

Cite this: *Chem. Sci.*, 2026, 17, 6055

All publication charges for this article have been paid for by the Royal Society of Chemistry

The interlocking process in molecular machines explained by a combined approach: the nudged elastic band method and machine learning potential

Lucio Peña-Zarate,  Alberto Vela * and Jorge Tiburcio *

Engineering molecular machines requires a precise knowledge of the mechanisms involved in programmed motions. Among artificial molecular machines, rotaxanes have emerged as a prominent model due to their ability to perform diverse and controlled motions, such as threading, shuttling, and pirouetting. In this work, we present a reliable theoretical framework to describe the threading motion during the assembly of rotaxane-like complexes. Our approach combines the climbing image nudged elastic band method with the ANI-1ccx neural network potential, trained with gold-standard data. Energetic and structural variations along a normalized displacement coordinate allowed an accurate atomistic description of the threading process of different dumbbell-shaped molecules (axles) through the cavity of two different macrocyclic hosts (tori). Using the methodology proposed herein, two key steps are identified: stabilization through hydrogen bonds, which we call the claw mechanism, and the expansion of the macrocycle. An energy decomposition analysis, performed by single-point calculations on selected structures, allows an analysis of the role of steric and electrostatic effects in the structural stabilization of supramolecular assemblies. We find that, although ANI-1ccx was not explicitly trained for charged systems, this neural network potential effectively discriminates among various charged states. Furthermore, calculated potential energy barriers are in good agreement with reported experimental free energy barriers. The featured methodology has the potential to become a fundamental artificial-intelligence-based tool for the study of diverse motions observed in supramolecular systems.

Received 27th October 2025
Accepted 12th January 2026

DOI: 10.1039/d5sc08303f

rsc.li/chemical-science

Introduction

Artificial molecular machines are synthetic systems designed to mimic the movement of macroscopic machines by responding to an external stimulus. These chemical systems share several similarities with macroscopic machines, such as the need for an energy input for operation, the relative movements of their components, the ability to monitor and control their functions, the capacity to repeat operation cycles, the time required to complete a full cycle, and the purpose of their operation.^{1,2}

Mechanically interlocked molecules (MIMs), like rotaxanes and catenanes, have emerged as prototypical artificial molecular machines because of their ability to perform large-amplitude motions in response to an energy source, while maintaining structural integrity.³ Rotaxanes, in particular, are a family of supramolecular host–guest complexes consisting of a dumbbell-shaped guest (axle) encircled by one or more macrocyclic hosts (tori), with both ends of the guest extending

outside the ring.⁴ A typical method to assemble rotaxanes involves threading the torus over one of the terminal groups on the axle, a process that is influenced by flexibility, size-complementarity, and steric and electrostatic effects between the components.⁵

Given the importance of MIMs, it is essential to understand the energetic profile underlying their dynamics through an atomistic model. The desired model must reveal how different non-covalent interactions and conformational changes interplay along the path of molecular motion. Additionally, estimating energy barriers could reveal how chemical modifications in the host and guest influence the assembly process. This method should also provide the molecular basis for rationalizing experimental data.

Computational tools have played a crucial role in understanding the mechanisms of shuttling, dissociation, aggregation, and conformational changes in supramolecular systems. These include methods such as molecular dynamics,^{6–10} metadynamics,^{11–14} nudged elastic band (NEB),^{15,16} and calculations of electronic structure.^{17–21} However, these techniques face limitations due to the complexity of supramolecular systems, including their size and dynamic processes,²² as well as

Department of Chemistry, Centro de Investigación y de Estudios Avanzados (Cinvestav), Avenida IPN 2508, Mexico City 07360, Mexico. E-mail: avela@cinvestav.mx; jtiburcio@cinvestav.mx



the challenge of selecting an appropriate collective variable^{23–25} that can describe a coherent reaction coordinate for the process. As a result, it remains challenging for computational studies to quickly and accurately identify these mechanisms.²²

One approach to addressing this challenge involves utilizing atomic neural network potentials (NNPs) as an alternative to conventional electronic structure approaches.²⁶ The majority of these potentials are developed based on datasets derived from DFT calculations. An illustrative example is the potential employing the Polarizable Atom Interaction Neural Network (PaiNN), which has been utilized in transition-state searches in conjunction with the NEB method.²⁷ This helps to obtain insights into the reaction mechanisms of various reactions, at speeds approximately three orders of magnitude faster than NEB/DFT calculations, demonstrating the capacity of PaiNN to expedite transition-state search techniques with NEB. The combination of NEB with reliable electronic structure calculations for transition-state analysis is markedly more computationally intensive than pairing NEB with a well-trained neural network.²⁸

Another set of neural network potentials is the ANAKIN-ME family, which has been utilized in a variety of applications, including high-precision calculations of thermochemical reaction values,^{29–32} determination of conformational change,^{33–37} molecular dynamics simulations,^{38–40} structural searches for new drugs,⁴¹ and the prediction of absorption, distribution, metabolism, and excretion (ADME).⁴² Among the ANAKIN-ME family, particular attention is given to ANI-1ccx, which was trained using the gold standard (CCSD(T)*/CBS) of molecular electronic structure theory. This model has produced reasonable predictions of the binding structures of ligands in relation to experimental data, even though electrically charged molecules are outside the scope of the ANI-1ccx training set.³⁸ This last study exemplifies the ability of the ANI-1ccx neural network to deliver coherent responses to conformational analysis, considering that the ANI-1ccx version cannot account for the charge and multiplicity of the systems under consideration.

The study described herein utilizes the ANI-1ccx potential in conjunction with the NEB method to elucidate the assembly processes involved in the formation of the rotaxanes depicted in Fig. 1. These rotaxanes feature asymmetric, dumbbell-shaped axles, with a bulky ester substituent on one end, acting as a stopper for the two investigated 24-membered macrocycles: 24-crown-8 (**24C8**) and dibenzo-24-crown-8 (**DB24C8**). The presence of one stopper on the axle allows macrocycle threading only over the other terminus. Guest family [1]²⁺ is composed of positively doubly charged axles, all of which have a pyridinium group linked *via* an ethyl bridge to cyclic ammonium substituents of different sizes, varying from piperidinium [**1a**]²⁺, to azepanium [**1b**]²⁺, and azocanium [**1c**]²⁺. An experimental study on the stability and association rates of these axles with the aforementioned macrocycles was previously described in the literature by our group.^{43,44} To explore the effect of charge on the formation of the [1]²⁺ complex, [2]⁺ and [3] axles with **24C8** were selected as model systems, in which nitrogen atoms in the axle were replaced with carbon atoms to produce both a singly charged and a neutral variant. The shape of the energy profile

during threading is dictated by non-covalent interactions and by the combined flexibility of the macrocycle and the guest substituents, factors determining whether the profile is smooth or rough. Furthermore, as our results show, even though ANI-1ccx neither accepts charge and multiplicity as inputs nor was trained on charged systems, its architecture allows it to indirectly capture charge effects, reproducing stability trends across neutral, and singly or doubly charged rotaxanes.

The proposed NEB/ANI-1ccx methodology allowed us to obtain detailed energy profiles consisting of plots of the association energy as a function of the normalized displacement coordinate (NDC). This allows a description of the conformational changes occurring in the macrocycles and the axles, and in conjunction with single-point DFT calculations together with non-covalent interaction (NCI) analysis and Shubin Liu's energy decomposition analysis (EDA-SBL), reveals the most important interactions that intervene in the ring translation over and between different end groups, allowing better comprehension of the interlocking mechanisms. These points are addressed in the Results and discussion section. The energy barriers for the interlocking processes are obtained as differences in potential energies calculated with ANI-1ccx, and they are found to be in good agreement with experimental activation energies. Finally, we draw general conclusions about the mechanisms involved in the formation of rotaxanes, using an atomistic model, as well as the implementation of NEB/ANI-1ccx as an alternative strategy to address association and translation processes in supramolecular systems.

Methodology

This section describes the methodological approach that combines the climbing image NEB with the NNP ANI-1ccx. The analysis tools used to characterize the interactions and energy contributions are explained in the SI.

Nudged elastic band (NEB)

This is a double-ended transition-state search method that approximates the reaction path on the potential energy surface by interpolating an initial path between the reactant and product, discretized by a system replica set called images.⁴⁵ With these images, an initial path is constructed, and subsequently, an optimization algorithm is used to iteratively relax a discrete path with the N images until convergence to an approximate minimum-energy path is achieved. The reaction coordinate is the normalized displacement coordinate, which is the accumulated distance obtained by adding the Euclidean distances between each consecutive pair of images, which is evaluated as $\tilde{s}_i = \sum_{j=1}^i \|\mathbf{R}_j - \mathbf{R}_{j-1}\|/s_N$, for $i = 1, 2, \dots, N$, where \mathbf{R}_j are the atomic coordinates of the j th image, N is the number of images, and s_N is the total length of the path. In this work, the normalized displacement coordinate (NDC) is used to build the energy profiles.

Several modifications to the NEB method exist; notably, the preconditioned chain method⁴⁶ is designed to handle poorly conditioned potentials, such as those in structures with



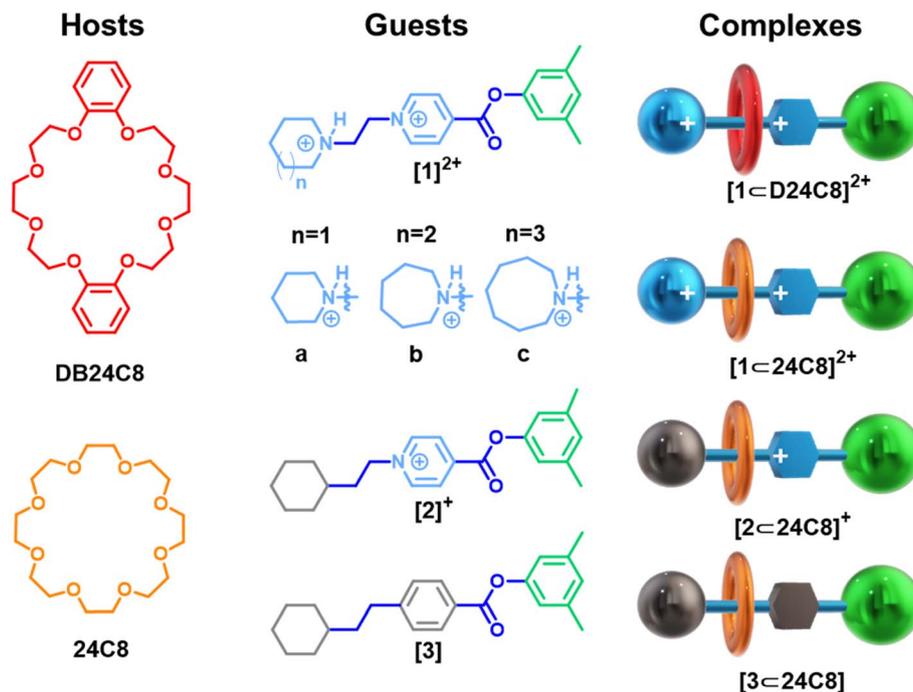


Fig. 1 Chemical and cartoon representations of hosts, guests, and complexes. The notation used to designate the complexes is [Guest \subset Host] where Guest denotes the axle: 1 corresponds to doubly charged axles, 2 to singly charged axles, and 3 to neutral axles; n indicates the size of the cyclic ammonium end-group on the axle: $n = 1$ (six-membered, piperidinium), $n = 2$ (seven-membered, azepanium), and $n = 3$ (eight-membered, azocanium). Hosts, the tori, are 24-crown-8 ether (24C8) or dibenzo-24-crown-8 ether (DB24C8).

significant variations in rigidity, enabling faster convergence. However, it does not guarantee that the highest-energy structure, the transition state, can be found. The climbing image nudged elastic band method addresses this by allowing a specific image to behave differently, following variable spring constants and moving towards the highest-energy structure.⁴⁷

ANI-1ccx. The ANI-1ccx potential was developed through transfer learning, initially utilizing DFT calculations obtained with ω B97X/6-31G* for systems composed of carbon, hydrogen, nitrogen, and oxygen, specifically from the ANI-1x dataset comprising five million out-of-equilibrium molecular conformations. Subsequently, it was retrained using a significantly smaller dataset: approximately 500 000 selected conformations from ANI-1x, using the CCSD(T)*/CBS level of accuracy, where the asterisk indicates that the domain-based local pair natural orbital approximation (DLPNO) was used.^{48,49} The resulting model operates at roughly nine orders of magnitude faster than DFT calculations. Through the transfer learning process, it not only refines the energy predictions but also improves the accuracy of force calculations. These achievements are enabled by extensive training and an advanced neural network architecture. The network consists of a series of hidden layers, with two of these layers being actively adjusted to optimize the transfer learning process, while the remaining two layers remain fixed to reduce the number of trainable parameters. This approach effectively prevents overfitting to the smaller CCSD(T)*/CBS dataset.⁵⁰

NEB/ANI-1ccx. Our methodology yields high-resolution reaction coordinates, specifically the normalized displacement

coordinate. This is because NEB constructs the reaction path directly in the full atomic configuration space through a chain of images representing complete geometries.⁵¹ Thus, unlike metadynamics, NEB does not require a reduction in dimensionality or the selection of collective variables, allowing an accurate description of the pathway and energy barriers without loss of information. This is achieved through the application of the climbing image NEB method, which is integrated *via* the Atomic Simulation Environment (ASE)⁵² into the ANI-1ccx neural network architecture and the LBFGS optimizer. The NDCs provide a thorough description of the assembly process in the rotaxane-type systems under investigation.

Shubin Liu's energy decomposition analysis (EDA-SBL). This analysis reveals how steric, electrostatic, and quantum effects govern supramolecular assembly. In the EDA-SBL framework, the total energy is partitioned as:

$$E_t = E_s + E_e + E_q$$

where E_s (steric energy) is associated with the Weizsäcker contribution to the kinetic energy functional, E_e (electrostatic energy) describes the direct interelectronic coulombic interaction, and E_q (quantum energy) represents the exchange–correlation and charge delocalization/transfer contributions.^{53,54}

In this work, we present energy variations relative to an internal reference, defined as the \mathbf{P}_1 structure of each system. Thus, for any selected point \mathbf{P}_i along the energy profile, the reported values correspond to the differences with respect to \mathbf{P}_1 ,



providing a clearer picture of the energetic contributions throughout the assembly process.

Computational details. All NEB calculations were performed by turning on the climbing image method and the string method⁴⁶ in ASE. The computations utilized 125 images and linear interpolation, using the ANI-1ccx neural network potential trained from CCSD(T)*/CBS calculations data through TorchANI, a PyTorch-based program for ANI (ANAKIN-ME) deep-learning model training/inference. The Limited memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) optimization algorithm was used, with a maximum force value of $\max(|F|) < 0.1 \text{ eV } \text{Å}^{-1}$, to achieve convergence of the calculations.

The reactant and product configurations used as input for the NEB calculations were aligned and optimized using the LBFGS method (ASE package), with the ANI-1ccx potential incorporated as an external calculator, using a convergence criterion of $\max(|F|) < 1 \times 10^{-4} \text{ eV } \text{Å}^{-1}$. Two NEB calculations were performed using three initial structures, with one serving as the product of the first NEB and the reactant of the second. To verify that the optimized structures of the reactant and product are local minima on the potential energy surface, a frequency analysis by finite differences was performed using ASE with ANI-1ccx as the calculator.

The structure used as input for the first NEB calculation, corresponding to the configuration with the axle and macrocycle with a very large separation, was also optimized. Its frequency analysis reveals the presence of imaginary modes that are the relative translational and rotational motions characteristic of supramolecular systems.⁵⁵

For the NCI and EDA-SBL analyses, single-point DFT calculations were performed, using a selected set of images generated from the NEB/ANI-1ccx calculation trajectories. The exchange–correlation energy functional used was the non-empirical global hybrid with dispersion PBE0-D3, and the orbital basis def2TZVP.^{56,57} In all cases, SCF convergence was achieved, using the quadratically convergent SCF procedure (SCF = QC).⁵⁸ All the DFT calculations were done using Gaussian 09.⁵⁹ The NCI and EDA analyses were done using Multiwfn.⁶⁰ For the NCI calculations, a 3D grid with 0.1 Å spacings along each Cartesian axis was used.

Results and discussion

Initially, we studied the threading process of the linear guest [1a]²⁺ through the cavity of the 24C8 macrocycle. The process is described by an NDC derived from successive images generated

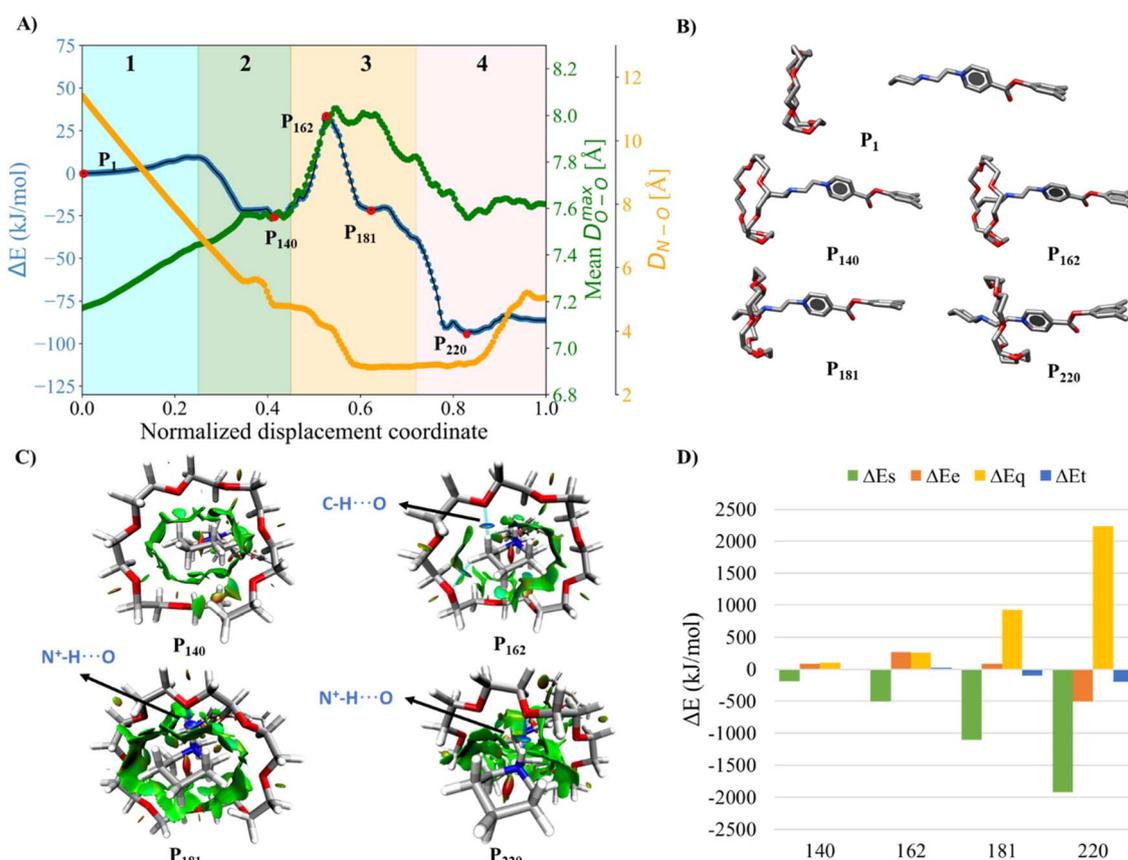


Fig. 2 (A) Energy profile (blue curve), average maximum distance between oxygen atoms in 24C8 (D_{O-O}^{\max} , green curve), and donor–acceptor distance of the hydrogen bond (D_{N-O} , orange curve) during the assembly process of [1a] + 24C8²⁺. (B) Structures P_1 , P_{140} , P_{162} , P_{181} , and P_{220} corresponding to selected points along the curve. For clarity, hydrogen atoms are omitted. (C) Isosurfaces of the reduced density gradient $s(\vec{r}) = 0.6$ colored according to $\text{sign}(\lambda_2)$ for the images P_{140} , P_{162} , P_{181} and P_{220} . (D) EDA analysis for the corresponding images, showing the contributions to the energy, using P_1 as reference.



by NEB. Trajectories were built from 250 images from two different NEB calculations. The energy of each structure was calculated using the ANI-1ccx potential. A graphical representation of the energy profile during the threading along the NDC is shown in Fig. 2A. This figure also depicts the average maximum distance between pairs of oxygen atoms in the macrocycle $D_{\text{OO}}^{\text{max}}$ (green curve), and the hydrogen-bond donor-acceptor distance (orange curve). These two geometrical parameters provide useful information about the conformational changes involved in the intertwining process, as well as the formation and breaking of non-covalent interactions.

The assembly process can be divided into four stages, labeled and colored in Fig. 2A, based on the energy behavior along the NCD.

(1) Initialization (Stage 1). The components, axle and torus, interact weakly, and the energy undergoes very small changes; neither molecule exhibits significant geometric distortion.

(2) Preparation (Stage 2). The two molecular species interact attractively; energy drops, and axle and torus align to facilitate molecular assembly.

(3) Activation (Stage 3). The torus slides over the axle terminus, overcoming unfavorable steric interactions. This step determines the energy barrier for the assembly of the rotaxane complex and contains the transition state (see structure \mathbf{P}_{162} in Fig. 2B).

(4) Stabilization (Stage 4). The torus reaches the positively charged region on the axle and encircles the ethylene group, providing the lowest energy conformation (see structure \mathbf{P}_{220} in Fig. 2B); the macrocycle undergoes significant geometrical rearrangements, which are reflected in the parameter $D_{\text{OO}}^{\text{max}}$.

In the initialization stage, the torus (macrocycle) and the axle (guest) are nearly non-interacting species. This was confirmed by comparing the energy of the NEB initial structure with the sum of the energies of the isolated macrocycle and guest. In all studied systems, the interaction energy at this initial stage is zero. This is further supported by the NCI analysis, which shows no isosurfaces between the two species (see Fig. S7).

During preparation, the torus approaches the axle, resulting in structures that are more stable than the non-interacting species, with an energy decrease of about -25 kJ mol^{-1} . This energy change is strongly linked to weak interactions caused by the proximity of the macrocycle to the guest, as can be seen by the NCI isosurfaces in the \mathbf{P}_{140} complex (Fig. 2C), which are attributed to van der Waals interactions between host and guest.

In the activation stage, the $24\text{C}8$ macrocycle slides over the cyclic ammonium located at the axle terminus, inducing significant conformational changes and cavity expansion. At the highest-energy point (\mathbf{P}_{162} in Fig. 2C), a C-H \cdots O hydrogen bond at 3.4 \AA is observed. After overcoming an energy barrier of 59 kJ mol^{-1} , the macrocycle approaches the nitrogen of the piperidinium group on the axle, resulting in the formation of an intermolecular hydrogen bond N $^+$ -H \cdots O (orange curve) with a donor-acceptor distance of 2.8 \AA . This set of non-covalent interactions stabilizes the complex formed during this activation stage. The EDA analysis indicates that in the \mathbf{P}_{162} complex, the electrostatic contribution is unfavorable (Fig. 2D), as it has

a positive value. This occurs because the macrocycle is still far from the region where the positive charge is mainly located. As the process continues, the macrocycle moves closer to the ammonium group, reducing the positive electrostatic energy in the \mathbf{P}_{181} structure. Additionally, proximity to the ammonium group on the axle facilitates a strong hydrogen bond of the N $^+$ -H \cdots O type between the macrocycle and the ammonium group, as shown by NCI analysis.

Finally, in the stabilization stage, the macrocycle reaches the minimum energy structure \mathbf{P}_{220} . The N $^+$ -H \cdots O hydrogen bond displays slight variations in distance compared to that observed in \mathbf{P}_{181} . By maintaining this bond, the upper part of the macrocycle remains anchored to the axle, while the lower part shifts its position toward the pyridinium group, adopting an S-type conformation, as shown in the \mathbf{P}_{220} complex in Fig. 2B. This combination of intermolecular interactions, along with the reduction in the size of the macrocycle cavity, explains the decrease in energy toward the most stable structure. The energy decomposition analysis in Fig. 2D clearly shows a significant reduction in steric hindrance in \mathbf{P}_{220} compared to earlier stages. At this point, the electrostatic contribution also becomes energetically favorable. Therefore, the formation of the lowest energy structure during the interlocking process results from the stabilization of both steric and electrostatic contributions.

We now address the issue of steric hindrance in the assembly process by comparing two different doubly charged complexes: $[\mathbf{1a} \subset 24\text{C}8]^{2+}$, with a piperidinium end-group, and $[\mathbf{1b} \subset 24\text{C}8]^{2+}$, featuring an azepanium terminal group. As previously discussed, for complex $[\mathbf{1a} \subset 24\text{C}8]^{2+}$, macrocycle expansion occurs at the preparation and activation stages, while the N $^+$ -H \cdots O interaction appears until the activation stage, whereas complex $[\mathbf{1b} \subset 24\text{C}8]^{2+}$ follows a more irregular path due to the bulkiness of the $[\mathbf{1b}]^{2+}$ terminal group. As shown in Fig. 3, the azepanium substituent markedly modifies the energy landscape, preventing the macrocycle from freely threading along the axle, unlike the situation with piperidinium (Fig. 2A and 3). As the macrocycle approaches the bulkier $[\mathbf{1b}]^{2+}$ terminus during the preparation stage, an important conformational change occurs, resulting in a claw-like arrangement

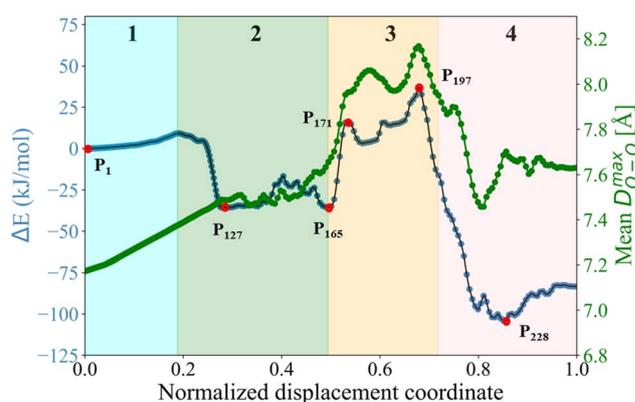


Fig. 3 Energy profile (blue curve) and average maximum distance between oxygen atoms in $24\text{C}8$ ($D_{\text{OO}}^{\text{max}}$, green curve) for the assembly of $[\mathbf{1b} \subset 24\text{C}8]^{2+}$.



(Fig. 4A), where the macrocycle surrounds the axle terminus, the azepanium end-group. As can be seen in Fig. 4B, this engulfed configuration aids threading by promoting the formation of $N^+-H\cdots O$ and $C-H\cdots O$ hydrogen bonds, as demonstrated by the NCI analysis.

This mechanism permits the macrocycle to anchor to the axle and slide along $[1b]^{2+}$. During the activation stage, it traverses the azepanium group, which adopts a boat conformation, as can be seen in structure P_{197} (Fig. 4C) stabilized by

interactions with the 24C8 ring. In P_{228} , only the $N^+-H\cdots O$ hydrogen bond remains, as ring relaxation allows it to reposition around the ethyl group (Fig. 4D), resulting in the loss of the $C-H\cdots O$ interaction previously maintained by the azepanium group with the macrocycle. Throughout the stages of initialization, preparation, and stabilization, azepanium consistently adopts a twisted-boat conformation, as shown in P_{228} . EDA-SBL analysis indicates that the steric contribution becomes negative and more significant compared to the reference; this is due to

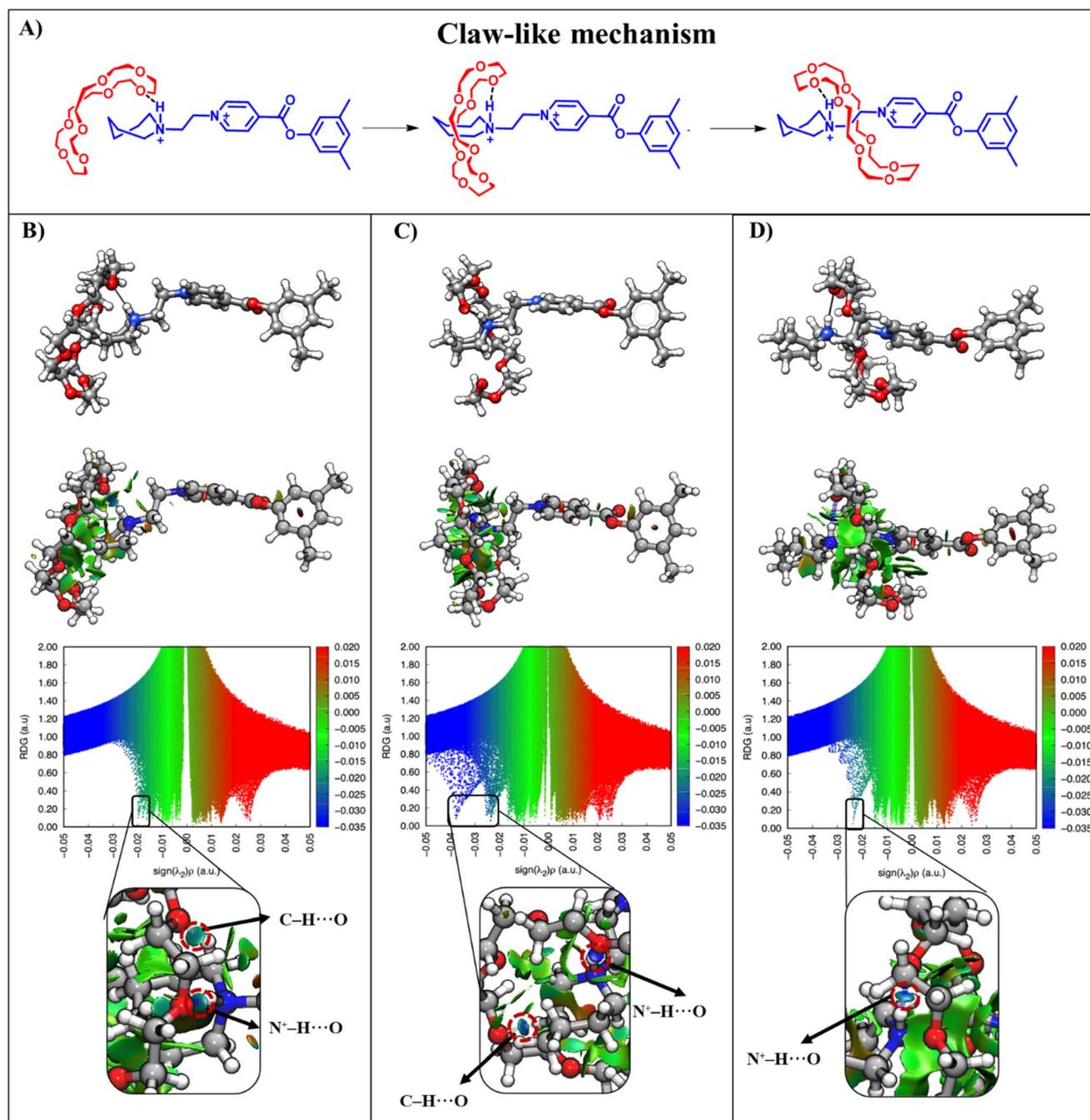


Fig. 4 (A) Claw-like mechanism for the assembly process of complex $[1b \subset 24C8]^{2+}$, at structures (B) P_{165} , (C) P_{197} , and (D) P_{228} of the energy profile; NCI analysis of each structure of $[1b \subset 24C8]^{2+}$ and ball-and-stick models depicting the isosurfaces of $s(\vec{r})$, with values of $s(\vec{r}) = 0.6$, colored according to the sign of $(\lambda_2)\rho(\vec{r})$.



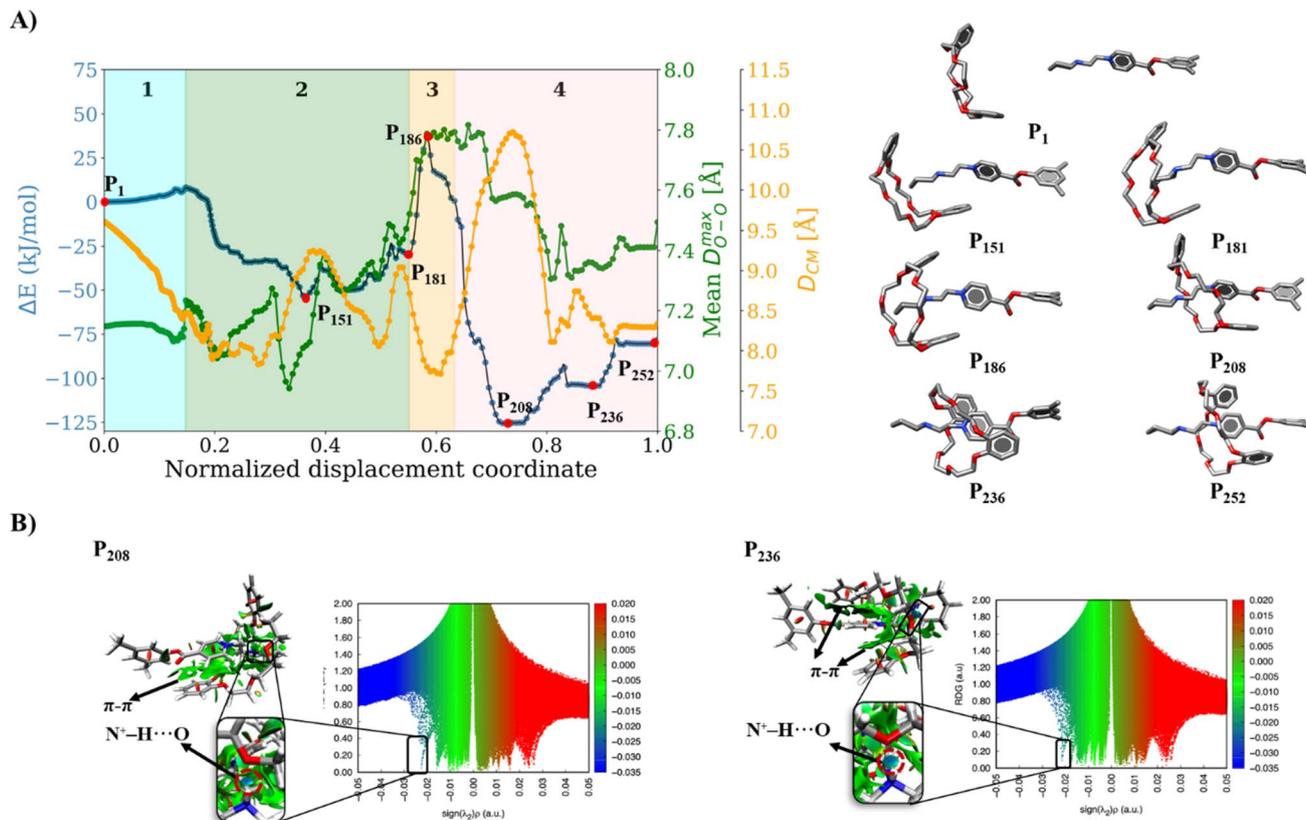


Fig. 5 (A) Energy profile (blue curve), average maximum distance between oxygen atoms in DB24C8 (D_{OO}^{\max} , green curve), and distances between the centers of mass of the aromatic rings (D_{CM} , orange curve) for the assembly of $[1a \subset DB24C8]^{2+}$ and structures P₁, P₁₅₁, P₁₈₁, P₁₈₆, P₂₀₈, P₂₃₆, and P₂₅₂ corresponding to selected points along the curve. For clarity, hydrogen atoms are omitted. (B) Isosurfaces of the reduced density gradient with a value of colored according to $\text{sign}(\lambda_2)$ for the images P₂₀₈ and P₂₃₆.

internal structural arrangements, where weak interactions and electrostatics relieve strain by relaxing both the macrocycle and the axle (Fig. S3C).

The claw-like mechanism, involving hydrogen-bond formation during the preparation stage, is not limited to systems with bulky terminal groups. A similar process also occurs in complexes with smaller terminal groups when a more rigid macrocycle is used. This is the case for the $[1a \subset DB24C8]^{2+}$ complex. In this system, the presence of aromatic substituents on DB24C8 restricts the cavity size, introducing steric hindrance despite the smaller size of the piperidinium terminal group. The maximum average oxygen–oxygen distance D_{OO}^{\max} in DB24C8 is approximately 0.2 Å shorter than that in 24C8 in complexes with guest $[1a]^{2+}$, based on the maximum values for each system. This reduction decreases the flexibility of the macrocycle, resulting in less smooth axle insertion. As a result, the ring is forced to approach the proton of the ammonium group, disrupting the original “sandwich-like” structure formed between the phenyl rings and the terminal group. This distortion allows hydrogen bonding, which stabilizes the macrocycle and helps it thread along the axle.

The conformational rearrangement of the macrocycle as it passes through the terminal group is confirmed by measuring the distances between the centers of mass of the phenyl rings

(D_{CM}). After overcoming the energy barrier, the pyridinium unit on the guest undergoes a rotation that allows π -stacking interaction with one of the catechol rings of DB24C8, while the other remains anchored by a hydrogen bond to the piperidinium group. This dual anchoring defines the maximum separation between the aromatic rings, as shown in the orange curve of Fig. 5A, and stabilizes the lowest-energy structure through the combined contribution of hydrogen bonding and π - π stacking interactions. This is further supported by the NCI analysis, which reveals a sheet-like isosurface between the catechol and pyridinium groups in P₂₀₈ (Fig. 5B). Additionally, the rotation of the piperidinium terminal group breaks the hydrogen bond present in the most stable conformation, allowing it to approach the pyridinium unit. This promotes the formation of new hydrogen bonds that favor the π - π stacking of both phenyl rings with the pyridinium group, as shown in P₂₃₆ (Fig. 5B).

For the claw-like mechanism to occur, both the size of the terminal group and the flexibility of the macrocycle are crucial factors. When a bulky group like azepanium is combined with a rigid macrocycle like DB24C8, a smoother displacement coordinate is observed compared to the $[1b \subset 24C8]^{2+}$ complex (Fig. 6A). This difference arises because the rigidity of the macrocycle and the steric bulk of the terminal group prevent the formation of a hydrogen bond, resulting in a higher energy



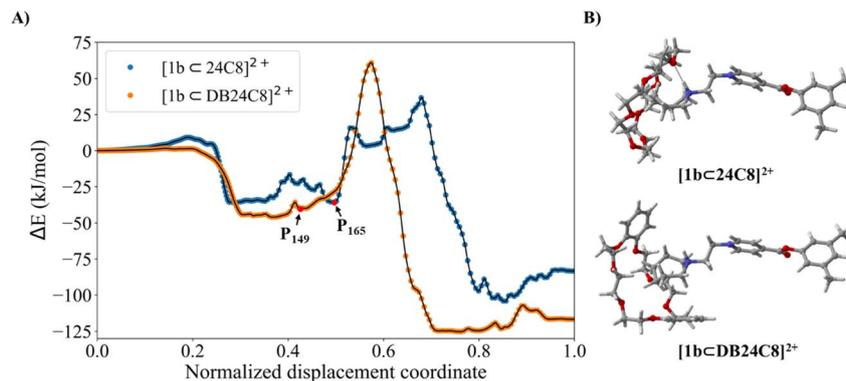


Fig. 6 (A) Energy profiles for the self-assembly normalized displacement coordinate of $[1b\subset 24C8]^{2+}$ and $[1b\subset DB24C8]^{2+}$. (B) representation of a structure corresponding to the $[1b\subset 24C8]^{2+}$ complex (P_{165}), showing the hydrogen-bond interaction of the macrocycle, and the structure of the $[1b\subset DB24C8]^{2+}$ complex at point P_{149} , showing the opening of the macrocycle of DB24C8.

pathway. For complex $[1b\subset DB24C8]^{2+}$, the macrocycle expands as the azepanium group approaches, adopting an L-shaped conformation (Fig. 6B), which restricts the maximum opening of DB24C8 (7.9 Å) compared to 24C8 (8.2 Å), highlighting the conformational constraint imposed by the catechol rings. As a result, the formation of a hydrogen bond between the azepanium proton and the phenolic oxygen atoms is hindered, favoring instead an interaction with one of the glycol chain oxygen atoms of DB24C8. This interaction enables guest rotation and promotes π - π stacking between the pyridinium group and one of the phenyl rings, contributing to the stabilization of the lowest-energy structure (Fig. S40).

The relationship between the rigidity of the macrocycle and the size of the terminal group becomes clearer when analyzing the assembly of the $[1c\subset 24C8]^{2+}$ complex. In this case, the assembly mechanism resembles that seen in systems with DB24C8 and azepanium; however, the increased flexibility of 24C8 allows for more efficient expansion to accommodate the steric hindrance caused by the terminal group. This is especially evident in the green curve of Fig. 7A, where the higher maximum value of 8.2 Å indicates the ability of the macrocycle to adapt more easily to the size and shape of the incoming azocanium group. The azocanium unit, which is even bulkier than azepanium, cannot form a hydrogen bond during the initial threading stage, preventing the claw-like mechanism and forcing the macrocycle to adopt the conformation shown in structure P_{169} (Fig. 7B). Along the displacement coordinate, the gradual adjustment of the O-C-C-O dihedral angles helps the macrocycle to expand, while the azocanium group maintains a twisted boat-chair geometry.

During the activation stage, an energy increase of 104 kJ mol^{-1} results from the expansion of the macrocycle in the absence of any significant stabilizing interactions. The P_{183} highest-energy structure is reached after a C-H \cdots O hydrogen bond has formed between a methylene group on the azocanium ring and an oxygen atom from 24C8 (Fig. 7C). The interaction, together with the proximity of the 24C8 macrocycle to the ammonium group, halts the increase in energy and promotes the formation of a more stable N $^+$ -H \cdots O hydrogen bond,

resulting in a significant decrease in the total energy of the supramolecular complex.

Subsequently, the system adopts its most stable structure along the displacement coordinate, forming an S-type conformation. This geometry is stabilized through a combination of hydrogen bonds and electrostatic interactions occurring between the azocanium unit and the lower region of the 24C8 macrocycle. These interactions are evident from the NCI isosurfaces shown in Fig. 7C and the negative contribution to electrostatic energy seen in structure P_{218} (Fig. 7D). Notably, this S-type conformation consistently appears across all complexes using 24C8, indicating that, although the terminal group size affects the assembly pathway, it does not prevent the formation of the final stabilized geometry, thanks to the inherent flexibility of the macrocycle.

In contrast, efforts to model the assembly of the $[1c\subset DB24C8]^{2+}$ complex using the same displacement-coordinate approach revealed significant limitations. The passage of the bulky azocanium group through the rigid DB24C8 macrocycle caused multiple bond reorganizations within the ring, even breaking some bonds in the macrocycle, and producing erratic energy profiles. These results indicate that the combined steric hindrance from the terminal group and the limited expansion capacity of DB24C8, where D_{OO}^{max} does not exceed 7.9 Å, hinders the effective accommodation of the guest. Consistently, there is no experimental evidence for the formation of the $[1c\subset DB24C8]^{2+}$ complex, which further supports the interpretation that its assembly is not feasible without rupture of the macrocycle. Overall, these findings emphasize that macrocycle flexibility and the size of the terminal group are critical factors governing supramolecular assembly. Moreover, they underscore the essential roles of steric effects and electrostatic contributions in enabling key hydrogen-bonding interactions that influence the viability and energetics of the interlocking process.

To examine the role of electrostatic interactions, energy curves were calculated for two additional axes, specifically, $[2]^+$ and $[3]$. Guest $[2]^+$ was built by replacing the piperidinium group with a cyclohexyl group, while guest $[3]$ involved



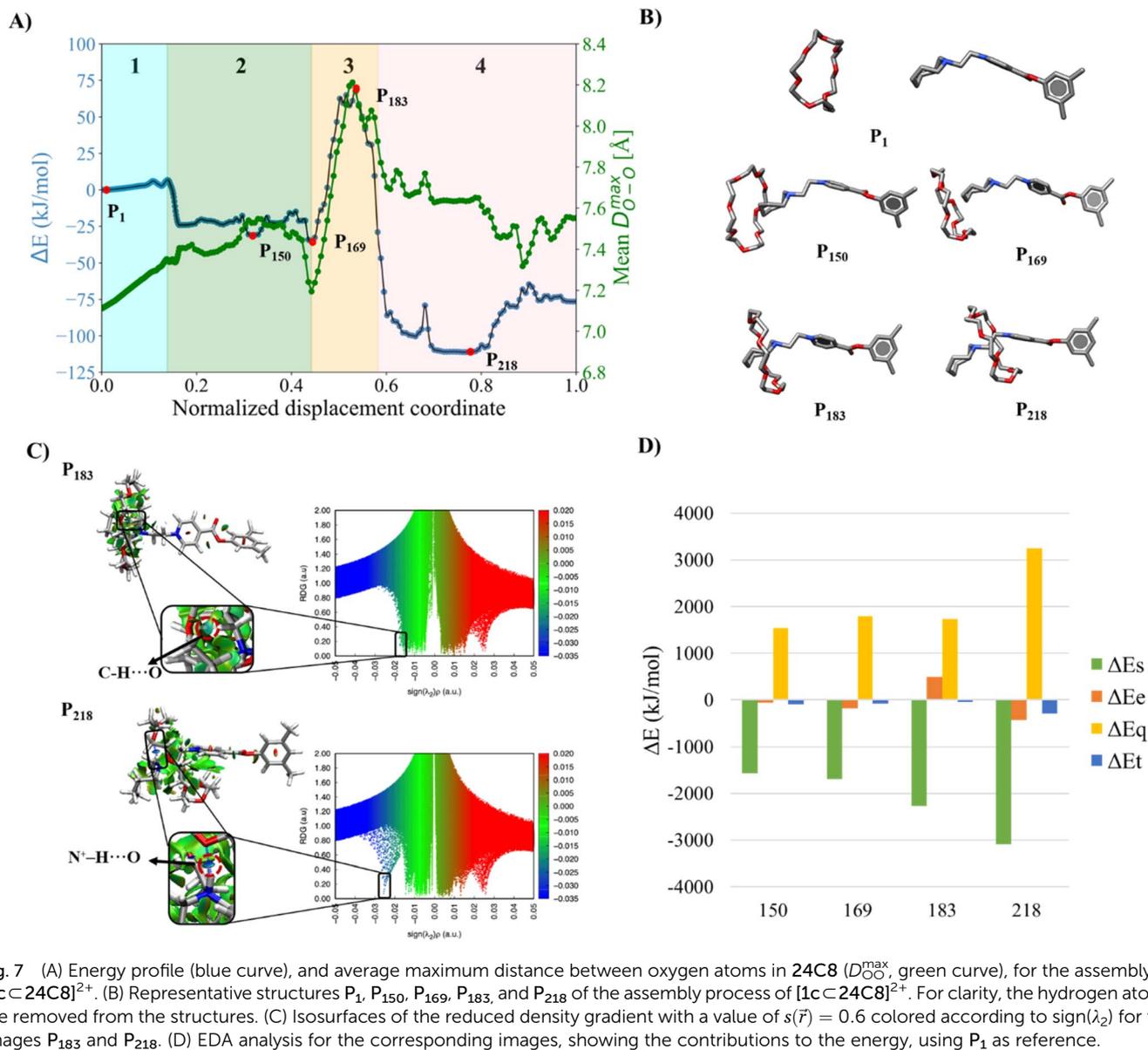


Fig. 7 (A) Energy profile (blue curve), and average maximum distance between oxygen atoms in $24C8$ (D_{O-O}^{max} , green curve), for the assembly of $[1c \subset 24C8]^{2+}$. (B) Representative structures P_1 , P_{150} , P_{169} , P_{183} , and P_{218} of the assembly process of $[1c \subset 24C8]^{2+}$. For clarity, the hydrogen atoms are removed from the structures. (C) Isosurfaces of the reduced density gradient with a value of $s(\vec{r}) = 0.6$ colored according to $sign(\lambda_2)$ for the images P_{183} and P_{218} . (D) EDA analysis for the corresponding images, showing the contributions to the energy, using P_1 as reference.

substituting both positively charged nitrogen atoms with neutral carbon atoms. Note that all axes and complexes are isoelectronic.

The NEB/ANI-1ccx profiles are shown in Fig. 8A. As can be seen, the activation energy barriers for the three systems do not exhibit significant changes until they reach the transition state. However, upon overcoming the transition state, the energy profiles for the complexes $[2 \subset 24C8]^+$ and $[3 \subset 24C8]$ differ significantly from that of complex $[1a \subset 24C8]^{2+}$. Although a valley is also formed for systems with cyclohexyl, for NDC values around 0.6, structures of higher energy than those obtained for the piperidinium complex were found. This is due to their inability to form stabilizing hydrogen bonds between the macrocycles and the axle. The energies of the structures in these valleys are approximately 42 kJ mol^{-1} higher than that corresponding to the $[1a \subset 24C8]^{2+}$ complex.

Recognition of the charge variations among the three rotaxanes is consistent with the changes in stability of the global minima. These minima correspond to the structures illustrated in Fig. 8B, which are located at points P_{220} , P_{236} , and P_{253} along the normalized displacement coordinate for each axle. The first feature to be noted is that the predicted most stable structure is the doubly charged axle, followed by the singly charged axle, and finally, the minimum with the highest energy is that with a neutral axle. Rotaxane $[2 \subset 24C8]^+$ is 19 kJ mol^{-1} higher in energy than the doubly charged complex $[1a \subset 24C8]^{2+}$, while neutral $[3 \subset 24C8]$ is 30 kJ mol^{-1} higher. The observed stability trend $[3 \subset 24C8] < [2 \subset 24C8]^+ < [1a \subset 24C8]^{2+}$ strongly suggests that the electrostatic interactions between axles and macrocycles are fundamental for the stabilization of these rotaxanes.



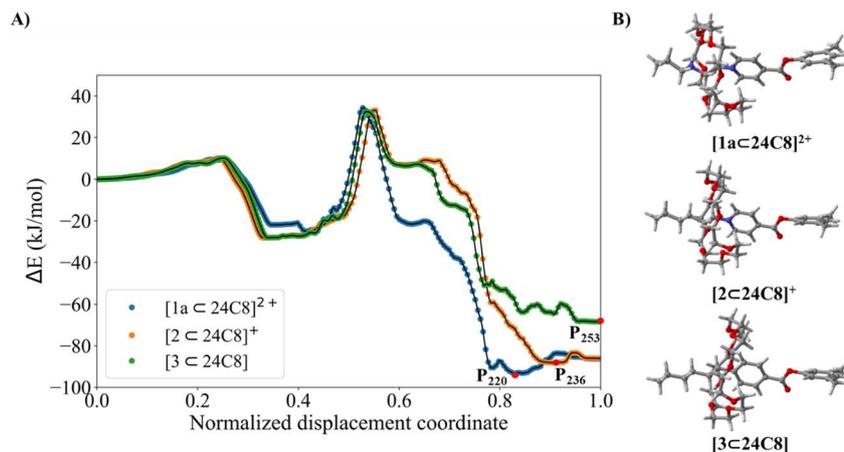


Fig. 8 (A) Energy profiles of the self-assembly process obtained with NEB/ANI-1ccx for $[1a \subset 24C8]^{2+}$, $[2 \subset 24C8]^+$, and $[3 \subset 24C8]$. (B) P_{220} , P_{236} and P_{253} global minima structures for each system, respectively.

The energy profiles for different axes demonstrate the ability of the ANI-1ccx potential to distinguish the charge variations in the three rotaxanes. This finding is significant, considering that ANI-1ccx does not include charged structures in its training set. We attribute this to the architecture and the training of the ANI-1ccx model. Regarding the architecture, one can mention that the input of the ANI models uses symmetry functions, based on the Behler–Parrinello scheme,⁶¹ to describe the local atomic environment in a translationally and rotationally invariant way, and that each element has its own neural network, mapping its atomic environment to an energy contribution. Regarding training, by keeping the DFT-trained early layers and training the final layers with the CCSD(T)/CBS global standard accuracy, overfitting is avoided due to the small size of the gold-standard training set.

The energy decomposition analyses EDA-SBL of structures P_{220} , P_{236} , and P_{253} , corresponding to the energy minima of the $[1a \subset 24C8]^{2+}$, $[2 \subset 24C8]^+$, and $[3 \subset 24C8]$ complexes, are shown in Fig. 2D, S1C, and S2C, respectively. These plots show that at the global minima, the electrostatic energy contributions in the charged complexes are negative, with the doubly charged

system having an electrostatic contribution of almost twice that of the singly charged rotaxane, and the neutral system having an almost vanishing (positive) electrostatic contribution. This result is perfectly aligned with the energy differences obtained from the ANI-1ccx energy profiles. Thus, we can conclude that the methodology NEB/ANI-1ccx correctly predicts the stability of rotaxanes as a function of their electrical charge, even though the training set of this NNP did not contain charged systems.

Finally, to complement these analyses with an energetic description of the association process, we evaluated the activation energy barriers for all systems. Table 1 shows the estimated barriers and the calculation times for convergence, demonstrating good agreement in the ordering and similarity with the experimental free energies, even though the barriers were calculated without considering the effects of temperature or solvent, effects that are certainly important and that could modify the trajectory and the energy profile and therefore the agreement shown in Table 1. These values were obtained directly from the NEB/ANI-1ccx profiles as the energy difference between the maximum in the activation stage and the minimum in the preparation stage.

Table 1 Experimental free energy values and potential energies obtained from NEB/ANI-1ccx calculations, along with the partial (NEB1 and NEB2) and total convergence times, in hours, of the NEB segments composing energy profiles. Note that there are no timings for macrocycle breaking because convergence is not achieved at certain points along the profile^a

Complex	$\Delta G_{\text{exp}}^{\ddagger}$ (kJ mol ⁻¹) ^b	ΔE^{\ddagger} (kJ mol ⁻¹)	t_{NEB1} (h)	t_{NEB2} (h)	t_{total} (h)
$[1a \subset 24C8]^{2+}$	<60.0	59.2	0.04	3.90	3.94
$[1b \subset 24C8]^{2+}$	71.7	72.3	0.02	16.8	16.8
$[1c \subset 24C8]^{2+}$	>100	105.0	0.06	19.7	19.7
$[1a \subset \text{DB}24\text{C}8]^{2+}$	74.2	89.6	0.41	39.5	39.9
$[1b \subset \text{DB}24\text{C}8]^{2+}$	92.8	107.0	0.02	3.80	3.82
$[1c \subset \text{DB}24\text{C}8]^{2+}$	NF	MB	—	—	—
$[2 \subset 24C8]^+$	NR	61.2	0.03	5.70	5.73
$[3 \subset 24C8]$	NR	60.4	0.03	5.50	5.53

^a All calculations were performed using an NVIDIA A100 GPU (80 GB PCIe); NF: No formation; MB: macrocycle breaking; NR: no report in the literature. ^b References $\Delta G_{\text{exp}}^{\ddagger}$ values.^{43,44}



The values in the table are accompanied by the total computational times required by each NEB segment forming the respective energy profile to reach the convergence criterion. The first NEB segment (P_1 to P_{127}), labeled NEB1, converges quickly, as it describes only the approach of the macrocycle to the guest, while the second NEB (P_{127} to P_{253}), labeled NEB2, is the most computationally intensive because it involves the assembly process, reflecting the complexity of the potential energy surface observed in previous profiles. This second step requires more images to obtain detailed mechanistic insight and improve the approximation to the minimum energy path.

Currently, the level of detail discussed in the previous paragraphs can only be achieved using the NEB/ANI-1ccx methodology presented in this work, because the large number of images needed makes it unfeasible with DFT, and even more challenging with the much more demanding CCSD(T). Therefore, the results presented here provide a detailed description of the assembly process and the interactions and energy contributions involved in forming the rotaxane-like complexes, at a very reasonable computational cost, as can be seen in Table 1. Therefore, this work demonstrates the potential of integrating well-trained and meticulously designed machine learning potentials with established yet computationally intensive algorithms to investigate the potential energy surface of complex systems, such as supramolecular assembly, and to elucidate the underlying mechanisms governing system dynamics.

Conclusions

This work shows that by combining the nudged elastic band (NEB) method with the deep neural network ANI-1ccx, it is possible to obtain energy profiles where the energy of a supramolecular complex is plotted against the normalized displacement coordinate (NDC) produced by the NEB algorithm. These profiles are very useful for understanding the interlocking mechanisms that drive the assembly of rotaxanes. The quality of the energies and forces supplied by the ANI-1ccx potential, together with the low computational times required to obtain them, opens up the possibility of generating energy profiles with a resolution that is unachievable using traditional methods like force fields or all-electron calculations, which are at least six orders of magnitude slower and more computationally demanding. In this study, we have presented energy profiles with 250 images, obtained in two segments, which are impossible to attain with DFT and even less feasible with highly correlated wave function methods.

The protocol proposed in this work reveals four energy profile stages: initialization, preparation, activation, and stabilization. Notably, ANI-1ccx effectively describes structural changes in interlocking mechanisms forming supramolecular complexes, especially rotaxanes. This transferable machine-learning model for predicting molecular energies and forces is effective, thanks to three features: it was trained on high-quality CCSD(T)*/CBS data from a smaller set selected from a larger DFT dataset; it encodes local atomic environments invariantly to rotation, translation, and permutation, essential

for transferability and physical accuracy—especially for charged atoms, despite not including charged species in training; and it features a deep network structure that captures complex non-linear relationships in the potential energy surface, like those in supramolecular assembly. These features enable rapid calculations many times faster than CCSD(T)*, with high accuracy, as shown by NEB calculations.

The information gathered from the NDC helped us understand how the size and charge of the terminal group on the guests and the rigidity of the macrocycle influence the energy barrier during assembly. It also helped identify the main interactions that promote macrocycle insertion. By comparing Shubin Liu's energy decomposition analysis of piperidinium and cyclohexyl complexes using **24C8** as the macrocycle, we have shown that electrostatic and steric contributions are essential for reaching the global minima. Additionally, this method allowed us to establish the existence of two key steps during the assembly process under the electrostatically assisted sliding approach, even though, as mentioned above, the ANI-1ccx potential does not include charged systems in its training set. This approach also enabled us to determine the potential energy barriers for threading processes, with an average difference of 8 kJ mol^{-1} compared to the free energy barriers reported in the literature.

In summary, this work paves the way for future research in using machine learning to address complex situations such as supramolecular assembly, where no chemical bonds are broken or formed, and elucidate different mechanisms that govern system dynamics. Important effects to consider in future work include solvent and temperature when predicting free energies.

Author contributions

All authors contributed equally to the conceptualization of the project and the manuscript's design, writing, editing, and revision.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the supplementary information (SI). Supplementary information: A description of (1) the non-interacting systems limit, (2) the energy profiles and EDA-SBL for some of the supramolecular complexes discussed in text, (3) the non-covalent interactions for some of the supramolecular complexes discussed in text, (4) plots of the energy and the distances between the center of mass of the phenyl rings in the $[1a \subset \text{DB24C8}]^{2+}$ complex, and (5) the S-type conformations of rotaxanes, are available in the file Supporting_Information-LPZ-AV-JT.pdf. The animations of the assembly of $[1a \subset \text{24C8}]^{2+}$ and $[1a \subset \text{DB24C8}]^{2+}$, together with the Cartesian coordinates of 8 representative structures discussed in the text, are available in the compressed file



Animation-Coordinates-LPZ-AV-JT.zip. See DOI: <https://doi.org/10.1039/d5sc08303f>.

Acknowledgements

Financial support from Conacyt through Basic Science project 255979 and Sinergias grant 1561802. LPZ thanks Conacyt/Secihti for Ph.D. fellowship 764412. We also acknowledge Laboratorio Nacional de Computación de Alto Rendimiento (LANCAD) for computer time. The authors acknowledge Oscar X. Guerrero-Gutierrez for his assistance in the technical implementation of the NEB/ANI-1ccx code.

References

- 1 S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan and A. L. Nussbaumer, *Chem. Rev.*, 2015, **115**, 10081–10206.
- 2 V. Balzani, A. Credi, F. M. Raymo and J. F. Stoddart, *Angew. Chem., Int. Ed.*, 2000, **39**, 3348–3391.
- 3 J. F. Stoddart, *Angew. Chem., Int. Ed.*, 2017, **56**, 11094–11125.
- 4 J.-P. Sauvage and C. Dietrich-Buchecker, *Molecular catenanes, rotaxanes and knots: a journey through the world of molecular topology*, John Wiley & Sons, 2008.
- 5 S. Mena-Hernando and E. M. Pérez, *Chem. Soc. Rev.*, 2019, **48**, 5016–5032.
- 6 G. Jana and J. L. Mendoza-Cortes, *J. Phys. Chem. A*, 2023, **127**, 2671–2687.
- 7 X. Grabuleda and C. Jaime, *J. Org. Chem.*, 1998, **63**, 9635–9643.
- 8 X. Grabuleda, P. Ivanov and C. Jaime, *J. Org. Chem.*, 2003, **68**, 1539–1547.
- 9 X. Grabuleda, P. Ivanov and C. Jaime, *J. Phys. Chem. B*, 2003, **107**, 7582–7588.
- 10 E. Kolodzeiski and S. Amirjalayer, *J. Chem. Theory Comput.*, 2021, **17**, 7010–7020.
- 11 J. Groppi, L. Casimiro, M. Canton, S. Corra, M. Jafari-Nasab, G. Tabacchi, L. Cavallo, M. Baroncini, S. Silvi, E. Fois and A. Credi, *Angew. Chem., Int. Ed.*, 2020, **59**, 14825–14834.
- 12 L. Leanza, C. Perego, L. Pesce, M. Salvalaglio, M. von Delius and G. M. Pavan, *Chem. Sci.*, 2023, **14**, 6716–6729.
- 13 G. Tabacchi, S. Silvi, M. Venturi, A. Credi and E. Fois, *ChemPhysChem*, 2016, **17**, 1913–1919.
- 14 H. Fu, X. Shao, C. Chipot and W. Cai, *Chem. Sci.*, 2017, **8**, 5087–5094.
- 15 R. F. Lalisce, R. Z. Pavlović, C. M. Hadad and J. D. Badjić, *Phys. Chem. Chem. Phys.*, 2022, **24**, 507–511.
- 16 R. Z. Pavlović, R. F. Lalisce, A. L. Hansen, C. A. Waudby, Z. Lei, M. Güney, X. Wang, C. M. Hadad and J. D. Badjić, *Angew. Chem., Int. Ed.*, 2021, **60**, 19942–19948.
- 17 T. Orlando, G. H. Weimer, P. R. dos, S. Salbego, A. Martinez-Cueva, J. Berna and M. A. P. Martins, *Chem. Eur J.*, 2024, **30**, e202403276.
- 18 Y. H. Jang, S. Hwang, Y.-H. Kim, S. S. Jang and W. A. Goddard, *J. Am. Chem. Soc.*, 2004, **126**, 12636–12645.
- 19 G. Jana and J. L. Mendoza-Cortes, *J. Phys. Chem. A*, 2023, **127**, 2671–2687.
- 20 C. Zazza, N. Sanna, S. Borocci and F. Grandinetti, *ChemPhysChem*, 2025, **26**, e202500074.
- 21 M. E. Foster and K. Sohlberg, *J. Chem. Theory Comput.*, 2007, **3**, 2221–2233.
- 22 H. Fu, X. Shao and W. Cai, *Phys. Chem. Chem. Phys.*, 2022, **24**, 1286–1299.
- 23 L. Bonati, V. Rizzi and M. Parrinello, *J. Phys. Chem. Lett.*, 2020, **11**, 2998–3004.
- 24 J. McCarty and M. Parrinello, *J. Chem. Phys.*, 2017, **147**, 204109.
- 25 G. Bussi and A. Laio, *Nat. Rev. Phys.*, 2020, **2**, 200–212.
- 26 G. A. Siddiqui, J. A. Stebani, D. Wragg, P.-S. Koutsourelakis, A. Casini and A. Gagliardi, *Chem. Eur J.*, 2023, **29**, e202302375.
- 27 M. Schreiner, A. Bhowmik, T. Vegge, J. Busk and O. Winther, *Sci. Data*, 2022, **9**, 779.
- 28 M. Schreiner, A. Bhowmik, T. Vegge, P. B. Jørgensen and O. Winther, *Mach. Learn. Sci. Technol.*, 2022, **3**, 45022.
- 29 P. Zheng, W. Yang, W. Wu, O. Isayev and P. O. Dral, *J. Phys. Chem. Lett.*, 2022, **13**, 3479–3491.
- 30 D. Mathieu, *Mol. Inf.*, 2022, **41**, 2100064.
- 31 L. I. Vazquez-Salazar, E. D. Boittier, O. T. Unke and M. Meuwly, *J. Chem. Theory Comput.*, 2021, **17**, 4769–4785.
- 32 D. L. Folmsbee, D. R. Koes and G. R. Hutchison, *J. Phys. Chem. A*, 2021, **125**, 1987–1993.
- 33 M. Wieder, J. Fass and J. D. Chodera, *Chem. Sci.*, 2021, **12**, 11364–11381.
- 34 N. V. Tkachenko, A. A. Tkachenko, B. Nebgen, S. Tretiak and A. I. Boldyrev, *Phys. Chem. Chem. Phys.*, 2023, **25**, 21173–21182.
- 35 S.-L. J. Lahey, T. N. Thien Phuc and C. N. Rowley, *J. Chem. Inf. Model.*, 2020, **60**, 6258–6268.
- 36 L. Kong and R. A. Bryce, *J. Comput. Chem.*, 2022, **43**, 2009–2022.
- 37 D. Hao, X. He, A. E. Roitberg, S. Zhang and J. Wang, *J. Chem. Theory Comput.*, 2022, **18**, 978–991.
- 38 S.-L. J. Lahey and C. N. Rowley, *Chem. Sci.*, 2020, **11**, 2362–2368.
- 39 B. M. Kriesche, L. E. Kronenberg, F. R. S. Purtscher and T. S. Hofer, *Front. Chem.*, 2023, **11**, 1100210.
- 40 E. Akkus, O. Tayfuroglu, M. Yildiz and A. Kocak, *J. Chem. Inf. Model.*, 2022, **62**, 4095–4106.
- 41 J. Zeng, Y. Tao, T. J. Giese and D. M. York, *J. Chem. Theory Comput.*, 2023, **19**, 1261–1275.
- 42 M. A. Lim, S. Yang, H. Mai and A. C. Cheng, *J. Chem. Inf. Model.*, 2022, **62**, 6336–6341.
- 43 A. C. Catalán and J. Tiburcio, *Chem. Commun.*, 2016, **52**, 9526–9529.
- 44 A. C. Catalán, A. A. Loreda, R. Cervantes and J. Tiburcio, *ChemistryOpen*, 2022, **11**, e202200112.
- 45 D. Sheppard, R. Terrell and G. Henkelman, *J. Chem. Phys.*, 2008, **128**, 134106.
- 46 S. Makri, C. Ortner and J. R. Kermode, *J. Chem. Phys.*, 2019, **150**, 094109.
- 47 G. Henkelman, B. P. Uberuaga and H. Jónsson, *J. Chem. Phys.*, 2000, **113**, 9901–9904.
- 48 C. Riplinger and F. Neese, *J. Chem. Phys.*, 2013, **138**, 034106.



- 49 C. Riplinger, B. Sandhoefer, A. Hansen and F. Neese, *J. Chem. Phys.*, 2013, **139**, 134101.
- 50 J. S. Smith, R. Zubatyuk, B. Nebgen, N. Lubbers, K. Barros, A. E. Roitberg, O. Isayev and S. Tretiak, *Sci. Data*, 2020, **7**, 134.
- 51 H. Jónsson, G. Mills and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, World Scientific, 1998.
- 52 A. Hjørth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, *J. Phys. Condens. Matter*, 2017, **29**, 273002.
- 53 S. Liu, *J. Chem. Phys.*, 2007, **126**, 244103.
- 54 Y. Huang, A.-G. Zhong, Q. Yang and S. Liu, *J. Chem. Phys.*, 2011, **134**, 84103.
- 55 S. Grimme, *Chem. Eur J.*, 2012, **18**, 9955–9964.
- 56 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
- 57 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
- 58 G. B. Bacskay, *Chem. Phys.*, 1981, **61**, 385–404.
- 59 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, *Gaussian 09, Revision D.01*, Gaussian, Inc., Wallingford CT, 2013.
- 60 T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580–592.
- 61 J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, **98**, 146401.

