博士論文

Construction and Application of High-Dimensional Neural Network Potentials for Large-Scale Modeling of Ziegler-Natta Catalysts

筑間　弘樹

主指導教員 谷池 俊明

北陸先端科学技術大学院大学

先端科学技術専攻

[マテリアルサイエンス]

令和７年３月

Referee-in-chief: Professor Toshiaki Taniike

Japan Advanced Institute of Science and Technology

Referees: Professor Yuki Nagao

Japan Advanced Institute of Science and Technology

Associate Professor Shun Nishimura

Japan Advanced Institute of Science and Technology

Associate Professor Kenta Hongo

Japan Advanced Institute of Science and Technology

Professor Akira Nakayama

The University of Tokyo

Abstract

Hiroki Chikuma (2220025)

Ziegler-Natta (ZN) catalysts are heterogeneous catalysts, essential for industrial olefin polymerization. Their functional units, called primary particles, consist of $MgCl_2$ nanoplates with chemisorbed $TiCl_4$, active site precursor, and internal donors (IDs), used to improve stereoregularity of resulting polymers. Insights of the primary particles are essential to understanding the functional origin of ZN catalysts. However, surface reconstruction due to chemisorption introduces structural heterogeneity, posing significant challenges for modeling the nano-scaled structure of ZN catalysts. The non-empirical structural determination of primary particles with such surface reconstructions explicitly considered was achieved using a combination of density functional theory (DFT)-based local geometry optimization and global exploration based on a genetic algorithm (GA). However, the computational cost of DFT limits the structure determination to scales smaller than realistic ones.

Accelerating the DFT geometry optimization is crucial to investigate catalyst structures at real catalyst sizes and compositions. In recent years, methods have been proposed to replace computationally expensive DFT calculations by constructing machine learning potentials (MLPs) that accurately reproduce DFT results. Here, the high-dimensional neural network potential (HDNNP) approach2 was introduced to accelerate structure determination of ZN catalyst primary particles. The HDNNP expresses the total energy as the sum of environmentally dependent atomic energies, where the environment of each atom is described based on the atom-centered symmetry functions (ACSFs).

In **Chapter 3**, HDNNPs were developed for the $MgCl_2/TiCl_4$ binary system, enabling rapid non-empirical structural determination. The reference datasets used for training HDNNPs were sampled from a DFT database, which was accumulated through past DFT-based GA calculations. The accuracy of the constructed HDNNPs was evaluated by comparing geometry optimization results with those obtained from DFT. Constructed HDNNP applied for $50MgCl_2/9TiCl_4$ with experimental size and $TiCl_4$ coverage.

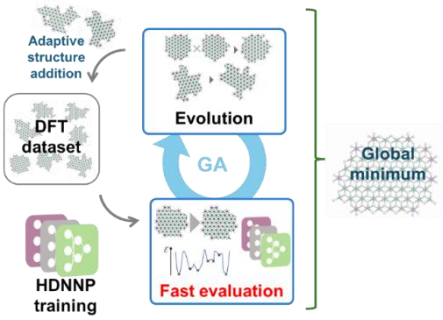


Figure 1. Scheme of constructing HDNNP and its application to structure determination.

Chapter 4 a HDNNP which applies across multiple systems was successfully established, where a reference dataset comprised multiple systems with different numbers of $MgCl_2$ units and $TiCl_4$ molecules. The established HDNNP was used to investigate the impact of the $TiCl_4$ coverage on the stability and active-site distribution. Analysis of metastable structures demonstrated a clear trend of increasing stereospecificity with higher TiCl4 coverage.

Chapter 5 details the development of a HDNNP for a $MgCl_2/TiCl_4$/ID system. While IDs are known to enhance polymer stereoselectivity, their precise impact on catalytic performance remains unclear. Previous studies using DFT and GA by da Silveira et al. identified key structural features for the $19MgCl_2/4TiCl_4$/5diethyl phthalate (DEP) system,[3] but their scope was limited by computational cost. he construction of the HDNNP for this system overcame this limitation, enabling a more comprehensive exploration of the parametric space. This expanded search identified many previously unreported stable and metastable structures, providing deeper insights. Analysis showed that DEP's diverse adsorption patterns significantly contribute to the geometric and electronic diversity of active sites.

In summary, the acceleration of structural determination using HDNNPs enabled the modeling of complex solid catalysts, providing new morphological insights into ZN catalysts that are critical for understanding catalytic systems. This approach demonstrates the importance for large-scale modeling of complex material systems, offering a deeper understanding of their properties.

(1) Takasao, G.et al., *ACS Catal.* **2019**, *9* (3), 2599–2609.
(2) Behler, J.et al., *Phys. Rev. Lett.* **2007**, *98* (14), 1–4.
(3) Da Silveira, J. M., et al. *ACS Catal.* **2024**, *14* (4), 2300–2312.

Keywords: machine learning potential, interatomic potential, density functional theory, genetic algorithm, structure determination, Ziegler-Natta catalyst, internal donor

Preface

The present thesis is submitted for the Degree of Doctor of Philosophy at Japan Advanced Institute of Science and Technology, Japan. The thesis is consolidation of results of the research work on the topic “Construction and application of high-dimensional neural network potentials for large-scale modeling of Ziegler-Natta catalysts” under the supervision of Prof. Toshiaki Taniike during April 2021– March 2024 at Graduate School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology.

Chapter 1 describes a general introduction and the purpose of this thesis. **Chapter 2** presents the numerical methods employed in this research, including the framework for constructing and applying high-dimensional neural network potentials (HDNNPs). In **Chapter 3**, I constructed HDNNPs to accelerate calculations and aimed to achieve structure determination of Ziegler-Natta (ZN) catalysts at realistic scales. **Chapter 4** focuses on the effect of $TiCl_4$ coverage, which is a precursor of the active site, on the morphology and electronic state of the active site of ZN catalysts. The systematic determination of ZN clusters with different coverages was carried out using a more generalized HDNNP. **Chapter 5** reports on the construction of HDNNP with internal donors, which are modifiers of ZN catalysts, and the application of this to structure determination. **Chapter 6** describes the summary and general conclusion of this thesis. The work is original and no part of this thesis has been plagiarized.

Hiroki Chikuma

Graduate School of Advanced Science and Technology

Japan Advanced Institute of Science and Technology

November 2024

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor, Professor Toshiaki Taniike, of the School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology (JAIST), for his invaluable guidance and support throughout my Ph.D. course. His dedication and insights have been instrumental in shaping my growth as a researcher.

I am also indebted to Professor Yuki Nagao, Associate Professor Shun Nishimura, Associate Prof. Kenta Hongo, and Professor Akira Nakayama to revise this thesis.

I would like to acknowledge my second supervisor, Professor Masayuki Yamaguchi for his guidance, Associate Professor Elena Groppo for her invaluable support and collaboration on my miner research project.

I would like to acknowledge the contributions of Professor Jörg Behler for supporting this research through collaboration, and Doctor Gentoku Takasao for providing essential program codes. I would also like to extend my sincere thanks to Assistant Professor Toru Wada and Doctor Patchanee Chammingkwan for their valuable feedback and suggestions throughout my research.

I am deeply grateful to all members of the Taniike Laboratory for their insightful discussions, and support.

This work was supported by JST SPRING, Japan Grant Number JPMJSP2102.

Finally, I would like to express my heartfelt gratitude to my family, who have always encouraged me and supported my dreams.

Hiroki Chikuma

Japan Advanced Institute of Science and Technology

Table of contents

Chapter 1

General Introduction

1.1. First-principles calculation

First-principles calculations are computational methods used to directly estimate physical properties from fundamental quantum mechanical quantities such as charge and mass. These methods provide a predictive capability rooted in the fundamental laws of physics, offering an advantage in accurately modeling complex systems where experimental data may be scarce or unavailable. As a result, first-principles calculations are widely applied in today's computational chemistry for tasks such as structure optimization—predicting stable structures by rearranging atoms along the potential energy surface (PES)—, degerming reaction pathways through transition state calculations, and the prediction of spectra, including infrared (IR), Raman, and ultraviolet (UV). A key feature of this approach is that it does not rely on empirically derived parameters, as the solutions are obtained through quantum mechanical theories and mathematical approximations. To achieve exact results, all quantum interactions within the system must be considered. However, since analytical solutions are not feasible for systems with more than three bodies, numerical solutions using approximations are necessarily introduced for simulations.

In the Schrödinger equation for multi-electron systems, the total Hamiltonian consists of terms representing the kinetic energy of electrons ($\widehat{T_e}$), Coulomb potential between nuclei and electrons ($\hat{V}_{ne}$), Coulomb potential between electrons (V_{ee}), the kinetic energy of nuclei ($\hat{T}_n$), and the Coulomb potential between nuclei ($\hat{V}_{nm}$). The total Hamiltonian of a system ($\widehat{H}_{\mathrm{tot}}$) with multiple nuclei and electrons, be expressed as

$$\widehat{H}_{\mathrm{tot}} = \widehat{H}_e + \widehat{H}_{en} + \widehat{H}_n = \widehat{T_e} + \hat{V}_{ee} + \hat{V}_{ne} + \hat{T}_n + \hat{V}_{nm} \quad (1.1).$$

$$= \frac{-\hbar^2}{2m_{\mathrm{e}}} \sum_i \boldsymbol{\nabla_i}^2 + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\boldsymbol{r_{\mathrm{i}}} - \boldsymbol{r_{\mathrm{j}}}|} - \frac{1}{2} \sum_{i,j} \frac{Z_j e^2}{|\boldsymbol{r_i} - \boldsymbol{R_j}|} - \sum_j \frac{\hbar^2}{2M_j} \boldsymbol{\nabla_j}^2 + \frac{1}{2} \sum_{j \neq k} \frac{Z_j Z_k e^2}{|\boldsymbol{R_j} - \boldsymbol{R_k}|}$$

Here, when the eigenfunction corresponding to the Hamiltonian is denoted as Φ and the eigenvalue as ε, the Schrödinger equation for the entire system is defined as

$$\widehat{H}_{\mathrm{tot}} \Phi(\boldsymbol{r_1}, \boldsymbol{r_2}, \ldots, \boldsymbol{R_1}, \boldsymbol{R_2}, \ldots) = \varepsilon \Phi(\boldsymbol{r_1}, \boldsymbol{r_2}, \ldots, \boldsymbol{R_1}, \boldsymbol{R_2}, \ldots) \quad (1.2).$$

The exact solution of Eq. (1.2) is practically impossible for most systems due to the high computational cost required to apply the Hamiltonian, which involves second-order partial derivatives and many-body interactions, to a multidimensional function. Thus, the Born-Oppenheimer approximation[1], assuming the nuclei are stationary due to their significantly larger mass compared to electrons (protons are 1840 times more massive than electrons) is common to apply. As a result,$\widehat{H}_{\mathrm{n}}$ in Eq. (1.1) is neglected, yielding

$$\widehat{H}_{\mathrm{tot}} = \widehat{H}_{\mathrm{e}} + \widehat{H}_{\mathrm{en}} \quad (1.3).$$

Based on the Pauli exclusion principle, electron pairs with the same quantum numbers must satisfy antisymmetry upon coordinate exchange. This means that the sign of the wave function is reversed, and is expressed as

$$\psi(\ldots, \mathbf{r}_i, \ldots \mathbf{r}_j, \ldots) = -\psi(\ldots, \mathbf{r}_j, \ldots \mathbf{r}_i, \ldots) \quad (1.4).$$

A representation corresponding to this in the form of a determinant was proposed as the Slater determinant, expressed as

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{\mathrm{N!}}} \begin{vmatrix} \Phi_1(1) & \ldots & \Phi_{\mathrm{N}}(1) \\ \vdots & \ddots & \vdots \\ \Phi_1(\mathrm{N}) & \ldots & \Phi_{\mathrm{N}}(\mathrm{N}) \end{vmatrix} \quad (1.5).$$

Hartree-Fock (HF) method, developed in 1930, approximates the multi-electron wavefunction using single-electron orbitals while satisfying this anti-symmetry[2]. However, HF theory has limitations in its chemical accuracy, as it fails to account for electron correlation, resulting in energy errors of about 1%, preventing it from achieving chemical accuracy (1 kcal/mol for energy and 0.1Å for bond length). To address this, methods like configuration interaction was proposed, which include both ground-state and excited-state electron configurations. However, the computational cost increases exponentially with the number of electrons. Thus, balancing computational cost and accuracy, especially in terms of electron correlation, remains a significant challenge in first-principles calculations. Today, density functional theory (DFT), explained in next section, is widely used for its effective balance between chemical accuracy and computational cost.

1.2. Density functional theory

By expressing the total energy as a functional of the electron density rather than directly solving the many-body wavefunction, density functional theory (DFT) offers an efficient method for calculating electronic states in many-body systems. The Hohenberg-Kohn theorem provides the theoretical foundation of DFT and establishes two key principles:[3]

- The external potential $V(\mathbf{r})$ is uniquely determined by the electron density $\rho(\mathbf{r})$.
- The total energy functional $E[\rho]$ satisfies a variational principle: the ground-state density minimizes $E[\rho]$.

In other words, according to the Hohenberg-Kohn theorem, if the exact energy functional for a system can be obtained, the electron density that minimizes the energy

is determined. The Kohn-Sham method reformulates the problem by introducing a system of non-interacting electrons that reproduce the same ground-state density as the interacting system. The electronic state is calculated using the Kohn-Sham equations[4]:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right]\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}) \quad (1.6),$$

$$V_{\text{eff}}(\mathbf{r}) = V(\mathbf{r}) + e^2\int\frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})} \quad (1.7),$$

$$\rho = \sum_i^n \rho_i = \sum_i^n |\phi_i(\mathbf{r})|^2 \quad (1.8).$$

Here, ϕ_i is the *i*-th Kohn-Sham orbital orbital in the non-interacting reference system, V_{eff} is the effective potential, and E_{xc} is the exchange-correlation energy functional. This term is also called simply the functional and includes information on electron correlation.

In practice, DFT calculations use Self-Consistent Field (SCF) iterations, leveraging the variational principle to optimize the electron density:

i. Define an initial set of molecular orbitals and establish the electron density for the target system.
ii. Construct the effective potential V_{eff} using Eq. (1.7).
iii. Solve the Kohn-Sham equations (1.6) to obtain the eigenvalues and molecular orbitals ϕ_i.
iv. Update the electron density using Eq. (1.8).
v. Repeat steps ii–iv until the convergence criteria is satisfied to optimize the electron density.

Since DFT includes electron correlation through the exchange-correlation functional E_{xc}, it typically achieves higher accuracy than the Hartree-Fock (HF) method for many systems. However, the exact form of E_{xc} remains unknown, necessitating approximations. Various functionals have been developed, such as:

- The Local Density Approximation (LDA), which depends solely on the electron density.
- The Generalized Gradient Approximation (GGA), which incorporates the gradient of the density to improve upon LDA.
- Hybrid functionals, which combine different methods or introduce empirical parameters to balance accuracy and computational efficiency.

The choice of functional is critical and should be guided by the system under study, the desired accuracy, and computational cost considerations.

The main computational bottleneck in DFT is solving the Kohn-Sham equations, which requires diagonalizing large matrices. This step scales as $O(N^3)$, where N is the number of electrons, making DFT computationally intensive for large systems.

1.3. Machine learning potential

One of the central challenges of first-principles calculations is the immense computational cost required for electronic state calculations. This demand becomes particularly critical for large-scale systems or simulations involving numerous molecules. As a result, empirical potentials have traditionally been used as an alternative, describing atomic interactions through simplified mathematical models. These models rely on parameters primarily chosen empirically to reproduce experimental results or quantum calculations. A representative example is the Lennard-

Jones potential[6]. Simulations based on empirical potentials are governed by classical mechanics equations and do not directly compute quantum mechanical effects, run faster than first principle calculation such as DFT. However, there are limitations in the systems they can be applied to and the accuracy they can achieve, such as the inability to describe changes in chemical bonds.

Machine learning potentials (MLP) aim to replace first-principles calculations using machine learning, targeting the accuracy of first-principles calculations at the computational speed of empirical potentials. Unlike empirical potentials, where mathematical models are manually defined by researchers, the predictive power of MLPs depends solely on the machine learning model. This allows for the construction of complex and high-dimensional potential energy surfaces (PES) that are not intuitively formulable by humans.

In recent years, MLP technology has rapidly advanced. Examples include neural network potentials (NNP) using neural networks[7–12], Gaussian Approximation Potentials (GAP)[13] based on Gaussian process regression, and Spectral Neighbor Analysis Potentials (SNAP)[14], which utilize linear regression.

The model structure is given by the three-dimensional coordinates of atoms, but the position vectors given by these coordinates do not possess symmetry under rotation or translation of the entire structure. Therefore, in typical MLP methods, descriptors that represent three-dimensional coordinates or other physical quantities as many-body functions are used as inputs to the learning model. Descriptor types include the Atom-Centered Symmetry Function (ACSF)[7,15], the Smooth Overlap of Atomic Positions (SOAP)[16], and the Coulomb Matrix (CM), which are selected based on the characteristics of the machine learning model and the system being studied. The

overview of MLP construction is performed in the following procedure (also shown in Figure 1.1).

i. Construct a dataset consisting of structures and first-principles calculation results (total energy, interatomic forces).

ii. Convert the structural dataset into descriptor vectors.

iii. Use machine learning to perform regression on the descriptors and the first-principles calculation results.

The accuracy of the constructed machine learning potential (MLP) is comparable to that of reference first-principles calculations. Moreover, MLPs bypass direct electronic structure calculations, allowing energy and forces to be obtained with exceptional speed. This efficiency makes MLPs highly valuable for large-scale simulations and for exploring the PES of complex systems, where traditional first-principles methods would be prohibitively time-consuming.

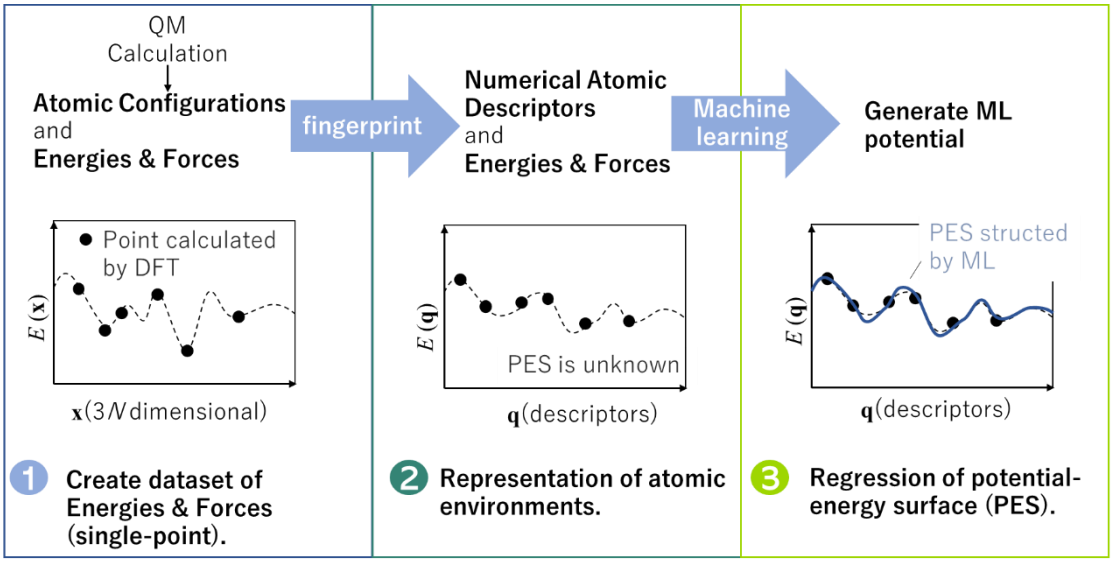


Figure 1.1. Schematic diagram of machine learning potential construction.

1.4. High-dimensional neural network potential

The high-dimensional neural network potential (HDNNP) is an MLP architecture introduced by Behler et al. in 2007[7]. A key feature of HDNNP is the introduction of ACSF and subnetworks. These elements allow HDNNP to achieve invariance to rotational and translational operations of the structure, as well as invariance to the ordering of atoms, while also maintaining scalability concerning the number of atoms in the system. The chemical environment of all atoms in the structure is described by a set of typically 20–100 ACSFs as an input to the corresponding atomic subnetworks[15,17]. Typically, the environment of atom iii is defined based on a combination of radial and angular ACSFs (G_{i}^{ang}) defined as

$$G_i^{\mathrm{rad}} = \sum_{j=1}^{N_{\mathrm{atoms}}} \mathrm{e}^{-\eta\left(R_{ij}-R_{\mathrm{s}}\right)^2} f_{\mathrm{c}}\left(R_{ij}\right) \tag{1.9}$$

$$G_i^{\mathrm{ang}} = 2^{1-\zeta} \sum_{j\neq i}^{N_{\mathrm{atoms}}} \sum_{k\neq i,j}^{N_{\mathrm{atoms}}} \Big\{ \left(1 + \lambda \cos\theta_{ijk}\right)^{\zeta} \mathrm{e}^{-\eta\left({R_{ij}}^2+{R_{jk}}^2+{R_{ik}}^2\right)} f_c\left(R_{ij}\right) f_c\left(R_{jk}\right) f_c\left(R_{ik}\right) \Big\} \tag{1.10}$$

where η, ζ, λ, and R_{s} are hyperparameters, R_{ij} is the distance between atoms i and j, and θ_{ijk} is the angle defined among atoms i, j, and k with i at the corner. The hyperparameters are chosen based on the system and are typically selected for each element pair (or triad). A cutoff function (f_{c}) is typically described using cosine or polynomial functions.

The architecture of the HDNNP is illustrated in Figure 1.2. Here, each atom in the system is assigned an individual subnetwork, which ensures invariance to atom permutation and scalability to systems with varying numbers of atoms. Each

subnetwork takes the ACSF vector of an atom as input and is constructed for each atomic species. The same subnetwork is applied to atoms of the same type. The atomic energy contribution E_i^{ν} is output from each subnetwork, and the total system energy E_{total} is represented as a linear sum of the individual atomic energies, defined as

$$E_{total} = \sum_{\nu}^{N_{elements}} \sum_{i}^{N_{atoms,\nu}} E_i^{\nu} \quad (1.11),$$

where N_{elements} is the number of atomic species in the system, and $N_{\text{atoms},\nu}$ is the number of atoms of species ν. The forces between atoms can be calculated as the analytical gradient of the energy contribution defined as

$$F_{\alpha,i} = -\frac{\partial E_{total}}{\partial \alpha_i} = -\sum_{j=1}^{N_{atoms}} \frac{\partial E_j}{\partial \alpha_i} = -\sum_{j=1}^{N_{atoms}} \sum_{\nu=1}^{N_{G,i}} \frac{\partial E_j}{\partial G_{j,\nu}} \frac{\partial G_{j,\nu}}{\partial \alpha_i} \quad (1.11),$$

where, α denotes any of the *X*, *Y*, or *Z* axes. Since the potential felt by atoms within the cutoff radius changes with atomic movement, the interaction is influenced by atomic environments up to twice the cutoff radius (Figure 1.3). The cutoff radius is determined according to the nature of the system, and is usually set to 6–10Å.[15]

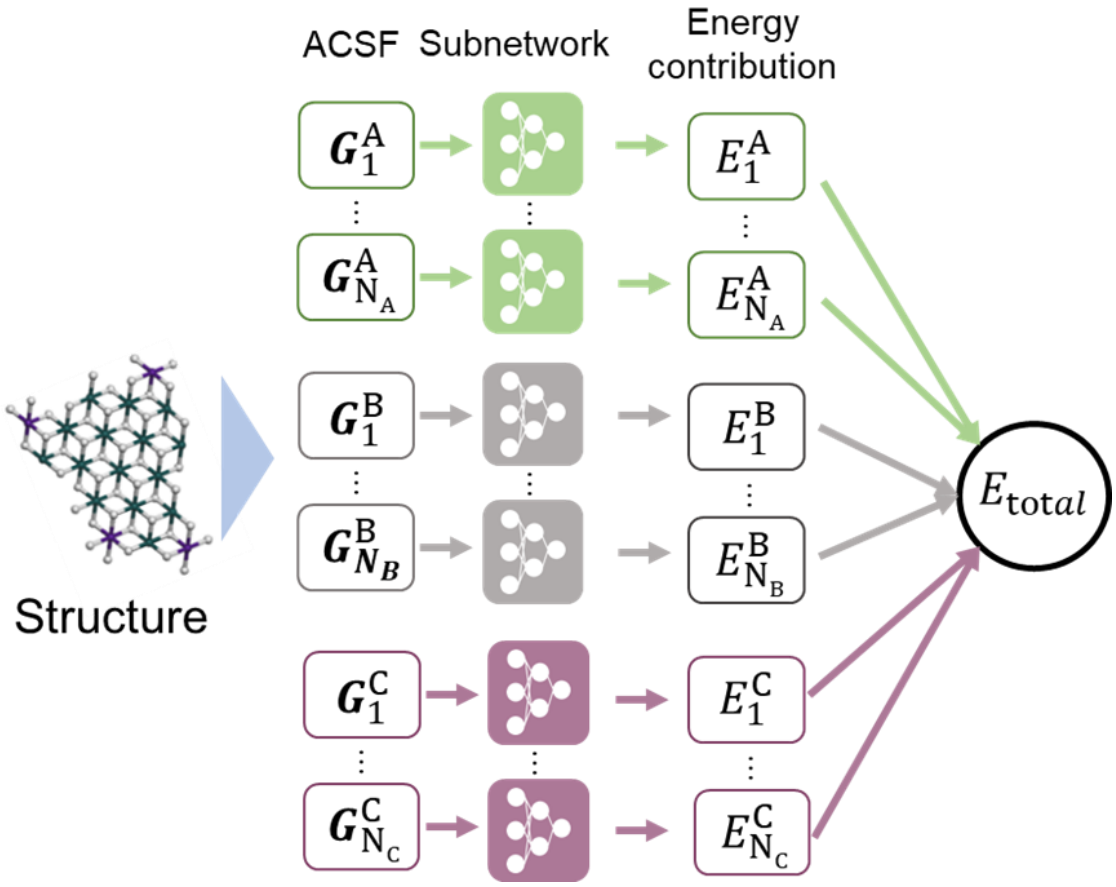


Figure 1.2. Typical architecture of high-dimensional neural network potentials for a system containing elements A, B, and C.

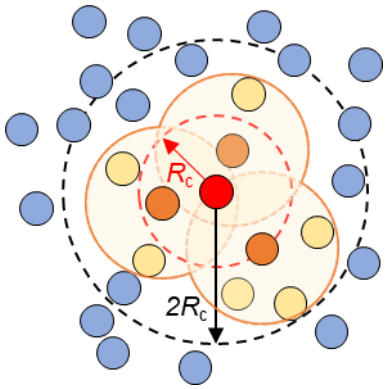


Figure 1.3. Schematic diagram showing the environment dependence of the force vector acting on the central red atom. Because the red atom is within the symmetry function cutoff radius (Rc) of all the orange atoms, the red atom is included in the chemical environment of the orange atoms. In addition, orange atoms affect the

potential of the yellow atoms that are beyond Rc. Therefore, the positions of the red and yellow atoms are mutually connected because they are both involved in the energy contribution of the orange atom. This extends the force's environment dependence to a radius of up to 2Rc.

1.5. Non-empirical structure determination

Clusters are aggregates composed of several to thousands of atoms or molecules, and their properties are influenced by their high surface area and the surrounding chemical and physical environment. In many cases, these physicochemical characteristics differ significantly from those of bulk materials, which are typically defined by periodic structures.[18,19] In small clusters, many constituent atoms are located on the surface, and changes in the cluster size lead to surface restructuring.

Computational approach is a powerful tool to understand cluster properties, yet determining the initial structure for simulations, which represent the nature of system, remains a significant challenge. In experiment, the structure of clusters is commonly discussed by combining techniques such as pair distribution function (PDF) or X-ray diffraction spectra, pore and surface area estimations from gas adsorption, and the analysis of surface chemical environments through IR spectroscopy[20–22]. However, to establish descriptors to link these observations to specific cluster structures is difficult.

In response, theoretical approaches using simulations have been developed to identify the most representable structure of clusters. To define a representative structure for real systems, the search for the global minimum is crucial. Geometry optimizations aim to obtain a stable configuration by minimizing the total energy, adjusting atomic positions toward local minima on the PES. However, the optimized stable structures (local

minima) may not always correspond to the most stable structure (global minimum), which means that consideration and comparison of various local minima are necessary to determine the global minimum (Figure 1.4).

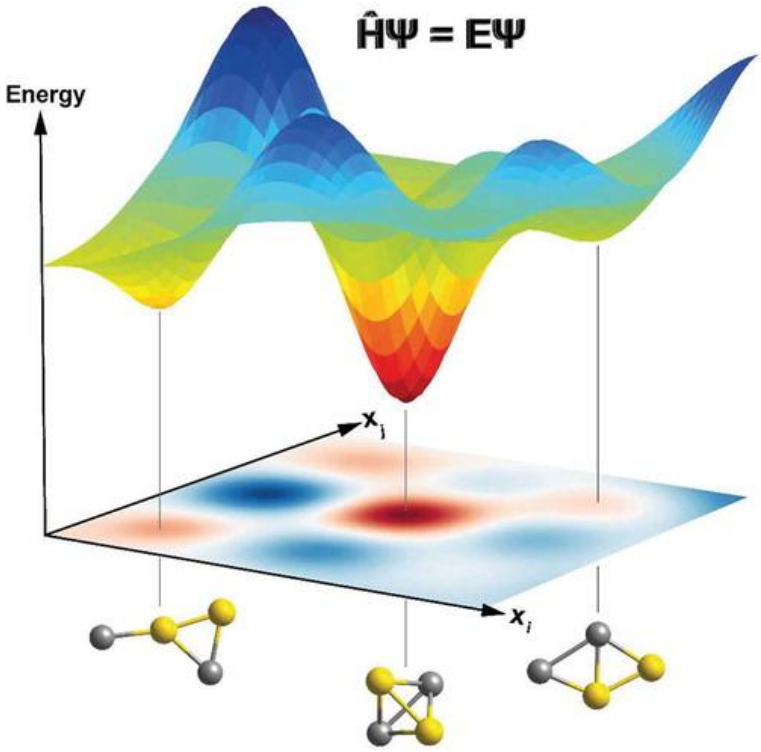


Figure 1.4. Image of minimum structures and the potential energy surface (PES) based on first-principles calculations. The three clusters are each located at a saddle point on the PES, with the most stable "global minimum" represented by the structure in the center. The saddle point obtained from the geometric optimization calculation depends on the initial structure used in the simulation. To identify the structure that corresponds to the global minimum, it is necessary to explore the entire global PES. Derived from Ref.[23]

To address this issue, various methods for global exploration of metastable structures on the PES, including simulated annealing[24], basin hopping[25], and evolutionary algorithms such as genetic algorithms[26,27], which repeatedly design initial configurations non-empirically and perform geometry optimizations, were proposed. The GA method, employed in this study, is widely used to determine the global minimum structure of a system. This algorithm, inspired by the natural processes of selection and evolution, creates new structures through operations such as crossover, which combines the structural features of different conformations, and mutation, which introduces random local structural changes. As a result, the algorithm efficiently facilitates hopping across different regions of the PES, enabling thorough exploration. Global exploration is ideally performed using accurate and versatile computational methods such as DFT[28,29]. However, as the number of atoms in the system increases, the computational cost and time required to find a solution increase rapidly. While global optimization of clusters containing a few dozen atoms has been achieved using DFT, determining the structures of larger clusters becomes exceedingly difficult. Consequently, in recent years, significant efforts have been made to develop more efficient exploration methods by improving optimization algorithms, as well as by replacing DFT calculations with interatomic potentials including MLPs. MLP enhanced structure determination works were reported on various materials such as supported small metal clusters such as $Pt_{10–13}$ on MgO(100)[30] and $Cu_{4–10}$ on ZnO($10\bar{1}0$),[31] as well as for unsupported clusters of a larger size such as Au_{58}[32], $Na_{20–40}$[33], and Cu–Zn alloys consisting of several hundred to thousand atoms.[34]

1.6. Ziegler-Natta catalysts

Ziegler-Natta (ZN) catalysts are polymerization catalysts that play a central role in polyolefin production and are widely used in industry.[35] The discovery of high isotactic propylene polymerization using $TiCl_4$ as a catalyst and AlR_3 as a co-catalyst by Natta in 1954 marked the beginning of various improvements. The industrial significance of ZN catalysts lies in their ability to control the polymerization process at the molecular level, accounting for 99% of global polypropylene (PP) production. By controlling the stereochemistry of polymer chains, it is possible to produce isotactic, syndiotactic, and atactic polymers, making it critical to control the microstructure of resulting polymers, which in turn determines material properties such as melting point, crystallinity, tensile strength, and elasticity.

Today, $MgCl_2$-supported heterogeneous system is widely used. The catalyst precursor is composed of co-adsorption of $MgCl_2$ as a support, $TiCl_4$ as an active species precursor, and organic Lewis bases (internal donors, ID) as modifiers. Internal donors exert steric and electronic effects on the activated Ti^{3+} species, influencing the stereospecificity of resulting polymers[36,37].

The basic structural unit, called primary particles, is a nanostructure of $MgCl_2$ terminated by chemisorbed species. These particles interact with one another through van der Waals forces, leading to their hierarchical aggregation along the c-axis. This aggregation results in a polycrystalline structure containing pores that range in size from micro to macro[38,39]. Regarding typical $MgCl_2$ surface structures, Figure 1.5a illustrates that they consist of 4-coordinated Mg^{2+} or 5-coordinated Mg^{2+} on {110} and {100} surfaces[40]. As shown in Figure 1.6, high-resolution TEM studies have observed the coexistence of (110) and (104) surfaces on activated $MgCl_2$, providing more insight into

the surface structure[22]. Additionally, the introduction of $TiCl_4$ resulted in the formation of an extremely disordered surface structure.[41] The epitaxial adsorption of $TiCl_4$ on different $MgCl_2$ surfaces, as shown in Figure 1.5b, has been supported by computational energy calculations[42–46]. The particle size varies significantly depending on the preparation method; however, primary particles are generally formed within the range of 2.4–4.0 nm,[39,47] with their surfaces believed to be almost entirely covered by adsorbates.[21] Although it is known that the Ti loading affects catalytic performance, increasing the surface exposure of Ti does not necessarily enhance activity.[48] The correlations between composition and surface adsorption structures, as well as between active structures and catalytic performance, remain subjects of ongoing debate despite extensive research over the years.

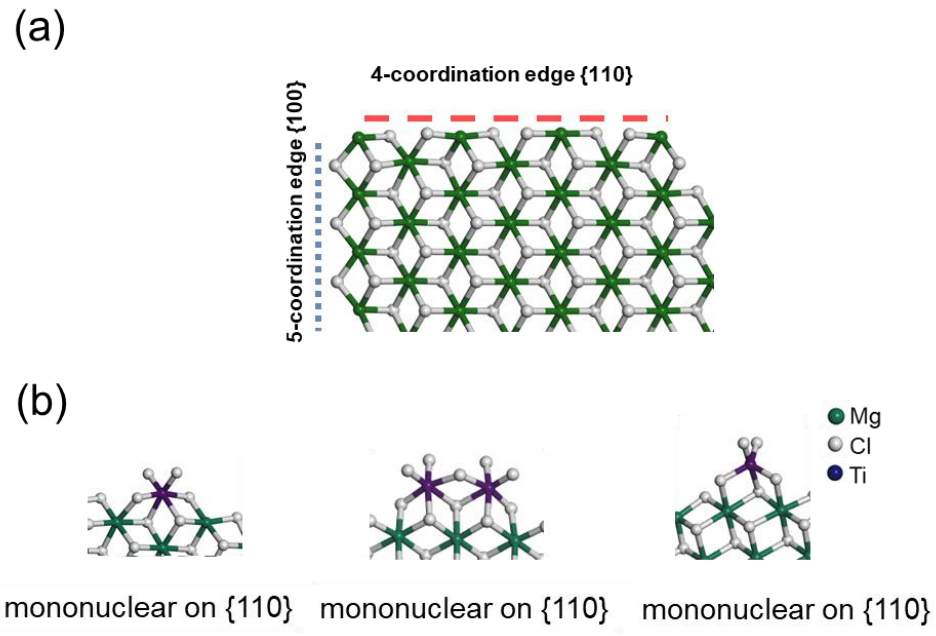


Figure 1.5. a) Surface structure model of a typical $MgCl_2$ plate mono layer. b) Typical $TiCl_4$ adsorption structure. From the left, a mononulear species adsorbed on the {110} plane, a dinuclear species adsorbed on the {100} plane, and a mononulear species of adsorbed on the {110} plane.

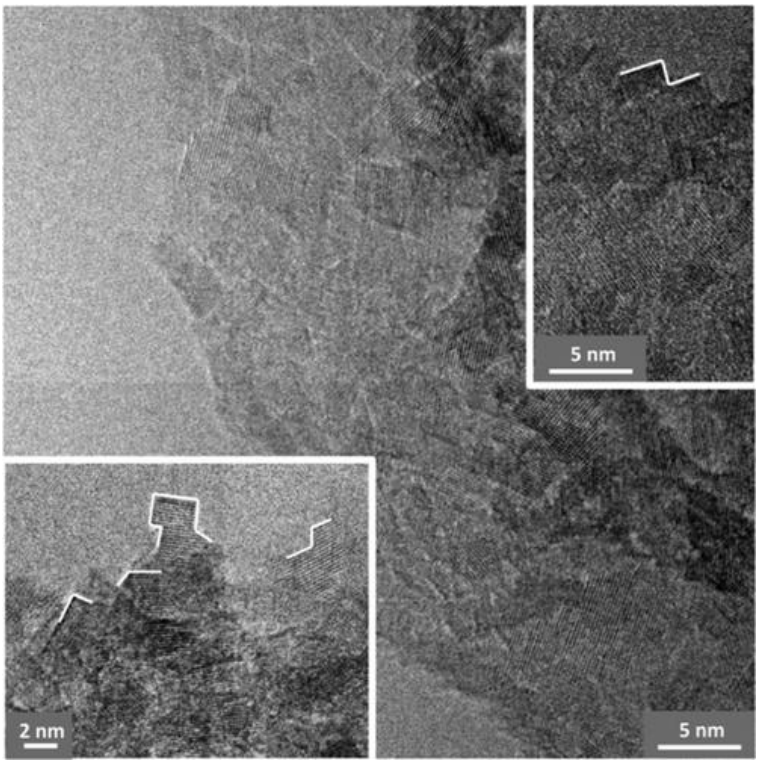


Figure 1.6. Representative high-resolution TEM images of activated $MgCl_2$. White lines highlight nanocrystal edges, revealing both 90° and 120° edge angles. Adapted from Ref[22].

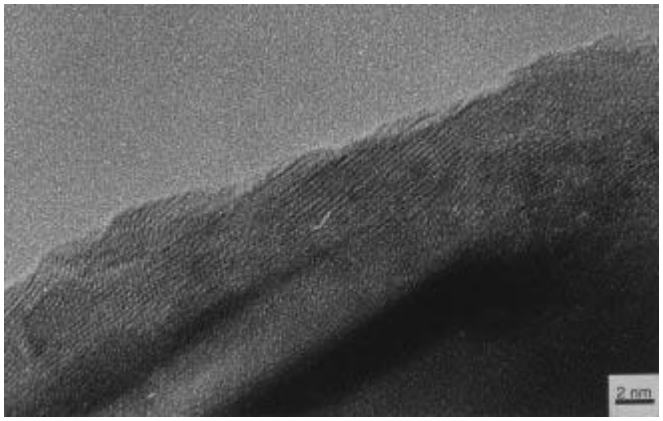


Figure 1.7. High-resolution TEM images of ZN catalyst particle obtained by ball-milling method. The linear surface, similar to that in Figure 1.6 is not visible. Reprinted from Ref.[41]. Copyright John Wiley & Sons.

One important finding regarding the morphology of ZN catalysts is the dynamic response of $MgCl_2$ to specific environments, where surface reconstruction of $MgCl_2$ due to adsorption has been experimentally demonstrated (Figure 1.7). These studies show that surface reconstruction is a critical factor in simulations of ZN primary particles. Consequently, using only slab models or cluster models to represent predefined surfaces such as {100} and {110} plane, or defects on these surfaces, is insufficient. To define the structure of ZN primary particles, it is essential to account for exploring various structure candidates while considering both the reconstruction of $MgCl_2$ surface and the state of adsorbates.

1.7. Internal donor

An important nature of ZN catalysts, particularly those supported on $MgCl_2$, is the use of ID and external donors (ED).[36] These organic compounds play a vital role in

improving the catalyst's stereoselectivity, which is essential for producing polymers with high isotacticity. IDs are introduced during catalyst preparation and commonly include ethyl benzoate (EB), phthalate esters (DEP), diethers (DE), and succinates. ID coordinates to the active sites of the catalyst, influencing the orientation of the monomers during polymerization. This coordination helps achieve the desired stereoregularity of the polymer chain, particularly in the production of polypropylene. EDs are added during the polymerization process, typically in the form of alkoxysilanes. The distribution of active sites is considered the origin of the molecular weight distribution of the resulting polymer, and it is known to vary significantly depending on the selected ID, while the impact of ED selection is minimal[49]. This observation suggests that during polymerization, IDs and EDs play similar roles in enhancing stereospecificity, while the formation and distribution of active sites during catalyst formation are strongly influenced by IDs. Therefore, understanding the adsorption structures of IDs and $TiCl_4$ is crucial for comprehending the function of ZN catalysts. Based on these facts, efforts have been made to determine the structure of ZN catalysts through theoretical calculations as explained in the next section.

1.8. Structure determination of ZN catalyst

Takasago et al. have reported on the non-empirical structure determination of $TiCl_4/MgCl_2$ systems. This is a molecular model of the $TiCl_4/MgCl_2$ system that was obtained by exploration for the thermodynamically most stable structure using DFT and GA[50–52] (Figure 1.8). In addition to the $MgCl_2$ cluster in the 6–19$MgCl_2$ system, non-empirical structure determination was carried out by changing the number of adsorbed $TiCl_4$ molecules, and the most stable structure for each composition and the distribution

of active species at the polymerization temperature were estimated[51]. From these results, it was found that $MgCl_2$ preferentially exposes the {100} plane, but that as the particle size increases due to geometric constraints, the distribution appears on the exposed surface, and that $TiCl_4$ preferentially adsorbs to the {110} plane. However, the effects of particle size and composition in these studies are extremely small compared to the scale of actual catalysts, and it has not been possible to successfully determine the structure of compositions close to actual catalysts. In more recent work, the structure of the $19MgCl_2/4TiCl_4/5ID$ system, in which two different types of ID (diethyl phthalate: DEP and 9,9-bis(methoxymethyl)fluorene) were introduced, was determined, and it was found that the co-adsorption of ID with $TiCl_4$ preferentially forms a sharp surface that includes edges and cavities[53]. This is a characteristic that is not seen in systems without ID, and the diverse adsorption modes play an important role in the reconstruction of $MgCl_2$.

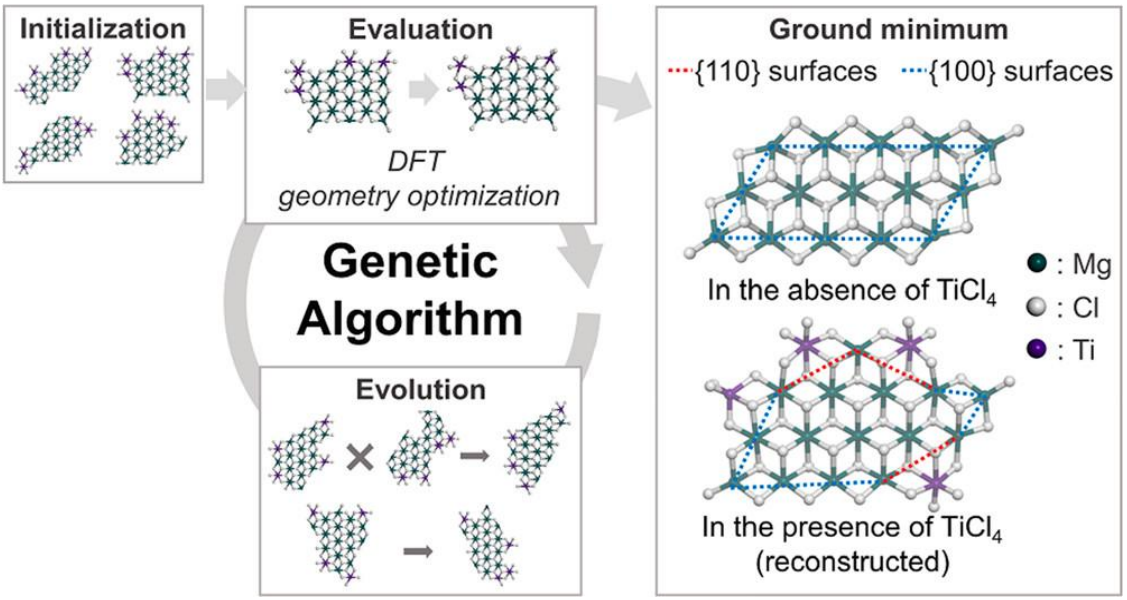


Figure 1.8. Overview of structure determination of ZN catalysts using DFT and GA. Reprinted from Ref[50] with permission. Copyright 2025 American Chemical Society.

The molar ratio of $MgCl_2$ to $TiCl_4$ in general heterogeneous Ziegler-Natta catalysts is thought to be around 4–20%. When the same molar ratio is reproduced using a molecular model with a small particle size, the specific surface area of $MgCl_2$ increases, and the exposed sites of $MgCl_2$ become larger than in the actual catalyst. In addition, the pattern of the surface structure of $MgCl_2$ is limited in small-scale clusters, so it is desirable to determine the structure in a system with a size comparable that of the experimental scale in order to accurately evaluate the physical properties. In addition, it is also important to search for a more comprehensive structure in order to estimate the distribution of active structures. However, due to the limitations of the calculation cost of DFT, it has been necessary to limit the particle size, composition, and space to be searched for in these structure determinations.

To access these computational cost problems, this study implemented a non-empirical structure determination method for more realistic systems by accelerating calculations using HDNNP. This is one of the first HDNNP studies implemented to describe practical and complex catalysts such as ZN catalysts.

1.9. Objective

Non-empirical structural determination of primary particle clusters in Ziegler-Natta (ZN) catalysts has been significantly limited by the computational cost of first-principles calculations. To achieve structural determination for catalysts with sizes and

compositions closer to those of real catalysts and to conduct more comprehensive explorations, accelerating the structural determination process is essential.

In this study, we introduced High-Dimensional Neural Network Potentials (HDNNPs) to accelerate calculations while maintaining high computational accuracy. Additionally, to reduce both the time required for training and the number of density functional theory (DFT) calculations needed for training data, it is desirable to achieve this with a minimal dataset. Therefore, efficient sampling methods play a crucial role. This work represents a rare implementation of machine learning potentials in practical and complex material systems.

In Chapter 2, we constructed HDNNPs for the $MgCl_2/TiCl_4$ binary system and achieved non-empirical structural determination using HDNNPs. I investigated sampling methods to cover the parameter space with fewer structures. And structure determinaion of $50MgCl_2/9TiCl_4$ cluster, which has realistic size and coverage, was performed.

In Chapter 3, I performed systematic structural determination for $MgCl_2/TiCl_4$ systems with varying coverage, aiming to estimate the effects of $TiCl_4$ coverage through realistic-scale simulations. To enable structural determination across multiple compositions, we developed HDNNPs applicable to various compositions. Unlike the previous chapter, where HDNNPs were constructed for each specific composition, this approach allowed the construction of a robust potential by efficiently sampling data from systems with diverse patterns, reflecting significant variations in chemical environments due to size and coverage differences.

In Chapter 4, I constructed a HDNNP and performed structural determination for ZN catalysts ($MgCl_2/TiCl_4$/ID) containing internal donors (ID), specifically focusing on

propylene polymerization catalysts. Using diethyl phthalate (DEP) as the ID, we applied accelerated HDNNP computations to explore a broader structural space for the 19$MgCl_2$/4$TiCl_4$/5DEP system.

This thesis successfully enabled modeling of Ziegler-Natta catalysts comparable to experimental observations of their morphology. The insights gained are indispensable for understanding the catalytic reactions of these materials. Modeling at realistic material scales remains a critical challenge for comprehending the properties of any material system.

REFERENCES

(1) Born, M.; Oppenheimer, R. Zur Quantentheorie Der Molekeln. *Ann. Phys.* **1927**, *389* (20), 457–484. https://doi.org/10.1002/andp.19273892002.

(2) Slater, J. C. A Simplification of the Hartree-Fock Method. *Phys. Rev.* **1951**, *81* (3), 385–390. https://doi.org/10.1103/PhysRev.81.385.

(3) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871. https://doi.org/10.1103/PhysRev.136.B864.

(4) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), A1133–A1138. https://doi.org/10.1103/PhysRev.140.A1133.

(5) Langreth, D. C.; Perdew, J. P. Theory of Nonuniform Electronic Systems. I. Analysis of the Gradient Approximation and a Generalization That Works. *Phys. Rev. B* **1980**, *21* (12), 5469–5493. https://doi.org/10.1103/PhysRevB.21.5469.

(6) On the Determination of Molecular Fields. —II. From the Equation of State of a Gas. *Proc. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character* **1924**, *106* (738), 463–477. https://doi.org/10.1098/rspa.1924.0082.

(7) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 1–4. https://doi.org/10.1103/PhysRevLett.98.146401.

(8) Artrith, N.; Morawietz, T.; Behler, J. High-Dimensional Neural-Network Potentials for Multicomponent Systems: Applications to Zinc Oxide. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2011**, *83* (15), 1–4. https://doi.org/10.1103/PhysRevB.83.153101.

(9) Eckhoff, M.; Behler, J. High-Dimensional Neural Network Potentials for Magnetic Systems Using Spin-Dependent Atom-Centered Symmetry Functions. *Npj Comput. Mater.* **2021**, *7* (1), 1–13. https://doi.org/10.1038/s41524-021-00636-z.

(10) Cooper, A. M.; Kästner, J.; Urban, A.; Artrith, N. Efficient Training of ANN Potentials by Including Atomic Forces via Taylor Expansion and Application to Water and a Transition-Metal Oxide. *Npj Comput. Mater.* **2020**, *6* (1), 1–14. https://doi.org/10.1038/s41524-020-0323-8.

(11) Takamoto, S.; Shinagawa, C.; Motoki, D.; Nakago, K.; Li, W.; Kurata, I.; Watanabe, T.; Yayama, Y.; Iriguchi, H.; Asano, Y.; Onodera, T.; Ishii, T.; Kudo, T.; Ono, H.; Sawada, R.; Ishitani, R.; Ong, M.; Yamaguchi, T.; Kataoka, T.; Hayashi, A.; Charoenphakdee, N.; Ibuka, T. Towards Universal Neural Network Potential for Material Discovery Applicable to Arbitrary Combination of 45 Elements. *Nat. Commun.* **2022**, *13* (1), 2991. https://doi.org/10.1038/s41467-022-30687-9.

(12) Klicpera, J.; Groß, J.; Günnemann, S. Directional Message Passing for Molecular Graphs. **2020**, 1–13.

(13) Deringer, V. L.; Csányi, G. Machine Learning Based Interatomic Potential for Amorphous Carbon. *Phys. Rev. B* **2017**, *95* (9), 1–15. https://doi.org/10.1103/PhysRevB.95.094203.

(14) Thompson, A. P.; Swiler, L. P.; Trott, C. R.; Foiles, S. M.; Tucker, G. J. Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials. *J. Comput. Phys.* **2015**, *285*, 316–330. https://doi.org/10.1016/j.jcp.2014.12.018.

(15) Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115* (16), 1032–1050. https://doi.org/10.1002/qua.24890.

(16) Bartók, A. P.; De, S.; Poelking, C.; Bernstein, N.; Kermode, J. R.; Csányi, G.; Ceriotti, M. Machine Learning Unifies the Modeling of Materials and Molecules. *Sci. Adv.* **2017**, *3* (12). https://doi.org/10.1126/sciadv.1701816.

(17) Behler, J. Perspective: Machine Learning Potentials for Atomistic Simulations. *J. Chem. Phys.* **2016**, *145* (17), 170901. https://doi.org/10.1063/1.4966192.

(18) De Heer, W. A. The Physics of Simple Metal Clusters: Experimental Aspects and Simple Models. *Rev. Mod. Phys.* **1993**, *65* (3), 611–676. https://doi.org/10.1103/RevModPhys.65.611.

(19) Alonso, J. A. Electronic and Atomic Structure, and Magnetism of Transition-Metal Clusters. *Chem. Rev.* **2000**, *100* (2), 637–678. https://doi.org/10.1021/cr980391o.

(20) Wada, T.; Takasao, G.; Piovano, A.; D'Amore, M.; Thakur, A.; Chammingkwan, P.; Bruzzese, P. C.; Terano, M.; Civalleri, B.; Bordiga, S.; Groppo, E.; Taniike, T.

Revisiting the Identity of δ-$MgCl_2$: Part I. Structural Disorder Studied by Synchrotron X-Ray Total Scattering. *J. Catal.* **2020**, *385*, 76–86. https://doi.org/10.1016/j.jcat.2020.03.002.

(21) Piovano, A.; D'Amore, M.; Wada, T.; Cleto Bruzzese, P.; Takasao, G.; Thakur, A.; Chammingkwan, P.; Terano, M.; Civalleri, B.; Bordiga, S.; Taniike, T.; Groppo, E. Revisiting the Identity of δ-$MgCl_2$: Part II. Morphology and Exposed Surfaces Studied by Vibrational Spectroscopies and DFT Calculation. *J. Catal.* **2020**, *387*, 1–11. https://doi.org/10.1016/j.jcat.2020.04.017.

(22) D'Amore, M.; Thushara, K. S.; Piovano, A.; Causà, M.; Bordiga, S.; Groppo, E. Surface Investigation and Morphological Analysis of Structurally Disordered $MgCl_2$ and $MgCl_2/TiCl_4$ Ziegler-Natta Catalysts. *ACS Catal.* **2016**, *6* (9), 5786–5796. https://doi.org/10.1021/acscatal.6b00871.

(23) Jäger, M.; Schäfer, R.; Johnston, R. L. First Principles Global Optimization of Metal Clusters and Nanoalloys. *Adv. Phys. X* **2018**, *3* (1), S100009. https://doi.org/10.1080/23746149.2018.1516514.

(24) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. Optimization by Simulated Annealing. *Science* **1983**, *220* (4598), 671–680. https://doi.org/10.1126/science.220.4598.671.

(25) Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101* (28), 5111–5116. https://doi.org/10.1021/jp970984n.

(26) Whitley, D. A Genetic Algorithm Tutorial. *Stat. Comput.* **1994**, *4* (2). https://doi.org/10.1007/BF00175354.

(27) Johnston, R. L. Evolving Better Nanoparticles: Genetic Algorithms for Optimising Cluster Geometries. *Dalton Trans.* **2003**, No. 22, 4193. https://doi.org/10.1039/b305686d.

(28) Heiles, S.; Johnston, R. L. Global Optimization of Clusters Using Electronic Structure Methods. *Int. J. Quantum Chem.* **2013**, *113* (18), 2091–2109. https://doi.org/10.1002/qua.24462.

(29) Zhao, J.; Shi, R.; Sai, L.; Huang, X.; Su, Y. Comprehensive Genetic Algorithm for *Ab Initio* Global Optimisation of Clusters. *Mol. Simul.* **2016**, *42* (10), 809–819. https://doi.org/10.1080/08927022.2015.1121386.

(30) Kolsbjerg, E. L.; Peterson, A. A.; Hammer, B. Neural-Network-Enhanced Evolutionary Algorithm Applied to Supported Metal Nanoparticles. *Phys. Rev. B* **2018**, *97* (19), 1–9. https://doi.org/10.1103/PhysRevB.97.195424.

(31) Paleico, M. L.; Behler, J. Global Optimization of Copper Clusters at the ZnO(1010) Surface Using a DFT-Based Neural Network Potential and Genetic Algorithms. *J. Chem. Phys.* **2020**, *153* (5), 054704. https://doi.org/10.1063/5.0014876.

(32) Ouyang, R.; Xie, Y.; Jiang, D. E. Global Minimization of Gold Clusters by Combining Neural Network Potentials and the Basin-Hopping Method. *Nanoscale* **2015**, *7* (36), 14817–14821. https://doi.org/10.1039/c5nr03903g.

(33) Chiriki, S.; Bulusu, S. S. Modeling of DFT Quality Neural Network Potential for Sodium Clusters: Application to Melting of Sodium Clusters (Na_{20} to Na_{40}). *Chem. Phys. Lett.* **2016**, *652*, 130–135. https://doi.org/10.1016/j.cplett.2016.04.013.

(34) Weinreich, J.; Paleico, M. L.; Behler, J. Properties of α-Brass Nanoparticles II: Structure and Composition. *J. Phys. Chem. C* **2021**, *125* (27), 14897–14909. https://doi.org/10.1021/acs.jpcc.1c02314.

(35) Chammingkwan, P.; Terano, M.; Taniike, T. High-Throughput Synthesis of Support Materials for Olefin Polymerization Catalyst. *ACS Comb. Sci.* **2017**, *19* (5), 331–342. https://doi.org/10.1021/acscombsci.7b00010.

(36) Taniike, T.; Terano, M. The Use of Donors to Increase the Isotacticity of Polypropylene. In *Polyolefins: 50 years after Ziegler and Natta I*; Kaminsky, W., Ed.; Advances in Polymer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; Vol. 257, pp 81–97. https://doi.org/10.1007/12_2013_224.

(37) Taniike, T.; Terano, M. Coadsorption Model for First-Principle Description of Roles of Donors in Heterogeneous Ziegler-Natta Propylene Polymerization. *J. Catal.* **2012**, *293*, 39–50. https://doi.org/10.1016/j.jcat.2012.06.001.

(38) Taniike, T.; Funako, T.; Terano, M. Multilateral Characterization for Industrial Ziegler–Natta Catalysts toward Elucidation of Structure–Performance Relationship. *J. Catal.* **2014**, *311*, 33–40. https://doi.org/10.1016/j.jcat.2013.10.023.

(39) Zannetti, R.; Marega, C.; Marigo, A.; Martorana, A. Layer-lattices in Ziegler–Natta Catalysts. *J. Polym. Sci. Part B Polym. Phys.* **1988**, *26* (12), 2399–2412. https://doi.org/10.1002/polb.1988.090261202.

(40) Bazhenov, A.; Linnolahti, M.; Pakkanen, T. A.; Denifl, P.; Leinonen, T. Modeling the Stabilization of Surface Defects by Donors in Ziegler-Natta Catalyst Support. *J. Phys. Chem. C* **2014**, *118* (9), 4791–4796. https://doi.org/10.1021/jp412386u.

(41) Mori, H.; Sawada, M.; Higuchi, T.; Hasebe, K.; Otsuka, N.; Terano, M. Direct Observation of MgCl2-Supported Ziegler Catalysts by High Resolution Transmission Electron Microscopy. *Macromol. Rapid Commun.* **1999**, *20* (5), 245–250. https://doi.org/10.1002/(SICI)1521-3927(19990501)20:5<245::AID-MARC245>3.0.CO;2-R.

(42) Boero, M.; Parrinello, M.; Terakura, K. First Principles Molecular Dynamics Study of Ziegler-Natta Heterogeneous Catalysis. *J. Am. Chem. Soc.* **1998**, *120* (12), 2746–2752. https://doi.org/10.1021/ja972367i.

(43) Boero, M.; Parrinello, M.; Terakura, K. Ziegler–Natta Heterogeneous Catalysis by First Principles Computer Experiments. *Surf. Sci.* **1999**, *438* (1–3), 1–8. https://doi.org/10.1016/S0039-6028(99)00537-3.

(44) Taniike, T.; Terano, M. Reductive Formation of Isospecific Ti Dinuclear Species on a $MgCl_2$ (110) Surface in Heterogeneous Ziegler-Natta Catalysts. *Macromol. Rapid Commun.* **2008**, *29* (17), 1472–1476. https://doi.org/10.1002/marc.200800310.

(45) D'Amore, M.; Credendino, R.; Budzelaar, P. H. M.; Causá, M.; Busico, V. A Periodic Hybrid DFT Approach (Including Dispersion) to MgCl2-Supported Ziegler–Natta Catalysts – 1: TiCl4 Adsorption on MgCl2 Crystal Surfaces. *J. Catal.* **2012**, *286*, 103–110. https://doi.org/10.1016/j.jcat.2011.10.018.

(46) Boero, M.; Parrinello, M.; Hüffer, S.; Weiss, H. First Principles Study of Propene Polymerization in Ziegler-Natta Heterogeneous Catalysis. *J. Am. Chem. Soc.* **2000**, *122* (3), 501–509. https://doi.org/10.1021/ja990913x.

(47) Marigo, A.; Marega, C.; Zannetti, R.; Morini, G.; Ferrara, G. Small- and Wide-Angle X-Ray Scattering Analysis of Ziegler–Natta Catalysts: Structural Disorder, Surface Area and Activity. *Eur. Polym. J.* **2000**, *36* (9), 1921–1926. https://doi.org/10.1016/S0014-3057(99)00250-5.

(48) Saelee, T.; Sitthijun, P.; Ngamlaor, C.; Kerdprasit, N.; Rittiruam, M.; Khajondetchairit, P.; Da Silva, J. L. F.; Buasuk, N.; Praserthdam, P.; Praserthdam, S. Experimental and First-Principles Investigation on How Support Morphology Determines the Performance of the Ziegler-Natta Catalyst during Ethylene

Polymerization. *Sci. Rep.* **2024**, *14* (1), 17835. https://doi.org/10.1038/s41598-024-68289-8.

(49) Potapov, A. G.; Bukatov, G. D.; Zakharov, V. A. DRIFTS Study of the Interaction of the AlEt3 Cocatalyst with the Internal Donor Ethyl Benzoate in Supported Ziegler–Natta Catalysts. *J. Mol. Catal. Chem.* **2009**, *301* (1–2), 18–23. https://doi.org/10.1016/j.molcata.2008.11.002.

(50) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Machine Learning-Aided Structure Determination for $TiCl_4$-Capped $MgCl_2$ Nanoplate of Heterogeneous Ziegler-Natta Catalyst. *ACS Catal.* **2019**, *9* (3), 2599–2609. https://doi.org/10.1021/acscatal.8b05080.

(51) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Insight into Structural Distribution of Heterogeneous Ziegler–Natta Catalyst from Non-Empirical Structure Determination. *J. Catal.* **2021**, *394*, 299–306. https://doi.org/10.1016/j.jcat.2020.11.005.

(52) Takasao, G.; Wada, T.; Chikuma, H.; Chammingkwan, P.; Terano, M.; Taniike, T. Preventing Premature Convergence in Evolutionary Structure Determination of Complex Molecular Systems: Demonstration in Few-Nanometer-Sized $TiCl_4$-Capped $MgCl_2$ Nanoplates. *J. Phys. Chem. A* **2022**, *126* (31), 5215–5221. https://doi.org/10.1021/acs.jpca.2c02112.

(53) Da Silveira, J. M.; Chikuma, H.; Takasao, G.; Wada, T.; Chammingkwan, P.; Taniike, T. Deciphering the Role of Internal Donors in Shaping Heterogeneous Ziegler–Natta Catalysts Based on Nonempirical Structural Determination. *ACS Catal.* **2024**, *14* (4), 2300–2312. https://doi.org/10.1021/acscatal.3c05480.

Chapter 2

Numerical method

2.1 General workflow

The primary challenges in structure determination on a realistic scale are the computational cost of DFT calculations and the exponential growth of the parameter space in larger systems. This thesis employed high-dimensional neural network potentials (HDNNPs)[1,2] to represent DFT geometry optimization, enabling significant acceleration in structure determination (Figure 2.1). HDNNPs consist of neural network (NN) subsets corresponding to each atom in the system. These subsets represent the total energy as the sum of atomic energies, ensuring permutational invariance and allowing for predictions across structures with varying numbers of atoms. The weight parameters of the NN subsets are shared among atoms of the same element, and interatomic forces are calculated as the analytical derivatives of atomic energies. Non-empirical structure determination was implemented based on a genetic algorithm (GA) framework.[3–5] In this approach, non-empirical structure determination targeted the primary particles of the Ziegler-Natta catalyst in the $MgCl_2/TiCl_4$ system or the $MgCl_2/TiCl_4$/internal donor (ID) system. Geometry optimization and energy evaluation were performed using HDNNPs instead of DFT, significantly accelerating the structure determination.

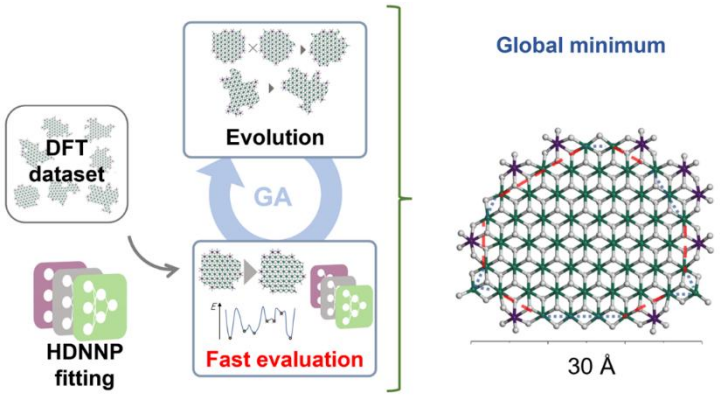


Figure 2.1. Overview of structure determination employing genetic algorithm (GA) and high-dimensional neural network potentials (HDNNPs).

2.2 Construction of high-dimensional neural network potentials

Here, the second-generation Behler-Parrinello-type HDNNPs were consistently employed. Atomic environments were described using Atom-Centered Symmetry Functions (ACSFs), [6,7] which guarantee rotational, translational, and permutational invariance. The environment of atom *i* was described based on a combination of radial ($G_{\mathrm{i}}^{\mathrm{rad}}$) and angular ACSFs ($G_{\mathrm{i}}^{\mathrm{ang1}}$, $G_{\mathrm{i}}^{\mathrm{ang2}}$), defined as,

$$G_i^{\mathrm{rad}} = \sum_{j=1}^{N_{\mathrm{atoms}}} \mathrm{e}^{-\eta\left(R_{ij}-R_{\mathrm{s}}\right)^2} f_{\mathrm{c}}\left(R_{ij}\right) \tag{2.1}$$

$$G_i^{\mathrm{ang1}} = 2^{1-\zeta} \sum_{j\neq i}^{N_{\mathrm{atoms}}} \sum_{k\neq i,j}^{N_{\mathrm{atoms}}} \left\{\left(1 + \lambda\cos\theta_{ijk}\right)^{\zeta} \mathrm{e}^{-\eta\left({R_{ij}}^2+{R_{jk}}^2+{R_{ik}}^2\right)} f_c\left(R_{ij}\right) f_c\left(R_{jk}\right) f_c\left(R_{ik}\right)\right\} \tag{2.2}$$

$$G_i^{\mathrm{ang2}} = 2^{1-\zeta} \sum_{j\neq i}^{N_{\mathrm{atoms}}} \sum_{k\neq i,j}^{N_{\mathrm{atoms}}} \left\{\left(1 + \lambda\cos\theta_{ijk}\right)^{\zeta} \mathrm{e}^{-\eta\left({R_{ij}}^2+{R_{ik}}^2\right)} f_c\left(R_{ij}\right) f_c\left(R_{ik}\right)\right\} \tag{2.3}$$

where η, ζ, λ, and R_{s} are hyperparameters, R_{ij} is the distance between atoms *i* and *j*, and θ_{ijk} is the angle defined among atoms *i*, *j*, and *k* with *i* at the central atom. A cutoff function (f_{c}) with a cutoff radius R_c is defined as,

$$f_{\mathrm{c}}\left(R_{ij}\right) = \begin{cases} 0.5\left[\cos\left(\dfrac{R_{ij}}{R_{\mathrm{c}}}\right) + 1\right] & \left(R_{ij} \leq R_c\right) \\ 0 & \left(R_{ij} > R_c\right) \end{cases} \tag{2.4}$$

Each vector component was normalized in the range of –0.5 to 0.5. Further details about the parameters and functions are provided in Chapters 3 and 5. The HDNNPs were constructed using the RuNNer code (version 1.2.0).[7,8] The number of hidden layers and nodes per layer were optimized based on the system requirements. A hyperbolic tangent function was employed as the activation function for the hidden layers, while a linear function was used for the output layer. The atomic neural networks were trained using a Kalman filter[9] to minimize the following loss function, defined as,

$$\Gamma = \frac{1}{N_{\text{structure}}} \sum_{i}^{N_{\text{structure}}} \left[\left(E^i_{\text{DFT}} - E^i_{\text{HDNNP}}\right)^2 + \frac{\beta}{3N^i_{\text{atom}}} \sum_{j}^{N^i_{atom}} \left(F^i_{j,\text{DFT}} - F^i_{j,\text{HDNNP}}\right)^2\right] \quad (2.5),$$

where $N_{\text{structure}}$ is the number of structures in the training set, N^i_{atom} is the number of atoms in structure i, E^i is the energy of structure i, and F^i_j is the force on atom j in structure i. Note that E^i_{HDNNP} is the sum of atomic energies, and $F^i_{j,\text{HDNNP}}$ is given as an analytical derivative of the atomic energies. The parameter β was adjusted based on the target system.

2.3 Reference dataset

In this thesis, reference datasets were created to train and validate the HDNNPs. These datasets consist of structures and their corresponding DFT calculation results. The structures were primarily obtained from two sources:

- Initial structures generated by GA, are the starting configurations used in the structure determination workflow.

- Trajectories of structure optimization simulations, which include intermediate and optimized structures from simulations performed on the GA-generated initial structures.

The structures in reference datasets were systematically sampled from these two sources. This dataset was then divided into a training set and a test set, with 70% of the data randomly allocated to the training set. Initial DFT data were sampled from the structures generated by the GA employing DFT calculations (DFT-GA).

Furthermore, the initial reference dataset was adaptively enriched by incorporating chemically diverse environments with uncertain predictions. This was achieved by geometry optimization using two HDNNPs with slightly different architectures trained on the same dataset. Geometry optimization was performed for a set of the same initial structures using these HDNNPs, and the following cases were listed for inclusion in the reference dataset after DFT single-point calculations:

- When geometry optimization encountered extrapolation failure, the structures at the time of suspension or the optimized structures with extrapolation warnings were included.
- When the structures optimized by the two HDNNPs differed beyond a threshold in energy or in the root mean square deviation (dRMS) of the interatomic distances. The thresholds were empirically adjusted for each system.

The dRMS between structures *X* and *Y*, both consisting of *N* atoms, is defined as,

$$dRMS(X,Y) = \sqrt{\frac{2}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j>i}^{N}\left(\left|r_{ij}^{X}\right| - \left|r_{ij}^{Y}\right|\right)^2} \qquad (2.8),$$

where $\left|r_{ij}^{X}\right|$ is the distance between atoms *i* and *j* in the structure *X*.

The DFT calculations were performed using the Perdew-Burke-Ernzerhof (PBE) functional and the double numerical basis functions (DNP)[10] with effective core potentials (ECP),[11] as implemented in DMol^3.[10]

2.4 Non-empirical structure determination

Structure determination for the primary particle of ZN catalyst systems was performed using the constructed HDNNPs. The methodology followed was identical to that previously reported by our group, substituting HDNNPs for DFT in local geometry optimization and energy evaluation steps. Starting structures of GA (first generation) were generated randomly, where $MgCl_2$ skeletons were formed using a self-avoiding random walk algorithm, Cl^- anions were placed at anion sites in descending order of coordination number, and adsorbates (including terminal Cl, $TiCl_4$, and ID) were randomly positioned at adsorption sites located on the lateral periphery of the $MgCl_2$ skeletons. Structures in subsequent generations were produced using genetic operators, such as mutation, crossover, and elitism, based on the energy evaluations of the preceding structures.

Initial structures of each generation were then subjected to geometry optimization using a HDNNP implemented by LAMMPS[12] with the n2p2 package.[13] The convergence criterion for optimization was set at 154 meV/Å in the magnitude of the global force vector. The energy of the optimized structures (E_{cluster}) was calculated as the sum of the surface energy and the adsorption energy of $TiCl_4$, expressed as,

$$E_{\text{cluster}} = E - N_{\text{MgCl}_2} \times E_{\text{MgCl}_2} - N_{\text{TiCl}_4} \times E_{\text{TiCl}_4} - N_{ID} \times E_{\text{ID}} \quad (2.6),$$

where E is the total energy of a structure, N_{MgCl_2} is the number of $MgCl_2$ units, E_{MgCl_2} is the energy of a $MgCl_2$ unit in α-$MgCl_2$, N_{TiCl_4} is the number of $TiCl_4$ molecules, and

E_{TiCl_4} is the energy of a $TiCl_4$ molecule in vacuum, N_{ID} is the number of ID molecules, and E_{ID} is the energy of a ID molecule in vacuum.

Parental structures used for the crossover and mutation are chosen by roulette selection based on the fitness (f) defined as,

$$f_i = e^{-3\left(\frac{E_i - E_{min}}{E_{max} - E_{min}}\right)} \quad (2.7),$$

where E_i, E_{min}, and E_{max} are the energy of the i^{th} structure in a generation, and the energies of the most and least stable structures in the same generation. The population is updated by repeating the same procedure, and the loop is terminated when multiple runs converged to the identical structure. Parameters in GA were empirically adjusted according to the characteristics of a target system.

REFERENCES

(1) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chem. Rev.* **2021**, *121* (16), 10037–10072. https://doi.org/10.1021/acs.chemrev.0c00868.

(2) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 1–4. https://doi.org/10.1103/PhysRevLett.98.146401.

(3) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Machine Learning-Aided Structure Determination for $TiCl_4$-Capped $MgCl_2$ Nanoplate

of Heterogeneous Ziegler-Natta Catalyst. *ACS Catal.* **2019**, *9* (3), 2599–2609. https://doi.org/10.1021/acscatal.8b05080.

(4) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Insight into Structural Distribution of Heterogeneous Ziegler–Natta Catalyst from Non-Empirical Structure Determination. *J. Catal.* **2021**, *394*, 299–306. https://doi.org/10.1016/j.jcat.2020.11.005.

(5) Da Silveira, J. M.; Chikuma, H.; Takasao, G.; Wada, T.; Chammingkwan, P.; Taniike, T. Deciphering the Role of Internal Donors in Shaping Heterogeneous Ziegler–Natta Catalysts Based on Nonempirical Structural Determination. *ACS Catal.* **2024**, *14* (4), 2300–2312. https://doi.org/10.1021/acscatal.3c05480.

(6) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134* (7), 074106. https://doi.org/10.1063/1.3553717.

(7) Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115* (16), 1032–1050. https://doi.org/10.1002/qua.24890.

(8) Behler, J. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chem. Int. Ed.* **2017**, *56* (42), 12828–12840. https://doi.org/10.1002/anie.201703114.

(9) Blank, T. B.; Brown, S. D. Adaptive, Global, Extended Kalman Filters for Training Feedforward Neural Networks. *J. Chemom.* **1994**, *8* (6), 391–407. https://doi.org/10.1002/cem.1180080605.

(10) Delley, B. An All-electron Numerical Method for Solving the Local Density Functional for Polyatomic Molecules. *J. Chem. Phys.* **1990**, *92* (1), 508–517. https://doi.org/10.1063/1.458452.

(11) Bergner, A.; Dolg, M.; Küchle, W.; Stoll, H.; Preuß, H. Ab Initio Energy-Adjusted Pseudopotentials for Elements of Groups 13–17. *Mol. Phys.* **1993**, *80* (6), 1431–1441. https://doi.org/10.1080/00268979300103121.

(12) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Comput. Phys. Commun.* **2022**, *271*, 108171. https://doi.org/10.1016/j.cpc.2021.108171.

(13) Singraber, A.; Behler, J.; Dellago, C. Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials. *J. Chem. Theory Comput.* **2019**, *15* (3), 1827–1840. https://doi.org/10.1021/acs.jctc.8b00770.

Chapter 3

Accelerating Non-Empirical Structure Determination of Ziegler-Natta Catalysts with a High-Dimensional Neural Network Potential

ABSTRACT

The determination of catalyst nanostructures with first-principles accuracy using genetic algorithms (GA) is very demanding due to the cubic scaling of the computational cost of density functional theory (DFT) calculations. Here, I demonstrate for the case of Ziegler-Natta $MgCl_2/TiCl_4$ nanoplates how this structure determination can be accelerated by employing a high-dimensional neural network potential (HDNNP) of essentially DFT accuracy. First, when building HDNNPs for $MgCl_2/TiCl_4$ clusters with computationally tractable sizes, I found that the structural diversity in the training set is crucial for obtaining HDNNPs reliably describing the large variety of structures generated by GA. The resulting HDNNPs dramatically accelerated the structure determination, while yielding results consistent with DFT. Subsequently, I developed a multistep adaptive procedure to construct a HDNNP for $MgCl_2/TiCl_4$ clusters consistent in size and $TiCl_4$ coverage with experiments where prior DFT results were scarcely collected. The structure determination and analyses underline the importance of system size and composition in order to predict some experimentally known facts such as the surface morphology and population of isospecific sites.

KEYWORDS: Neural network potential, genetic algorithm, structure determination, Ziegler-Natta catalyst

3.1 INTRODUCTION

Detailed atomistic information about the structure of materials, which is the key for gaining deeper insights into their functions, is difficult to obtain by experiments for complex systems such as solid catalysts. Thus, non-empirical structure determination, combining a global search algorithm and local geometry optimization, especially at the level of a DFT accuracy, has been actively researched.[1–7] A major challenge lies in the fact that the parameter space to be explored and the cost of DFT calculations increase rapidly and simultaneously with the system size, which prevents structure determination at experimentally relevant sizes and complexities.

Machine learning potentials (MLPs) nowadays offer a promising approach to achieve the accuracy of underlying first-principle calculations with a scalability close to that of molecular mechanics.[8–12] Various methods were proposed for construction of MLPs, including neural network potentials (NNPs),[13,14] Gaussian approximation potentials,[15] spectral neighborhood analysis potentials,[16] and moment tensor potentials.[17] MLPs can provide more generalized and expressive potential energy surfaces than classical empirical potentials. Hence, they have been successfully applied to a wide variety of molecular systems, including organic molecules,[18] water,[19,20] metals,[21–24] and metal oxides,[25–29] in the contexts of geometry optimization, molecular dynamics, and structure determination. Of particular relevance, NNP-based structure determination has been reported for supported small metal clusters such as $Pt_{10–13}$ on MgO(100)[30] and $Cu_{4–10}$ on ZnO($10\bar{1}0$),[26] as well as for unsupported clusters of a larger size such as Au_{58},[31] $Na_{20–40}$,[32] and Cu–Zn alloys consisting of several hundred to thousand atoms.[33]

The heterogeneous Ziegler-Natta catalyst is a catalyst responsible for the industrial olefin polymerization.[34] Its solid component has a morphology of micrometer-sized

spheroidal particles, which consist of hierarchical and irregular agglomeration of primary particles as the structural and functional unit of the catalyst.[35] Various experimental results including combined total X-ray scattering and vibrational spectroscopies[36–38] well established that the primary particles are $MgCl_2$ nanoplates whose lateral edges are terminated by chemisorption of catalytic components such as $TiCl_4$ and organic Lewis basic modifiers (called donors). In more detail, truly nanosized $MgCl_2$ plates cannot be formed without the existence of strongly coordinating molecules (i.e., $TiCl_4$ and donors),[36] as the energy of the lateral surfaces is much higher than that of the coordinatively saturated basal {001} plane. Besides, it is known that $MgCl_2$ flexibly reconstructs its structure in the presence of strongly coordinating molecules.[39] These facts suggest that the structure determination for the Ziegler-Natta primary particles necessitates exploring $MgCl_2$ nanostructures in the presence of adsorbates, although, where most of computational studies in the literature presumed a static support surface for clarifying the structure-performance relationship.[40–49] Takasao et al. realized such structure determination, where a combination of a genetic algorithm (GA) and DFT geometry optimization allowed us to simultaneously explore the $MgCl_2$ nanostructure, and the structure and location of $TiCl_4$ adsorbates. The method was successfully applied to structure determination of $x$$MgCl_2$/$y$$TiCl_4$ (x = 6–19, y = 0–4), uncovering the critical role of $TiCl_4$ chemisorption in shaping the $MgCl_2$ nanoplate morphology and in diversifying the structures lying in a thermodynamically accessible energy range with plausible relevance to the multisite nature of the Ziegler-Natta catalyst.[6,49] More recently, the premature convergence of GA, which is rooted to the increased diversity of metastable structures for a larger molecular system, was overcome by implementing a distributed GA based on asynchronous migration from a

structure database. This enabled structure determination of $50MgCl_2/3TiCl_4$ with an experimentally comparable size.[50,51] However, to date the issue of the computational cost of DFT has not yet been solved for this system as no atomistic potentials are available offering efficient alternatives. For example, the structure determination of $50MgCl_2/3TiCl_4$ took more than a year even using a high-performance computing cluster, and more demanding calculations, such as increasing the number of $TiCl_4$ to reach coverages comparable to experiments and Hessian calculations required for thermodynamic analysis, etc., are hardly feasible.

Here, I construct high-dimensional neural network potentials (HDNNPs) for the $MgCl_2/TiCl_4$ system, which I combine with GA. HDNNPs, which have been introduced by Behler and Parrinello in 2007, allow to significantly accelerate the structure determination, while keeping the precision of DFT.[14] As mentioned above, introduced structure determination method explores a vast configurational space composed of the structures of both adsorbates and the support, so it was challenging to establish HDNNPs applicable to a wide variety of atomic environments. First, HDNNPs were constructed by using the results of previously performed DFT calculations, and HDNNP-based structure determination was demonstrated for relatively small $MgCl_2/TiCl_4$ systems. Then, HDNNP construction and structure determination were extended to a $50MgCl_2/9TiCl_4$ system with a realistic size and coverage, for which structure determination has not been performed due to prohibitively high computational cost. The resulting metastable structures were also analyzed to investigate the effect of the $MgCl_2$ size and surface coverage on the distribution of the structures, especially the structure of $TiCl_4$. Note that $TiCl_4$ is not an active form. During polymerization, it is reduced, alkylated, and redistributed upon contact with alkylaluminum.[34,52] Due to the

transient nature of the active sites, such as formation, deactivation and transformation throughout the polymerization process, quantification of the working active site concentration is challenging and is estimated to be quite small relative to the total Ti amount.[53]

3.2 Numerical method

3.2.1 Construction of high-dimensional neural network potentials

Here, the second-generation high-dimensional neural network potentials (HDNNPs) of the Behler and Parrinello type were adopted. HDNNPs express the total energy as a sum of environment-dependent atomic energies provided by individual atomic feed-forward neural networks, and the atomic environment is described based on atom-centered symmetry functions (ACSFs),[54,55] which assure rotational, translational, and permutation invariance. In this study, the environment of atom *i* was described based on a combination of radial ($G_{\mathrm{i}}^{\mathrm{rad}}$) and angular ACSFs ($G_{\mathrm{i}}^{\mathrm{ang1}}, G_{\mathrm{i}}^{\mathrm{ang2}}$), defined as,

$$G_i^{\mathrm{rad}} = \sum_{j=1}^{N_{\mathrm{atoms}}} \mathrm{e}^{-\eta\left(R_{ij}-R_{\mathrm{s}}\right)^2} f_{\mathrm{c}}\left(R_{ij}\right) \tag{3.1}$$

$$G_i^{\mathrm{ang1}} = 2^{1-\zeta} \sum_{j\neq i}^{N_{\mathrm{atoms}}} \sum_{k\neq i,j}^{N_{\mathrm{atoms}}} \left\{\left(1 + \lambda\cos\theta_{ijk}\right)^{\zeta} \mathrm{e}^{-\eta\left({R_{ij}}^2+{R_{jk}}^2+{R_{ik}}^2\right)} f_c\left(R_{ij}\right) f_c\left(R_{jk}\right) f_c\left(R_{ik}\right)\right\} \tag{3.2}$$

$$G_i^{\mathrm{ang2}} = 2^{1-\zeta} \sum_{j\neq i}^{N_{\mathrm{atoms}}} \sum_{k\neq i,j}^{N_{\mathrm{atoms}}} \left\{\left(1 + \lambda\cos\theta_{ijk}\right)^{\zeta} \mathrm{e}^{-\eta\left({R_{ij}}^2+{R_{ik}}^2\right)} f_c\left(R_{ij}\right) f_c\left(R_{ik}\right)\right\} \tag{3.3}$$

where η, ζ, λ, and R_{s} are hyperparameters, R_{ij} is the distance between atoms *i* and *j*, and θ_{ijk} is the angle defined among atoms *i*, *j*, and *k* with *i* at the center. A cutoff function (f_{c}) with a cutoff radius R_c is defined as follows,

$$f_c(R_{ij}) = \begin{cases} 0.5\left[\cos\left(\frac{R_{ij}}{R_c}\right) + 1\right] & (R_{ij} \leq R_c) \\ 0 & (R_{ij} > R_c) \end{cases} \quad (3.4)$$

Note that G^{ang2} was employed when the elemental triad contained two or more Ti atoms (which tend to be far from each other), while the rest of the triads was described by G^{ang1}. The number of radial ACSFs was set to 9 for each element pair, and that of angular ACSFs was 14 for each element triad. For example, the environment of an atom in a $MgCl_2/TiCl_4$ system was described by 27 radial and 84 angular ACSFs, corresponding to a 111-dimensional ACSF vector. Each vector component was normalized in the range of –0.5 to 0.5. Parameters relevant to ACSFs were empirically determined, as described in Tables 1.1–1.3. To describe the atomic environments, a cutoff radius R_c of 16 Bohr has been used due to the ionic nature of the system and the rather large interatomic distances.

The HDNNPs were constructed by using the RuNNer code (version 1.2.0).[10,55] The numbers of hidden layers and nodes per layer were optimized between 2 and 3, and between 10 and 20, respectively. A hyperbolic tangent function was employed as the activation function of hidden layers, and a linear function as that of the output layer. The atomic neural networks were trained using a Kalman filter.[56] The atomic neural networks were trained using a Kalman filter[56] to minimize a loss function, defined as,

$$\Gamma = \frac{1}{N_{\mathrm{structure}}} \sum_{i}^{N_{\mathrm{structure}}} \left[\left(E^i_{\mathrm{DFT}} - E^i_{\mathrm{HDNNP}}\right)^2 + \frac{\beta}{3N^i_{\mathrm{atom}}} \sum_{j}^{N^i_{atom}} \left(F^i_{j,\mathrm{DFT}} - F^i_{j,\mathrm{HDNNP}}\right)^2\right] \quad (3.5),$$

where $N_{\mathrm{structure}}$ is the number of structures in the training set, N_{atom}^{i} is the number of atoms in structure i, E^{i} is the energy of structure i, and F_{j}^{i} is the force on atom j in structure i. Note that E_{HDNNP}^{i} is the sum of the energy of the relevant atoms and that $F_{j,\mathrm{HDNNP}}^{i}$ is given as an analytical derivative of the atomic energies. The parameter β was set to 1.0 in this chapter. More details about the method can be found in recent reviews.[57]

2.2.2 Determination of ACSF parameters

Table 3.1–3.2 include the parameters for the radial and angular ACSFs of the high-dimensional neural network potentials (HDNNPs), respectively. To describe the atomic environments, a cutoff radius R_c of 16 Bohr has been used due to the ionic nature of the system and the rather large interatomic distances.

Three different ACSF sets were compared: Set 1 in which all atomic triads were described by G^{ang1}, set 2 in which all atomic triads were described by G^{ang2}, and set 3 in which G^{ang2} was employed when the elemental triads contain two or more Ti atoms, while the other triads are described by G^{ang1}. Table 3.3 shows the RMSE values of the HDNNPs constructed for 19$MgCl_2$/4$TiCl_4$ using the three ACSF sets, where the HDNNPs were trained on the FPS training set and validated on the FPS test set in Table 3.1–3.2. The ASCF set 3 was found to provide the most accurate HDNNP.

Table 2.1. Parameters of the radial ACSFs. The same parameter set was applied to all the elemental diads.

N	η (1/Bohr^2)	R_s (Bohr)	R_c (Bohr)
1	0.000	0.0	16.0
2	0.002	0.0	16.0
3	0.005	0.0	16.0
4	0.009	0.0	16.0
5	0.016	0.0	16.0
6	0.029	0.0	16.0
7	0.050	3.0	16.0
8	0.050	6.0	16.0
9	0.050	10.5	16.0

Table 3.2. Parameters of the radial ACSFs. The same parameter set was applied to all the elemental triads.

N	η (1/Bohr^2)	λ	ζ	R_c (Bohr)
1	0.001	1.0	1.0	16.0
2	0.001	1.0	2.0	16.0
3	0.001	1.0	4.0	16.0
4	0.001	1.0	16.0	16.0
5	0.001	-1.0	1.0	16.0
6	0.001	-1.0	2.0	16.0

7	0.001	-1.0	4.0	16.0
8	0.001	-1.0	16.0	16.0
9	0.003	1.0	1.0	16.0
10	0.003	1.0	2.0	16.0
11	0.003	1.0	4.0	16.0
12	0.003	-1.0	1.0	16.0
13	0.003	-1.0	2.0	16.0
14	0.003	-1.0	4.0	16.0

Table 3.3. Accuracy of HDNNPs constructed for $19MgCl_2/4TiCl_4$ using different ACSF sets[a,b]

Training dataset	Training		Testing	
	E (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)
Set 1	0.424	67.0	0.494	66.6
Set 2	0.452	72.4	0.538	72.8
Set 3	0.232	50.3	0.215	50.0

[a] Three ACSF sets were prepared: Set 1, in which all atomic triads were described by G^{ang1}, set 2, in which all atomic triads were described by G^{ang2}, and set 3, in which G^{ang2} was employed when the elemental triads contain two or more Ti atoms, while the other triads are described by G^{ang1}.

2.2.3 Reference datasets

In this study, HDNNPs were constructed for four different systems: 19$MgCl_2$, 19$MgCl_2$/4$TiCl_4$, 50$MgCl_2$/3$TiCl_4$, and 50$MgCl_2$/9$TiCl_4$. For the former three systems, structure determination was previously performed at a DFT-generalized gradient approximation (GGA) level, and hence, the results of DFT geometry optimization were accessible for a huge number of structures.[6,49] The reference datasets were obtained in two steps: Sampling a specified number of structures from the DFT databases, where the structures included not only the optimized ones but also the ones obtained in the trajectory of geometry optimization, followed by the addition of structures which exhibited a large deviation in the prediction between two preliminary HDNNPs. Contrary, the last system had scarce prior DFT results so that the construction of the reference dataset and DFT calculations were performed synchronously, whose detail is described later. The numbers of the structures in the final training were 10,000–13,000. The method of the DFT calculations was performed using the Perdew-Burke-Ernzerhof (PBE) functional[58] and the double numerical basis functions (DNP)[59] with effective core potentials (ECP)[60] were implemented by using DMol^3.[59] These setting were selected refer to previous liteatures,[6,49,50] to balance between the computational cost and accuracy.

2.2.4 Non-empirical structure determination

Structure determination was performed for the $MgCl_2$/$TiCl_4$ systems using the constructed HDNNPs. The method of the structure determination was identical to that previously reported by some of us, except that HDNNPs were used instead of DFT for

the local geometry optimization and the energy evaluation. Briefly, starting structures are generated using random numbers, where $MgCl_2$ skeletons are formed based on a self-avoiding random walk, Cl^- ions are placed at anion sites with the descending order of the coordination number, and $TiCl_4$ is randomly placed at adsorption sites defined on the lateral periphery of the given $MgCl_2$ skeletons. These structures are subjected to geometry optimization implemented in LAMMPS with the n2p2 package[61] (an interface that adds the HDNNP method in LAMMPS[62]) and a convergence criterion of 154 meV/Å in the length of the global force vector. The energy of the optimized structures is expressed by E_{cluster}, which is a sum of the surface energy and the adsorption energy of $TiCl_4$, defined as,

$$E_{\text{cluster}} = E - N_{\text{MgCl}_2} \times E_{\text{MgCl}_2} - N_{\text{TiCl}_4} \times E_{\text{TiCl}_4} \quad (3.6)$$

where E is the total energy of a structure, N_{MgCl_2} is the number of $MgCl_2$ units, E_{MgCl_2} is the energy of a $MgCl_2$ unit in α-$MgCl_2$, N_{TiCl_4} is the number of $TiCl_4$ molecules, and E_{TiCl_4} is the energy of a $TiCl_4$ molecule in vacuum. The structures of the descendant generations are generated by crossover, mutation of skeletons and adsorbates, and elitism. If needed, a migration operator can also be implemented, which randomly extracts structures within the energy range of elite structures from a database at a specified frequency of generations, and adds them to the population. This is an effective strategy for preventing premature convergence of the GA, particularly for larger systems.[50] Parental structures used for the crossover and mutation are chosen by roulette selection based on the fitness (f) defined as,

$$f_i = e^{-3\left(\frac{E_i - E_{min}}{E_{max} - E_{min}}\right)} \quad (3.7),$$

where E_i, E_{min}, and E_{max} are the energy of the i^{th} structure in a generation, and the energies of the most and least stable structures in the same generation. The population is updated by repeating the same procedure, and the loop is terminated when multiple runs converged to the identical structure. Parameters in GA were empirically adjusted according to the characteristics of a target system (Table 3.4).

Table 3.4. Parameters used in the genetic algorithm

	$19MgCl_2$	$19MgCl_2/4TiCl_4$	$50MgCl_2/3TiCl_4$	$50MgCl_2/9TiCl_4$
Population	52	90	105	125
Crossover (%)	38	40	38	40
Mutation (skeleton) (%)	15	20	17	16
Mutation (adsorbate) (%)	12	13	11	16
Elitism (%)	35	27	34	28

3.3 RESULT AND DISCUSSION

2.3.1 Construction of the high-dimensional neural network potentials

The number of structures stored in the DFT databases was too huge (1,000,000−3,000,000 structures) to use all of them for training the HDNNPs. Here, I investigate a method of down sampling in order to represent a parametric space with a much smaller number of structures for $19MgCl_2/4TiCl_4$. First, two such small datasets,

respectively termed random and FPS sets, were prepared based on random sampling of 15,000 structures from the DFT database and farthest point sampling (FPS)[63] of the same number of structures. FPS was performed on the ACSF vector space for each element, and it actually selected a structure that included the atomic environment specified by FPS. The random and FPS sets were randomly split into 70% for the training and 30% for the testing. The accuracy of the trained HDNNPs was evaluated based on the root mean square errors (RMSE) of the energy and forces, which represents the deviation from the corresponding DFT values. Moreover, given that trained HDNNPs are to be used for the geometry optimization of various structures produced in GA, additional testing, called geometry optimization testing, was also performed, where 3,000 initial structures (before the geometry optimization), randomly selected from the DFT database, were subjected to the geometry optimization by using the trained HDNNPs. The robustness of the HDNNPs was evaluated in terms not only of the deviation between the HDNNP and DFT geometry optimization results but also of the percentage of successful geometry optimization, which was not terminated by the occurrence of ACSF extrapolation or did not show extrapolation warning at the end of the geometry optimization. A high probability of extrapolation would lead to the loss of certain structural features in the evolution.

The results are summarized in Table 3.5. One can see that the HDNNP trained on the random set was more accurate in the training than those trained on the FPS set, while their accuracy was significantly worsened in the testing, especially for the FPS test set. This is due to an inherent bias present in the DFT database: GA directs sampling towards structures with lower energies in the course of its evolution. As a result, random sampling from the biased database failed to ensure the diversity of ACSF vectors,

leading to overfitting. The diversity of the ACSF vectors in the training set was found even more critical for the geometry optimization testing: Extrapolation failure occurred in about half of the geometry optimizations in the random set, compared with only 7% in the FPS set. Here, one can notice that the RMSE values of energy in the geometry optimization testing was far larger than those in the single-point testing. This is due to the deviation between the minima determined by the HDNNPs and those by DFT.[31] When the structure optimized by the HDNNP is geometry-optimized by DFT, this energy difference is reduced. Note that the deviation was found to be rather systematic, and hence, it hardly affected the energy order of different structures. Figure 3.1 indicates the energy deviation between the HDNNP and DFT for $19MgCl_2/4TiCl_4$. The structures for the geometry optimization testing were geometry-optimized by the HDNNP that was trained on FPS+, and the optimized structures were evaluated by single-point DFT calculations.

Table 3.5 Accuracy of HDNNPs constructed by using differently sampled structures for $19MgCl_2/4TiCl_4$[a,b]

Training dataset	Training		Testing (random)		Testing (FPS)		Testing (geometry optimization)		
	E (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	Success rate (%)

Random	0.356	61.0	0.732	61.2	2.779	134.3	4.758	476.4	50.2
FPS	0.430	69.6	0.428	52.1	0.483	65.2	3.505	56.00	93.0
FPS+	0.501	63.7	0.432	49.3	0.525	61.1	1.549	43.38	99.8

[a] Three reference datasets were prepared by sampling structures from the DFT database: A random set, in which 15,000 structures were selected randomly; a FPS set, in which 15,000 structures were selected by FPS; a FPS+ set, in which 3,098 structures were added adaptively to the FPS training set (see the details in the text). The random and FPS sets were split into 70% for training and 30% for testing.

[b] Root mean squared errors (RMSE) of the energies *E* and forces *F* between the HDNNP predictions and the DFT results. Trained HDNNPs were validated on the random and FPS test sets. Moreover, geometry optimization testing was performed on 3,000 initial structures that were randomly sampled from the DFT database. They were geometry-optimized using the trained HDNNPs, followed by single-point calculations of DFT. The percentage of successful geometry optimization without extrapolation failure was also recorded to measure the robustness of the HDNNPs.

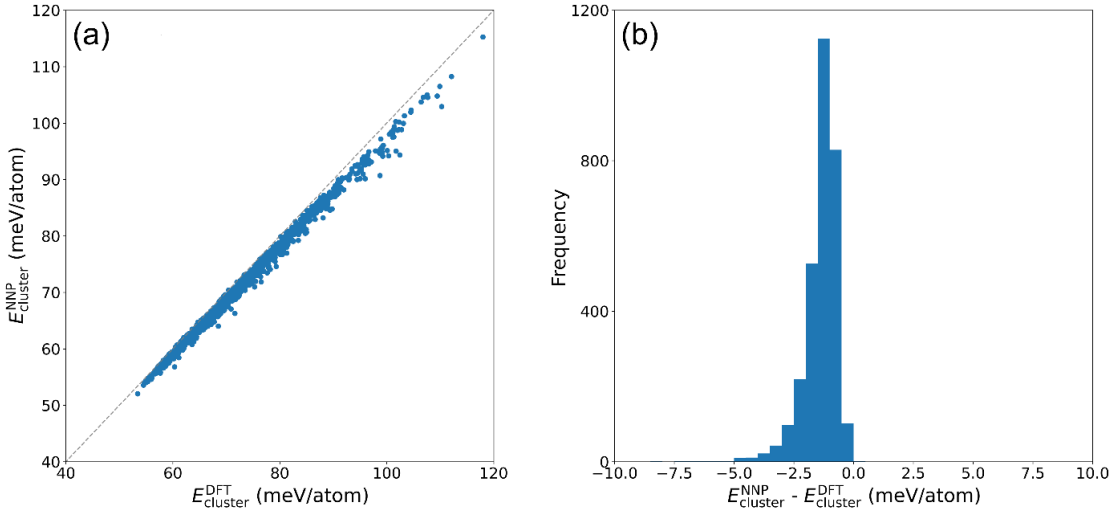


Figure 3.1. a) Comparison of the energy between the HDNNP and DFT, and b) the distribution of the energy deviation. $19MgCl_2/4TiCl_4$ structures in the geometry optimization test set of Table 1 were geometry-optimized by the HDNNP trained on FPS+. The DFT energy was obtained by single-point calculations for the optimized structures.

In order to improve the accuracy and robustness of the HDNNPs, we adopted an adaptive to reinforce the FPS set. In doing so, 5,000 initial structures which were not used in the geometry optimization testing were randomly sampled from the DFT database. They were subjected to geometry optimization by using two HDNNPs having different architectures but with comparable accuracy. When the structures optimized by the two HDNNPs differed from each other beyond 1.0 kcal/mol (i.e., 0.558 meV/atom) in energy or beyond 0.20 Å in root mean square deviation (dRMS) of the distance between all atom pairs, the two structures after the geometry optimization were listed. The dRMS between structures *X* and *Y*, both consisting of *N* atoms, is defined as,

$$dRMS(X,Y) = \sqrt{\frac{2}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j>i}^{N}\left(\left|r_{ij}^{X}\right| - \left|r_{ij}^{Y}\right|\right)^2} \quad (3.8),$$

where $\left|r_{ij}^{X}\right|$ is the distance between atoms i and j in the structure X. Also, when the geometry optimization posed extrapolation failure, the structures at the timing of the suspension or the optimized structures with extrapolation warning were listed. In this way, 3,098 structures were identified and added to the FPS set to obtain the FPS+ training set. As can be seen in Table 1, the HDNNP trained on the FPS+ set exhibited improved accuracy for the geometry optimization testing and negligible occurrence of the extrapolation, while keeping comparable accuracy in the single-point testing. The final accuracy of the HDNNP corresponded to 1.5 meV/atom in energy, which is superior or at least comparable to several meV/atom typically reported in literature.[10]

Having determined how to construct the reference dataset, HDNNP training and testing were applied to the other systems with sufficient prior DFT results (19$MgCl_2$ and 50$MgCl_2$/3$TiCl_4$). As can be seen in Table 3.6, HDNNPs having comparable accuracy and success rate of the geometry optimization were obtained. This suggests the effectiveness of the method based on the FPS and adaptive reinforcement in order to obtain accurate and robust HDNNPs irrespective of the $MgCl_2$ size and $TiCl_4$ coverage. The most accurate HDNNPs consisted of three hidden layers with 15 nodes for the three systems.

Table 3.6. Accuracy of HDNNPs constructed for different $MgCl_2$/$TiCl_4$ systems[a]

System	Training		Testing (FPS)		Testing (geometry optimization)		
	E (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	Success rate (%)
$19MgCl_2$	0.830	61.8	0.852	61.6	1.867	45.9	99.9
$19MgCl_2/4TiCl_4$	0.501	63.7	0.525	61.1	1.549	43.3	99.8
$50MgCl_2/3TiCl_4$	0.329	60.9	0.359	61.2	1.626	50.1	99.9

[a] The training sets were obtained in the same manner as that of FPS+ in Table 1. Note that 30% of the structures after FPS were split for testing, and that 3,000 initial structures were randomly sampled from the corresponding DFT database for additional geometry optimization testing.

Next, I constructed the HDNNPs for $50MgCl_2/9TiCl_4$, where the structure determination had not been realized due to the computational cost, i.e., there were a quite limited amount of prior DFT-GA results (approx. 200,000 structures, corresponding to ca. a few thousands initial structures). The reference dataset was constructed in three steps. First, 12,000 structures were sampled from the limited DFT database based on FPS (termed the FPS set). Second, the trained HDNNP was used in combination with GA to collect structures exhibiting extrapolation. From these

structures, 1,200 structures were sampled based on FPS. They were subjected to DFT single-point calculations, and the results were added to the FPS training set to prepare the FPS+ set. Third, initial structures created in the second step were randomly sampled and subjected to the geometry optimization by using two HDNNPs that were trained on the FPS+ set and possessed different architectures with comparable accuracy. The structures which showed deviations beyond the criteria or extrapolation (similar to the process to prepare the FPS+ set in Table 1) were computed by DFT and the results were added to the FPS+ set to prepare the FPS++ set. The results of the training and testing at each step are shown in Table 3.7 It can be seen that the scores of HDNNPs in the training and single-point testing did not significantly change among the three steps. On the other hand, in the geometry optimization testing, which was performed for 2,000 initial structures selected from the HDNNP-GA database excluding the structures contained in the reference dataset, the success rate of the FPS set was as low as 50%, and it improved stepwise to 95.1% for FPS+, and to 98.4% for FPS++. This again signified the importance of an adaptive procedure in constructing a robust HDNNP. The most accurate HDNNP at FPS++ consisted of two hidden layers with 20 nodes.

Table 3.7. Three-step construction of the HDNNPs for $50MgCl_2/9TiCl_4$[a,b]

Training dataset	Training		Testing (FPS)		Testing (geometry optimization)		
	E (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	Success rate (%)
FPS	0.178	54.0	0.206	54.2	2.214	57.9	50.0
FPS+	0.222	53.3	0.229	51.5	1.441	49.4	95.1
FPS++	0.232	50.3	0.215	50.0	1.017	48.4	98.4

[a] For $50MgCl_2/9TiCl_4$, where prior DFT results were scarcely present, the reference dataset was constructed in three steps: First based on FPS from the limited DFT database (FPS set); Second, based on the addition of structures which showed extrapolation failure during geometry optimization by using the HDNNPs trained on the FPS set (FPS+ set); Third, based on the addition of structures, similar to that of FPS+ for the other systems (FPS++ set). See the details in the text.

[b] The structures for the single-point testing were obtained by splitting 30% from the FPS set, while the structures for the geometry optimization testing were collected by newly creating initial structures based on GA.

2.3.2 Genetic algorithm for structure determination

The constructed HDNNPs were used for GA-based structure determination instead of DFT. The results of two GA runs for each of the systems $19MgCl_2$, $19MgCl_2/4TiCl_4$, $50MgCl_2/3TiCl_4$, and $50MgCl_2/9TiCl_4$ are shown in Figure 3.2 as evolutionary progress plots. The obtained most stable structures are shown in Figures 3.3 and 3.4 The hyperparameters used in GA are summarized in Table 3.4. One can see that the two GA runs converged to the same structure with the identical energy for each system.

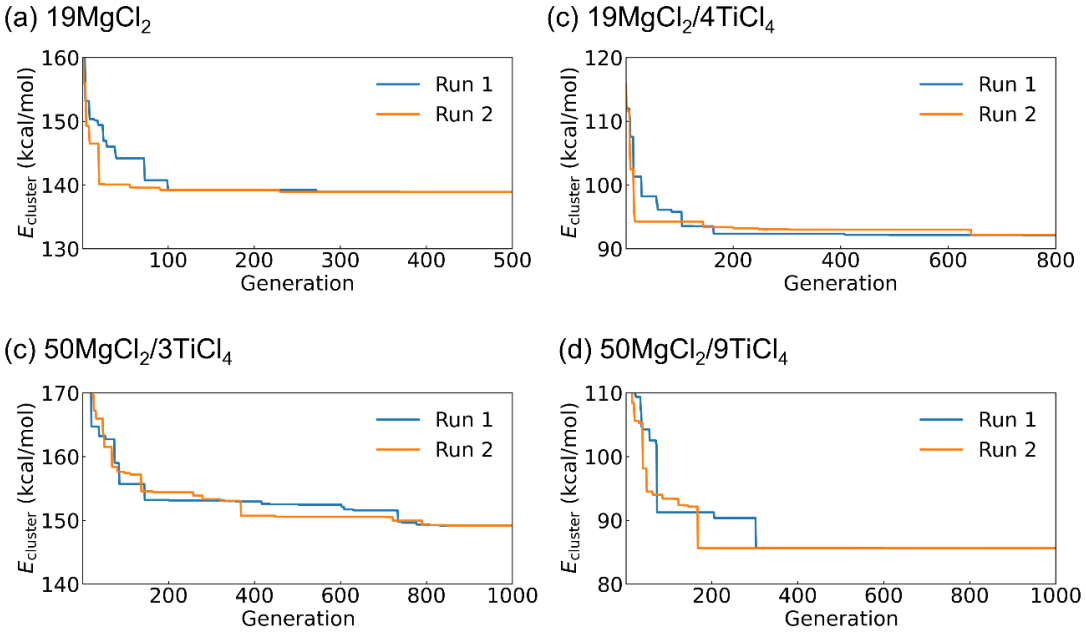


Figure 3.2. Evolutionary progress plots of HDNNP-GA for the structure determination of a)$19MgCl_2$, b) $19MgCl_2/4TiCl_4$, c) $50MgCl_2/3TiCl_4$, and d) $50MgCl_2/9TiCl_4$. The energy of the most stable structure in the population is plotted along with the generation. $E_{cluster}$ is a sum of the surface energy and the adsorption energy of $TiCl_4$, as defined in Eq. 3.7.

For $19MgCl_2$ and $50MgCl_2/3TiCl_4$, the obtained most stable structures were identical to those previously determined by DFT-GA.[49,50] On the other hand, the most stable structure determined by HDNNP-GA for $19MgCl_2/4TiCl_4$ differed from that by DFT-GA. However, the former corresponded to the second most stable structure in DFT-GA, and HDNNP-GA found the most stable structure in DFT-GA during the evolution. The energy difference between these two structures was only 0.1 kcal/mol, which was hard to distinguish based on the given RMSE of HDNNP-GA. A similar reversal in the energy ranking also occurred when the exchange-correlational functional of DFT was changed from PBE to BP. Hence, it was concluded that HDNNP-GA can reach almost the same conclusion as DFT-GA at a dramatically reduced computational cost. For example, the time required for geometry optimization of $50MgCl_2/3TiCl_4$ was typically reduced from 10,000 seconds for DFT to 7 seconds for the HDNNP.

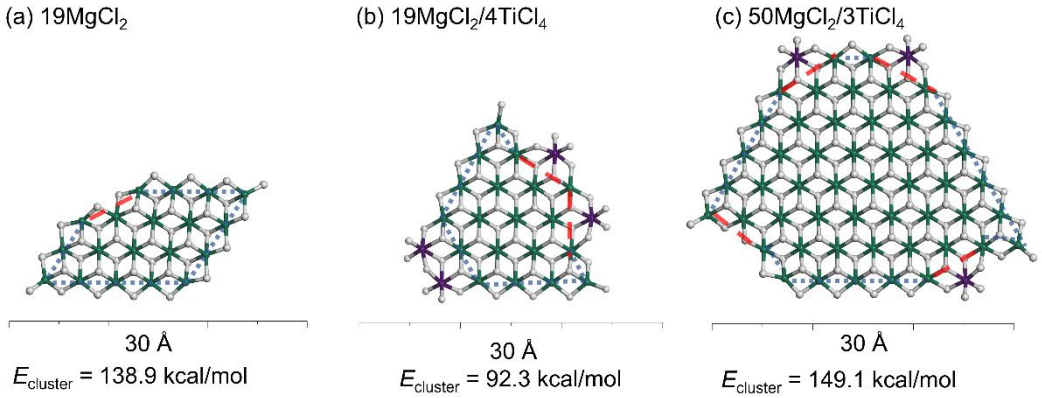


Figure 3.3. The most stable structures and their energies for a) $19MgCl_2$, b) $19MgCl_2/4TiCl_4$, and c) $50MgCl_2/3TiCl_4$ determined by HDNNP-GA. The dotted blue and dashed red lines indicate the $MgCl_2$ {100} and {110} surfaces, respectively.

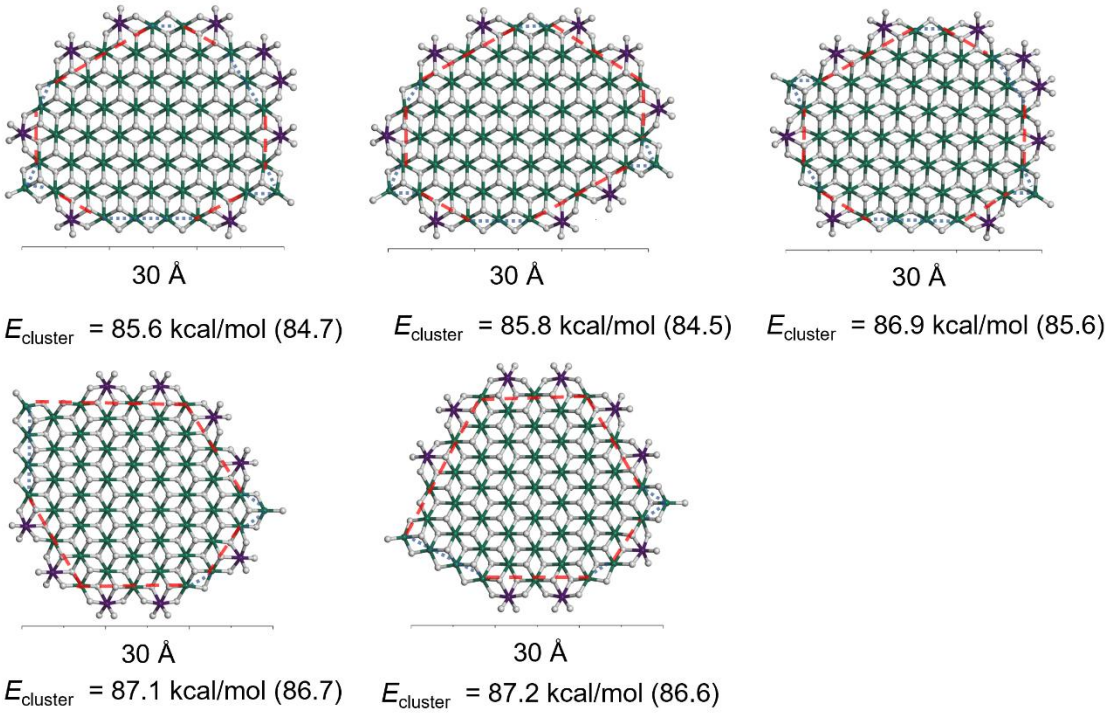


Figure 3.4. Five most stable structures and their energies for 50MgCl2/9TiCl4 determined by HDNNP-GA. The energies of the DFT geometry optimizations are given in the parentheses.

The most stable structure of $50MgCl_2/9TiCl_4$ in DFT-GA was not known in advance. Therefore, we listed the five lowest energy structures obtained by HDNNP-GA, and performed geometry optimization by DFT. The results are shown in Figure 3. One can see that the energies of the most stable and the second most stable structures were reversed between the HDNNP and DFT. However, the energy difference was only 0.2 kcal/mol, indicating successful structure determination in HDNNP-GA. The five most stable structures possessed similar structural features: Most of the $TiCl_4$ molecules were bound to the {110} terraces as mononuclear species, while bare {110} terraces were not exposed. This is because the stronger adsorption of $TiCl_4$ on {110} than that on {100}

overcomes the energetic penalty of exposing the {110} terraces, whose surface energy is higher than that of the {100} terraces. Structural comparison between $50MgCl_2/3TiCl_4$ and $50MgCl_2/9TiCl_4$ revealed the effect of the coverage on the formation of $MgCl_2$ clusters at a realistic size scale. The lateral surfaces of $50MgCl_2/3TiCl_4$ were mainly composed by relatively long {100} terraces with short {110} terraces at their edges, whereas in $50MgCl_2/9TiCl_4$, the lateral surfaces were composed of short {110} and {100} terraces, thus exposing much more edges. These results are likely in line with past high-resolution TEM observations that the lateral surfaces of bare $MgCl_2$ were flat, while those of a catalyst were round-like (suggesting an abundance of edges).[64]

2.3.3 Distribution of $TiCl_4$ species

The dramatic acceleration based on HDNNP-GA resulted in a database much larger in size than that previously acquired by DFT-GA (e.g., in the case of $19MgCl_2/4TiCl_4$, the number of structures explored by HDNNP-GA was 8.6 times that by DFT-GA). This improved the coverage of low-energy structures, i.e., thermodynamically important structures. Figure 3.5 compares the histograms of the relative energies of local minima in $19MgCl_2/4TiCl_4$ system within 6 kcal/mol of $\Delta E_{\text{cluster}}$ collected by HDNNP-GA and DFT-GA. This improved the coverage of low-energy structures, i.e., thermodynamically important structure.

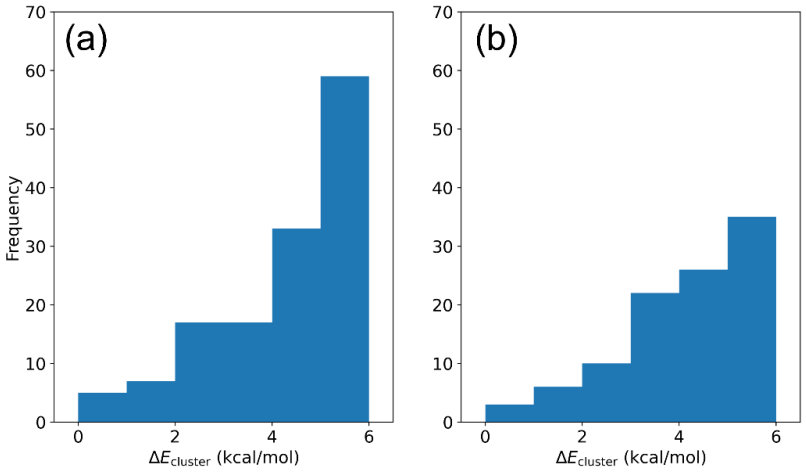


Figure 3.5. Histograms of the low-energy structures for $19MgCl_4/4TiCl_4$, collected by a) HDNNP-GA and b) DFT-GA. The values of $\Delta E_{\text{cluster}}$ in a) were evaluated using HDNNP, and while in b) were evaluated using DFT.

Here, the database was analyzed in order to get an insight into the distribution of catalytically relevant features. The population (p) of each structure was estimated based on the Boltzmann factor,

$$p = e^{-\frac{\Delta E_{\text{cluster}}}{RT}} \qquad (3.9),$$

where $\Delta E_{\text{cluster}}$ is the cluster energy of a structure relative to that of the most stable structure, R is the gas constant, and the temperature (T) was set to 350 K.

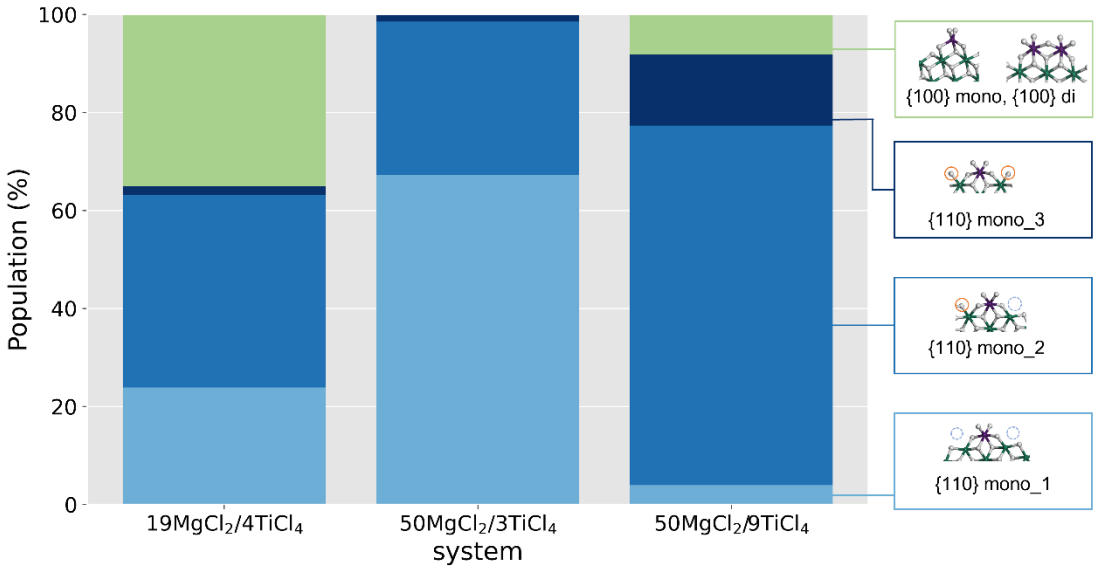


Figure 3.6. Distribution of different $TiCl_4$ species on $MgCl_2$ surfaces. The adsorbed $TiCl_4$ molecules are classified into mononuclear and dinuclear species on the {100} surfaces and mononuclear species on the {110} surfaces. The mononuclear species on the {110} surfaces are further subdivided into three species according to the presence or absence of stereo-controlling ligands based on the three-site model.[65] The distribution of individual species was derived by using Boltzmann weights applied to the identified local minima at 350 K.

In general, the structure of Ti species is closely related to their catalytic properties, and their distribution is considered to be the main cause of the experimentally observed primary structure distribution of polymers.[34] Three major types of $TiCl_4$ species on $MgCl_2$ surfaces are known: Mononuclear species on {110} terraces, and mononuclear and dinuclear species on {100} terraces. Among these, the mononuclear species on {110} is known to be thermodynamically more stable,[66] and many advanced characterization techniques support this species as the most representative type.[36,67,68]

Furthermore, according to the widely accepted three-site model, where the presence of bulky ligands at the two adjacent Mg^{2+} cations is essential for high isospecificity, the mononuclear species on {110} were subdivided into three species, mono_1−3.[65] In the framework of the three-site model, Ti centers without adjacent ligands are assigned to non-stereospecific (aspecific) sites.[40,47,65,69] The stereo-controlling ligands can be either donors or Cl^- anions, but only the latter is applicable in the current study. Indeed, donor-free catalysts produce isotactic polypropylene, although not as isotactic as that produced by donor-containing catalysts.[70]

The distribution of $TiCl_4$ species was determined for each system according to the above-described three major and three minor classifications (Figure 4). Note that $TiCl_4$ is not the active form. Upon contact with alkylaluminum, it is reduced and alkylated into the active form,[53]which accompanies redistribution of Ti species, e.g., forming agglomerates. Such the transient nature of the active sites is not included in the present research. It can be seen that regardless of the system, different $TiCl_4$ species coexist, or the structure of $TiCl_4$ is distributed. As seen in Figure 3, this originates from the thermodynamic distribution due to many energetically-accessible structures, as well as from the non-ideality of lateral surfaces, leading to different species within each structure. In all the three systems, the mononuclear species on {110} were found to be dominant, consistent with the consensus in literature.[35,66–68] The much larger population of the mononuclear species on {110} in $50MgCl_2/3TiCl_4$ and $50MgCl_2/9TiCl_4$ compared to $19MgCl_2/4TiCl_4$ is due to the ease of reconstructing larger clusters for exposing the {110} terraces.[49] On the other hand, as the $TiCl_4$ coverage increased on $50MgCl_2$, the population of mono_2,3, which possess one or two stereo-controlling ligands, increased drastically along with a reduction in the population of mono_1, which

is known to be aspecific. This is likely consistent with an experimental observation that a larger Ti content increased the population of isospecific active sites.[71]

3.4 Conclusions

High-dimensional neural network potentials (HDNNPs) were established for nanosized $MgCl_2/TiCl_4$ ternary halides, in order to accelerate the genetic algorithm (GA)-based structure determination of Ziegler-Natta catalysts. For constructing the HDNNP applicable to a wide variety of structures explored by GA, we found it is essential to ensure a sufficient diversity of ACSF vectors, i.e., atomic environments, in the training set based on farthest point sampling (FPS) and to adaptively reinforce the training set. A three-step procedure of preparing a training set was successfully proposed for a system where prior DFT data are not sufficiently collected due to the computational cost. By combining the HDNNP with GA, structure determination was demonstrated for $MgCl_2/TiCl_4$ of different sizes and Ti contents. The HDNNP-GA was found to give results consistent with DFT-GA, approximately 1,400 times faster for $50MgCl_2/3TiCl_4$. This allowed the structure determination for a larger system ($50MgCl_2/9TiCl_4$) that had not previously been handled by DFT-GA, and a more comprehensive collection of metastable structures. The structures determined for $50MgCl_2/9TiCl_4$ were featured with a rounder morphology of $MgCl_2$, comprising of short {110} and {100} terraces. $TiCl_4$ were dominantly bound to the {110} terraces without exposing bare {110} terraces. Thermodynamic analyses for metastable structures signified distributed $TiCl_4$ structures, where an increase in the $MgCl_2$ size and that in the surface coverage enhanced the population of the mononuclear species on

{110} and that of isospecific species, respectively, in perfect agreements with the experimental observations.

Rapid and more exhaustive structure exploration has brought us closer to tackling the essence of practical heterogeneous catalysis solely based on computational chemistry. Realistic structural models obtained in this way are useful for obtaining deeper insights into the structures and functions of the catalysts. For example, the models give a realistic interpretation of experimental spectroscopic results.[36,51]

REFERENCES

(1) Wang, J.; Hao, D.; Ye, J.; Umezawa, N. Determination of Crystal Structure of Graphitic Carbon Nitride: Ab Initio Evolutionary Search and Experimental Validation. *Chem. Mater.* **2017**, *29* (7), 2694–2707. https://doi.org/10.1021/acs.chemmater.6b02969.

(2) Buendía, F.; Vargas, J. A.; Johnston, R. L.; Beltrán, M. R. Study of the Stability of Small AuRh Clusters Found by a Genetic Algorithm Methodology. *Comput. Theor. Chem.* **2017**, *1119*, 51–58. https://doi.org/10.1016/j.comptc.2017.09.008.

(3) Lysgaard, S.; Mýrdal, J. S. G.; Hansen, H. A.; Vegge, T. A DFT-Based Genetic Algorithm Search for AuCu Nanoalloy Electrocatalysts for CO_2 Reduction. *Phys. Chem. Chem. Phys.* **2015**, *17* (42), 28270–28276. https://doi.org/10.1039/c5cp00298b.

(4) Aslan, M.; Davis, J. B. A.; Johnston, R. L. Global Optimization of Small Bimetallic Pd-Co Binary Nanoalloy Clusters: A Genetic Algorithm Approach at the DFT Level. *Phys. Chem. Chem. Phys.* **2016**, *18* (9), 6676–6682. https://doi.org/10.1039/c6cp00342g.

(5) Hussein, H. A.; Davis, J. B. A.; Johnston, R. L. DFT Global Optimisation of Gas-Phase and MgO-Supported Sub-Nanometre AuPd Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18* (37), 26133–26143. https://doi.org/10.1039/c6cp03958h.

(6) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Machine Learning-Aided Structure Determination for $TiCl_4$-Capped $MgCl_2$ Nanoplate of Heterogeneous Ziegler-Natta Catalyst. *ACS Catal.* **2019**, *9* (3), 2599–2609. https://doi.org/10.1021/acscatal.8b05080.

(7) Zhai, H.; Alexandrova, A. N. Fluxionality of Catalytic Clusters: When It Matters and How to Address It. *ACS Catal.* **2017**, *7* (3), 1905–1911. https://doi.org/10.1021/acscatal.6b03243.

(8) Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine Learning for Molecular Simulation. *Annu. Rev. Phys. Chem.* **2020**, *71* (1), 361–390. https://doi.org/10.1146/annurev-physchem-042018-052331.

(9) Behler, J. Perspective: Machine Learning Potentials for Atomistic Simulations. *J. Chem. Phys.* **2016**, *145* (17), 170901. https://doi.org/10.1063/1.4966192.

(10) Behler, J. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chem. Int. Ed.* **2017**, *56* (42), 12828–12840. https://doi.org/10.1002/anie.201703114.

(11) Watanabe, S.; Li, W.; Jeong, W.; Lee, D.; Shimizu, K.; Mimanitani, E.; Ando, Y.; Han, S. High-Dimensional Neural Network Atomic Potentials for Examining Energy Materials: Some Recent Simulations. *J. Phys. Energy* **2021**, *3* (1), 012003. https://doi.org/10.1088/2515-7655/abc7f3.

(12) Deringer, V. L.; Caro, M. A.; Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **2019**, *31* (46), 1–16. https://doi.org/10.1002/adma.201902765.

(13) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. A Fourth-Generation High-Dimensional Neural Network Potential with Accurate Electrostatics Including Non-Local Charge Transfer. *Nat. Commun.* **2021**, *12* (1), 398. https://doi.org/10.1038/s41467-020-20427-2.

(14) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 1–4. https://doi.org/10.1103/PhysRevLett.98.146401.

(15) Deringer, V. L.; Csányi, G. Machine Learning Based Interatomic Potential for Amorphous Carbon. *Phys. Rev. B* **2017**, *95* (9), 1–15. https://doi.org/10.1103/PhysRevB.95.094203.

(16) Thompson, A. P.; Swiler, L. P.; Trott, C. R.; Foiles, S. M.; Tucker, G. J. Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials. *J. Comput. Phys.* **2015**, *285*, 316–330. https://doi.org/10.1016/j.jcp.2014.12.018.

(17) Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Model. Simul.* **2016**, *14* (3), 1153–1173. https://doi.org/10.1137/15M1054183.

(18) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. *Chem. Sci.* **2017**, *8* (4), 3192–3203. https://doi.org/10.1039/C6SC05720A.

(19) Morawietz, T.; Singraber, A.; Dellago, C.; Behler, J. How van Der Waals Interactions Determine the Unique Properties of Water. *Proc. Natl. Acad. Sci.* **2016**, *113* (30), 8368–8373. https://doi.org/10.1073/pnas.1602375113.

(20) Daru, J.; Forbert, H.; Behler, J.; Marx, D. Coupled Cluster Molecular Dynamics of Condensed Phase Systems Enabled by Machine Learning Potentials: Liquid Water Benchmark. *Phys. Rev. Lett.* **2022**, *129* (22), 226001. https://doi.org/10.1103/PhysRevLett.129.226001.

(21) Boes, J. R.; Kitchin, J. R. Modeling Segregation on AuPd(111) Surfaces with Density Functional Theory and Monte Carlo Simulations. *J. Phys. Chem. C* **2017**, *121* (6), 3479–3487. https://doi.org/10.1021/acs.jpcc.6b12752.

(22) Shimizu, K.; Arguelles, E. F.; Li, W.; Ando, Y.; Minamitani, E.; Watanabe, S. Phase Stability of Au-Li Binary Systems Studied Using Neural Network Potential. *Phys. Rev. B* **2021**, *103* (9), 1–10. https://doi.org/10.1103/PhysRevB.103.094112.

(23) Artrith, N.; Hiller, B.; Behler, J. Neural Network Potentials for Metals and Oxides - First Applications to Copper Clusters at Zinc Oxide. *Phys. Status Solidi B Basic Res.* **2013**, *250* (6), 1191–1203. https://doi.org/10.1002/pssb.201248370.

(24) Goryaeva, A. M.; Dérès, J.; Lapointe, C.; Grigorev, P.; Swinburne, T. D.; Kermode, J. R.; Ventelon, L.; Baima, J.; Marinica, M. C. Efficient and Transferable Machine Learning Potentials for the Simulation of Crystal Defects in Bcc Fe and W. *Phys. Rev. Mater.* **2021**, *5* (10), 1–20. https://doi.org/10.1103/PhysRevMaterials.5.103803.

(25) Artrith, N.; Morawietz, T.; Behler, J. High-Dimensional Neural-Network Potentials for Multicomponent Systems: Applications to Zinc Oxide. *Phys. Rev. B* -

Condens. Matter Mater. Phys. **2011**, *83* (15), 1–4. https://doi.org/10.1103/PhysRevB.83.153101.

(26) Paleico, M. L.; Behler, J. Global Optimization of Copper Clusters at the ZnO($10\bar{1}0$) Surface Using a DFT-Based Neural Network Potential and Genetic Algorithms. *J. Chem. Phys.* **2020**, *153* (5), 054704. https://doi.org/10.1063/5.0014876.

(27) Li, W.; Ando, Y.; Minamitani, E.; Watanabe, S. Study of Li Atom Diffusion in Amorphous Li_3PO_4 with Neural Network Potential. *J. Chem. Phys.* **2017**, *147* (21), 214106. https://doi.org/10.1063/1.4997242.

(28) Artrith, N.; Urban, A. An Implementation of Artificial Neural-Network Potentials for Atomistic Materials Simulations: Performance for TiO_2. *Comput. Mater. Sci.* **2016**, *114*, 135–150. https://doi.org/10.1016/j.commatsci.2015.11.047.

(29) Li, W.; Ando, Y.; Watanabe, S. Cu Diffusion in Amorphous Ta_2O_5 Studied with a Simplified Neural Network Potential. *J. Phys. Soc. Jpn.* **2017**, *86* (10), 1–7. https://doi.org/10.7566/JPSJ.86.104004.

(30) Kolsbjerg, E. L.; Peterson, A. A.; Hammer, B. Neural-Network-Enhanced Evolutionary Algorithm Applied to Supported Metal Nanoparticles. *Phys. Rev. B* **2018**, *97* (19), 1–9. https://doi.org/10.1103/PhysRevB.97.195424.

(31) Ouyang, R.; Xie, Y.; Jiang, D. E. Global Minimization of Gold Clusters by Combining Neural Network Potentials and the Basin-Hopping Method. *Nanoscale* **2015**, *7* (36), 14817–14821. https://doi.org/10.1039/c5nr03903g.

(32) Chiriki, S.; Bulusu, S. S. Modeling of DFT Quality Neural Network Potential for Sodium Clusters: Application to Melting of Sodium Clusters (Na_{20} to Na_{40}). *Chem. Phys. Lett.* **2016**, *652*, 130–135. https://doi.org/10.1016/j.cplett.2016.04.013.

(33) Weinreich, J.; Paleico, M. L.; Behler, J. Properties of α-Brass Nanoparticles II: Structure and Composition. *J. Phys. Chem. C* **2021**, *125* (27), 14897–14909. https://doi.org/10.1021/acs.jpcc.1c02314.

(34) Busico, V. Giulio Natta and the Development of Stereoselective Propene Polymerization. In *Polyolefins: 50 years after Ziegler and Natta I*; Kaminsky, W., Ed.; Advances in Polymer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; Vol. 257, pp 37–57. https://doi.org/10.1007/12_2013_213.

(35) Chammingkwan, P.; Terano, M.; Taniike, T. High-Throughput Synthesis of Support Materials for Olefin Polymerization Catalyst. *ACS Comb. Sci.* **2017**, *19* (5), 331–342. https://doi.org/10.1021/acscombsci.7b00010.

(36) Piovano, A.; D'Amore, M.; Wada, T.; Cleto Bruzzese, P.; Takasao, G.; Thakur, A.; Chammingkwan, P.; Terano, M.; Civalleri, B.; Bordiga, S.; Taniike, T.; Groppo, E. Revisiting the Identity of δ-$MgCl_2$: Part II. Morphology and Exposed Surfaces Studied by Vibrational Spectroscopies and DFT Calculation. *J. Catal.* **2020**, *387*, 1–11. https://doi.org/10.1016/j.jcat.2020.04.017.

(37) Piovano, A.; Signorile, M.; Braglia, L.; Torelli, P.; Martini, A.; Wada, T.; Takasao, G.; Taniike, T.; Groppo, E. Electronic Properties of Ti Sites in Ziegler-Natta Catalysts. *ACS Catal.* **2021**, *11* (012), 9949–9961. https://doi.org/10.1021/acscatal.1c01735.

(38) Piovano, A.; Wada, T.; Amodio, A.; Takasao, G.; Ikeda, T.; Zhu, D.; Terano, M.; Chammingkwan, P.; Groppo, E.; Taniike, T. Formation of Highly Active Ziegler-Natta Catalysts Clarified by a Multifaceted Characterization Approach. *ACS Catal.* **2021**, *11* (22), 13782–13796. https://doi.org/10.1021/acscatal.1c03067.

(39) Andoni, A.; Chadwick, J.; Niemantsverdriet, H.; Thune, P. The Role of Electron Donors on Lateral Surfaces of $MgCl_2$-Supported Ziegler–Natta Catalysts: Observation by AFM and SEM. *J. Catal.* **2008**, *257* (1), 81–86. https://doi.org/10.1016/j.jcat.2008.04.020.

(40) Ratanasak, M.; Parasuk, V. Roles of Malonate Donor on Activity and Stereoselectivity of Ziegler-Natta Catalyzed Propylene Polymerization. *J. Organomet. Chem.* **2015**, *775*, 6–11. https://doi.org/10.1016/j.jorganchem.2014.10.018.

(41) Vittoria, A.; Meppelder, A.; Friederichs, N.; Busico, V.; Cipullo, R. Demystifying Ziegler–Natta Catalysts: The Origin of Stereoselectivity. *ACS Catal.* **2017**, *7* (7), 4509–4518. https://doi.org/10.1021/acscatal.7b01232.

(42) Credendino, R.; Liguori, D.; Fan, Z.; Morini, G.; Cavallo, L. Toward a Unified Model Explaining Heterogeneous Ziegler-Natta Catalysis. *ACS Catal.* **2015**, *5* (9), 5431–5435. https://doi.org/10.1021/acscatal.5b01076.

(43) Xie, K.; Xu, S.; Hao, W.; Wang, J.; Huang, A.; Zhang, Y. Surface Effect of the $MgCl_2$ Support in Ziegler–Natta Catalyst for Ethylene Polymerization: A Computational Study. *Appl. Surf. Sci.* **2022**, *589*, 153002. https://doi.org/10.1016/j.apsusc.2022.153002.

(44) Masoori, M.; Nekoomanesh, M.; Posada-Pérez, S.; Rashedi, R.; Bahri-Laleh, N. Exploring Cocatalyst Type Effect on the Ziegler–Natta Catalyzed Ethylene Polymerizations: Experimental and DFT Studies. *J. Polym. Res.* **2022**, *29* (5). https://doi.org/10.1007/s10965-022-03050-1.

(45) Guo, X.; Cui, L.; Yi, J.; Liu, Z.; Liu, B. Understanding the Role of Sulfonyl Amine Donors in Propylene Polymerization Using $MgCl_2$-Supported Ziegler-Natta

Catalyst. *J. Phys. Chem. C* **2022**, *126* (20), 8655–8666. https://doi.org/10.1021/acs.jpcc.2c02566.

(46) Ratanasak, M.; Rungrotmongkol, T.; Saengsawang, O.; Hannongbua, S.; Parasuk, V. Towards the Design of New Electron Donors for Ziegler–Natta Catalyzed Propylene Polymerization Using QSPR Modeling. *Polymer* **2015**, *56*, 340–345. https://doi.org/10.1016/j.polymer.2014.11.022.

(47) Guo, X.; Cui, L.; Wang, Y.; Yi, J.; Sun, J.; Liu, Z.; Liu, B. Mechanistic Study on Effect of Electron Donors in Propylene Polymerization Using the Ziegler–Natta Catalyst. *J. Phys. Chem. C* **2021**, *125* (16), 8533–8542. https://doi.org/10.1021/acs.jpcc.0c11273.

(48) Fallah, M.; Bahri-Laleh, N.; Didehban, K.; Poater, A. Interaction of Common Cocatalysts in Ziegler–Natta-Catalyzed Olefin Polymerization. *Appl. Organomet. Chem.* **2020**, *34* (2). https://doi.org/10.1002/aoc.5333.

(49) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Insight into Structural Distribution of Heterogeneous Ziegler–Natta Catalyst from Non-Empirical Structure Determination. *J. Catal.* **2021**, *394*, 299–306. https://doi.org/10.1016/j.jcat.2020.11.005.

(50) Takasao, G.; Wada, T.; Chikuma, H.; Chammingkwan, P.; Terano, M.; Taniike, T. Preventing Premature Convergence in Evolutionary Structure Determination of Complex Molecular Systems: Demonstration in Few-Nanometer-Sized $TiCl_4$-Capped $MgCl_2$ Nanoplates. *J. Phys. Chem. A* **2022**, *126* (31), 5215–5221. https://doi.org/10.1021/acs.jpca.2c02112.

(51) D'amore, M.; Takasao, G.; Chikuma, H.; Wada, T.; Taniike, T.; Pascale, F.; Ferrari, A. M. Spectroscopic Fingerprints of $MgCl_2/TiCl_4$ Nanoclusters Determined by

Machine Learning and DFT. *J. Phys. Chem. C* **2021**, *125* (125), 20048−20058. https://doi.org/10.1021/acs.jpcc.1c05712.

(52) Taniike, T.; Terano, M. Reductive Formation of Isospecific Ti Dinuclear Species on a $MgCl_2$ (110) Surface in Heterogeneous Ziegler-Natta Catalysts. *Macromol. Rapid Commun.* **2008**, *29* (17), 1472–1476. https://doi.org/10.1002/marc.200800310.

(53) Thakur, A.; Wada, T.; Chammingkwan, P.; Terano, M.; Taniike, T. Development of Large-Scale Stopped-Flow Technique and Its Application in Elucidation of Initial Ziegler-Natta Olefin Polymerization Kinetics. *Polymers* **2019**, *11* (6). https://doi.org/10.3390/polym11061012.

(54) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134* (7), 074106. https://doi.org/10.1063/1.3553717.

(55) Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115* (16), 1032–1050. https://doi.org/10.1002/qua.24890.

(56) Blank, T. B.; Brown, S. D. Adaptive, Global, Extended Kalman Filters for Training Feedforward Neural Networks. *J. Chemom.* **1994**, *8* (6), 391–407. https://doi.org/10.1002/cem.1180080605.

(57) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chem. Rev.* **2021**, *121* (16), 10037–10072. https://doi.org/10.1021/acs.chemrev.0c00868.

(58) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(59) Delley, B. An All-electron Numerical Method for Solving the Local Density Functional for Polyatomic Molecules. *J. Chem. Phys.* **1990**, *92* (1), 508–517. https://doi.org/10.1063/1.458452.

(60) Bergner, A.; Dolg, M.; Küchle, W.; Stoll, H.; Preuß, H. Ab Initio Energy-Adjusted Pseudopotentials for Elements of Groups 13–17. *Mol. Phys.* **1993**, *80* (6), 1431–1441. https://doi.org/10.1080/00268979300103121.

(61) Singraber, A.; Behler, J.; Dellago, C. Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials. *J. Chem. Theory Comput.* **2019**, *15* (3), 1827–1840. https://doi.org/10.1021/acs.jctc.8b00770.

(62) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Comput. Phys. Commun.* **2022**, *271*, 108171. https://doi.org/10.1016/j.cpc.2021.108171.

(63) Eldar, Y.; Lindenbaum, M.; Porat, M.; Zeevi, Y. Y. The Farthest Point Strategy for Progressive Image Sampling. *IEEE Trans. Image Process.* **1997**, *6* (9), 1305–1315. https://doi.org/10.1109/83.623193.

(64) Mori, H.; Sawada, M.; Higuchi, T.; Hasebe, K.; Otsuka, N.; Terano, M. Direct Observation of MgCl2-Supported Ziegler Catalysts by High Resolution Transmission Electron Microscopy. *Macromol. Rapid Commun.* **1999**, *20* (5), 245–250. https://doi.org/10.1002/(SICI)1521-3927(19990501)20:5<245::AID-MARC245>3.0.CO;2-R.

(65) Busico, V.; Cipullo, R.; Monaco, G.; Talarico, G.; Vacatello, M.; Chadwick, J. C.; Segre, A. L.; Sudmeijer, O. High-Resolution ^{13}C NMR Configurational Analysis of Polypropylene Made with $MgCl_2$-Supported Ziegler−Natta Catalysts. 1. The "Model" System $MgCl_2/TiCl_4$−2,6-Dimethylpyridine/$Al(C_2H_5)_3$. *Macromolecules* **1999**, *32* (13), 4173–4182. https://doi.org/10.1021/ma981941n.

(66) Bazhenov, A.; Linnolahti, M.; Pakkanen, T. A.; Denifl, P.; Leinonen, T. Modeling the Stabilization of Surface Defects by Donors in Ziegler-Natta Catalyst Support. *J. Phys. Chem. C* **2014**, *118* (9), 4791–4796. https://doi.org/10.1021/jp412386u.

(67) D'Amore, M.; Thushara, K. S.; Piovano, A.; Causà, M.; Bordiga, S.; Groppo, E. Surface Investigation and Morphological Analysis of Structurally Disordered $MgCl_2$ and $MgCl_2/TiCl_4$ Ziegler-Natta Catalysts. *ACS Catal.* **2016**, *6* (9), 5786–5796. https://doi.org/10.1021/acscatal.6b00871.

(68) Groppo, E.; Seenivasan, K.; Barzan, C. The Potential of Spectroscopic Methods Applied to Heterogeneous Catalysts for Olefinpolymerization. *Catal Sci Technol* **2013**, *3* (4), 858–878. https://doi.org/10.1039/C2CY20559A.

(69) Antinucci, G.; Vittoria, A.; Cipullo, R.; Busico, V. Regioirregular Monomeric Units in Ziegler-Natta Polypropylene: A Sensitive Probe of the Catalytic Sites. *Macromolecules* **2020**, *53* (10), 3789–3795. https://doi.org/10.1021/acs.macromol.0c00433.

(70) Wada, T.; Taniike, T.; Kouzai, I.; Takahashi, S.; Terano, M. Propylene Polymerization Performance of Isolated and Aggregated Ti Species Studied Using a Well-Designed $TiCl_4/MgCl_2$ Ziegler-Natta Model Catalyst. *Macromol. Rapid Commun.* **2009**, *30* (11), 887–891. https://doi.org/10.1002/marc.200900015.

(71) Takahashi, S.; Wada, T.; Chammingkwan, P.; Taniike, T.; Terano, M. Origin of Chemical Composition of Ethylene/Propylene Copolymer Produced with Ziegler–Natta Catalyst. *Macromol. React. Eng.* **2017**, *11* (1), 1600038.

Chapter 4

Compositional Effect on Ziegler-Natta Catalysts in Realistic Scale Studied with a Scalable High-Dimensional Neural Network Potential

ABSTRACT

A comprehensive understanding of catalytic activity requires not only macro- to microscale observations from experiments but also atomic- and electronic-level insights. In general, key chemical environments are highly sensitive to system composition and local structure, with significant changes in different compositions. Evaluating the effect of composition on active structures is particularly crucial for complex systems such as solid catalysts. In this chapter, the effect of $TiCl_4$ coverage on Ziegler-Natta (ZN) catalysts was systematically evaluated through non-empirical structure determination of various compositions at the actual catalyst scale. The adoption of a high-dimensional neural network potential (HDNNP) significantly accelerated calculations, enabling the efficient and systematic determination of structures across a range of compositions in large-scale systems. Analysis of multiple metastable structures revealed a consistent trend of increasing steric specificity with higher $TiCl_4$ coverage. On the other hand, it was suggested that many of the TiCl4 molecules were not activated in highly covered clusters. This chapter presents the first systematic investigation of the effect of $TiCl_4$ coverage on the cluster geometry and the electronic state of the active site at a realistic scale.

KEYWORDS: Neural network potential, genetic algorithm, structure determination, Ziegler-Natta catalyst, Density functional theory

4.1. Introduction

Quantum calculations, which provide insights into material structures and properties at the electronic level, are powerful tools for understanding complex materials.[1,2] However, many simulation studies rely on limited assumptions and empirical initial models that restrict simulation reliability. Real systems often feature non-uniform surface structures that vary with adsorbate presence, whereas most empirical models depict only periodic, ideal surfaces or highly localized environments.[3–8] To overcome these limitations, combining machine learning with quantum calculations offers a non-empirical approach to structural determination, free from experimental constraints. By integrating machine learning, researchers can model structural variations more accurately and explore complex atomic interactions, extending the applicability of quantum methods to a broader range of materials, including single element clusters,[9] alloys[10,11], and supported catalysts.[5,12,13] As the number of atoms in a system increases, the computational cost and time required to reach a solution escalate rapidly. While global optimization of clusters with a few dozen atoms has been achieved using density functional theory (DFT), quantum calculations for realistic systems, commonly containing hundreds or thousands of atoms, become prohibitively time-intensive. Determining the structures of large nanoparticles requires the exploration of extensive potential energy surfaces (PES) to account for diverse atomic interactions within both adsorbate and support. Thus, a precise and efficient potential capable of covering this vast parameter space is essential.

Machine learning potentials (MLPs), which construct potential energy surfaces (PESs) via machine learning, present a promising solution to this challenge.[14–16] MLPs are developed based on reference datasets of representative conformations obtained

through quantum calculations, such as DFT. These models can predict energies and forces with first-principles accuracy comparable to the reference calculation methods, yet without requiring electronic state calculations, allowing for computational speeds approaching those of force fields. Several MLP algorithms have been developed, including high-dimensional neural network potentials (HDNNP),[17,18] graph neural network potentials,[19] and other neural network-based approaches,[20] as well as Gaussian approximation potentials,[21] spectral neighbor analysis potentials,[22] and moment tensor potentials.[23]

The solid component of $MgCl_2$-supported Ziegler-Natta (ZN) catalysts, a system widely used for industrial olefin polymerization,[24] forms through aggregation of $MgCl_2$ nanoplate primary particles. These primary particles are terminated by $TiCl_4$, the active site precursor, and are modified by organic Lewis bases (internal donors, IDs), which enable stereoselective olefin polymerization.[25–27] Activation is achieved with alkylaluminium co-catalysts, while additional Lewis bases (external donors), introduced during the reaction, improve stereospecificity and activity by co-adsorbing with Ti species to form a steric barrier.[25,28–30] Despite extensive research and development in this area, the primary particle structure and its distribution remain poorly understood due to the system's complexity and surface reconfigurations induced by adsorbates.

Many computational approaches have sought to elucidate structure-performance relationships in Ziegler-Natta catalysts,[27,31–39] however, most assume static surfaces. In a notable advancement, Takasao et al. achieved non-empirical structural determination of $MgCl_2/TiCl_4$ primary particles, exploring both the support structure and adsorbates simultaneously through a combination of genetic algorithm (GA) and DFT optimization.[5,39,40] Their study revealed that $TiCl_4$ chemisorption in Ziegler-Natta

catalysts contributes to the emergence of various metastable states. The most stable structures of different $x\mathrm{MgCl_2}/y\mathrm{TiCl_4}$ (x = 6–19, y = 0–4) systems with particle diameters under 2.0 nm were determined, estimating the effects of particle size and $\mathrm{TiCl_4}$ coverage on structural distribution.[39] Although the system size is constrained by the high computational cost of DFT, their models qualitatively match experimental observations. Further HDNNP-based structure determinations for larger $\mathrm{MgCl_2}/\mathrm{TiCl_4}$ systems (discussed in Chapter 2) revealed that morphological features at experimentally relevant particle scales (approximately 4.0 nm)[41,42] differ significantly from those of smaller systems. This finding suggests that studying the effects of composition at realistic scales could enhance understanding of catalytic structure distributions.

In this chapter, I focused on understanding the effects of composition on cluster morphology and active structures at realistic catalyst scales by introducing HDNNP acceleration. To achieve this, I constructed an HDNNP model based on a dataset containing a range of compositions and sizes, designed to capture representative chemical environments for interpolation across systems with varying atom counts. HDNNP approach is adaptable to systems of different sizes as long as the elemental composition is consistent, and thus, it is typically trained on datasets that span multiple structural configurations.[15,18] However, constructing a robust and generalized HDNNP requires samples that thoroughly represent the diverse local chemical environments within $\mathrm{MgCl_2}/\mathrm{TiCl_4}$ systems, which change significantly with size and coverage. In Chapter 2, separate reference datasets were generated for each composition to simplify data selection, necessitating the construction of individual potentials for each unique composition.[43] This chapter goes further, addressing the challenges of efficiently sampling data from clusters with a range of compositional patterns, which strongly

influence chemical environment distribution. I summarize the effective criteria for composition selection in HDNNP construction and discuss the effect of compositional variations on the morphology of large-scale systems, thereby providing insights into the structural diversity and catalytic performance of $MgCl_2/TiCl_4$ catalysts.

4.2. Numerical method

4.2.1. Construction of reference dataset for HDNNP training

The second-generation high-dimensional neural network potentials (HDNNPs)[18] were adopted, and local chemical environments were described by atom centered symmetry functions (ACSFs).[44] The employed function form and parameters of ACSFs are completely same as Chapter 2. The atomic neural network subset has 2 layers with 20 nodes for each.

In Chapter 2, reference datasets and HDNNPs were developed for four distinct compositions. These datasets were generated using farthest point sampling (FPS) combined with adaptive reinforcement to enhance their overall effectiveness. To create a comprehensive reference database incorporating multiple compositions, these datasets were merged. Here, I utilized systems that included $TiCl_4$-covered compositions ($19MgCl_2/4TiCl_4$, $50MgCl_2/3TiCl_4$, $50MgCl_2/9TiCl_4$) from the dataset obtained in Chapter 2.[43] These compositions were chosen to increase the proportion of Ti atoms within the reference dataset, ensuring a more comprehensive coverage of Ti-related chemical environments. A reference dataset was subsequently created by sampling structures using the FPS method based on ACSFs. FPS was applied within the ACSF vector space for each element to ensure the selection of structures containing representative atomic environments. This approach ensured the training data included

diverse and representative configurations to enhance generalizability. 70% of the structures in the reference dataset were allocated for training, while the remaining 30% were reserved for testing. To further strengthen the reference dataset, adaptive sampling was employed. Geometry optimizations were performed on $19MgCl_2/9TiCl_4$ and $40MgCl_2$ clusters, representing high $TiCl_4$ coverage and moderate-sized, $TiCl_4$-free clusters, respectively. Thousands of initial structures, randomly sampled from the DFT database, were optimized using two HDNNPs with different architectures but comparable accuracy. Structures showing significant differences in geometry or predicted energy between the two HDNNPs were recalculated with DFT and added to the reference dataset.

The constructed HDNNP was evaluated based on its accuracy in single-point calculations and geometry optimization for structures with compositions not included in the training dataset. The HDNNPs were constructed by using the RuNNer code (version 1.2.0).[45,46] All DFT calculations were performed using the Perdew-Burke-Ernzerhof (PBE) functional[47] and the double numerical basis functions (DNP)[48] with effective core potentials (ECP)[49] were implemented by using DMol^3.[48]

4.2.2. Non-empirical structure determination

Non-empirical structure determination was performed for the $50MgCl_2$ cluster with different numbers (4–13) of $TiCl_4$ adsorbates using HDNNP and genetic algorithm (GA). These compositions were selected to cover a range of $TiCl_4$ coverage from very low to almost full. The structures are subjected to geometry optimization implemented in LAMMPS with the n2p2 package[50] which integrates the HDNNP method into LAMMPS[51]. The optimization process employed a conjugate gradient method with a

convergence criterion of 154 meV/Å for the global force vector, within a maximum of 1,000 steps. The energy of the optimized structures is expressed by E_{cluster}, which is defined as,

$$E_{\text{cluster}} = E - N_{\text{MgCl}_2} \times E_{\text{MgCl}_2} - N_{\text{TiCl}_4} \times E_{\text{TiCl}_4} \quad (4.1)$$

where E is the total energy of a structure, N_{MgCl_2} is the number of $MgCl_2$ units, E_{MgCl_2} is the energy of a $MgCl_2$ unit in α-$MgCl_2$, N_{TiCl_4} is the number of $TiCl_4$ molecules, and E_{TiCl_4} is the energy of a $TiCl_4$ molecule in vacuum. For each system, one independent run was performed. For 3–6 additional runs, migration operators were introduced starting from the 300th generation of the GA cycle. This operator involved randomly selecting structures within the energy range of elite structures from a stable configuration database and adding them to the population to enhance diversity. The details of GA algorithm and each operations can be found in Ref.[5,40]. Roulette selection was used for its capacity to probabilistically favor high-fitness structures for crossover and mutation. Fitness (f) of the i^{th} structure defined as

$$f_i = e^{-3\left(\frac{E_i - E_{min}}{E_{max} - E_{min}}\right)} \quad (4.2)$$

where E_i, E_{min}, and E_{max} are the energy of the i^{th} structure in a generation, and the energies of the most and least stable structures in the same generation.

Table 4.1. Parameters used in the genetic algorithm

Parameter	Value
Population (structure)	105
Crossover (%)	38.1
Mutation (skeleton) (%)	14.3
Mutation (adsorbate) (%)	14.3
Elitism (%)	33.3

4.3. RESULT AND DISCUSSI

4.3.1. Evaluation of HDNNP

The reference data set was selected from a population containing multiple compositions using the farthest-point sampling (FPS) method based on the atomic center symmetry function (ACSF). In addition, structures were added through adaptive learning. Table 4.2 summarizes the composition of the final reference dataset. The first three systems were selected using FPS, with a dominant selection of the $19MgCl_2/4TiCl_4$ cluster. This suggests that clusters with relatively small and medium coverages more significantly reflect changes in the local atomic environment within the ACSF framework. The latter two systems represent structures added by adaptive sampling. The $19MgCl_2/9TiCl_4$ system is cluster with extremely high coverage that is not included in the initial reference data set, necessitating addition of more structures to improve representation.

Table 4.2. Structure of the reference dataset [a]

System	N_{train}	N_{test}
$19MgCl_2/4TiCl_4$	3398	1468
$50MgCl_2/3TiCl_4$	2225	1011
$50MgCl_2/9TiCl_4$	2705	1154
$19MgCl_2/9TiCl_4$	1187	484
$40MgCl_2$	74	32

[a] The reference data set is a mixture of five systems with different compositions. The first three systems were selected by FPS, and the structures of the remaining two systems were added to adaptive. 70% was used as the training set and 30% as the testing set.

Table 4.3. HDNNP prediction errors [a]

Training		Single-point test		Geometry optimization test ($19MgCl_2/9TiCl_4$)		
E (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	Success rate (%)
0.416	52.9	0.611	53.6	1.30	51.4	99.0

[a] The root mean squared errors (RMSE) of energies (*E*) and forces (*F*) were calculated relative to DFT results. During the training process, 70% of the reference dataset was used for training, while the remaining 30% was allocated for single-point testing. Additionally, geometry optimization tests were performed using 2,480 initial structures

of $19MgCl_2/9TiCl_4$ cluster, characterized by high $TiCl_4$ coverage. RMSE values were calculated for 100 structures picked randomly from 2,455 successfully optimized using the HDNNP. Success rate means percentage of structures for which no extrapolation occurred during the geometry optimization.

The accuracy of the trained HDNNPs was evaluated based on the root mean square error (RMSE) of energies and forces relative to the corresponding reference data. Ensuring the robustness of the HDNNP during simulation steps is critically important. Therefore, an additional test, referred to as the geometry optimization test, was performed. Given the need to determine cluster structures across varying coverage levels, this test focused on systems with extreme coverage conditions. Specifically, 2,480 initial structures of $19MgCl_2/9TiCl_4$ cluster were optimized using the HDNNP, and 100 resulting structures were randomly sampled. The energies and forces of these sampled structures were recalculated using DFT single-point calculations, and the deviations from the HDNNP predictions were evaluated using RMSE. Additionally, the success rate of the geometry optimizations, defined as the percentage of optimizations that either did not terminate due to ACSF extrapolation or showed no extrapolation warnings upon completion, was also evaluated.

Table 4.2 summarizes the accuracy evaluation of the HDNNP. The RMSE for energy was 0.416 meV/atom for the training set and 0.611 meV/atom for the testing set. Although this was inferior to the HDNNP constructed in Chapter 2 (approximately 0.2 meV/atom), these values remain within the typical HDNNP accuracy range (0.5–10 meV/atom). For forces, the RMSE was 52.9 meV/Å for the training set and 53.6 meV/Å for the testing set, close to 50 meV/Å reported in Chapter 2,[43] indicating negligible

impact from dataset generalization. In the geometry optimization test, 99% success rate was achieved, and the energy predictions for the optimized structures showed a favorable RMSE of 1.30 meV/atom. The robustness of the HDNNP for the exploration of unknown systems is evaluated in also the first part of the next section.

4.3.2. Structure determination

The constructed HDNNP was applied to non-empirical structure determination using GA. First, in order to confirm that the constructed HDNNP has sufficient accuracy for the target system, I performed DFT validation on several structures obtained by GA. GA was run for $50MgCl_2/4TiCl_4$ and $50MgCl_2/13TiCl_4$, which have lowest and highest coverage in trials, and the top 5 stable structures at around the 300th generation were picked out. DFT geometry optimization was performed on each of these structures to evaluate the deviation in prediction between the HDNNPs and DFT (Figure 4.1). The difference between the DFT and the predicted values for the obtained structures was 1–2 kcal/mol (equivalent to 0.2–0.4 meV/atom), indicating that the HDNNP has sufficient accuracy.

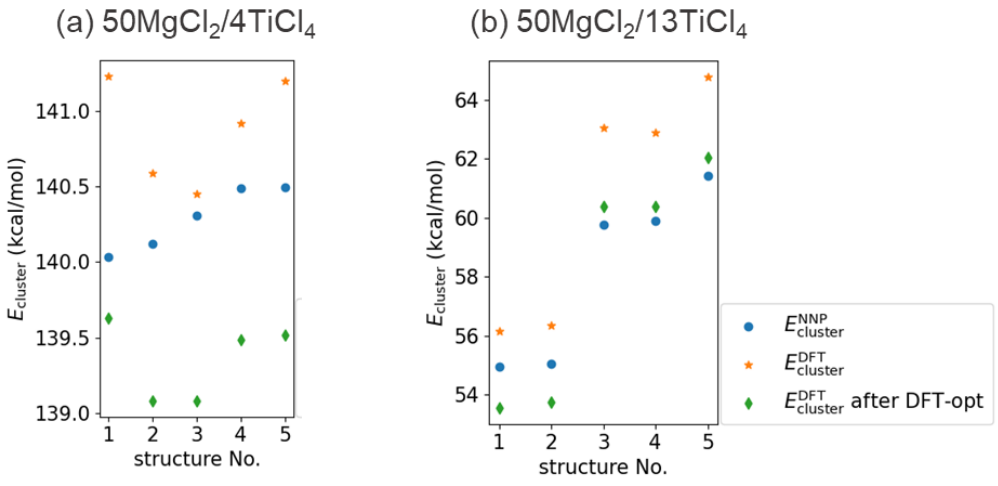


Figure 4.1. Comparison of energies after optimization using HDNNP and DFT. The energies referred to by the GA using HDNNP are indicated by the blue circles. The orange stars show the energies of the structures optimized using HDNNP, as evaluated by DFT, and the deviation between the orange and blue circles indicates the prediction error for the optimized structures. The green diamonds show the energies of the structures re-optimized using DFT, and the deviation between the green and blue dots indicates the error in energy evaluation in the GA.

Next, structure determination using HDNNP was performed on a $50MgCl_2$ cluster with different $TiCl_4$ coverage: $50MgCl_2/4TiCl_4$, $50MgCl_2/7TiCl_4$, $50MgCl_2/11TiCl_4$, and $50MgCl_2/13TiCl_4$. Each GA run were performed until the 1000^{th} generation. The energy evolution during GA is summarized in Figure 4.2. Due to the vastness of the parameter space, more than 500 generations of GA iterations were required to reach the global minimum. The most stable structure for each composition was determined after searching for 100,000 to 50,000 structure candidates.

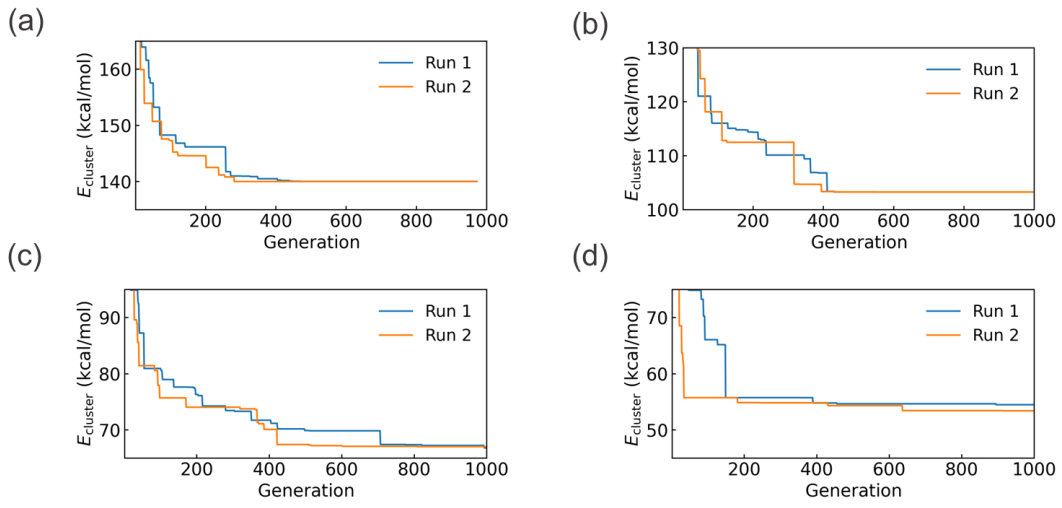
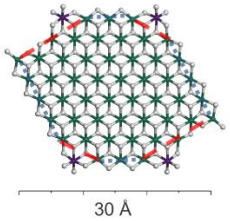


Figure 4.2. Evolutionary progress plots of HDNNP-GA for the structure determination of a)50$MgCl_2$/4$TiCl_4$, b) 50$MgCl_2$/7$TiCl_4$, c) 50$MgCl_2$/11$TiCl_4$, and d) 50$MgCl_2$/13$TiCl_4$. The energy of the most stable structure in the population is plotted along with the generation. $E_{cluster}$ is a sum of the surface energy and the adsorption energy of $TiCl_4$, as defined in Eq. 4.2.

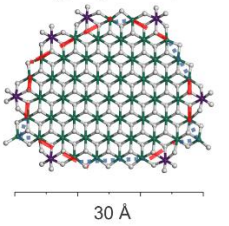
{100} surface {110} surface

(a) $50MgCl_2/4TiCl_4$



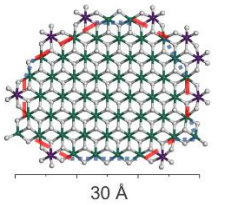
140.0 kcal/mol (139.6) 140.1 kcal/mol (139.1) 140.2 kcal/mol (139.9) 140.3 kcal/mol (139.1)

(b) $50MgCl_2/7TiCl_4$



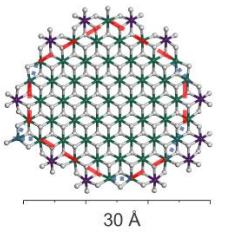
103.3 kcal/mol (104.0) 104.7 kcal/mol (104.8) 106.5 kcal/mol (107.9) 106.6 kcal/mol (106.6)

(c) $50MgCl_2/9TiCl_4$



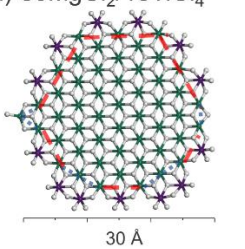
84.7 kcal/mol (84.7) 85.8 kcal/mol (84.5) 86.9 kcal/mol (85.6) 87.1 kcal/mol (86.7)

(d) $50MgCl_2/11TiCl_4$



66.8 kcal/mol (67.9) 67.1 kcal/mol (67.9) 67.4 kcal/mol (68.3) 67.4 kcal/mol (68.1)

(d) $50MgCl_2/13TiCl_4$



50.0 kcal/mol (50.1) 51.5 kcal/mol (53.2) 53.4 kcal/mol (52.4) 54.3 kcal/mol (53.5)

Figure 4.3. The most stable structures and their energies of a)50$MgCl_2$/4$TiCl_4$, b) 50$MgCl_2$/7$TiCl_4$, c) 50$MgCl_2$/11$TiCl_4$, and d) 50$MgCl_2$/13$TiCl_4$ cluster determined by HDNNP-GA. The dotted blue and dashed red lines indicate the $MgCl_2$ {100} and {110} surfaces, respectively. The energies obtained by DFT-geometry optimization are given in parentheses.

Figure 4.3 shows the top 4 stable structures, and their energies obtained on GA using HDNNP (HDNNP-GA). The structures and energies in 50$MgCl_2$/9$TiCl_4$ cluster are cited from Chapter 2. The energies were further evaluated through geometry optimization using DFT, with DFT results shown in parentheses in Figure 4.3. The cluster energies predicted by HDNNP closely matched those obtained from DFT geometry optimization, confirming that the PES exploration was conducted with high accuracy.

Table 4.4 summarizes the average coverage rates of the obtained clusters. The coverage rate is calculated as the ratio of Mg^{2+} cation sites occupied by adsorbates (terminal Cl and $TiCl_4$) to the total coordination vacancies of surface Mg^{2+} cations in the $MgCl_2$ cluster, excluding all adsorbates. Specifically, a 5-coordinated Mg^{2+} ion was counted as having one vacancy, while a 4-coordinated Mg^{2+} ion was counted as having two vacancies. Contributions from each cluster were weighted by the following Boltzmann factor:

$$p = e^{-\frac{\Delta E_{\mathrm{cluster}}}{RT}} \quad (4.3),$$

where $\Delta E_{\text{cluster}}$ is the cluster energy of a structure relative to that of the most stable structure, R is the gas constant, and the temperature (T) was set to 350 K. Based on this analysis, the structure determination for $50MgCl_2/4–13TiCl_4$ clusters corresponds to coverage rates ranging from 34% to 84%.

Table 4.4. Average coverage of clusters in each composition [a]

System	Coverage
$50MgCl_2/4TiCl_4$	34.32%
$50MgCl_2/7TiCl_4$	51.65%
$50MgCl_2/9TiCl_4$	62.48%
$50MgCl_2/11TiCl_4$	72.65%
$50MgCl_2/13TiCl_4$	83.93%

[a] Coverages were estimated as the ratio of surface Mg^{2+} cation coordination vacancies occupied by adsorbates. Specifically, a 5-coordinated Mg^{2+} ion is counted as having one vacancy, while a 4-coordinated Mg^{2+} ion is counted as having two vacancies. The population of each cluster was weighted by the Boltzmann factor in Eq. 2.9.

In terms of general morphological characteristics, the cluster surface structures consist of $TiCl_4$-covered {110} terraces and uncovered {100} terraces, consistent with previous structural studies on ZN catalysts[3,6,9]. Furthermore, $TiCl_4$ adsorption preferentially exposes {110} terraces, reducing the exposure of {100} terraces as the coverage increases. The most stable structures for each composition had different structures of

$MgCl_2$ skeleton, suggesting that surface reconstruction due to $TiCl_4$ adsorption occurs regardless of the coverage.

Figure 3.4 plots the stabilizing effect of $TiCl_4$ adsorption on different compositions using the most stable structures. The stabilizing effect is calculated as the energy difference between the most stable structure of $50MgCl_2$ cluster ($E_{cluster}$ = 139.9 kcal/mol, reported in Ref.[40]) and the most stable structures of $50MgCl_2/4–13TiCl_4$ cluster, divided by the number of adsorbed $TiCl_4$ molecules. This value corresponds to the adsorption energy, accounting for the surface reconstruction of $50MgCl_2$ skeleton. The plot is based on the most stable structures and energies obtained from DFT geometry optimization. The results show that the adsorption energy decreases consistently as the coverage of $TiCl_4$ increases. For clusters with high coverage, the formation of {110} terraces become less favorable, leading to adsorption on the {100} terraces. Consequently, the average adsorption energy decreases. Consequently, the average adsorption energy decreases. A similar trend was observed for $50MgCl_2/4TiCl_4$ and $50MgCl_2/7TiCl_4$ clusters, which contain only {110} $TiCl_4$ monomers. This indicates that the stabilization effect is influenced not only by the adsorption configuration but also by charge redistribution during $TiCl_4$ adsorption.

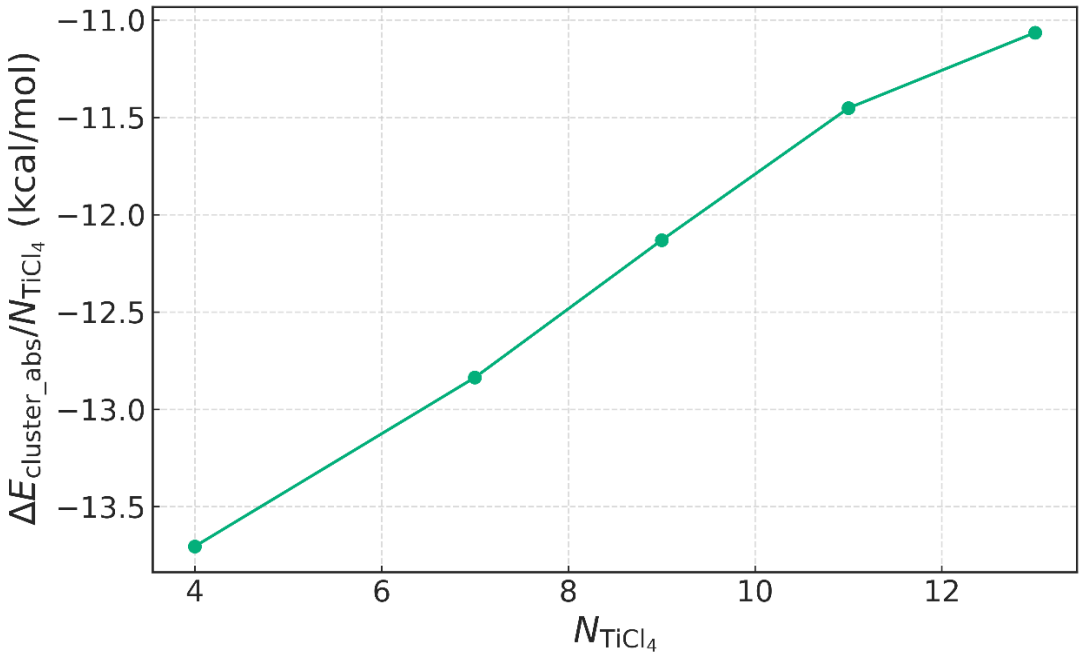


Figure 4.4. Comparison of the stabilization effect from $TiCl_4$ adsorption on the most stable structures of the $50MgCl_2$/4–13$TiCl_4$ cluster. The stabilization is evaluated as the energy difference between the most stable structure of each composition and the most stable structure of the bare $50MgCl_2$ cluster, normalized by the number of $TiCl_4$ molecules ($\Delta E_{cluster\_abs}/N_{TiCl_4}$). The energy of the bare $50MgCl_2$ is set to 139.9 kcal/mol.[40]

Next, I investigate the adsorption behavior based on the obtained stable structures. Specifically, the Gibbs free energy change of the system as a function of number of $TiCl_4$ on a $50MgCl_2$ cluster were compared particles of different compositions by analyzing. The thermodynamic stability of the $50MgCl_2/nTiCl_4$ cluster was evaluated using the equilibrium:

$$50MgCl_2 + (TiCl_4)_{13} \rightarrow 50MgCl_2/nTiCl_4 + (TiCl_4)_{13-n} \quad (4.4$$

),

Both the electronic energy and Gibbs free energy were computed at the GGA-PBE level, consistent with the methodology used for structural optimization. The most stable configurations of $50MgCl_2/4–13TiCl_4$ and bare $50MgCl_2$, as well as the $TiCl_4$ monomer, were optimized using density functional theory (DFT). Vibrational frequencies were calculated within the harmonic approximation, and Gibbs free energies were obtained at 298.15 K.

To account for the entropy loss in the condensed phase, the entropy contribution of $TiCl_4$ in the Gibbs free energy calculation was scaled by a factor of 2/3, a widely accepted and commonly employed correction.[52–54] The dependence of Gibbs free energy on the number of adsorbed $TiCl_4$ molecules is shown in Fig. 4.4. The results demonstrate a continuous decrease in the system's free energy with increasing $TiCl_4$ coverage, indicating that $TiCl_4$ adsorption is thermodynamically favorable. These results are consistent with experimental observations, which suggest that $MgCl_2$ nanoparticles are nearly fully covered by adsorbed $TiCl_4$.[55]

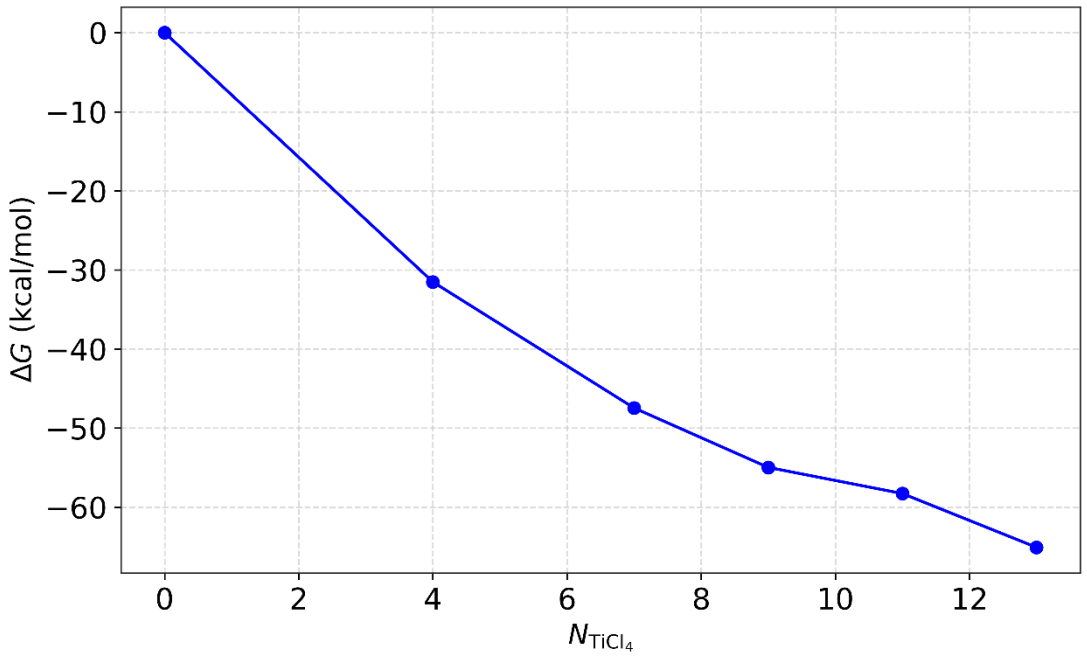


Figure 4.5. Gibbs free energy change (ΔG) of the reaction shown at Eq.4.4 as a function of the number of adsorbed $TiCl_4$ molecules on a $50MgCl_2$ cluster.

4.3.3. Distribution analysis

The charge of the $TiCl_4$ active site precursor on the cluster is sensitive to its chemical environment, and this affects the performance of the active site that is generated. Here, the effect of the coverage on the structural and charge distributions of $TiCl_4$ is discussed. Because the second-generation HDNNP cannot predict charges, the most stable structures within a 6 kcal/mol range were selected from those obtained by HDNNP-GA and recalculated using DFT geometry optimization for these analyses. The estimation of the structural distribution was performed based on weighting by the Boltzmann factor defined as Eq.4.3, and the temperature (T) was set to 350 K. Note that distribution analysis was performed using the energies obtained by DFT-geometry optimization. As

shown in Figure 4.2, the deviation between HDNNP on DFT in the rankings of stable structures and their relative energies was minimal.

The distribution of Ti species for each composition is summarized in Figure 4.6. Ti species are largely classified into mononuclear species on the {110} terrace, mononuclear species on the {100} terrace, and dinuclear species, with the mononuclear species on the {110} terrace being known to be thermodynamically more stable and majority of the active site. Furthermore, according to the three-site model, {110} mononuclear species is subdivided, based on presence/absence of ligands present on two adjacent Mg^{2+} cations, into three types: mono_1 (without ligands), mono_2 (with one ligand), and mono_3 (with two ligands).[56] Coordination species adjacent to the active sites limit monomer insertion, resulting in enhanced stereospecificity. Consistent with previous studies, the {110} mononuclear species dominated in all compositions. Furthermore, as the $TiCl_4$ coverage increased, the population of the non-stereospecific mono_1 decreased, while the populations of the stereospecific mono_2 and 3 increased. A roughly similar trend has been reported in previous structural determination studies, and here, for the first time, it was shown to be a continuous trend by continuously varying the coverage. As $TiCl_4$ coverage increases, the surface exposure pattern becomes increasingly restricted, with {110} terrace exposure decreasing and $TiCl_4$ adsorption on {100} terraces being promoted. Computational calculations indicate that $TiCl_4$ clusters on {100} terraces form steric barriers, limiting activation to a single Ti atom. Experimental and computational studies have revealed that mere Ti exposure is not a guarantee of high catalytic activity,[57] and this analysis provides strong support for those findings. The simultaneous increase of mono_3 and {100} species further

underscores that stereospecific active sites are not predominant in the absence of internal donors.

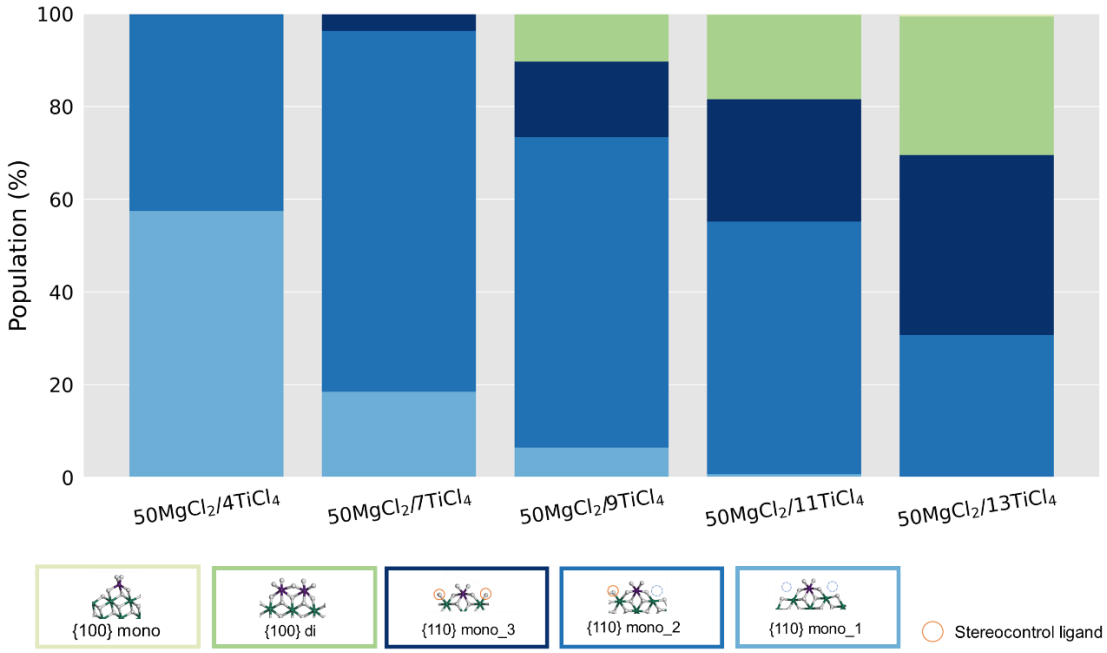


Figure 4.6. Distribution of different $TiCl_4$ species on $MgCl_2$ surfaces. The adsorbed $TiCl_4$ molecules are classified into mononuclear and dinuclear species on the {100} surfaces and mononuclear species on the {110} surfaces. The mononuclear species on the {110} surfaces are further subdivided into three species according to the presence or absence of stereo-controlling ligands based on the three-site model.[56] The distribution of individual species was derived by using Boltzmann weights applied to the identified local minima at 350 K.

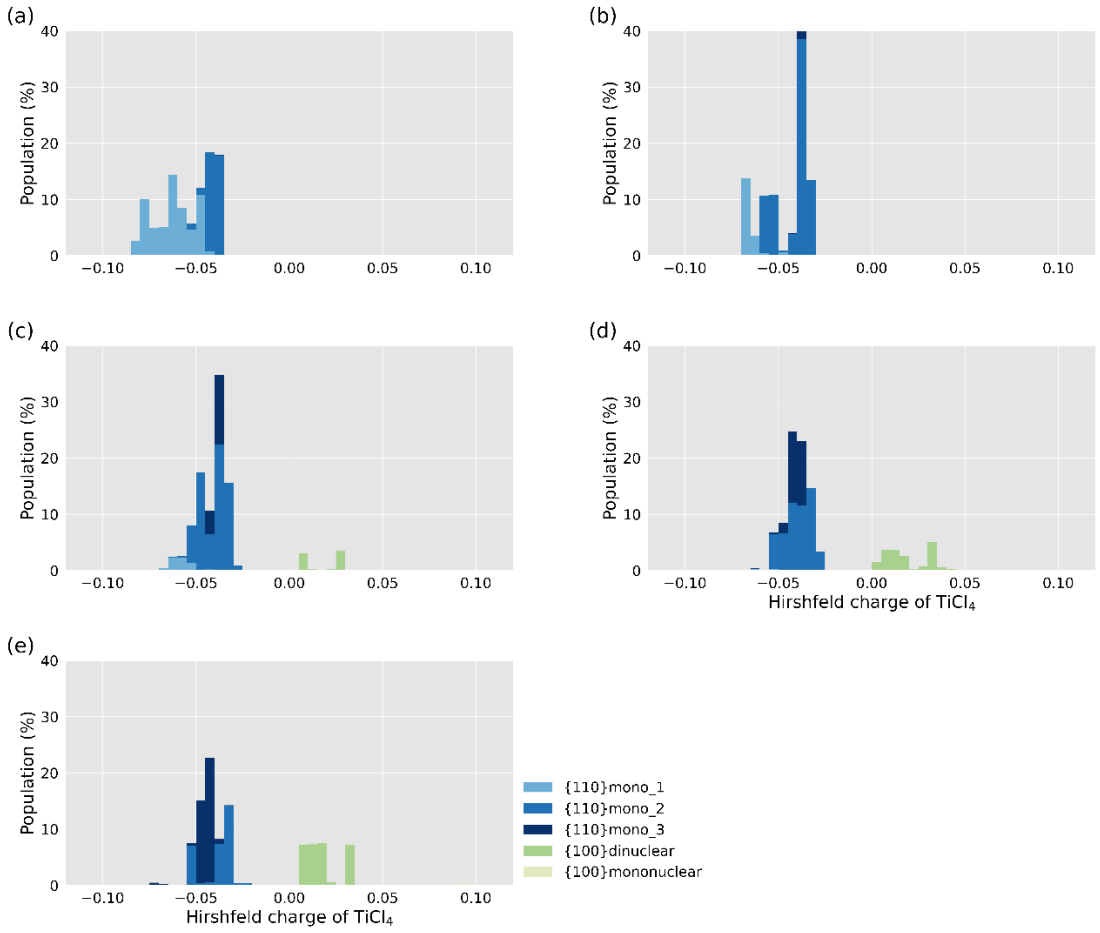


Figure 3.7. Population of $TiCl_4$ species and these Hirshfeld charge: a)$50MgCl_2/4TiCl_4$, b)$50MgCl_2/7TiCl_4$, c)$50MgCl_2/9TiCl_4$, d)$50MgCl_2/11TiCl_4$, and e)$50MgCl_2/13TiCl_4$. The percentage of presence was weighted by the Boltzmann factor.

The properties of Ti active centers are influenced not only by steric hindrance in the surrounding environment, but also by electrostatic interactions. For example, electron-deficient Ti species can form π complexes favorably, resulting in low activation energies.[39] The charge state of $TiCl_4$ is closely related to its adsorption structure. On {100} surfaces, mononuclear and dinuclear species are positively charged, while mononuclear species on {110} surfaces exhibit a negative charge distribution, ranging

from -0.08 (without internal donors) to between -0.03 and 0 (with internal donors), as reported by Hirschfeld charge analysis. Figure 4.5 summarizes the charge states of $TiCl_4$ as estimated using Hirschfeld charges. Mononuclear species on the {110} surface are negatively charged, whereas dinuclear species on the {100} surface are positively charged. Although {110} mononuclear species were rarely observed, those that were found were positively charged, consistent with previously reported characteristics. In the $50MgCl_2/4–9TiCl_4$ cluster containing {110} mono_1, the species exhibited a more negative charge compared to {110} mono_2 and 3, with the overall charge distribution shifting positively as coverage increased. Conversely, {110} mono_3 exhibited a shift toward negative charge as coverage increased, which corresponds to an increase in {100} dinuclear species, charged in positive. This is because the Cl of {100} dinuclear is coordinated to {110} mono_3, and the interaction makes the charge of {110} mono_3 more negative. Moreover, while the distribution of {110} mono_3 is narrow and shifts with changes in composition. This implies that in systems lacking internal donors, the charge state distribution at the highly stereospecific active site remains narrow.

4.4. Conclusion

This chapter systematically evaluates the effect of $TiCl_4$ coverage on Ziegler-Natta (ZN) catalysts without internal donors by determining the structures of multiple compositions at a realistic catalyst scale. Due to the high computational cost of DFT, which limits systematic structural determination, a high-dimensional neural network potential (HDNNP) was developed to describe a wide range of $MgCl_2/TiCl_4$ compositions.

By combining HDNNP with a genetic algorithm (GA), new stable structures for 50$MgCl_2$/4–13$TiCl_4$ clusters (corresponding to coverages of 30–84%) were identified. This revealed a consistent trend: as the $TiCl_4$ coverage increases, stereospecificity improves, aligning with experimental and computational findings that confirm the dominance of mononuclear $TiCl_4$ species on {110} surfaces. However, the increase in {100} dinuclear species was observed at higher coverages, negatively affects the overall catalytic activity. The adsorption behavior of $TiCl_4$ on $MgCl_2$ was estimated using Gibbs free energy, showing that clusters with higher coverage are thermodynamically more stable. Charge distribution analysis further revealed direct and indirect interactions between Ti species, highlighting the influence of the chemical environment. These findings demonstrate that traditional infinite-plane models or localized active site models are insufficient for accurately simulating catalytic behavior.

While these points have been qualitatively discussed in the past, this study provides quantitative insights by establishing molecular models. For the first time, this work systematically elucidates how $TiCl_4$ coverage affects cluster morphology and the electronic state of active sites at a realistic scale.

REFERENCES

(1) Stamatakis, M.; Vlachos, D. G. Unraveling the Complexity of Catalytic Reactions via Kinetic Monte Carlo Simulation: Current Status and Frontiers. *ACS Catal.* **2012**, *2* (12), 2648–2663. https://doi.org/10.1021/cs3005709.

(2) Bahri-Laleh, N.; Hanifpour, A.; Mirmohammadi, S. A.; Poater, A.; Nekoomanesh-Haghighi, M.; Talarico, G.; Cavallo, L. Computational Modeling of

Heterogeneous Ziegler-Natta Catalysts for Olefins Polymerization. *Prog. Polym. Sci.* **2018**, *84*, 89–114. https://doi.org/10.1016/j.progpolymsci.2018.06.005.

(3) D'Amore, M.; Thushara, K. S.; Piovano, A.; Causà, M.; Bordiga, S.; Groppo, E. Surface Investigation and Morphological Analysis of Structurally Disordered $MgCl_2$ and $MgCl_2/TiCl_4$ Ziegler-Natta Catalysts. *ACS Catal.* **2016**, *6* (9), 5786–5796. https://doi.org/10.1021/acscatal.6b00871.

(4) Taniike, T.; Chammingkwan, P.; Thang, V. Q.; Goto, K.; Fujitani, T.; Terano, M. Chemisorption-Induced Activation of $MgCl_2$ Film as Realistic Route for Heterogeneous Ziegler-Natta Surfaces under Ultrahigh Vacuum. *J. Phys. Chem. C* **2017**, *121* (43), 24085–24092. https://doi.org/10.1021/acs.jpcc.7b08242.

(5) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Machine Learning-Aided Structure Determination for $TiCl_4$-Capped $MgCl_2$ Nanoplate of Heterogeneous Ziegler-Natta Catalyst. *ACS Catal.* **2019**, *9* (3), 2599–2609. https://doi.org/10.1021/acscatal.8b05080.

(6) Wada, T.; Takasao, G.; Terano, M.; Chammingkwan, P.; Taniike, T. Structure Determination of the δ-$MgCl_2$ Support in Ziegler-Natta Catalysts. *J. Jpn. Pet. Inst.* **2022**, *65* (3), 88–96. https://doi.org/10.1627/jpi.65.88.

(7) De Heer, W. A. The Physics of Simple Metal Clusters: Experimental Aspects and Simple Models. *Rev. Mod. Phys.* **1993**, *65* (3), 611–676. https://doi.org/10.1103/RevModPhys.65.611.

(8) Alonso, J. A. Electronic and Atomic Structure, and Magnetism of Transition-Metal Clusters. *Chem. Rev.* **2000**, *100* (2), 637–678. https://doi.org/10.1021/cr980391o.

(9) Bao, K.; Goedecker, S.; Koga, K.; Lançon, F.; Neelov, A. Structure of Large Gold Clusters Obtained by Global Optimization Using the Minima Hopping Method.

Phys. Rev. B - Condens. Matter Mater. Phys. **2009**, *79* (4), 1–4. https://doi.org/10.1103/PhysRevB.79.041405.

(10) Clausen, C. M.; Nielsen, M. L. S.; Pedersen, J. K.; Rossmeisl, J. Ab Initio to Activity: Machine Learning-Assisted Optimization of High-Entropy Alloy Catalytic Activity. *High Entropy Alloys Mater.* **2022**, *1*, 0–3. https://doi.org/10.1007/s44210-022-00006-4.

(11) Lysgaard, S.; Mýrdal, J. S. G.; Hansen, H. A.; Vegge, T. A DFT-Based Genetic Algorithm Search for AuCu Nanoalloy Electrocatalysts for CO_2 Reduction. *Phys. Chem. Chem. Phys.* **2015**, *17* (42), 28270–28276. https://doi.org/10.1039/c5cp00298b.

(12) Wang, J.; Hao, D.; Ye, J.; Umezawa, N. Determination of Crystal Structure of Graphitic Carbon Nitride: Ab Initio Evolutionary Search and Experimental Validation. *Chem. Mater.* **2017**, *29* (7), 2694–2707. https://doi.org/10.1021/acs.chemmater.6b02969.

(13) Buendía, F.; Vargas, J. A.; Johnston, R. L.; Beltrán, M. R. Study of the Stability of Small AuRh Clusters Found by a Genetic Algorithm Methodology. *Comput. Theor. Chem.* **2017**, *1119*, 51–58. https://doi.org/10.1016/j.comptc.2017.09.008.

(14) Mishin, Y. Machine-Learning Interatomic Potentials for Materials Science. *Acta Mater.* **2021**, *214*, 116980. https://doi.org/10.1016/j.actamat.2021.116980.

(15) Behler, J. Perspective: Machine Learning Potentials for Atomistic Simulations. *J. Chem. Phys.* **2016**, *145* (17), 170901. https://doi.org/10.1063/1.4966192.

(16) Behler, J.; Csányi, G. Machine Learning Potentials for Extended Systems: A Perspective. *Eur. Phys. J. B* **2021**, *94* (7), 1–11. https://doi.org/10.1140/epjb/s10051-021-00156-1.

(17) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. A Fourth-Generation High-Dimensional Neural Network Potential with Accurate Electrostatics Including Non-

Local Charge Transfer. *Nat. Commun.* **2021**, *12* (1), 398. https://doi.org/10.1038/s41467-020-20427-2.

(18) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 1–4. https://doi.org/10.1103/PhysRevLett.98.146401.

(19) Takamoto, S.; Shinagawa, C.; Motoki, D.; Nakago, K.; Li, W.; Kurata, I.; Watanabe, T.; Yayama, Y.; Iriguchi, H.; Asano, Y.; Onodera, T.; Ishii, T.; Kudo, T.; Ono, H.; Sawada, R.; Ishitani, R.; Ong, M.; Yamaguchi, T.; Kataoka, T.; Hayashi, A.; Charoenphakdee, N.; Ibuka, T. Towards Universal Neural Network Potential for Material Discovery Applicable to Arbitrary Combination of 45 Elements. *Nat. Commun.* **2022**, *13* (1), 2991. https://doi.org/10.1038/s41467-022-30687-9.

(20) Pun, G. P. P.; Batra, R.; Ramprasad, R.; Mishin, Y. Physically Informed Artificial Neural Networks for Atomistic Modeling of Materials. *Nat. Commun.* **2019**, *10* (1). https://doi.org/10.1038/s41467-019-10343-5.

(21) Deringer, V. L.; Csányi, G. Machine Learning Based Interatomic Potential for Amorphous Carbon. *Phys. Rev. B* **2017**, *95* (9), 1–15. https://doi.org/10.1103/PhysRevB.95.094203.

(22) Thompson, A. P.; Swiler, L. P.; Trott, C. R.; Foiles, S. M.; Tucker, G. J. Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials. *J. Comput. Phys.* **2015**, *285*, 316–330. https://doi.org/10.1016/j.jcp.2014.12.018.

(23) Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Model. Simul.* **2016**, *14* (3), 1153–1173. https://doi.org/10.1137/15M1054183.

(24) Busico, V. Giulio Natta and the Development of Stereoselective Propene Polymerization. In *Polyolefins: 50 years after Ziegler and Natta I*; Kaminsky, W., Ed.; Advances in Polymer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; Vol. 257, pp 37–57. https://doi.org/10.1007/12_2013_213.

(25) Taniike, T.; Terano, M. The Use of Donors to Increase the Isotacticity of Polypropylene. In *Polyolefins: 50 years after Ziegler and Natta I*; Kaminsky, W., Ed.; Advances in Polymer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; Vol. 257, pp 81–97. https://doi.org/10.1007/12_2013_224.

(26) Soga, K. Ziegler-Natta Catalysts for Olefin Polymerizations. *Prog. Polym. Sci.* **1997**, *22* (7), 1503–1546. https://doi.org/10.1016/S0079-6700(97)00003-8.

(27) Ratanasak, M.; Rungrotmongkol, T.; Saengsawang, O.; Hannongbua, S.; Parasuk, V. Towards the Design of New Electron Donors for Ziegler–Natta Catalyzed Propylene Polymerization Using QSPR Modeling. *Polymer* **2015**, *56*, 340–345. https://doi.org/10.1016/j.polymer.2014.11.022.

(28) Toto, M.; Morini, G.; Guerra, G.; Corradini, P.; Cavallo, L. Influence of 1,3-Diethers on the Stereospecificity of Propene Polymerization by Supported Ziegler-Natta Catalysts. A Theoretical Their Adsorption on (110) and (100) Lateral Cuts of $MgCl_2$ Platelets. *Macromolecules* **2000**, *33* (4), 1134–1140. https://doi.org/10.1021/ma990959a.

(29) Taniike, T.; Terano, M. Coadsorption and Support-Mediated Interaction of Ti Species with Ethyl Benzoate in $MgCl_2$-Supported Heterogeneous Ziegler-Natta Catalysts Studied by Density Functional Calculations. *Macromol. Rapid Commun.* **2007**, *28* (18–19), 1918–1922. https://doi.org/10.1002/marc.200700363.

(30) Taniike, T.; Terano, M. Coadsorption Model for First-Principle Description of Roles of Donors in Heterogeneous Ziegler-Natta Propylene Polymerization. *J. Catal.* **2012**, *293*, 39–50. https://doi.org/10.1016/j.jcat.2012.06.001.

(31) Ratanasak, M.; Parasuk, V. Roles of Malonate Donor on Activity and Stereoselectivity of Ziegler-Natta Catalyzed Propylene Polymerization. *J. Organomet. Chem.* **2015**, *775*, 6–11. https://doi.org/10.1016/j.jorganchem.2014.10.018.

(32) Vittoria, A.; Meppelder, A.; Friederichs, N.; Busico, V.; Cipullo, R. Demystifying Ziegler–Natta Catalysts: The Origin of Stereoselectivity. *ACS Catal.* **2017**, *7* (7), 4509–4518. https://doi.org/10.1021/acscatal.7b01232.

(33) Credendino, R.; Liguori, D.; Fan, Z.; Morini, G.; Cavallo, L. Toward a Unified Model Explaining Heterogeneous Ziegler-Natta Catalysis. *ACS Catal.* **2015**, *5* (9), 5431–5435. https://doi.org/10.1021/acscatal.5b01076.

(34) Xie, K.; Xu, S.; Hao, W.; Wang, J.; Huang, A.; Zhang, Y. Surface Effect of the $MgCl_2$ Support in Ziegler–Natta Catalyst for Ethylene Polymerization: A Computational Study. *Appl. Surf. Sci.* **2022**, *589*, 153002. https://doi.org/10.1016/j.apsusc.2022.153002.

(35) Masoori, M.; Nekoomanesh, M.; Posada-Pérez, S.; Rashedi, R.; Bahri-Laleh, N. Exploring Cocatalyst Type Effect on the Ziegler–Natta Catalyzed Ethylene Polymerizations: Experimental and DFT Studies. *J. Polym. Res.* **2022**, *29* (5). https://doi.org/10.1007/s10965-022-03050-1.

(36) Guo, X.; Cui, L.; Yi, J.; Liu, Z.; Liu, B. Understanding the Role of Sulfonyl Amine Donors in Propylene Polymerization Using $MgCl_2$-Supported Ziegler-Natta Catalyst. *J. Phys. Chem. C* **2022**, *126* (20), 8655–8666. https://doi.org/10.1021/acs.jpcc.2c02566.

(37) Guo, X.; Cui, L.; Wang, Y.; Yi, J.; Sun, J.; Liu, Z.; Liu, B. Mechanistic Study on Effect of Electron Donors in Propylene Polymerization Using the Ziegler–Natta Catalyst. *J. Phys. Chem. C* **2021**, *125* (16), 8533–8542. https://doi.org/10.1021/acs.jpcc.0c11273.

(38) Fallah, M.; Bahri-Laleh, N.; Didehban, K.; Poater, A. Interaction of Common Cocatalysts in Ziegler–Natta-Catalyzed Olefin Polymerization. *Appl. Organomet. Chem.* **2020**, *34* (2). https://doi.org/10.1002/aoc.5333.

(39) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Insight into Structural Distribution of Heterogeneous Ziegler–Natta Catalyst from Non-Empirical Structure Determination. *J. Catal.* **2021**, *394*, 299–306. https://doi.org/10.1016/j.jcat.2020.11.005.

(40) Takasao, G.; Wada, T.; Chikuma, H.; Chammingkwan, P.; Terano, M.; Taniike, T. Preventing Premature Convergence in Evolutionary Structure Determination of Complex Molecular Systems: Demonstration in Few-Nanometer-Sized $TiCl_4$-Capped $MgCl_2$ Nanoplates. *J. Phys. Chem. A* **2022**, *126* (31), 5215–5221. https://doi.org/10.1021/acs.jpca.2c02112.

(41) Marigo, A.; Marega, C.; Zannetti, R.; Morini, G.; Ferrara, G. Small- and Wide-Angle X-Ray Scattering Analysis of Ziegler–Natta Catalysts: Structural Disorder, Surface Area and Activity. *Eur. Polym. J.* **2000**, *36* (9), 1921–1926. https://doi.org/10.1016/S0014-3057(99)00250-5.

(42) Zannetti, R.; Marega, C.; Marigo, A.; Martorana, A. Layer-lattices in Ziegler–Natta Catalysts. *J. Polym. Sci. Part B Polym. Phys.* **1988**, *26* (12), 2399–2412. https://doi.org/10.1002/polb.1988.090261202.

(43) Chikuma, H.; Takasao, G.; Wada, T.; Chammingkwan, P.; Behler, J.; Taniike, T. Accelerating Non-Empirical Structure Determination of Ziegler–Natta Catalysts with a High-Dimensional Neural Network Potential. *J. Phys. Chem. C* **2023**, *127* (24), 11683–11691. https://doi.org/10.1021/acs.jpcc.3c01511.

(44) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134* (7), 074106. https://doi.org/10.1063/1.3553717.

(45) Behler, J. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chem. Int. Ed.* **2017**, *56* (42), 12828–12840. https://doi.org/10.1002/anie.201703114.

(46) Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115* (16), 1032–1050. https://doi.org/10.1002/qua.24890.

(47) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(48) Delley, B. An All-electron Numerical Method for Solving the Local Density Functional for Polyatomic Molecules. *J. Chem. Phys.* **1990**, *92* (1), 508–517. https://doi.org/10.1063/1.458452.

(49) Bergner, A.; Dolg, M.; Küchle, W.; Stoll, H.; Preuß, H. Ab Initio Energy-Adjusted Pseudopotentials for Elements of Groups 13–17. *Mol. Phys.* **1993**, *80* (6), 1431–1441. https://doi.org/10.1080/00268979300103121.

(50) Singraber, A.; Behler, J.; Dellago, C. Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials. *J. Chem. Theory Comput.* **2019**, *15* (3), 1827–1840. https://doi.org/10.1021/acs.jctc.8b00770.

(51) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Comput. Phys. Commun.* **2022**, *271*, 108171. https://doi.org/10.1016/j.cpc.2021.108171.

(52) Zorve, P.; Linnolahti, M. Adsorption of Titanium Tetrachloride on Magnesium Dichloride Clusters. *ACS Omega* **2018**, *3* (8), 9921–9928. https://doi.org/10.1021/acsomega.8b01878.

(53) Zaccaria, F.; Ehm, C.; Budzelaar, P. H. M.; Busico, V. Accurate Prediction of Copolymerization Statistics in Molecular Olefin Polymerization Catalysis: The Role of Entropic, Electronic, and Steric Effects in Catalyst Comonomer Affinity. *ACS Catal.* **2017**, *7* (2), 1512–1519. https://doi.org/10.1021/acscatal.6b03458.

(54) Zaccaria, F.; Cipullo, R.; Budzelaar, P. H. M.; Busico, V.; Ehm, C. Backbone Rearrangement during Olefin Capture as the Rate Limiting Step in Molecular Olefin Polymerization Catalysis and Its Effect on Comonomer Affinity. *J. Polym. Sci. Part Polym. Chem.* **2017**, *55* (17), 2807–2814. https://doi.org/10.1002/pola.28685.

(55) Piovano, A.; D'Amore, M.; Wada, T.; Cleto Bruzzese, P.; Takasao, G.; Thakur, A.; Chammingkwan, P.; Terano, M.; Civalleri, B.; Bordiga, S.; Taniike, T.; Groppo, E. Revisiting the Identity of δ-$MgCl_2$: Part II. Morphology and Exposed Surfaces Studied

by Vibrational Spectroscopies and DFT Calculation. *J. Catal.* **2020**, *387*, 1–11. https://doi.org/10.1016/j.jcat.2020.04.017.

(56) Busico, V.; Cipullo, R.; Monaco, G.; Talarico, G.; Vacatello, M.; Chadwick, J. C.; Segre, A. L.; Sudmeijer, O. High-Resolution ^{13}C NMR Configurational Analysis of Polypropylene Made with $MgCl_2$-Supported Ziegler−Natta Catalysts. 1. The "Model" System $MgCl_2/TiCl_4$−2,6-Dimethylpyridine/$Al(C_2H_5)_3$. *Macromolecules* **1999**, *32* (13), 4173–4182. https://doi.org/10.1021/ma981941n.

(57) Saelee, T.; Sitthijun, P.; Ngamlaor, C.; Kerdprasit, N.; Rittiruam, M.; Khajondetchairit, P.; Da Silva, J. L. F.; Buasuk, N.; Praserthdam, P.; Praserthdam, S. Experimental and First-Principles Investigation on How Support Morphology Determines the Performance of the Ziegler-Natta Catalyst during Ethylene Polymerization. *Sci. Rep.* **2024**, *14* (1), 17835. https://doi.org/10.1038/s41598-024-68289-8.

Chapter 5

Application of a High-Dimensional Neural Network Potential to Ziegler-Natta Catalysts with Internal Donors

ABSTRACT

The influence of internal donors (IDs) on the morphological characteristics of Ziegler-Natta catalysts is crucial for understanding propylene polymerization. A comprehensive structural distribution analysis is essential to investigate the complex interactions between the active sites, chemical adsorption, and the surface state of the $MgCl_2$ support. However, traditional structure searches that combine genetic algorithms (GA) with density functional theory (DFT) are often limited by the high computational cost of DFT. In this chapter, we employed High-Dimensional Neural Network Potentials (HDNNP) to significantly accelerate the computational process, enabling the exploration of 12.4 times more structural candidates. This approach led to the identification of a new, most stable structure, which is 2.6 kcal/mol more stable than previously known configurations. Furthermore, HDNNP-based GA discovered 200 thermodynamically exchangeable metastable structures that were inaccessible via the DFT-based GA approach, providing novel insights into the distribution of active species.

KEYWORDS: Neural network potential, genetic algorithm, structure determination, Ziegler-Natta catalyst, internal donor, global minimization

5.1. INTRODCUTION

Ziegler-Natta (ZN) catalysts, developed by Karl Ziegler and Giulio Natta in the 1950s, revolutionized the production of high-performance plastics.[1] These catalysts are essential for industrial polyolefin production and are used in 99% of polypropylene (PP) manufacturing. $TiCl_4$ serves as a precursor for the active sites, and $MgCl_2$ supports, are typically employed.[2] An important nature of ZN catalysts for propylene polymerization is the use of internal (ID) and external (ED) donors, composed of organic compounds, to improve the stereo regularity and molecular weight distribution of the resulting polymer.[3–5] Controlling the stereochemistry of the polymer chain allows the synthesis of isotactic, syndiotactic, and atactic polymers, which in turn determines material properties such as melting point, crystallinity, tensile strength, and elastic modulus. IDs, such as ethyl benzoate (EB), phthalate esters (DEP), diethers (DE), and succinates, are added during catalyst preparation and coordinate with the active sites, influencing monomer orientation during polymerization.[5,6] This binding is crucial for achieving the desired stereoregularity, especially in polypropylene production. EDs, typically alkoxysilanes, are added during polymerization. While the choice of ID significantly impacts the distribution of active sites and the resultant molecular weight distribution of the polymer, the effect of ED selection is minor.[7] These observations suggest that while both IDs and EDs enhance stereospecificity during polymerization, IDs have a substantial influence on the formation and distribution of active sites during catalyst preparation.

The functional unit of ZN catalysts consists of $MgCl_2$ nanoplates terminated with adsorbed $TiCl_4$ and internal donors (IDs), called primary particle. $MgCl_2$ nanoplates are composed of {110} and {100} surfaces, corresponding to 4-fold and 5-fold coordinated

Mg^{2+} cation sites, respectively.[8,9] The primary particles stack disorderly along the c-axis, weakly bound by van der Waals forces.[9] Understanding the morphology of primary particles and the local chemical environment surrounding active sites provides crucial insights into catalytic functionality. On the other hand, the $MgCl_2$ support structure is highly sensitive to its surrounding chemical environment and is prone to reconfiguration upon adsorption.[10,11] Additionally, X-ray scattering spectra tend to be broad, making it difficult to identify specific surface exposures.[12]

Quantum chemical simulations provide powerful tools for interpreting experimental observations at the atomic and molecular levels. For instance, in ZN catalysts, a combination of spectroscopy techniques, such as infrared (IR), ultraviolet-visible (UV-Vis), or other measurement techniques with simulations is applied to decipher structural complexities.[10,13–15] On the other hand, accurately correlating simulations with the properties of real catalysts requires a molecular model that reliably represents the system, such as the most stable structure. The integration of machine learning with DFT has enabled "non-empirical" structure determination, allowing representative structures to be proposed without prior experimental knowledge. This approach has been applied across various fields, from single-atom clusters and alloys to supported catalysts.[16–19] Structure determination involves exploring the potential energy surface (PES) over a vast configuration space, which is computationally expensive when using DFT. Machine learning potentials (MLPs) have emerged as a solution to this challenge.[20–27]

To investigate the detailed structure of ZN catalysts, Takasao et al. combined genetic algorithms (GA) with DFT geometry optimization, enabling simultaneous exploration of the support structure and adsorbates, which led to the determination of $MgCl_2/TiCl_4$ primary particle structures.[28–30] This work revealed that the chemisorption of $TiCl_4$

promotes structural diversity in Ziegler-Natta catalysts by forming $MgCl_2$ nanoplate morphologies. Furthermore, Chapters 2 and 3 of this thesis reported non-empirical structure determination for experimentally relevant sizes using high-dimensional neural network potentials (HDNNPs),[20] a classical MLP strategy introduced by Behler and Parinello.[27] More recently, the structure determination of $MgCl_2/TiCl_4$/ID systems with two different IDs (diethyl phthalate (DEP) and 9,9-bis(methoxymethyl)fluorene) has shown that the co-adsorption of IDs and $TiCl_4$ preferentially forms sharp surfaces with edges and cavities, essential for $MgCl_2$ reconstruction.[31] Incorporating IDs in structure determination increases the DFT computation time and rapidly expands the configuration space due to the diversity of adsorption patterns, making investigating enough conformational space difficult.

Here, I explored a more comprehensive configuration space of $MgCl_2/TiCl_4$/ID cluster using HDNNPs to identify the most stable structures while revealing diverse stable configurations. The constructed HDNNP must capture both the covalent nature of IDs and the ionic characteristics of $MgCl_2/TiCl_4$. This chapter represents one of the few cases applying HDNNPs to such a complex system. Also, accelerating these calculations is a crucial step toward simulating $MgCl_2/TiCl_4$/ID clusters at realistic scales.

5.2. Numerical method

5.2.1. Composition of the $MgCl_2/TiCl_4$/ID cluster

Here, a comprehensive exploration of orientation space for a $19MgCl_2/4TiCl_4$/5diethyl phthalate (DEP) cluster was conducted using high-dimensional neural network potentials (HDNNPs). The HDNNP was trained on a DFT database from a structural

determination study for the same system, reported by da Silveira et al. 19$MgCl_2$/4$TiCl_4$/5ID composition was selected for its balance between computational feasibility and realistic catalyst coverage. This chapter focuses on DEP, which has a distinct carbonyl (C=O) peak in IR spectroscopy,[32] suitable for spectroscopic simulation study in the future.

5.2.2. HDNNP construction for $MgCl_2$/$TiCl_4$/DEP system

The HDNNP input was created by combining radial ($G_{\mathrm{i}}^{\mathrm{rad}}$) and angular ($G_{\mathrm{i}}^{\mathrm{ang1}}, G_{\mathrm{i}}^{\mathrm{ang2}}$) atom centered symmetry functions (ACSFs)[21,27,33] and their cutoff function (f_{c}) with a cutoff radius R_c are defined as,

$$G_i^{\mathrm{rad}} = \sum_{j=1}^{N_{\mathrm{atoms}}} \mathrm{e}^{-\eta\left(R_{ij}-R_{\mathrm{s}}\right)^2} f_{\mathrm{c}}\left(R_{ij}\right) \tag{5.1}$$

$$G_i^{\mathrm{ang1}} = 2^{1-\zeta} \sum_{j\neq i}^{N_{\mathrm{atoms}}} \sum_{k\neq i,j}^{N_{\mathrm{atoms}}} \left\{\left(1 + \lambda\cos\theta_{ijk}\right)^{\zeta} \mathrm{e}^{-\eta\left({R_{ij}}^2+{R_{jk}}^2+{R_{ik}}^2\right)} f_c\left(R_{ij}\right) f_c\left(R_{jk}\right) f_c(R_{ik})\right\} \tag{5.2}$$

$$G_i^{\mathrm{ang2}} = 2^{1-\zeta} \sum_{j\neq i}^{N_{\mathrm{atoms}}} \sum_{k\neq i,j}^{N_{\mathrm{atoms}}} \left\{\left(1 + \lambda\cos\theta_{ijk}\right)^{\zeta} \mathrm{e}^{-\eta\left({R_{ij}}^2+{R_{ik}}^2\right)} f_c\left(R_{ij}\right) f_c(R_{ik})\right\} \tag{5.3}$$

$$f_{\mathrm{c}}\left(R_{ij}\right) = \begin{cases} 0.5\left[\cos\left(\dfrac{R_{ij}}{R_{\mathrm{c}}}\right) + 1\right] & \left(R_{ij} \leq R_c\right) \\ 0 & \left(R_{ij} > R_c\right) \end{cases} \tag{5.4}$$

where η, ζ, λ, and R_{s} are hyperparameters, R_{ij} is the distance between atoms *i* and *j*, and θ_{ijk} is the angle defined among atoms *i*, *j*, and *k* with *i* at the center. The ACSF parameters were determined empirically. Different ACSF parameters were applied for ionic (pairs or triads consisting of Mg, Cl, and Ti), covalent (O, H, and C), and mixed

environments, which include both types. ACSF parameters to describe ionic environment were selected based on Chapter 2, with slight adjustments. The radial ACSF parameters for covalent environments referenced those used in ANI-2x, which is the HDNNP package representing the organic system provided by Devereux et al.[34] For mixed interactions, only essential functions within an effective interaction range were included, omitting descriptors (e.g., H-Mg-Ti interactions) with minimal impact on the entire energy. The details of hyperparameters is discussed in supporting information part.

The HDNNPs were constructed by using the RuNNer code (version 1.2.0).[22,35] The atomic neural networks were trained using a Kalman filter.[36] The atomic neural networks were trained using a Kalman filter to minimize a loss function (Γ), defined as,

$$\Gamma = \frac{1}{N_{\text{structure}}} \sum_{i}^{N_{\text{structure}}} [\left(E^i_{\text{DFT}} - E^i_{\text{HDNNP}}\right)^2 + \frac{\beta}{3N^i_{\text{atom}}} \sum_{j}^{N^i_{atom}} (F^i_{j,\text{DFT}} - F^i_{j,\text{HDNNP}})^2] \quad (5.5),$$

where $N_{\text{structure}}$ is the number of structures in the training set, N^i_{atom} is the number of atoms in structure i, E^i is the energy of structure i, and F^i_j is the force on atom j in structure i. E^i_{HDNNP} is computed as the sum of atomic energies and Since F^i_j is obtained as the analytical derivative of the atomic energy. The network is thus optimized to output atomic energies that minimize the loss function. Here, the beta value was set to 10.0 to ensure that the optimizer focused on force fitting because the distribution of force magnitude significantly depending on the nature of the chemical bonding. The final HDNNP atomic network was composed of two layers, each containing 15 nodes.

5.2.3. Dataset construction

An initial reference dataset of 1,800 structures from the DFT database of the 19$MgCl_2$/4$TiCl_4$/5DEP cluster was selected, including geometry optimization trajectories. Furthest point sampling (FPS) was used to efficiently select representative environments for the initial dataset, with 70% allocated for training and 30% for testing. The initial adaptive sampling cycle added extrapolated structures to expand the potential energy surface (PES). Geometry optimization was performed on 10,000 initial structures, with 300–1,500 selected from extrapolating structures using FPS and recalculated with DFT. The collection process continued until a 95% success rate (defined as the proportion of trials that avoided extrapolation at the final simulation step) in avoiding extrapolation was achieved. Subsequently, geometric optimization was performed using two HDNNPs with different architectures but comparable accuracy. Structures with inconsistent energy and geometry results between the two HDNNPs were recalculated using DFT and added to the reference dataset. For more details on the adaptive sampling approach, refer to Chapter 2.

The constructed HDNNP was evaluated based on its accuracy in single-point calculations and in geometry optimization for structures beyond the training set. DFT calculations at GGA-PBE[37] level employing the double numerical basis functions (DNP)[38] basis sets, implemented with effective core potentials (ECP)[39] in $DMol^3$.[38]

5.2.4. Structure determination

The HDNNP and genetic algorithm (GA) were combined for an extensive structural exploration of the 19$MgCl_2$/4$TiCl_4$/5DEP cluster, guided by the structure-determination

algorithm in Ref[31]. Each candidate structure was optimized using LAMMPS with an HDNNP interface, applying a force convergence criterion of 154 meV/Å, and the energy for optimized structures was calculated as E_{cluster} defined as,

$$E_{\text{cluster}} = E - N_{\text{MgCl}_2} \times E_{\text{MgCl}_2} - N_{\text{TiCl}_4} \times E_{\text{TiCl}_4} - N_{\text{ID}} \times E_{\text{ID}} \quad (5.6),$$

where E is the total energy of a structure, N_{MgCl_2} is the number of $MgCl_2$ units, E_{MgCl_2} is the energy of a $MgCl_2$ unit in α-$MgCl_2$, N_{TiCl_4} is the number of $TiCl_4$ molecules, and E_{TiCl_4} is the energy of a $TiCl_4$ molecule in vacuum, N_{ID} is the number of ID molecules, and E_{ID} is the energy of a ID molecule in vacuum. The GA was performed to minimize E_{cluster}, and the features of the structures with high fitness *(f)* were preferentially inherited through roulette selection. f of the i^{th} structure is defined as,

$$f_i = e^{-3\left(\frac{E_i - E_{min}}{E_{max} - E_{min}}\right)} \quad (5.7),$$

where E_i, E_{min}, and E_{max} are the energy of the i^{th} structure in a generation, and the energies of the most and least stable structures in the same generation.

The primary GA parameters used are detailed in Table 5.1. In this system, total energy is primarily influenced by adsorbate interactions, meaning that even structures with similar $MgCl_2$ skeleton can differ significantly in energy. To ensure diversity in adsorption patterns, a relatively large number of elite structures were retained in the GA process. In order to improve structural diversity, 9 parallel runs were performed, and in 8 of these runs, an migration operator,[30] that imports elite structures from the database based on relative energy, was introduced after 200 generations.

Table 5.1. Parameters used in the genetic algorithm

Parameter	Value
Population (structure)	72
Crossover (%)	27.8
Mutation (skeleton) (%)	16.7
Mutation (adsorbate) (%)	11.1
Elitism (%)	44.4

5.3. RESULT AND DISCUSSION

5.3.1. HDNNP construction

By incorporating adaptive sampling from the structural optimization trajectory, I expanded the dataset to obtain structures of 19$MgCl_2$/4$TiCl_4$/5DEP The accuracy of the DFT calculations for the HDNNP trained on this dataset was evaluated using the root-mean-square-error (RMSE), as shown in Table 5.2. Due to the complexity of fitting this system, the RMSEs for energy and force in both training and single-point tests were higher than those for the MgCl2/TiCl4 system. However, these values, around 1 meV/atom, remained within acceptable limits.

A geometry optimization test was conducted using 1,000 initial structures randomly extracted from the dataset, with 300 structures, randomly sampled, further analyzed using DFT single-point calculations for verification. As indicated in Table 5.2, the energy RMSE was 1.54 meV, reflecting the systematic underestimation commonly observed in MLP-based geometry optimization. Since the optimization algorithm inherently follows the PES toward a more stable structure, underestimation errors can

accumulate over successive optimization cycles. Nevertheless, deviations of a few meV/atom during structure optimization are typical for HDNNP accuracy.[40]

Table 5.1. Prediction errors of the HDNNP of $19MgCl_2/4TiCl_4/5DEP$ [a]

Training		Single-point test		Geometry-optimization test		
E (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	*E* (meV/atom)	*F* (meV/Å)	Success rate (%)
0.434	72.5	0.747	74.1	1.54	47.8	96.8

[a] The energies *E* and forces *F* indicate root mean squared error (RMSE) from DFT. In training, 70% of the reference dataset was used as the training set and remaining 30% as the testing set. Single-point test shows RMSE for the testing set. Geometry optimization tests were performed using 10,000 initial structures of $19MgCl_2/4TiCl_4/5DEP$ cluster, having extremally high $TiCl_4$ coverage. The RMSE values were calculated for the 100 structures optimized using the HDNNP. Success rate means percentage of structures for which no extrapolation occurred during the geometry optimization.

5.3.2. Structure determination

Structure determination was performed using the HDNNP and GA (HDNNP-GA) was applied to $19MgCl_2/4TiCl_4/5DEP$ cluster. Figure 4 presents the energy evolution plot of GA runs, which were run for 1,000 generations. This plot illustrates the energy of the most ststable structure in the generation, highlighting the optimization process

and convergence behavior of the GA toward more stable structures. The GA was terminated after 1,000 generations as 7 out of 9 runs—excluding Run 1, which did not include migration operations, and Run 3—converged to the same most stable structure. All runs resulted close to E_{cluster} = -52.6 kcal/mol. The HDNNP-GA explored a total of 178,377 structures (12.4 times the search space of DFT-GA), and obtained 621 stable structures ranked within 6 kcal/mol of relative energy $\Delta E_{\text{cluster}}$ from the most stable one.

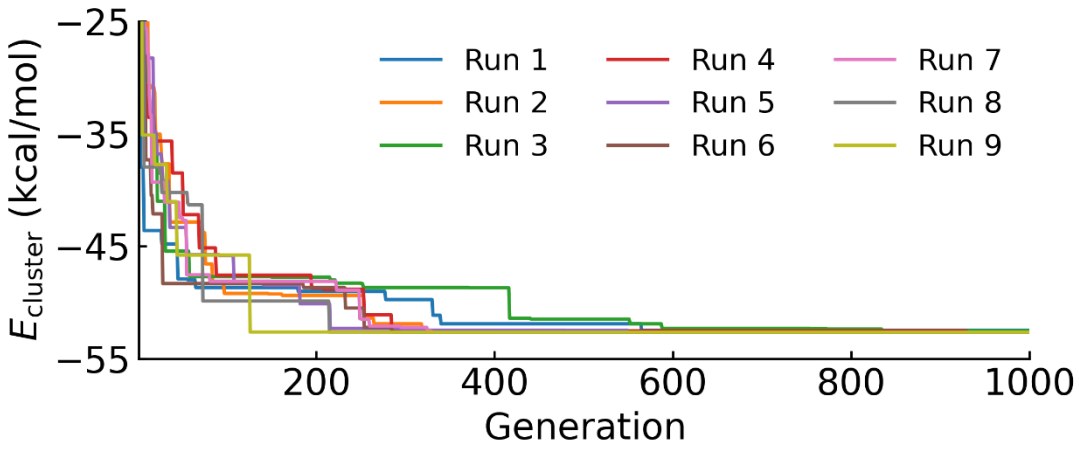


Figure 5.1. Evolutionary progress plots of HDNNP-GA for the structure determination of 19$MgCl_2$/4$TiCl_4$/5DEP. The plot shows the energy of the most stable structure in the population of the generation. E_{cluster} is a sum of the surface energy and the adsorption energy of $TiCl_4$, as defined in Eq. 5.6.

The most stable structure obtained by HDNNP-GA and one that obtained by GA using DFT (DFT-GA) obtained in Ref.[31] are shown in Figure 5.3a. Here, the most stable structure obtained using HDNNP-GA (denoted as HDNNP-stable) with that obtained using DFT-GA from Ref.[31] (denoted as DFT-stable). Both structures, , which were re-

evaluated via DFT-geometry optimization, have nearly identical energies, indistinguishable within the chemical accuracy limit of 1 kcal/mol. Although the DFT-stable structure was identified during the HDNNP-GA search, the HDNNP-stable structure was not found in the DFT-GA. Despite having nearly the same energy, HDNNP-stable features a different DEP adsorption distribution, characterized by three chelates and two bridges.

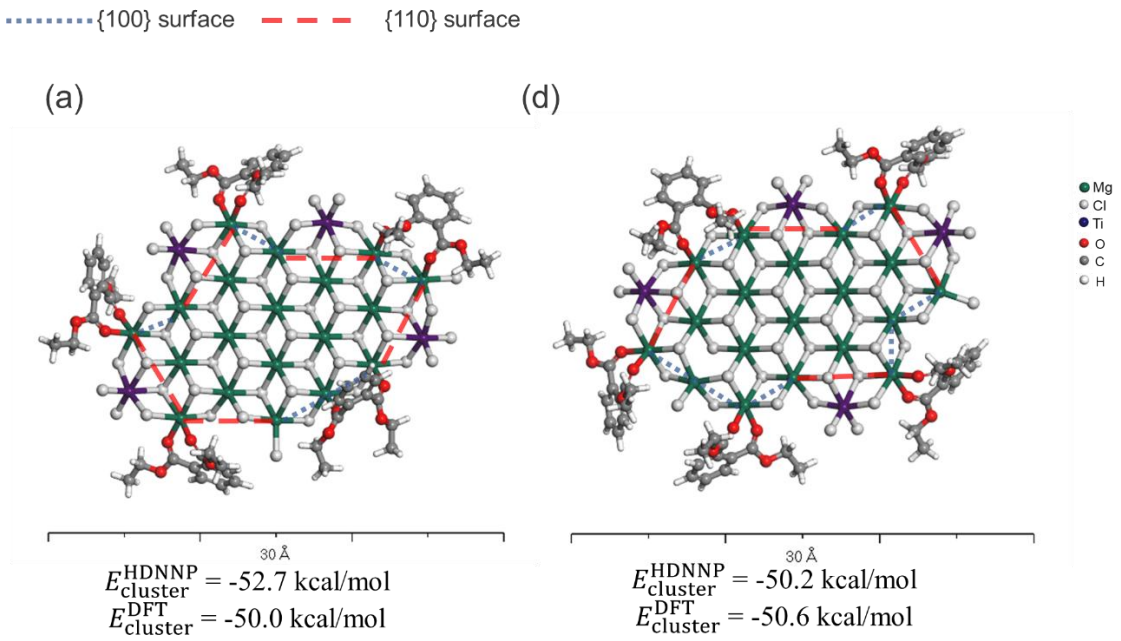


Figure 5.2. The most stable structures obtained by a) HDNNP-GA and b) DFT-GA. The energies of each structure were evaluated using both DFT geometry optimization ($E_{cluster}^{DFT}$) and HDNNP structure optimization ($E_{cluster}^{HDNNP}$). Note that the HDNNP geometry optimization was applied to the initial structures generated by the GA, while the DFT geometry optimization for structure a) was conducted following the HDNNP geometry optimization.

HDNNP estimated an energy difference of 4 kcal/mol between these closely related structures. Although this difference falls within the expected range based on 1.5 meV/atom, it can influence thermodynamic distribution estimations. To refine this analysis, 621 of the top stable structures identified by HDNNP-GA were optimized using DFT, and their energies were re-evaluated. Of these, 336 structures fell within 6 kcal/mol of the most stable energy (Fig. 5.4a), and some had even lower energies than the DFT-stable structure (Fig. 5.3c).

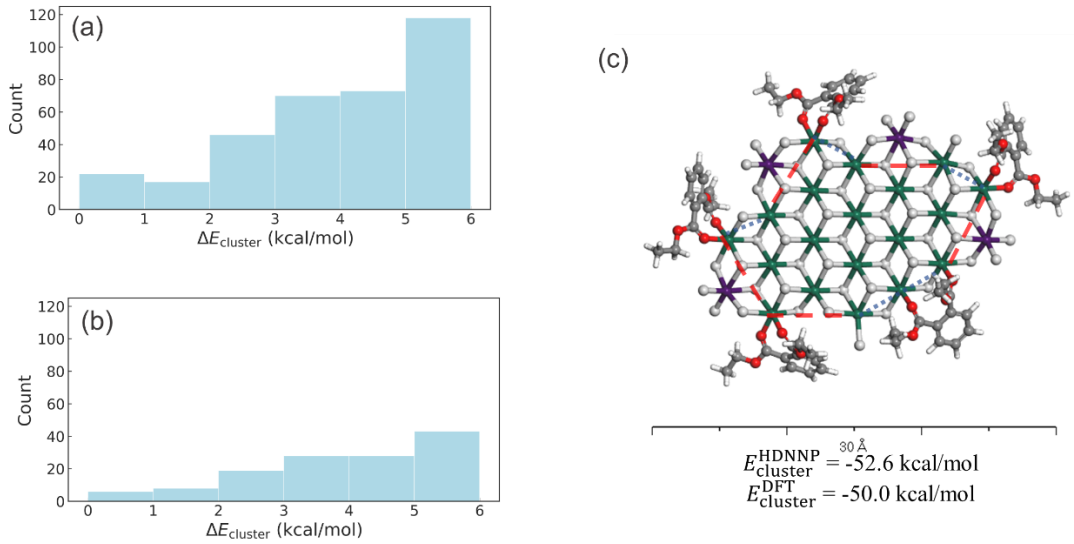


Figure 5.3. Distribution of relative energy to the most stable one ($\Delta E_{\text{cluster}}$) in a) HDNNP-GA and b) DFT-GA. c) The most stable structure obtained from metastable structures in HDNNP-GA evaluated using DFT-geometry optimization. The dotted blue and dashed red lines indicate the $MgCl_2$ {100} and {110} surfaces, respectively.

Although this analysis required DFT optimization of hundreds of structures, the computational cost remained significantly lower than a full search of tens of thousands of candidates for structure determination. Figure 5.4a shows that HDNNP-GA

uncovered more than 20 unique stable structures within 1 kcal/mol, indicating that the inclusion of DEP led to a considerable diversification of stable configurations. This exploration reached areas that were insufficiently covered by DFT-GA.

The most stable structure identified by DFT (Fig. 5.4c) shared a motif similar to the DFT-GA result, particularly in the distribution of DEP adsorption patterns and terminal Cl adsorption sites, but the overall cluster energy differed by 2.6 kcal/mol. A notable feature of the HDNNP-stable structure was a $TiCl_4$ species situated between two chelating DEPs, which also had an identical $MgCl_2$ framework as shown in Figure 5.3a. This suggests that certain stable skeleton found in one GA run were reinforced by genetic operations, implying that the variety of DEP adsorption structures plays a key role in the overall structural distribution.

5.3.3. Distribution analysis

The structure of Ti species is closely tied to their catalytic properties and plays a critical role in determining the molecular weight distribution of resulting polymers. On the $MgCl_2$ surface, $TiCl_4$ species are generally categorized into three main types: mononuclear species on {110} terraces, mononuclear species on {100} terraces, and dinuclear species on {100} terraces. Among these, the mononuclear species on {110} terraces are known to be thermodynamically more stable, supported by various advanced characterization techniques, which consistently identify this species as the most representative. Within the framework of the three-site model,[41] bulky ligands located at two adjacent Mg^{2+} cations are essential for forming high stereo specific active site. Consequently, the mononuclear species on {110} terraces have been further subdivided into three subtypes: mono_1–3. In this model, 'ligands' refer to Cl^- anions, or

oxygen atoms from donor molecules that coordinate to Mg^{2+} cations.[41–44] The distribution of each species was evaluated using Boltzmann weighting, defined as,

$$p = e^{-\frac{\Delta E_{\text{cluster}}}{RT}} \quad (5.8)$$

where p is the population of structure, $\Delta E_{\text{cluster}}$ is the cluster energy of a structure relative to that of the most stable structure, R is the gas constant, at a temperature T = 350 K. Energies of HDNNP-GA-derived structures were evaluated using DFT-geometry optimization to ensure consistency. This section provides a detailed chemical understanding of the $TiCl_4$ distribution and insights into how the breadth of the search space impacts distribution estimates.

I independently estimated and compared the $TiCl_4$ distributions based on the stable structures obtained via DFT-GA and HDNNP-GA. Figure 5.4 shows the distribution of $TiCl_4$ species corresponding to these stable structures. Both methods consistently identified {110} mono_3 as the dominant $TiCl_4$ species, with negligible contributions from {100} species. Although minor differences in the occupancy of {110} mono_3 were observed between the two methods, these variations were not statistically significant due to the sensitivity of Boltzmann factors to relative energy differences. This result highlights the intrinsic role of IDs in enhancing the stereospecificity of active sites. As discussed in Chapter 3, clusters without IDs exhibit a lower proportion of stereospecific $TiCl_4$ species, even when sufficient surface area is available. These insights explain the experimental results that show that, isotactic polypropylene can be produced using a catalyst without IDs but less than that of a catalyst that does contain a donor.[45]

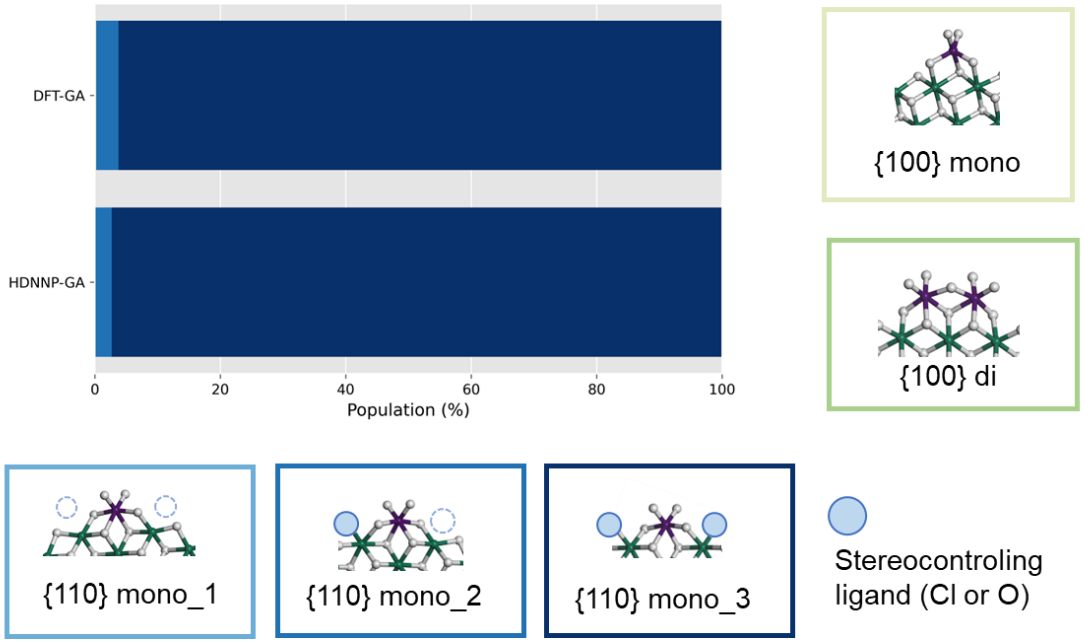


Figure 5.4. Comparison of the $TiCl_4$ species distribution based on stable structures obtained from DFT-GA and HDNNP-GA. The population of each species is expressed as a percentage, calculated using Boltzmann weighting at 350 K. The insects display representative structures of each species, including {110} mononuclear species (mono_1–3), {100} mononuclear species ({100} mono), and {100} dinuclear species ({100} di) based on three-site model.[41]

To further refine the analysis, the $TiCl_4$ mono_3 species were classified based on the types of ligands present, namely Cl^- anions, bridge-adsorbed donors on {100} terraces, or chelate-adsorbed donors on {110} terraces (Figure 5.5).[46,47] The adsorption mode of the donor on the {110} terrace could potentially include bridge sites; however, no such configurations were observed among the stable structures within 6 kcal/mol of the stable structures obtained via both DFT and HDNNP-GA. Three pairs of ligand classifications yielded six different subtypes, and their distributions were estimated for structures

derived from both methods, shown in Figure 5.6. Interestingly, while broader classifications revealed minimal differences between the two GA approaches, finer classifications based on ligand types exposed significant disparities. The DFT-GA-derived structures exhibited a large population of $TiCl_4$ species coordinated between chelate and bridge ligands, whereas the occurrence of $TiCl_4$ species sandwiched by two Cl ligands or two chelate ligands was underestimated. These differences in distribution arose from the more diverse structural patterns sampled through exhaustive configurational searches. Notably, the extensive configurational search enabled by HDNNP-GA identified $TiCl_4$ species sandwiched between two Cl ligands, a configuration less observed in the DFT-GA results. This finding highlights that, even in systems containing IDs, Cl anions can act as steric control ligands, influencing the stereospecificity of the active sites. These facts indicate that the active center is not governed by a specific chemical environment, and that a variety of active structures coexist.

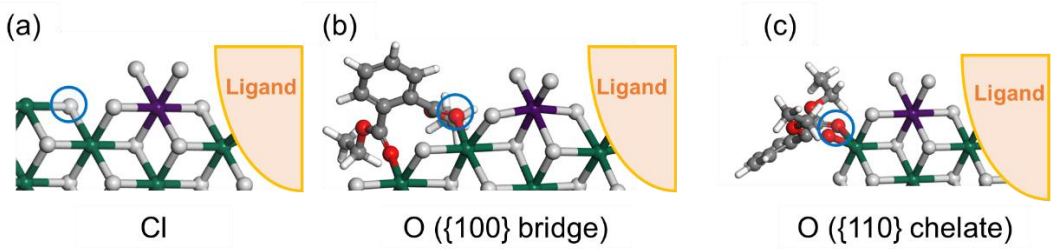


Figure 5.5. Representative structures of {110} mono_3 $TiCl_4$ species classified based on ligand types. a) Cl^- anions, b) bridge-adsorbed donor on a {100} terrace, c) chelate-adsorbed donor on a {110} terrace. The ligand is surrounded by a blue circle. The chelate-adsorbed ID has two O atoms in the vicinity of $TiCl_4$, but here they are defined as a single ligand.

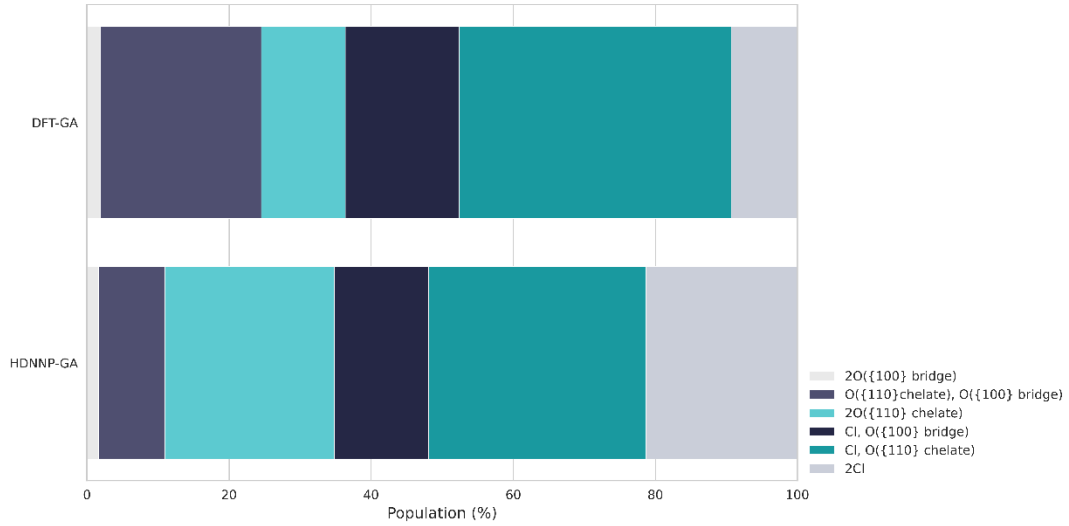


Figure 5.6. Distribution of $TiCl_4$ species (T = 350 K) based on stable structures obtained from DFT-GA and HDNNP-GA. The {110} mono_3 species, which has ligands adjacent on both sides, was further subdivided into six types depending on the combination of the types of ligands shown in Figure 5.5.

Next, the discussion focuses on the charge distribution of $TiCl_4$ species. The introduction of IDs contributes both sterically and electronically to the Ti active center, altering its characteristics. Consequently, the electronic state of $TiCl_4$ provides critical insights into catalyst performance. For instance, even Ti species with similar steric environments exhibit a range of charge states, and species with different charges may exhibit varying catalytic activities.[29,31] Figure 5.7 presents the charge distribution for structures obtained via HDNNP-GA. The charge of $TiCl_4$ was evaluated using Hirshfeld charges calculated through DFT at GGA-PBE level. Based on the three-site model (Figure 5.5a), the mono_3 structures, which make up the majority of the stable

structures, exhibit charges ranging from -0.25 to -0.1. This is significantly more negatively charged and widely distributed than the distribution of donor-free clusters reported in Chapter 3 (-0.08 to -0.03). This trend is consistent with findings from DFT-GA,[31] which demonstrated that the Lewis basicity of donors increases the negative charge of all $TiCl_4$ species, aligning well with DFT[29] and experimental observations.[48,49]

Notably, the charge of the stereospecific {110} mono_3 species spans a wide range, indicating the presence of diverse active species. Figure 5.7b illustrates the charge distribution of {110} mono_3, categorized by the ligand types shown in Figure 5.5. Interestingly, the charge distribution varies depending on the ligand type, suggesting that the donor makes a direct electronic contribution to the adjacent Ti species. Chelate-adsorbed donors, likely due to the two oxygen atoms functioning as ligands, induce the most negative charge on $TiCl_4$. Furthermore, compared to bridge-adsorbed donors, chelate-adsorbed donors result in a broader charge distribution for $TiCl_4$. This broader distribution can be attributed to the diverse range of stable conformations, reflecting their chemical environments.

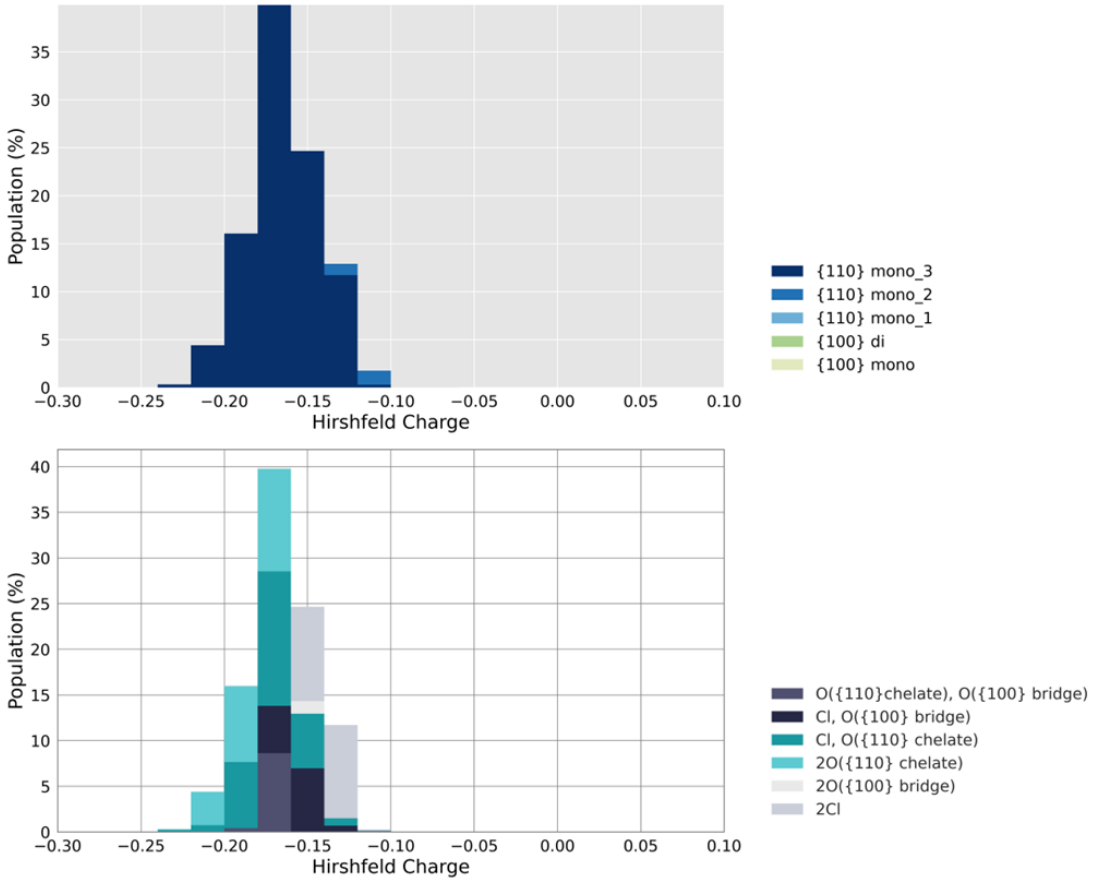


Figure 5.7. Hirshfeld charge distributions for different $TiCl_4$ species. a) Distribution of Hirshfeld charges for general active site classifications based on three-site model.[41] b) Detailed Hirshfeld charge distribution for {110} mono_3 species, further categorized by ligand types shown in Figure 5.5.

5.4. Conclusion

A high-dimensional neural network potential (HDNNP) was developed to enable a comprehensive exploration of the structural space of $MgCl_2$/$TiCl_4$/diethyl phthalate (DEP) clusters. A key aspect of HDNNP development was the iterative incorporation of conformations of flexible organic molecules through HDNNP-based simulations,

allowing for the efficient sampling of diverse chemical environments associated with DEP. This approach resulted in a potential capable of accurately reproducing DFT-level structural optimizations.

The non-empirical structure determination was performed using a genetic algorithm (GA), and the significant computational speedup provided by HDNNP enabled the exploration of a structure space 12 times larger than what could be achieved using DFT alone. This expanded search uncovered hundreds of previously unexplored metastable structures. Subsequent re-evaluation of the energies of these structures via DFT geometric optimization identified configurations that are 2.6 kcal/mol more stable than those previously reported.

Additionally, DEP-containing clusters can form a wide range of metastable structures within a narrow energy range, primarily due to variations in chemical adsorption patterns. Notably, these structural variations occur without requiring substantial rearrangement of the $MgCl_2$ surface. The ability to sample a broader spectrum of stable structures provides crucial insights into the distribution of chemical environments surrounding the active center, emphasizing the significance of the explored parametric space.

Furthermore, an evaluation of the charge distribution in $TiCl_4$ across different ligand types revealed that chelate-adsorbed donors induced a more negative charge on the Ti species. The adsorption mode of DEP broadened the distribution of charge states, a discovery that provides an explanation for the wide molecular weight distribution of the resulting polymer.

The HDNNP approach described here represents a significant step toward realistic simulations of the $MgCl_2/TiCl_4$/DEP system and, more broadly, all $MgCl_2/TiCl_4$/internal donor systems.

REFERENCE

(1) Busico, V. Giulio Natta and the Development of Stereoselective Propene Polymerization. In *Polyolefins: 50 years after Ziegler and Natta I*; Kaminsky, W., Ed.; Advances in Polymer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; Vol. 257, pp 37–57. https://doi.org/10.1007/12_2013_213.

(2) Chammingkwan, P.; Terano, M.; Taniike, T. High-Throughput Synthesis of Support Materials for Olefin Polymerization Catalyst. *ACS Comb. Sci.* **2017**, *19* (5), 331–342. https://doi.org/10.1021/acscombsci.7b00010.

(3) Taniike, T.; Terano, M. The Use of Donors to Increase the Isotacticity of Polypropylene. In *Polyolefins: 50 years after Ziegler and Natta I*; Kaminsky, W., Ed.; Advances in Polymer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; Vol. 257, pp 81–97. https://doi.org/10.1007/12_2013_224.

(4) Soga, K. Ziegler-Natta Catalysts for Olefin Polymerizations. *Prog. Polym. Sci.* **1997**, *22* (7), 1503–1546. https://doi.org/10.1016/S0079-6700(97)00003-8.

(5) Ratanasak, M.; Rungrotmongkol, T.; Saengsawang, O.; Hannongbua, S.; Parasuk, V. Towards the Design of New Electron Donors for Ziegler–Natta Catalyzed Propylene Polymerization Using QSPR Modeling. *Polymer* **2015**, *56*, 340–345. https://doi.org/10.1016/j.polymer.2014.11.022.

(6) Dang, X.; Li, Q.; Li, H.; Yang, Y.; Zhang, L.; Hu, Y. Ziegler-Natta Catalysts with Novel Internal Electron Donors for Propylene Polymerization. *J. Polym. Res.* **2014**, *21* (12), 619. https://doi.org/10.1007/s10965-014-0619-9.

(7) Potapov, A. G.; Bukatov, G. D.; Zakharov, V. A. DRIFTS Study of the Interaction of the AlEt3 Cocatalyst with the Internal Donor Ethyl Benzoate in Supported

Ziegler–Natta Catalysts. *J. Mol. Catal. Chem.* **2009**, *301* (1–2), 18–23. https://doi.org/10.1016/j.molcata.2008.11.002.

(8) Mori, H.; Sawada, M.; Higuchi, T.; Hasebe, K.; Otsuka, N.; Terano, M. Direct Observation of MgCl2-Supported Ziegler Catalysts by High Resolution Transmission Electron Microscopy. *Macromol. Rapid Commun.* **1999**, *20* (5), 245–250. https://doi.org/10.1002/(SICI)1521-3927(19990501)20:5<245::AID-MARC245>3.0.CO;2-R.

(9) Zannetti, R.; Marega, C.; Marigo, A.; Martorana, A. Layer-lattices in Ziegler–Natta Catalysts. *J. Polym. Sci. Part B Polym. Phys.* **1988**, *26* (12), 2399–2412. https://doi.org/10.1002/polb.1988.090261202.

(10) D'Amore, M.; Thushara, K. S.; Piovano, A.; Causà, M.; Bordiga, S.; Groppo, E. Surface Investigation and Morphological Analysis of Structurally Disordered $MgCl_2$ and $MgCl_2/TiCl_4$ Ziegler-Natta Catalysts. *ACS Catal.* **2016**, *6* (9), 5786–5796. https://doi.org/10.1021/acscatal.6b00871.

(11) Taniike, T.; Chammingkwan, P.; Thang, V. Q.; Goto, K.; Fujitani, T.; Terano, M. Chemisorption-Induced Activation of $MgCl_2$ Film as Realistic Route for Heterogeneous Ziegler-Natta Surfaces under Ultrahigh Vacuum. *J. Phys. Chem. C* **2017**, *121* (43), 24085–24092. https://doi.org/10.1021/acs.jpcc.7b08242.

(12) Wada, T.; Takasao, G.; Terano, M.; Chammingkwan, P.; Taniike, T. Structure Determination of the δ-$MgCl_2$ Support in Ziegler-Natta Catalysts. *J. Jpn. Pet. Inst.* **2022**, *65* (3), 88–96. https://doi.org/10.1627/jpi.65.88.

(13) Piovano, A.; Signorile, M.; Braglia, L.; Torelli, P.; Martini, A.; Wada, T.; Takasao, G.; Taniike, T.; Groppo, E. Electronic Properties of Ti Sites in Ziegler-Natta

Catalysts. *ACS Catal.* **2021**, *11* (012), 9949–9961. https://doi.org/10.1021/acscatal.1c01735.

(14) D'Amore, M.; Piovano, A.; Vottero, E.; Piovano, A.; Rudić, S.; Erba, A.; Groppo, E.; Civalleri, B. Inelastic Neutron Scattering Investigation of $MgCl_2$ Nanoparticle-Based Ziegler-Natta Catalysts for Olefin Polymerization. *ACS Appl. Nano Mater.* **2020**, *3* (11), 11118–11128. https://doi.org/10.1021/acsanm.0c02296.

(15) Piovano, A.; D'Amore, M.; Wada, T.; Cleto Bruzzese, P.; Takasao, G.; Thakur, A.; Chammingkwan, P.; Terano, M.; Civalleri, B.; Bordiga, S.; Taniike, T.; Groppo, E. Revisiting the Identity of δ-$MgCl_2$: Part II. Morphology and Exposed Surfaces Studied by Vibrational Spectroscopies and DFT Calculation. *J. Catal.* **2020**, *387*, 1–11. https://doi.org/10.1016/j.jcat.2020.04.017.

(16) Aslan, M.; Davis, J. B. A.; Johnston, R. L. Global Optimization of Small Bimetallic Pd-Co Binary Nanoalloy Clusters: A Genetic Algorithm Approach at the DFT Level. *Phys. Chem. Chem. Phys.* **2016**, *18* (9), 6676–6682. https://doi.org/10.1039/c6cp00342g.

(17) Bao, K.; Goedecker, S.; Koga, K.; Lançon, F.; Neelov, A. Structure of Large Gold Clusters Obtained by Global Optimization Using the Minima Hopping Method. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2009**, *79* (4), 1–4. https://doi.org/10.1103/PhysRevB.79.041405.

(18) Hussein, H. A.; Davis, J. B. A.; Johnston, R. L. DFT Global Optimisation of Gas-Phase and MgO-Supported Sub-Nanometre AuPd Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18* (37), 26133–26143. https://doi.org/10.1039/c6cp03958h.

(19) Ganguly Neogi, S.; Chaudhury, P. Structural, Spectroscopic Aspects, and Electronic Properties of $(TiO_2)_n$ Clusters: A Study Based on the Use of Natural

Algorithms in Association with Quantum Chemical Methods. *J. Comput. Chem.* **2014**, *35* (1), 51–61. https://doi.org/10.1002/jcc.23465.

(20) Chikuma, H.; Takasao, G.; Wada, T.; Chammingkwan, P.; Behler, J.; Taniike, T. Accelerating Non-Empirical Structure Determination of Ziegler–Natta Catalysts with a High-Dimensional Neural Network Potential. *J. Phys. Chem. C* **2023**, *127* (24), 11683–11691. https://doi.org/10.1021/acs.jpcc.3c01511.

(21) Behler, J. Perspective: Machine Learning Potentials for Atomistic Simulations. *J. Chem. Phys.* **2016**, *145* (17), 170901. https://doi.org/10.1063/1.4966192.

(22) Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115* (16), 1032–1050. https://doi.org/10.1002/qua.24890.

(23) Ouyang, R.; Xie, Y.; Jiang, D. E. Global Minimization of Gold Clusters by Combining Neural Network Potentials and the Basin-Hopping Method. *Nanoscale* **2015**, *7* (36), 14817–14821. https://doi.org/10.1039/c5nr03903g.

(24) Lilienfeld, O. A. V.; Rupp, M.; Tkatchenko, A.; Mu, K. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. **2012**, *058301* (February), 1–5. https://doi.org/10.1103/PhysRevLett.108.058301.

(25) Bart, A. P. Gaussian Approximation Potentials : A Brief Tutorial Introduction arXiv : 1502 . 01366v2 [Cond-Mat . Mtrl-Sci] 5 Feb 2020. **2020**, 1–20.

(26) Seko, A.; Ishiwata, S. Prediction of Perovskite-Related Structures in $ACuO_{3-x}$ (A = Ca, Sr, Ba, Sc, Y, La) Using Density Functional Theory and Bayesian Optimization. *Phys. Rev. B* **2020**, *101* (13), 134101. https://doi.org/10.1103/PhysRevB.101.134101.

(27) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 1–4. https://doi.org/10.1103/PhysRevLett.98.146401.

(28) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Machine Learning-Aided Structure Determination for $TiCl_4$-Capped $MgCl_2$ Nanoplate of Heterogeneous Ziegler-Natta Catalyst. *ACS Catal.* **2019**, *9* (3), 2599–2609. https://doi.org/10.1021/acscatal.8b05080.

(29) Takasao, G.; Wada, T.; Thakur, A.; Chammingkwan, P.; Terano, M.; Taniike, T. Insight into Structural Distribution of Heterogeneous Ziegler–Natta Catalyst from Non-Empirical Structure Determination. *J. Catal.* **2021**, *394*, 299–306. https://doi.org/10.1016/j.jcat.2020.11.005.

(30) Takasao, G.; Wada, T.; Chikuma, H.; Chammingkwan, P.; Terano, M.; Taniike, T. Preventing Premature Convergence in Evolutionary Structure Determination of Complex Molecular Systems: Demonstration in Few-Nanometer-Sized $TiCl_4$-Capped $MgCl_2$ Nanoplates. *J. Phys. Chem. A* **2022**, *126* (31), 5215–5221. https://doi.org/10.1021/acs.jpca.2c02112.

(31) Da Silveira, J. M.; Chikuma, H.; Takasao, G.; Wada, T.; Chammingkwan, P.; Taniike, T. Deciphering the Role of Internal Donors in Shaping Heterogeneous Ziegler–Natta Catalysts Based on Nonempirical Structural Determination. *ACS Catal.* **2024**, *14* (4), 2300–2312. https://doi.org/10.1021/acscatal.3c05480.

(32) Piovano, A.; Wada, T.; Amodio, A.; Takasao, G.; Ikeda, T.; Zhu, D.; Terano, M.; Chammingkwan, P.; Groppo, E.; Taniike, T. Formation of Highly Active Ziegler-Natta Catalysts Clarified by a Multifaceted Characterization Approach. *ACS Catal.* **2021**, *11* (22), 13782–13796. https://doi.org/10.1021/acscatal.1c03067.

(33) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134* (7), 074106. https://doi.org/10.1063/1.3553717.

(34) Devereux, C.; Smith, J. S.; Huddleston, K. K.; Barros, K.; Zubatyuk, R.; Isayev, O.; Roitberg, A. E. Extending the Applicability of the ANI Deep Learning Molecular Potential to Sulfur and Halogens. *J. Chem. Theory Comput.* **2020**, *16* (7), 4192–4202. https://doi.org/10.1021/acs.jctc.0c00121.

(35) Behler, J. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chem. Int. Ed.* **2017**, *56* (42), 12828–12840. https://doi.org/10.1002/anie.201703114.

(36) Blank, T. B.; Brown, S. D. Adaptive, Global, Extended Kalman Filters for Training Feedforward Neural Networks. *J. Chemom.* **1994**, *8* (6), 391–407. https://doi.org/10.1002/cem.1180080605.

(37) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(38) Delley, B. An All-electron Numerical Method for Solving the Local Density Functional for Polyatomic Molecules. *J. Chem. Phys.* **1990**, *92* (1), 508–517. https://doi.org/10.1063/1.458452.

(39) Bergner, A.; Dolg, M.; Küchle, W.; Stoll, H.; Preuß, H. Ab Initio Energy-Adjusted Pseudopotentials for Elements of Groups 13–17. *Mol. Phys.* **1993**, *80* (6), 1431–1441. https://doi.org/10.1080/00268979300103121.

(40) Weinreich, J.; Römer, A.; Paleico, M. L.; Behler, J. Properties of α-Brass Nanoparticles I: Neural Network Potential Energy Surface. **2020**.

(41) Busico, V.; Cipullo, R.; Monaco, G.; Talarico, G.; Vacatello, M.; Chadwick, J. C.; Segre, A. L.; Sudmeijer, O. High-Resolution ^{13}C NMR Configurational Analysis of Polypropylene Made with $MgCl_2$-Supported Ziegler−Natta Catalysts. 1. The "Model" System $MgCl_2/TiCl_4$−2,6-Dimethylpyridine/$Al(C_2H_5)_3$. *Macromolecules* **1999**, *32* (13), 4173–4182. https://doi.org/10.1021/ma981941n.

(42) Ratanasak, M.; Parasuk, V. Roles of Malonate Donor on Activity and Stereoselectivity of Ziegler-Natta Catalyzed Propylene Polymerization. *J. Organomet. Chem.* **2015**, *775*, 6–11. https://doi.org/10.1016/j.jorganchem.2014.10.018.

(43) Guo, X.; Cui, L.; Wang, Y.; Yi, J.; Sun, J.; Liu, Z.; Liu, B. Mechanistic Study on Effect of Electron Donors in Propylene Polymerization Using the Ziegler–Natta Catalyst. *J. Phys. Chem. C* **2021**, *125* (16), 8533–8542. https://doi.org/10.1021/acs.jpcc.0c11273.

(44) Antinucci, G.; Vittoria, A.; Cipullo, R.; Busico, V. Regioirregular Monomeric Units in Ziegler-Natta Polypropylene: A Sensitive Probe of the Catalytic Sites. *Macromolecules* **2020**, *53* (10), 3789–3795. https://doi.org/10.1021/acs.macromol.0c00433.

(45) Wada, T.; Taniike, T.; Kouzai, I.; Takahashi, S.; Terano, M. Propylene Polymerization Performance of Isolated and Aggregated Ti Species Studied Using a Well-Designed $TiCl_4/MgCl_2$ Ziegler-Natta Model Catalyst. *Macromol. Rapid Commun.* **2009**, *30* (11), 887–891. https://doi.org/10.1002/marc.200900015.

(46) Cavallo, L.; Del Piero, S.; Ducéré, J.-M.; Fedele, R.; Melchior, A.; Morini, G.; Piemontesi, F.; Tolazzi, M. Key Interactions in Heterogeneous Ziegler−Natta Catalytic Systems: Structure and Energetics of $TiCl_4$ −Lewis Base Complexes. *J. Phys. Chem. C* **2007**, *111* (11), 4412–4419. https://doi.org/10.1021/jp0678204.

(47) Breuza, E.; Antinucci, G.; Budzelaar, P. H. M.; Busico, V.; Correa, A.; Ehm, C. $MgCl_2$ -Supported Ziegler–Natta Catalysts: A DFT-D "Flexible-Cluster" Approach to Internal Donor Adducts. *J. Phys. Chem. C* **2018**, *122* (16), 9046–9053. https://doi.org/10.1021/acs.jpcc.8b01500.

(48) Mori, H.; Hasebe, K.; Terano, M. XPS Study of the Interaction of Titanium Species with Internal Electron Donors on $MgCl_2$-Supported Ziegler Catalysts. *J. Mol. Catal. Chem.* **1999**, *140* (2), 165–172. https://doi.org/10.1016/S1381-1169(98)00225-8.

(49) Zhao, X.; Zhang, Y.; Song, Y.; Wei, G. XPS STUDY OF INTERACTION OF TITANIUM SPECIES WITH DIFFERENT INTERNAL ELECTRON DONORS ON Z – N CATALYSTS. *Surf. Rev. Lett.* **2007**, *14* (05), 951–955. https://doi.org/10.1142/S0218625X07010482.

Supporting information

Tables S5.1–5.6 show the parameter set for the atom-centered symmetry functions (ACSFs) introduced in this chapter. The parameters differ in the element set, so these are indicated separately in each table. It was adjusted empirically based on the results obtained in Chapter 2 and other literature. Those with a maximum value of less than 0.2 in the training set were deleted. For the angular ACSF of H, the triad function that includes Ti, Cl, and Mg is not used. Also, the Ti-H-H, Mg-H-H, and Cl-H-H functions were not used for angular ACSF of Ti, Cl, and Mg.

Table S5.1. Parameters of the radial ACSFs. The same parameter set was applied to describe diad in set of H, C, and O.

N	η (1/Bohr^2)	R_s (Bohr)	R_c (Bohr)
1	5.517	1.512	9.638
2	5.517	2.02	9.638
3	5.517	2.528	9.638
4	5.517	3.055	9.638
5	5.517	3.543	9.638
6	5.517	4.051	9.638
7	5.517	4.559	9.638
8	5.517	5.067	9.638
9	5.517	5.575	9.638
10	5.517	6.083	9.638

11	5.517	6.590	9.638
12	5.517	7.098	9.638
13	5.517	8.114	9.638
14	5.517	8.622	9.638

Table S5.2. Parameters of the angular ACSFs. The same parameter set was applied to describe triad in set of H, C, and O. The function type was G_i^{ang1} in Eq.5.2.

N	η (1/Bohr^2)	λ	ζ	R_c (Bohr)
1	0.000	1.0	1.0	6.614
2	0.000	1.0	2.0	6.614
3	0.000	1.0	4.0	6.614
4	0.000	1.0	16.0	6.614
5	0.000	-1.0	1.0	6.614
6	0.000	-1.0	2.0	6.614
7	0.000	-1.0	4.0	6.614
8	0.000	-1.0	16.0	6.614

Table S5.3. Parameters of the radial ACSFs. The same parameter set was applied to describe diad containing both set of H, C, O and set of Mg, Cl, Ti.

N	η (1/Bohr^2)	R_s (Bohr)	R_c (Bohr)
1	0.002	0.0	12.00
2	0.005	0.0	12.00
3	0.009	0.0	12.00
4	0.016	0.0	12.00
6	0.029	0.0	12.00

Table S5.4. Parameters of the radial ACSFs. The same parameter set was applied to describe triad containing both set of H, C, O and set of Mg, Cl, Ti. The function type was G_i^{ang2} in Eq.5.3.

N	η (1/Bohr^2)	λ	ζ	R_c (Bohr)
1	0.000	1.0	1.0	12.0
2	0.000	1.0	2.0	12.0
3	0.000	1.0	8.0	12.0
4	0.000	-1.0	1.0	12.0
5	0.000	-1.0	2.0	12.0
6	0.000	-1.0	4.0	12.0

Table S5.5. Parameters of the radial ACSFs. The same parameter set was applied to describe diad in set of Mg, Cl, Ti.

N	η (1/Bohr^2)	R_s (Bohr)	R_c (Bohr)
1	0.000	0.0	16.0
2	0.002	0.0	16.0
3	0.005	0.0	16.0
4	0.009	0.0	16.0
5	0.016	0.0	16.0
6	0.029	0.0	16.0
7	0.050	3.0	16.0
8	0.050	6.0	16.0
9	0.050	10.5	16.0

Table S5.6. Parameters of the angular ACSFs. The same parameter set was applied to describe triad in set of Mg, Cl, Ti. The function type was G_i^{ang2} in Eq.5.3.

N	η (1/Bohr^2)	λ	ζ	R_c (Bohr)
1	0.001	1.0	1.0	16.0
2	0.001	1.0	2.0	16.0
3	0.001	1.0	4.0	16.0
4	0.001	1.0	16.0	16.0
5	0.001	-1.0	1.0	16.0
6	0.001	-1.0	2.0	16.0
7	0.001	-1.0	4.0	16.0
8	0.001	-1.0	16.0	16.0

Chapter 6

General conclusion

The Ziegler-Natta (ZN) catalyst, widely utilized in olefin polymerization, is a heterogeneous catalyst whose primary structural unit (primary particle) comprises $MgCl_2$ nanoplates chemically modified with $TiCl_4$, the precursor of the active site, and an internal donor (ID) to enhance performance. Surface reconstruction induced by chemical adsorption introduces significant structural heterogeneity, posing a major challenge in designing accurate molecular models. Empirical structure determination of primary particles, accounting for such reconstructions, has been achieved by combining local optimization based on density functional theory (DFT) with global searches using genetic algorithms (GA). However, the computational cost of DFT has restricted these studies to small scales, making it challenging to model structures at a more realistic scale.

To explore catalyst structures that more closely resemble the size and composition of actual catalysts, it is crucial to accelerate the modeling process. Here I employed the high-dimensional neural network potential (HDNNP) approach, which significantly speeds up calculations while preserving high accuracy. Moreover, to minimize the number of DFT calculations required for training, efficient sampling methods are essential to construct accurate potentials using a minimal dataset. This study highlights a practical application of machine learning potentials in complex material systems.

In **Chapter 3**, HDNNPs were developed for the $MgCl_2/TiCl_4$ binary system, enabling non-empirical structural determination. By employing farthest-point sampling (FPS) and adaptive structure addition, efficient sampling was achieved, covering the parameter space with fewer structures. This strategy effectively identifies crucial training data from the vast potential energy surface (PES). Accelerated structure searches using HDNNPs successfully determined the structure of a $50MgCl_2/9TiCl_4$ cluster,

representing realistic sizes and coverage. The analysis revealed morphological features consistent with experimental observations, such as a round shape with numerous edges and an increase in sterically specific active sites with higher Ti coverage.

Chapter 4 investigates the effect of $TiCl_4$ coverage through realistic-scale simulations. Systematic structural determination was conducted for $MgCl_2/TiCl_4$ systems with varying coverage. HDNNPs applicable to diverse compositions were developed, allowing for structural determination across a wide compositional range. A comprehensive reference dataset was generated using FPS, and the robustness of HDNNP predictions was carefully validated on unknown systems. Structure determination for $MgCl_2/TiCl_4$ cluster with different coverage unveiled systematic changes in the steric and electronic properties of active sites due to coverage variations. The findings clarified the experimental observation that increasing Ti species on the $MgCl_2$ surface does not always enhance activity. Furthermore, charge distribution analyses emphasized the importance of Ti–Ti interactions, underscoring the need for simulations using realistic cluster models.

In **Chapter 5**, HDNNPs were developed for ZN catalysts containing internal donors, focusing on propylene polymerization. Using diethyl phthalate (DEP) as the internal donor, accelerated HDNNP calculations were applied to the $19MgCl_2/4TiCl_4/5DEP$ system, enabling exploration of a broader structural space. This facilitated the reliable determination of the most stable structures and provided deeper insights into the structural distribution of active sites.

This study successfully modeled Ziegler-Natta catalysts in a manner consistent with experimental morphological observations. The results from each chapter underscore the importance of simulating realistic scales and sampling diverse chemical environments

for accurate modeling of complex material systems. The accelerated HDNNP approach offers a powerful tool for advancing the understanding of catalytic systems at practical scales.

List of Publications and Other Achievements

Hiroki Chikuma

A) PABLICATION

1. Accelerating Non-Empirical Structure Determination of Ziegler–Natta Catalysts with a High-Dimensional Neural Network Potential, **Hiroki Chikuma**, Gentoku Takasao, Toru Wada, Patchanee Chammingkwan, Jörg Behler, Toshiaki Taniike, *The Journal of Physical Chemistry C* 2023, *127*, 11683–11691.

2. Preventing Premature Convergence in Evolutionary Structure Determination of Complex Molecular Systems: Demonstration in Few-Nanometer-Sized $TiCl_4$-Capped $MgCl_2$ Nanoplates, Gentoku Takasao, Toru Wada, **Hiroki Chikuma**, Patchanee Chammingkwan, Minoru Terano, Toshiaki Taniike, *The Journal of Physical Chemistry A* 2022, *126*, 5215-5221.

3. Deciphering the Role of Internal Donors in Shaping Heterogeneous Ziegler-Natta Catalysts Based on Nonempirical Structural Determination, João da Silveira, **Hiroki Chikuma**, Gentoku Takasao, Toru Wada, Patchanee Chammingkwan, Toshiaki Taniike, *ACS Catalysis* 2024, *14*, 2300–2312.

B) INTERNATIONAL CONFERENCE

1. Accelerating structure determination of $MgCl_2/TiCl_4$ nanoplate for Ziegler-Natta catalysts with a highdimensional neural network potential, **Hiroki Chikuma**, Gentoku Takasao, Jörg Behler, Toshiaki Taniike, First Combined Blue Sky - INCOREP Polyolefin Conference, Sorrento, Italy, Jun. 12–16, 2023, poster.

2. Efficient Structure Determination of Ziegler-Natta Catalysts with a High-Dimensional Neural Network Potential and Genetic Algorithm, **Hiroki Chikuma**, Gentoku Takasao, Jörg Behler, Toshiaki Taniike, The 13th SPSJ International Polymer Conference (IPC2023), Sapporo, Japan, Jul. 18–21, 2023, oral.

C) DOMESTIC CONFERENCE

1. 機械学習ポテンシャルによるZiegler-Natta触媒ナノ構造の非経験的構造決定, **筑間 弘樹**, 高棹 玄徳, Behler Jörg, 谷池 俊明, 第52回石油・石油化学討論会, 長野, 2022年10月27–28日, 一般口頭.