limiting ideal case. For each molecule, we compare the lowest-energy conformations found by our algorithm with those found by both parallel tempering MC (PTMC)^{27,35} and a simple local search method.

Preliminaries

We give some preliminaries before presenting the details of our proposed conformational search method.

Problem definition. As discussed above, we consider the conformational search problem as a special case of molecular structure analysis, where the structure is kept fixed except for rotations around selected bonds. Each of these torsional degrees of freedoms we hereafter refer to as a *torsion* for simplicity. We denote the i -th rotatable bond by T_i and assign its rotational angle a variable t_i , with i representing the torsion index. It is convenient to identify a conformation of a molecule with M torsions by a *torsion vector* $\mathbf{t} = [t_1, \dots, t_M]$. Without loss of generality, we assume $t_i \in [0, 2\pi)$, for all i , knowing that the method remains unchanged if each torsion has its own range chosen based on prior knowledge, experimental data, or known symmetries. For simplicity, let us assume all torsion angle values are chosen from the same set of d values $\Theta = \{\theta_1, \dots, \theta_d\}$. The theoretical precision of this discretization scheme increases with d , while the size of the search space d^M grows exponentially with the number of torsions.

Although it is natural to describe a molecule using a molecular graph, where the atoms and their bonds are represented by vertices and edges, respectively, we find it helpful to use the torsions to partition the molecule into $M + 1$ subsets called *rigid bodies*. The partitioning is performed such that all atoms within a rigid body are interconnected through non-torsion bonds. As a result, the relative positions of the atoms within a rigid body, denoted by R_r , remain invariant under rotation and are therefore independent of \mathbf{t} . This simplified representation of the molecule is now easily described by a *rigid-body graph* $G = (\mathcal{R}, \mathcal{T})$, where \mathcal{R} is the set of $M + 1$ vertices and \mathcal{T} is the set of M edges. In G , each vertex represents a rigid body and each edge represents a rotatable bond. Two vertices are connected by an edge if their associated rigid bodies are connected by the rotatable bond that the edge represents. We will therefore use T_i to refer to both torsion i and its associated edge in the rigid-body graph. We further assume that each torsion is free to rotate independently of others, thus restricting the presence of ring systems or other cycles in the molecular graph to individual rigid bodies. Under this assumption, the rigid-body graph has no cycles and is a tree. An example of a simple molecule and its rigid-body graph is shown in Fig. 1.

The search space of the conformational search problem is a hypersurface described by an energy model (function) $U: \Theta^M \rightarrow \mathbb{R}$. For a given \mathbf{t} , $U(\mathbf{t})$ is the molecular energy consisting of the sum of all interatomic potentials in the molecule (e.g., van der Waals, torsional, and bending), which are dependent on the relative coordinates of the atoms. Various energy models or effective *force fields* can be used for our purpose, such as the widely used “Universal force field” (UFF)³⁶. The conformational search problem, with the objective of finding the global minimum-energy conformer of a given molecule, can then be formulated as

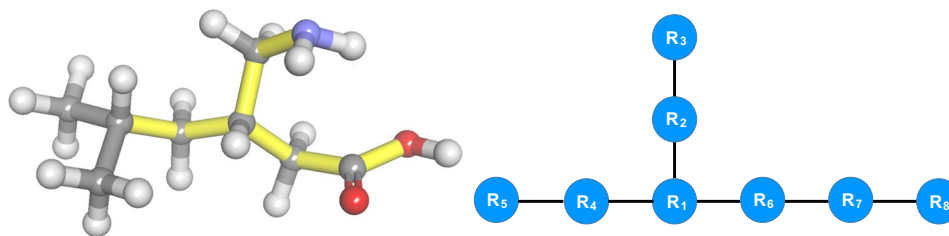


Figure 1. Illustration of the rigid-body graph for a simple organic molecule (pregabalin). The molecular structure, with seven rotatable bonds highlighted in yellow, is shown on the left. The rotatable bonds connected to the two methyl groups have been discarded for visual clarity. The associated rigid-body graph is depicted on the right.

$$\begin{aligned} \min_{\mathbf{t}} \quad & U(\mathbf{t}) \\ \text{s.t.} \quad & \mathbf{t} \in \Theta^M. \end{aligned} \quad (1)$$

Upon changing the torsion angles of the molecule, some of the interatomic potential contributions will remain unchanged, while other contributions will change depending on the torsion angle vector \mathbf{t} . To be more specific, let us denote the values of the torsion angles on the path connecting R_a to R_b on the rigid-body graph by a vector \mathbf{t}_{ab} . The length of this path is represented by m_{ab} , meaning that \mathbf{t}_{ab} has m_{ab} elements. The relative positions of the atoms in R_a with respect to the atoms in R_b depend only on the torsions on this path. Now,

$$U(\mathbf{t}) = \sum_{a:R_a \in \mathcal{R}} U_a + \sum_{\substack{a,b:R_a,R_b \in \mathcal{R} \\ a \neq b}} U_{ab}(\mathbf{t}_{ab}), \quad (2)$$

where U_a is the sum of the interatomic potentials of the atoms within rigid body R_a , which is invariant under rotation, while $U_{ab}: \Theta^{m_{ab}} \rightarrow \mathbb{R}$ such that for a given \mathbf{t}_{ab} , $U_{ab}(\mathbf{t}_{ab})$ is the sum of the interatomic potentials of all pairs of atoms where one atom is in R_a and the other is in R_b .

Binary optimization formulation for the conformational search problem. In order to use a quantum annealer to solve the conformational search problem, one needs to reformulate it as a quadratic unconstrained binary optimization (QUBO) problem (see the Supplementary Information for more details). To this end, we start by applying a one-hot encoding to the discrete values of the torsion angles, establishing a mapping between the torsion angle vector space and a binary solution space. That is, for each t_i , we assign a binary variable x_{ik} , $1 \leq k \leq d$, such that

$$x_{ik} = \begin{cases} 1 & \text{if } t_i = \theta_k; \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

As a result, t_i can be expressed as

$$t_i = \sum_{k=1}^d \theta_k x_{ik}, \quad (4)$$

where we add a constraint

$$\sum_{k=1}^d x_{ik} = 1 \quad (5)$$

to ensure t_i takes one and only one value at a time. The constraint (5) is commonly referred to as a one-hot encoding constraint. Note that after applying the one-hot encoding, any arbitrary function $f(t_i)$ can be written as

$$f(t_i) = \sum_{k=1}^d f(\theta_k) x_{ik}. \quad (6)$$

Similar to Eq. (6), a binary representation for $U(\mathbf{t}_{ab})$ can be found. For simplicity of presentation, let us assume that the torsion angles in \mathbf{t}_{ab} are indexed sequentially from 1 to m , that is, $\mathbf{t}_{ab} = [t_1, t_2, \dots, t_m]$. Now,

$$U_{ab}(\mathbf{t}_{ab}) = \sum_{k_1=1}^d \cdots \sum_{k_m=1}^d U_{ab}(t_1 = \theta_{k_1}, \dots, t_m = \theta_{k_m}) \prod_{i=1}^m x_{ik_i}, \quad (7)$$

where $U_{ab}(t_1 = \theta_{k_1}, \dots, t_m = \theta_{k_m})$ can be pre-evaluated for all possible d^m values of \mathbf{t}_{ab} .

Input: Initial solution \mathbf{t}_{init}

$\mathbf{t}_c \leftarrow \mathbf{t}_{\text{init}}$ ▷ Initialize the current solution

Select an arbitrary $N_k \in \{N_1, \dots, N_K\}$

while termination criteria have not been met **do**

$\mathbf{t}' \leftarrow \arg \min_{\mathbf{t} \in N_k(\mathbf{t}_c)} U(\mathbf{t})$ ▷ Find the best neighbour in $N_k(\mathbf{t}_c)$

if $U(\mathbf{t}') < U(\mathbf{t}_c)$ **then**

$\mathbf{t}_c \leftarrow \mathbf{t}'$ ▷ Update the current solution

end if

$N_k \leftarrow \text{NeighbourhoodChange}()$ ▷ Change neighbourhoods

end while

Table 1. Variable neighbourhood descent method.

Substituting $U_{ab}(\mathbf{t}_{ab})$ from Eq. (7) into Eq. (2) results in a representation of the molecular energy $U(\mathbf{t})$ in terms of the binary vector \mathbf{x} . We denote this representation of the energy function by $E: \{0, 1\}^{Md} \rightarrow \mathbb{R}$ and write the conformational search problem (1) as

$$\begin{aligned} \min_{\mathbf{x}} \quad & E(\mathbf{x}) \\ \text{s.t.} \quad & \sum_{k=1}^d x_{ik} = 1, \quad \forall i \in \{1, \dots, M\}, \\ & \mathbf{x} \in \{0, 1\}^{Md}. \end{aligned} \quad (8)$$

To solve the above binary optimization problem using a quantum annealer, one faces three challenges. First, the objective function in formulation (8) is not necessarily quadratic as $U_{ab}(\mathbf{t}_{ab})$ may depend on more than two torsions, that is, $m_{ab} > 2$. Second, it is a constrained binary optimization problem. These two challenges indicate that the problem cannot be solved directly on a quantum annealer. The third challenge is that if $\max_{a,b:R_a,R_b \in \mathcal{R}} m_{ab}$ is not much smaller than M , constructing an instance of formulation (8) becomes very computationally expensive due to the pre-evaluation of the coefficients in Eq. (7). In the following section, we propose a method that addresses these challenges in order to be able to use a quantum annealer for solving the conformational search problem.

Variable Neighbourhood Descent for the Conformational Search Problem

Neighbourhood search, or local search (LS), is known to be an effective heuristic algorithm for solving a large number of combinatorial optimization problems. In defining a neighbourhood relation between solutions of a problem, local search begins from an initial solution and iteratively explores the neighbourhood of the current solution for improvement. It has been shown that a solution produced by a local search algorithm will often not be globally optimal, but will be suboptimal with respect to another neighbourhood relation³⁷. When multiple neighbourhood relations are considered, the algorithm is often referred to as *variable neighbourhood search*³⁸. In the context of conformational search, a solution refers to a vector of torsion angles \mathbf{t} .

Let N_k , for $k \in \{1, \dots, K\}$, denote a finite set of neighbourhood structures and $N_k(\mathbf{t})$ be the set of all solutions in the k -th neighbourhood of \mathbf{t} . Starting from an initial solution and a neighbourhood structure, in each iteration, *variable neighbourhood descent* finds the best solution in the neighbourhood of the current solution. It then updates the current solution with the best solution found, and changes the neighbourhood structure before proceeding with the next iteration. The VND method is summarized in Table 1.

A simple neighbourhood structure is obtained by considering two solutions as neighbours if and only if they differ by exactly one torsion angle value. Such a VND heuristic is exactly the LS heuristic described above. While computationally inexpensive, the performance of LS can suffer in cases where a decrease in the molecular energy cannot be achieved by changing only a single torsion angle value in an iteration. Our proposed VND method improves upon LS by exploring more-complex neighbourhoods. In the following, we describe the components of the method.

Initial Solution. The initial solution in Table 1 can be selected in a variety of ways. One may simply choose a randomly generated torsion angle vector for the given molecule as the initial solution. Alternatively, one can use a greedy construction method. Another approach is to start from a known high-quality solution. This applies when using our VND method in conjunction with another conformational search method or by exploiting some prior knowledge about a given molecule Table 2.

Neighbourhood Structures. We now describe a more powerful neighbourhood structure which, to our knowledge, has not been previously studied. Let $G = (\mathcal{R}, \mathcal{T})$ be a rigid-body graph $\mathcal{T} \subseteq \mathcal{T}$, and G' be the graph resulting from contracting all edges in $\mathcal{T} \setminus \mathcal{T}$ (see Fig. 2 for an example). If G' is a star graph (a tree graph with at most one vertex of degree > 1), then we say that \mathcal{T} has the property of *2-torsion dependency*. The motivation for using this terminology is that any two vertices of G' are connected with at most two edges (torsions). We label the maximal 2-torsion-dependent subsets of \mathcal{T} as $\mathcal{T}_1, \dots, \mathcal{T}_K$ and their associated star graphs as G_1, \dots, G_K . Then, the neighbourhood structure N_k defines neighbourhoods containing all solutions which differ only in torsion angle values corresponding to edges in \mathcal{T}_k , for $k = 1, \dots, K$. Solutions in neighbourhoods defined by an arbitrary N_k are also called *neighbours under N_k* .

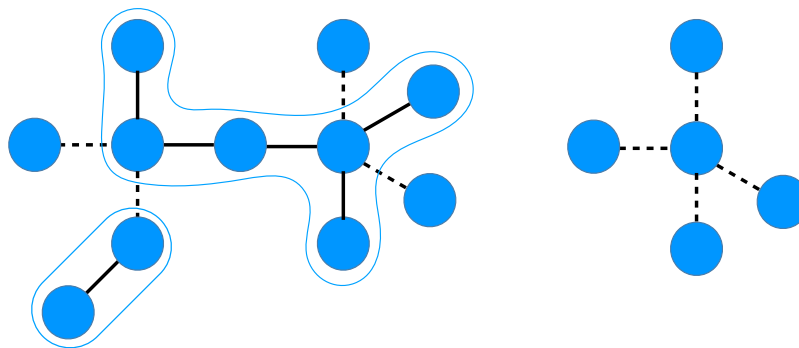


Figure 2. Example of a 2-torsion-dependent set of edges in a rigid-body graph (left). The star graph G' (right) results from selecting a 2-torsion-dependent set \mathcal{T} (dashed lines) and contracting all edges not in \mathcal{T} .

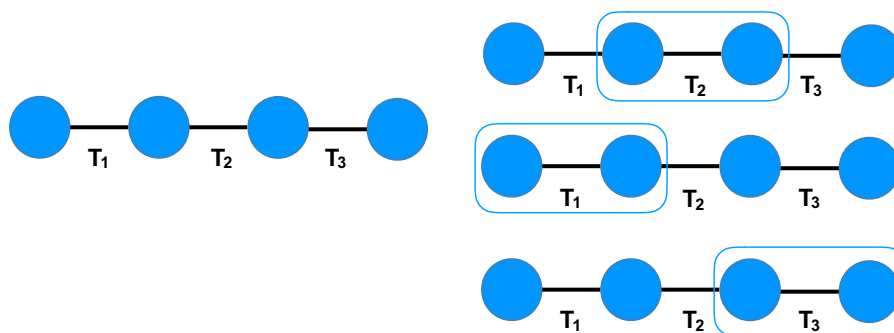


Figure 3. Example of finding different neighbourhood structures of a rigid-body graph. The original rigid-body graph, $G = (\mathcal{R}, \mathcal{T})$, is shown on the left. All neighbourhood structures defined by the maximal 2-torsion-dependent subsets of \mathcal{T} are depicted on the right.

Neighbourhood structures are illustrated for an example rigid-body graph $G = (\mathcal{R}, \mathcal{T})$ in Fig. 3 that represents a molecule with three torsions. The maximal 2-torsion-dependent subsets of the torsions are $\mathcal{T}_1 = \{T_1, T_3\}$, $\mathcal{T}_2 = \{T_2, T_3\}$, and $\mathcal{T}_3 = \{T_1, T_2\}$. The neighbourhood structure N_1 defines neighbourhoods containing all torsion vectors that differ only in torsion angle values for T_1 and T_3 . For example, $\mathbf{t} = [5^\circ, 10^\circ, 20^\circ]$ and $\mathbf{t}' = [0^\circ, 10^\circ, 90^\circ]$ are neighbours under N_1 while $\mathbf{t} = [5^\circ, 10^\circ, 20^\circ]$ and $\mathbf{t}' = [5^\circ, 15^\circ, 20^\circ]$ are not because they differ in the torsion angle value for T_2 .

Neighbourhood Search. Based on the discussion in Section 2.2, the problem of finding the best solution in $N_k(\mathbf{t})$ can be formulated as a QUBO problem by restricting the optimization problem in (8) to the binary variables corresponding to the torsions in \mathcal{T}_k and moving the one-hot encoding constraints to the objective function using the quadratic penalty method as follows:

$$\begin{aligned} \min \quad & \sum_{i,j:T_i,T_j \in \mathcal{T}_k} \sum_{k_i=1}^d \sum_{k_j=1}^d U_{ij}(\theta_{k_i}, \theta_{k_j}) x_{ik_i} x_{jk_j} \\ & + \sum_{i:T_i \in \mathcal{T}_k} \sum_{k_i=1}^d U_i(\theta_{k_i}) x_{ik_i} + p \sum_{i:T_i \in \mathcal{T}_k} \left(\sum_{k_i=1}^d x_{ik_i} - 1 \right)^2, \\ \text{s. t. } \quad & x_{ik_i} \in \{0, 1\}, \quad \forall i, k_i. \end{aligned} \quad (9)$$

Here, the $U_{ij}(\theta_{k_i}, \theta_{k_j})$ terms represent the interaction energy of the two vertices of G_k that are connected by T_i and T_j when $t_i = \theta_{k_i}$ and $t_j = \theta_{k_j}$. The $U_i(\theta_{k_i})$ terms represent the interaction energy of the two vertices connected by T_i on G_k when $t_i = \theta_{k_i}$ and p is a sufficiently large penalty coefficient that enforces the one-hot encoding constraints. The above QUBO problem can be solved using various methods^{39–41}, as well as specialized hardware devices such as quantum annealers.

Neighbourhood change. At each iteration, the neighbourhood is selected based on a random ordering, $\phi: \mathcal{T} \rightarrow \mathcal{T}$, of the torsions. The pseudocode for the neighbourhood change function is given below.

```

Generate a random ordering,  $\phi = (\phi(1), \phi(2), \dots, \phi(M))$ , of  $\mathcal{T}$ 
 $\mathcal{T}' \leftarrow \emptyset$ 
for  $i = 1, \dots, M$  do
  if  $\mathcal{T}' \cup \{\phi(i)\}$  is 2-torsion dependent then
     $\mathcal{T}' \leftarrow \mathcal{T}' \cup \{\phi(i)\}$  ▷ Add  $\phi(i)$  to  $\mathcal{T}'$ 
  end if
end for
return  $N_k$  corresponding to  $\mathcal{T}'$ 

```

Table 2. Neighbourhood change procedure.

Practical considerations. Here, we detail the practical considerations of the proposed VND method to accommodate the use of existing and future quantum annealers.

The formulated QUBO problem in (9) is fully connected, meaning that for all i and j ($i \neq j$), the term $x_{ik_i}x_{jk_j}$ appears in the objective function. On the other hand, the connectivity of the qubits on the D-Wave 2000Q follows a “Chimera graph”; thus, the problem in (9) must be embedded onto the hardware graph using an embedding strategy⁴². There is a limit on the number of variables a fully connected QUBO problem that can be embedded onto the graph can have. We take this limitation into consideration by imposing a limit on the number of variables in the formulated QUBO problem for each selected subset of torsion angles \mathcal{T}_k . We denote this parameter by s . In the following, we explain how this limit is imposed on the formulated QUBO problem.

For the selected neighbourhood structure at each iteration, we (randomly) select a total of s discrete values. That is, for each T_i in \mathcal{T}_k , we randomly choose a set of $\Theta_i \subseteq \Theta$ discrete values such that $s = \sum_{i: T_i \in \mathcal{T}_k} |\Theta_i|$. We have already defined the $N_k(\mathbf{t})$ neighbourhood as the set of torsion angle vectors \mathbf{t}' that are different from \mathbf{t} only in the angle values of the torsions in \mathcal{T}_k . In addition, for any T_i in \mathcal{T}_k , t'_i (i.e., the value associated with T_i in \mathbf{t}') takes on values only from Θ_i .

With the above choice of neighbours, for an arbitrary \mathbf{t} , $N_k(\mathbf{t})$ contains

$$S_k = \prod_{i: T_i \in \mathcal{T}_k} |\Theta_i| \quad (10)$$

solutions. However, to find the best solution in $N_k(\mathbf{t})$, we need to pre-evaluate only

$$\sum_{\substack{i, j: T_i, T_j \in \mathcal{T}_k \\ i \neq j}} |\Theta_i| |\Theta_j| + s \quad (11)$$

energy terms to formulate the QUBO problem. As seen above, S_k grows linearly with the product of $|\Theta_i|$, for all i , whereas the growth of (11) is quadratic. This means that the number of energy pre-evaluations grows more slowly than the size of the neighbourhood as $|\mathcal{T}_k|$ increases.

Another practical consideration for our proposed VND method is the stopping criteria. The first stopping criterion sets a limit on the computational effort of the method by introducing a maximum number of iterations, denoted by B . The second stopping criterion aims to terminate the method early if it becomes stuck at a local minimum or finds the global minimum of the problem. For this purpose, we introduce a parameter called A that represents the maximum number of consecutive iterations to can be performed without decreasing the energy.

With the above-mentioned practical considerations, the implemented VND method is summarized in Table 3.

Two important features of the VND method is its scalability and ease of adaptation to the size and connectivity of the quantum annealer. More specifically, one can simply increase s in the described method to take advantage of improvements in the number of qubits and their connectivity. The rest of the method remains intact. It is worth mentioning that although the focus here has been on quantum annealers, other QUBO problem solvers could be used.

Effect of the molecular structure. As previously discussed, due to the limitations of current quantum annealers, we are able to jointly optimize only a subset of torsions that are 2-torsion dependent at each iteration of VND. The potential speedup of the quantum annealer over a naïve exhaustive QUBO problem solver depends on the cardinality of the selected subset of torsions.

Since $\sum_{i: T_i \in \mathcal{T}_k} |\Theta_i|$ is fixed, the number of solutions in $N_k(\mathbf{t})$, S_k , is maximized when the cardinality of each Θ_i is the same. That is, when $|\Theta_i| = s/|\mathcal{T}_k|$ for all $i \in \{1, \dots, |\mathcal{T}_k|\}$, resulting in the maximum number of solutions

$$S_k = \left(\frac{s}{|\mathcal{T}_k|} \right)^{|\mathcal{T}_k|}. \quad (12)$$

S_k is an increasing function of $|\mathcal{T}_k|$ if $s/|\mathcal{T}_k| \geq e$, where e is the base of the natural logarithm.

As the potential speedup of the quantum annealer over an exhaustive QUBO solver is dependent on S_k , it is favourable to select each \mathcal{T}_k with a large cardinality. The speedup diminishes when $|\mathcal{T}_k| = 2$, which is the case for molecules with linear rigid-body graphs. The above discussion suggests that star-like molecules that have rigid-body graphs with high-degree nodes stand to benefit more (than those that do not) from using a quantum annealer to solve the QUBO problem at each VND iteration.

Input: A random initial conformation of the molecule, denoted by a torsion vector \mathbf{t}_{init}

```

 $a, b \leftarrow 0$                                 ▷ Initialize the counters
Select an arbitrary  $N_k \in \{N_1, \dots, N_K\}$     ▷ Initialize the neighbourhood structure
 $\mathbf{t}_c \leftarrow \mathbf{t}_{\text{init}}$                     ▷ Initialize the current solution  $\mathbf{t}_c$ 
while  $a < A$  and  $b < B$  do
  Formulate the QUBO problem to find the best solution in  $N_k(\mathbf{t}_c)$ 
   $\mathbf{t}' \leftarrow \operatorname{argmin}_{\mathbf{t} \in N_k(\mathbf{t}_c)} U(\mathbf{t})$     ▷ Find the best solution in  $N_k(\mathbf{t}_c)$ 
  if  $U(\mathbf{t}') < U(\mathbf{t}_c)$  then
     $\mathbf{t}_c \leftarrow \mathbf{t}'$                                 ▷ Update the current solution
     $b \leftarrow 0$                                     ▷ Reset the no-improvement counter
  else
     $b \leftarrow b + 1$                                 ▷ Update the no-improvement counter
  end if
   $a \leftarrow a + 1$                                     ▷ Update the counter
   $N_k \leftarrow \text{NeighbourhoodChange}()$             ▷ Change neighbourhoods
end while

```

Table 3. Conformational search VND method.

Experimental Results

In this section, we evaluate the performance of the proposed VND method. We first provide details about the molecules used in our experiments and then present the experimental results.

Trial molecules. The performance of VND was evaluated using a testbed containing the nine molecules depicted in Fig. 4. These model systems include three organometallic molecules useful for catalyzing reactions relevant to industry (labelled “A”^{43,44}, “B”^{45,46}, and “C”^{47,48}), a set of three n-alkanes whose basic structural motif appears in fuels, lubricants, solvents⁴⁹, and resins⁵⁰ (labelled by D, E, and F), as well as a set of three ortho-phenylene oligomers that are of interest as electronic materials and nanomaterials (labelled “G”, “H”, and “I”)^{51,52}.

The choice of model systems A–I was motivated by several considerations. First, it is important to show that our method can be applied to a wide variety of conformational search problems of relevance to industry. Second, the model systems are representative of a diversity of molecular graphs: systems A–C have star-like graphs, whereas D–I have linear graphs, albeit with different structures. Third, the model systems represent a significant variety of active torsions, ranging from very modest (e.g., A) to very substantial (e.g., F and I). Another motivation for choosing these systems was the existence of experimental data pertaining to their three-dimensional structure (see the Supplementary Information for details).

Energy model. In our experiments, the Lennard-Jones 6–12 potential is used to model the interaction energy of two atoms α and β as

$$V(\alpha, \beta) = \varepsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right)^{12} - 2 \left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right)^6 \right], \quad (13)$$

where $\varepsilon_{\alpha\beta}$ is the depth of the potential well, $\sigma_{\alpha\beta}$ is the van der Waals bond length, and $r_{\alpha\beta}$ is the distance between the two atoms. An expression for Eq. (2) is then obtained by summing over all pairs of atoms. Values for the parameters $\varepsilon_{\alpha\beta}$ and $\sigma_{\alpha\beta}$ are taken from the UFF³⁶ without modification. One should note that while we use the Lennard-Jones potential in what follows to provide a proof of concept for the proposed VND method, other force fields, that allow for representing the molecular energy as Eq. (2), can also be used.

Results. We present the results of solving the conformational search problem for the selected molecules using the proposed VND method. Since the D-Wave 2000Q quantum annealer is not guaranteed to find an optimal solution, we first present results for VND, where an exact QUBO problem solver is used. These results focus solely on the performance of the VND method as a conformational search approach by preventing any deterioration of the results attributable to the use of a device that has imperfections. For comparison purposes, we also present the results achieved by performing a local search heuristic (LS), and a hybrid of the two methods (LS–VND) in which a random conformation is first optimized using LS and then passed to VND for further optimization. The comparison with LS is helpful in understanding the improvement that can be achieved through the use of a more complex neighbourhood than what is used for LS. We also present the VND results with the quantum annealer used as the underlying QUBO problem solver. Finally, we compare the results with those found by our implementation of the PTMC algorithm for the conformational search problem, a state-of-the-art metaheuristic for conformational sampling. Details of the parameters used for VND, the quantum annealer, and PTMC are presented in the Supplementary Information.

Reference Conformations. We compare our results against *reference* conformations found using a PTMC conformational search method. More specifically, we use an *initial* conformation for each of the selected molecules as input to a PTMC algorithm and let it run with a sufficiently large number of sweeps, such that the resulting

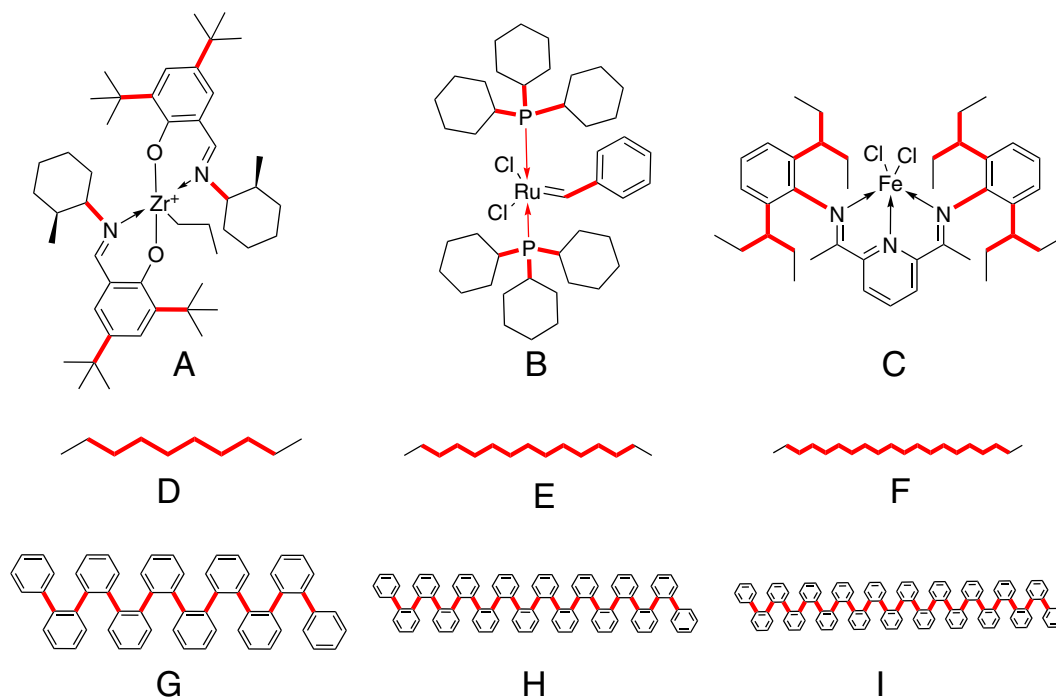


Figure 4. Graphical representation of the six model systems studied: three organometallic compounds (A–C), three n-alkanes (D–F), and three ortho-phenylenes (G–I). The thick (red) lines represent the torsion bonds. See the Supplementary Information for additional details about each model system.

reference geometry can be assumed to represent the absolute minimum conformation with a high degree of confidence. Details of this procedure are given in the Supplementary Information. To ensure the fairness and accuracy of this comparison, the reference conformations are generated using the same potential energy model as the one we used for VND.

Performance Metrics. The following metrics have been used for performance evaluation.

- Success rate: the fraction of runs that found a conformation with an energy within 1 kcal/mol (roughly chemical accuracy) of the reference conformation's energy.
- Number of energy evaluations: the number of molecular energy evaluations needed to arrive at the best found conformation, averaged over all runs.
- Residual: the energy difference between the best conformation found in a run and the reference conformation. The normalized residual, when reported, refers to the ratio of the residual to the number of atoms in the system.
- Time to solution (TTS): the time it took to find the best solution (conformation) in a single run.

Each run was terminated once it found a conformation within 0.1 kcal/mol of the reference conformation, or if it reached some other stopping criterion.

VND vs. LS vs. LS–VND. As a baseline for comparison, we used an exact QUBO problem solver to optimize the selected set of torsions at each VND iteration in runs of both VND and LS–VND. The success rate, residual, and TTS for each of the three methods, for all nine model systems, are presented in Table 4. Here, the TTS for VND includes the time needed for selecting the neighbourhoods, formulating their associated QUBO problems, solving the QUBO problems with an exact solver, and translating the solutions of the QUBO problems back to torsion vectors. The number of energy evaluations reported for the VND method is based on the number of energy evaluations required to find the $U_{ij}(\theta_{k_i}, \theta_{k_j})$ and $U_i(\theta_{k_i})$ coefficients of the QUBO problem in (9). As shown, whereas LS is faster than VND and LS–VND, its success rate and residual results are generally inferior to the other two methods. Our implementation of LS terminates when there is no neighbouring solution which has a lower energy value, ensuring that the algorithm terminates at a locally optimal solution with respect to the single-torsion neighbourhood. Further, since LS–VND combines both the LS and VND methods, it is not surprising that its success probability and residual are generally at least as good as those of LS and VND.

VND Using a Quantum Annealer. The results for VND using a quantum annealer as the underlying solver are presented in Table 5. Here, the TTS includes the time spent on all pre- and postprocessing steps needed to solve the QUBO problems on the quantum annealer. Note that the actual time spent using the quantum

Method	Model system	Success rate	Num. energy evaluations	Normalized residual			TTS (seconds)		
				Min	50th	75th	Min	50th	75th
VND	A	1.00	4.6×10^4	0.00	0.00	0.00	19.7	90.2	110.2
	B	0.60	1.42×10^5	0.00	0.01	0.01	90.9	271.6	320.3
	C	0.04	2.11×10^5	0.00	0.13	0.16	127.3	418.7	498.3
	D	1.00	1.96×10^4	0.00	0.00	0.00	3.0	10.0	13.8
	E	1.00	4.22×10^4	0.00	0.00	0.00	11.7	25.9	30.7
	F	1.00	7.25×10^4	0.00	0.00	0.00	34.0	64.1	78.5
	G	0.16	1.23×10^5	0.00	0.14	0.21	37.0	101.1	153.5
	H	0.00	1.69×10^5	0.02	0.25	0.31	67.5	211.6	243.4
	I	0.00	1.87×10^5	0.01	0.30	0.38	155.7	374.1	391.4
LS	A	1.00	1.45×10^4	0.00	0.00	0.00	7.8	9.6	10.6
	B	0.03	1.01×10^5	0.01	0.28	0.39	35.0	69.7	91.6
	C	0.02	2.47×10^5	0.00	0.16	0.21	84.4	162.4	198.9
	D	1.00	2.21×10^4	0.00	0.00	0.00	5.7	8.9	9.9
	E	1.00	7.17×10^4	0.00	0.00	0.00	26.5	38.4	41.8
	F	0.99	1.53×10^5	0.00	0.00	0.00	70.9	105.2	115.7
	G	0.03	1.87×10^5	0.00	0.15	0.18	36.3	112.9	148.2
	H	0.00	4.76×10^5	0.05	0.28	0.37	168.5	468.1	580.7
	I	0.00	8.10×10^5	0.09	0.29	0.39	354.6	784.2	1078.2
LS-VND	A	1.00	1.45×10^4	0.00	0.00	0.00	7.5	10.3	11.4
	B	0.44	2.23×10^5	0.00	0.01	0.13	68.9	294.5	421.7
	C	0.04	3.58×10^5	0.00	0.12	0.16	135.6	305.8	474.3
	D	1.00	2.16×10^4	0.00	0.00	0.00	5.7	8.7	10.0
	E	1.00	7.19×10^4	0.00	0.00	0.00	25.8	38.1	40.9
	F	1.00	1.48×10^5	0.00	0.00	0.00	76.6	101.0	109.3
	G	0.17	2.30×10^5	0.00	0.15	0.18	26.3	142.0	199.0
	H	0.03	5.88×10^5	0.00	0.23	0.31	272.6	576.6	742.7
	I	0.02	8.42×10^5	0.00	0.30	0.39	435.1	1048.3	1284.8

Table 4. Results for VND, LS, and LS-VND, when using an exact solver as the underlying QUBO solver. For each model system we report the minimum, median, and 75th percentile of the residual and TTS, over 100 runs.

annealer, referred to as the annealing time, is much smaller than the reported TTS. In addition, to aid in visually assessing the found lowest-energy conformations, their graphical representations are provided in the Supplementary Information. A comparison of these results with the results of VND when using an exact QUBO problem solver (see Table 4) shows that those of the former are of lower quality. It is expected, however, that spending more effort on tuning the quantum annealer's parameters would improve these results (See Supplementary Information for more details). Another observation from Table 5 is that VND employing a quantum annealer to solve QUBO problems has a larger TTS than VND using an exact QUBO problem solver. The quantum annealer solves a QUBO problem much faster than the exact solver, so one might think that its TTS should also be lower.

To explain this observation, one should note that it is not sufficient to solve QUBO problems merely fast, because if they are not solved optimally by the quantum annealer, the VND method may take longer to converge. Further, a significant portion of time is spent on transforming the QUBO problems into Ising problems (see the Supplementary Information), communicating with the quantum annealer, and mapping the results from the quantum annealer back to the logical bits. It is worth noting that elapsed real time is a fair measure of the time required to solve actual problems using a quantum annealer. This is in contrast to the customary approach of reporting only the annealing time, which is very small in comparison (on the order of a few microseconds per annealing cycle).

There are several known factors that make the QUBO problems generated at the VND iterations challenging for the current generation of quantum annealers. First, the problem graphs are fully connected, whereas the quantum annealer's connectivity graph is extremely sparse, resulting in each logical bit being embedded onto chains of 17 qubits on the quantum annealer. It is difficult to maintain identical states for those qubits, resulting in a higher error rate. Second, these problems have a very large range of coefficients, due to the $1/r^{12}$ and $1/r^6$ terms in the Lennard-Jones 6-12 energy model. On the other hand, the couplers of the existing quantum annealers have a limited bit precision and a fixed range. The large range of coefficients results in a loss of precision, which manifests itself in a lower success probability. Third, the current generation of quantum annealers has a high level of noise, referred to as intrinsic control error (ICE), leading to a significant loss in precision. Future quantum annealers are expected to mitigate these factors, with more-dense hardware graphs, higher bit precision, and lower ICE levels.

Model system	Success rate	Num. energy evaluations	Normalized residual			TTS (seconds)		
			Min	50th	75th	Min	50th	75th
A	0.00	4.56×10^4	0.01	0.05	0.06	156.1	330.2	366.4
B	0.00	4.99×10^4	0.12	0.47	0.76	136.0	288.9	395.5
C	0.00	9.03×10^4	0.14	0.38	0.62	300.9	507.3	726.9
D	1.00	3.86×10^4	0.00	0.00	0.01	299.0	516.4	608.0
E	0.72	8.24×10^4	0.00	0.01	0.02	235.8	639.1	815.6
F	0.60	1.17×10^5	0.00	0.02	0.02	533.7	961.4	1038.8
G	0.00	9.28×10^4	0.08	0.24	0.34	385.8	788.3	990.1
H	0.00	1.27×10^5	0.19	0.43	0.64	445.7	982.2	1107.3
I	0.00	1.7×10^5	0.25	0.52	0.75	511.1	1221.2	1648.8

Table 5. Results for VND using a quantum annealer as the underlying QUBO solver. For each model system we report the minimum, median, and 75th percentile of the residual and TTS, over 25 runs.

Model system	Success rate	Num. energy evaluations	Normalized residual			TTS (seconds)		
			Min	50th	75th	Min	50th	75th
A	1.00	1.12×10^5	0.00	0.00	0.00	40.1	107.5	143.3
B	1.00	3.83×10^6	0.00	0.00	0.00	301.0	787.5	1092.8
C	1.00	3.19×10^6	0.00	0.00	0.00	585.5	2068.3	3025.0
D	1.00	4.05×10^4	0.00	0.00	0.00	52.1	87.9	108.0
E	1.00	1.03×10^5	0.00	0.00	0.00	195.6	290.5	326.3
F	1.00	2.00×10^5	0.00	0.00	0.00	457.4	685.3	776.0
G	1.00	2.52×10^6	0.00	0.00	0.00	154.8	710.4	1085.0
H	0.60	5.54×10^6	0.00	0.00	0.05	1051.4	4711.6	5975.8
I	0.12	3.58×10^6	0.00	0.07	0.09	3025.9	5758.3	6631.8

Table 6. Results for PTMC. For each model system, we report the minimum, median, and 75th percentile of the residual and TTS, over 100 runs.

Neighbourhood Size (s)	Model System	Success Rate	Residual (kcal/mol)		
			Min	50th	75th
30	B	0.07	0.2	16.7	26.0
	C	0.00	1.1	19.4	25.6
60	B	0.39	0.0	1.3	1.9
	C	0.04	0.1	13.4	17.1
90	B	0.55	0.0	0.9	1.5
	C	0.09	0.0	8.7	14.8
120	B	0.58	0.0	0.9	1.5
	C	0.10	0.1	8.7	14.1

Table 7. Effect of the neighbourhood size on the performance of VND. For each model system, we report the minimum, median, and 75th percentile of the residual, over 500 runs.

Comparison with PTMC. The results for PTMC are presented in Table 6. PTMC had a high success rate for all model systems except for system F. It is worth noting that PTMC's TTS and number of energy evaluations are generally significantly higher than that of VND.

Effect of the Neighbourhood Size on VND's Performance. In the VND experiments discussed above, we restricted the neighbourhood size, s , in each VND iteration to 63, the largest size of QUBO problem, whose underlying graph is complete, that can be solved using an equal-length embedding on the quantum annealer. In what follows, we report on the effect of s on the performance of VND. This is useful in predicting the performance improvement achievable by increasing either the number of qubits or their connectivity.

Table 7 presents the results for molecules B and C for different neighbourhood sizes when an exact solver is used to solve the QUBO problem at each iteration. The reason for choosing these two molecules is that they have a star-like structure and, as discussed in Section 3.6, the advantage of using a quantum annealer over an exact solver is expected to be more pronounced for these molecules.

As shown, increasing the neighbourhood size from 30 to 60 and then from 60 to 90 noticeably improves the results for both molecules. However, the improvement exhibits diminishing returns when the neighbourhood size is increased beyond 90. We expect similar behaviour to occur at different neighbourhood sizes for different families of molecules.

Conclusion

In this paper, we have presented a variable neighbourhood descent (VND) method for conformational search. We introduced the concept of a rigid-body graph and used this simplified molecular structure to carefully define a neighbourhood structure to allow for efficient optimization using a binary quadratic optimizer. Based on current quantum annealing hardware, we selected a 2-torsion-dependent neighbourhood at each iteration such that finding the best solution in the selected neighbourhood could then be formulated as a QUBO problem. The size of the neighbourhood can be chosen such that the method can be adapted to the number of available qubits as well as to their connectivity on the quantum annealer. As a result, the proposed method is not only well-suited for current hardware, but can easily be adapted to take advantage of hardware improvements. Whereas the proposed method can be used as a standalone conformational search approach, it can be combined with existing conformational search methods for potentially improved performance.

Beyond a simple presentation of the method, we also conducted a preliminary case study based on an implementation of the VND method using the D-Wave 2000Q quantum annealer for two families of molecules. In this exploration, we compared the results of our method with those of PTMC, a state-of-the-art solver for conformational search. To understand how much of the gap between the results of PTMC and those of VND used with a quantum annealer can be attributed to the imperfections of the quantum annealer, we replaced it with an exact QUBO problem solver. VND used along with the exact solver was able to find noticeably better conformations than those found using VND and the quantum annealer together. This observation points to the potential improvement achievable in the short term through more-advanced tuning of the existing quantum annealer, and in the long term using improved hardware.

This work suggests a number of possible future research directions. For example, investigating refined neighbourhood change functions rather than using a random ordering of the torsion angles to choose the neighbourhood could lead to significant improvements. This could involve further exploitation of the molecular graph or the solutions from previous VND iterations. In addition, an improved selection of the s discrete points at each VND iteration to account for hardware limitations could further improve the results. We leave these improvements for future work.

We believe that the proposed method, based on careful hardware-aware neighbourhood selection, holds the potential to provide promising solutions to important optimization problems. While PTMC shows better performance over the molecules studied, our method has opened a scalable path forward for leveraging emerging quantum technologies for conformational search, a critically important problem in the field of chemical and materials science.

References

- Perola, E. & Charifson, P. S. Conformational analysis of drug-like molecules bound to proteins: an extensive study of ligand reorganization upon binding. *Journal of Medicinal Chemistry* **47**, 2499–2510 (2004).
- Kirchmair, J., Laggner, C., Wolber, G. & Langer, T. Comparative analysis of protein-bound ligand conformations with respect to Catalyst's conformational space subsampling algorithms. *Journal of Chemical Information and Modeling* **45**, 422–430 (2005).
- Dill, K. A. & MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **338**, 1042–1046 (2012).
- Boyd, R. H. & Phillips, P. J. *The Science of Polymer Molecules* (Cambridge University Press, 1996).
- Miao, Y. & McCammon, J. A. Unconstrained enhanced sampling for free energy calculations of biomolecules: A review. *Molecular Simulation* **42**, 1046–1055 (2016).
- Copeland, R. A. Conformational adaptation in drug–target interactions and residence time. *Future Medicinal Chemistry* **3**, 1491–1501 (2011).
- Izgorodina, E. I., Lin, C. Y. & Coote, M. L. Energy-directed tree search: an efficient systematic algorithm for finding the lowest energy conformation of molecules. *Physical Chemistry Chemical Physics* **9**, 2507–2516 (2007).
- Hawkins, P. C. Conformation generation: The state of the art. *Journal of Chemical Information and Modeling* **57**, 1747–1756 (2017).
- Leach, A. R. A survey of methods for searching the conformational space of small and medium-sized molecules. *Reviews in Computational Chemistry* 1–55 (1991).
- Hatfield, M. & Lovas, S. Conformational sampling techniques. *Current Pharmaceutical Design* **20**, 3303–3313 (2014).
- Loferer, M. J., Kolosváry, I. & Aszódi, A. Analyzing the performance of conformational search programs on compound databases. *Journal of Molecular Graphics and Modelling* **25**, 700–710 (2007).
- Bruccoleri, R. E. & Karplus, M. Prediction of the folding of short polypeptide segments by uniform conformational sampling. *Biomolecules: Original Research on Biomolecules* **26**, 137–168 (1987).
- Goodman, J. M. & Still, W. C. An unbounded systematic search of conformational space. *Journal of Computational Chemistry* **12**, 1110–1117 (1991).
- Christen, M. & Van Gunsteren, W. F. On searching in, sampling of, and dynamically moving through conformational space of biomolecular systems: A review. *Journal of Computational Chemistry* **29**, 157–166 (2008).
- Goedecker, S. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *The Journal of Chemical Physics* **120**, 9911–9917 (2004).
- Vainio, M. J. & Johnson, M. S. Generating conformer ensembles using a multiobjective genetic algorithm. *Journal of Chemical Information and Modeling* **47**, 2462–2474 (2007).
- Supady, A., Blum, V. & Baldauf, C. First-principles molecular structure search with a genetic algorithm. *Journal of Chemical Information and Modeling* **55**, 2338–2348 (2015).
- Lee, J., Scheraga, H. A. & Rackovsky, S. New optimization method for conformational energy calculations on polypeptides: conformational space annealing. *Journal of Computational Chemistry* **18**, 1222–1232 (1997).
- Heo, S., Lee, J., Joo, K., Shin, H.-C. & Lee, J. Protein loop structure prediction using conformational space annealing. *Journal of Chemical Information and Modeling* **57**, 1068–1078 (2017).
- Morales, L. B., Garduño-Juárez, R., Aguilar-Alvarado, J. & Riveros-Castro, F. A parallel tabu search for conformational energy optimization of oligopeptides. *Journal of Computational Chemistry* **21**, 147–156 (2000).
- Stepanenko, S. & Engels, B. Tabu search based strategies for conformational search. *The Journal of Physical Chemistry A* **113**, 11699–11705 (2009).
- Li, Z., Laidig, K. E. & Daggett, V. Conformational search using a molecular dynamics–minimization procedure: Applications to clusters of coulombic charges, lennard–jones particles, and waters. *Journal of Computational Chemistry* **19**, 60–70 (1998).
- Doshi, U. & Hamelberg, D. Towards fast, rigorous and efficient conformational sampling of biomolecules: Advances in accelerated molecular dynamics. *Biochimica et Biophysica Acta (BBA)–General Subjects* **1850**, 878–888 (2015).

24. Gehrke, R. & Reuter, K. Assessing the efficiency of first-principles basin-hopping sampling. *Physical Review B* **79**, 085412 (2009).
25. Wilson, S. R., Cui, W., Moskowitz, J. W. & Schmidt, K. E. Applications of simulated annealing to the conformational analysis of flexible molecules. *Journal of Computational Chemistry* **12**, 342–349 (1991).
26. Sperandio, O., Souaille, M., Delfaud, F., Miteva, M. A. & Villoutreix, B. O. Med-3dmc: A new tool to generate 3D conformation ensembles of small molecules with a Monte Carlo sampling of the conformational space. *European Journal of Medicinal Chemistry* **44**, 1405–1409 (2009).
27. Thachuk, C., Shmygelska, A. & Hoos, H. H. A replica exchange Monte Carlo algorithm for protein folding in the HP model. *BMC Bioinformatics* **8**, 342 (2007).
28. Rhee, Y. M. & Pande, V. S. Multiplexed-replica exchange molecular dynamics method for protein folding simulation. *Biophysical Journal* **84**, 775–786 (2003).
29. Santoro, G. E., Martoňák, R., Tosatti, E. & Car, R. Theory of quantum annealing of an ising spin glass. *Science* **295**, 2427–2430 (2002).
30. Morita, S. & Nishimori, H. Mathematical foundation of quantum annealing. *Journal of Mathematical Physics* **49**, 125210 (2008).
31. Perdomo-Ortiz, A., Dickson, N., Drew-Brook, M., Rose, G. & Aspuru-Guzik, A. Finding low-energy conformations of lattice protein models by quantum annealing. *Scientific Reports* **2**, 571 (2012).
32. The D-Wave 2000Q quantum computer: Technology overview, https://www.dwavesys.com/sites/default/files/D-Wave%202000Q%20Tech%20Collateral_0117F.pdf.
33. Van der Ploeg, S. *et al.* Adiabatic quantum computation with flux qubits, first experimental results. *IEEE Transactions on Applied Superconductivity* **17**, 113–119 (2007).
34. Johnson, M. W. *et al.* Quantum annealing with manufactured spins. *Nature* **473**, 194 (2011).
35. Swendsen, R. H. & Wang, J.-S. Replica Monte Carlo simulation of spin-glasses. *Physical Review Letters* **57**, 2607 (1986).
36. 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. *Journal of the American Chemical Society* **114**, 10024–10035 (1992).
37. Glover, F. W. & Kochenberger, G. A. *Handbook of metaheuristics*, vol. 57 (Springer Science & Business Media, 2006).
38. Mladenović, N. & Hansen, P. Variable neighborhood search. *Computers & operations research* **24**, 1097–1100 (1997).
39. Kochenberger, G. *et al.* The unconstrained binary quadratic programming problem: a survey. *Journal of Combinatorial Optimization* **28**, 58–81 (2014).
40. Wang, Y., Lü, Z., Glover, F. & Hao, J.-K. Path relinking for unconstrained binary quadratic programming. *European Journal of Operational Research* **223**, 595–604 (2012).
41. Shi, J., Zhang, Q., Derbel, B. & Liefvooghe, A. A parallel tabu search for the unconstrained binary quadratic programming problem. In *IEEE Congress on Evolutionary Computation (CEC 2017)*, 557–564 (2017).
42. Biswas, R. *et al.* A NASA perspective on quantum computing: Opportunities and challenges. *Parallel Computing* **64**, 81–98 (2017).
43. Arriola, D. J., Carnahan, E. M., Hustad, P. D., Kuhlman, R. L. & Wenzel, T. T. Catalytic production of olefin block copolymers via chain shuttling polymerization. *Science* **312**, 714–719 (2006).
44. Makio, H., Kashiwa, N. & Fujita, T. FI Catalysts: A New Family of High Performance Catalysts for Olefin Polymerization. *Advanced Synthesis and Catalysis* **344**, 477–493 (2002).
45. Torker, S., Müller, A., Sigrist, R. & Chen, P. Tuning the Steric Properties of a Metathesis Catalyst for Copolymerization of Norbornene and Cyclooctene toward Complete Alternation. *Organometallics* **29**, 2735–2751 (2010).
46. Schwab, P., France, M. B., Ziller, J. W. & Grubbs, R. H. A series of well-defined metathesis catalysts- synthesis of $[RuCl_2(=CHR')(PR_3)_2]$ and its reactions. *Angewandte Chemie International Edition* **34**, 2039–2041 (1995).
47. Kawakami, T., Ito, S. & Nozaki, K. Iron-catalysed homo- and copolymerisation of propylene: Steric influence of bis(imino)pyridine ligands. *Dalton Transactions* **44**, 20745–20752 (2015).
48. Small, B. L., Brookhart, M. & Bennett, A. M. A. Highly Active Iron and Cobalt Catalysts for the Polymerization of Ethylene. *Journal of the American Chemical Society* **120**, 4049–4050 (1998).
49. Schmidt, R. *et al.* Hydrocarbons. In *Ullmann's Encyclopedia of Industrial Chemistry* (Wiley-VCH Verlag GmbH & Co. KGaA., 2002).
50. Spalding, M. A. & Chatterjee, A. M. (eds.) *Handbook of Industrial Polyethylene and Technology: Definitive Guide to Manufacturing, Properties, Processing, Applications and Markets* (Scrivener Publishing LLC, 2018).
51. He, J. *et al.* Ortho-Phenylenes: Unusual conjugated oligomers with a surprisingly long effective conjugation length. *Journal of the American Chemical Society* **132**, 13848–13857 (2010).
52. Berresheim, A. J., Müller, M. & Müllen, K. Polyphenylene nanostructures. *Chemical Reviews* **99**, 1747–1785 (1999).

Acknowledgements

The authors thank Jamie Cohen, Alejandro Garza, Steve Arturo, Maritza Hernandez, Helmut G. Katzgraber, Kausar N. Samli, Takeshi Yamazaki, and Arman Zaribafiyani for their insightful comments, Varinia Bernales for assistance in preparing molecular geometries, and Marko Bucyk for reviewing and editing the manuscript.

Author Contributions

The method proposed in this paper was developed by M.N. and B.W. The majority of the paper was written by M.N. and P.M., with contributions from G.R., D.M., B.W. and A.R. The majority of the experiments' code was written by G.R. and M.N., with contributions from B.W. and D.M. Chemistry domain knowledge and guidance on the experiments' design were provided by P.M., M.C. and D.D. The experiments and their analyses were conducted by M.N., G.R., A.R. and U.Y.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-47298-y>.

Competing Interests: D.M., M.N., A.R., G.R., B.W. and U.Y. were employed by IQBit during their contribution to this work. IQBit is a company active in optimization and quantum computing. M.C., D.D. and P.M. were employed by The Dow Chemical Company during their contribution to this work. The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019