

# LEARNING MOLECULAR CHIRALITY VIA CHIRAL DE-TERMINANT KERNELS

**Runhan Shi<sup>1</sup>, Zhicheng Zhang<sup>1</sup>, Letian Chen<sup>1,2</sup>, Gufeng Yu<sup>1</sup>, Yang Yang<sup>1\*</sup>**

<sup>1</sup>AGI Institute, School of Computer Science, Shanghai Jiao Tong University

<sup>2</sup>Shanghai Innovation Institute

{han.run.jiangming, zhicheng.zhang, clt2001, jm5820zz}@sjtu.edu.cn  
yangyang@cs.sjtu.edu.cn

## ABSTRACT

Chirality is a fundamental molecular property that governs stereospecific behavior in chemistry and biology. Capturing chirality in machine learning models remains challenging due to the geometric complexity of stereochemical relationships and the limitations of traditional molecular representations that often lack explicit stereochemical encoding. Existing approaches to chiral molecular representation primarily focus on central chirality, relying on handcrafted stereochemical tags or limited 3D encodings, and thus fail to generalize to more complex forms such as axial chirality. In this work, we introduce **ChiDeK** (**Chiral Determinant Kernels**), a framework that systematically integrates stereogenic information into molecular representation learning. We propose the chiral determinant kernel to encode the SE(3)-invariant chirality matrix and employ cross-attention to integrate stereochemical information from local chiral centers into the global molecular representation. This design enables explicit modeling of chiral-related features within a unified architecture, capable of jointly encoding central and axial chirality. To support the evaluation of axial chirality, we construct a new benchmark for electronic circular dichroism (ECD) and optical rotation (OR) prediction. Across four tasks, including R/S configuration classification, enantiomer ranking, ECD spectrum prediction, and OR prediction, ChiDeK achieves substantial improvements over state-of-the-art baselines, most notably yielding over 7% higher accuracy on axially chiral tasks on average.

## 1 INTRODUCTION

Chirality, the property of molecules that are non-superimposable mirror images, is a fundamental concept in chemistry, biology, and materials science (Moss, 1996). It plays a central role in shaping molecular behavior across diverse applications, including biomolecular recognition (Ceramella et al., 2022), stereoselective synthesis (Gaucherand et al., 2024), and enantioselective catalysis (He et al., 2023). Chiral molecules, often existing as pairs of enantiomers, can exhibit drastically different chemical and biological properties despite sharing identical compositions (Peluso & Chankvetadze, 2022). Such differences are especially critical in drug efficacy (Liu et al., 2023b), toxicity (McVicker & O’Boyle, 2024), and protein-binding affinity (Schneider et al., 2018; Shen et al., 2013), underscoring the paramount importance of accurately modeling stereochemical information in molecular representation learning.

Despite this significance, capturing chirality in machine learning models remains challenging. Existing molecular representation learning (MRL) approaches predominantly focus on central chirality and often rely on implicit encoding strategies that fail to extract explicit stereochemical features. While recent advances in SE(3)-invariant architectures have shown promise through the incorporation of torsion angles (Klicpera et al., 2021; Liu et al., 2022) and explicitly chirality-aware designs (Adams et al., 2022; Yan et al., 2025), research on comprehensive chirality modeling within MRL remains limited. More complex stereogenic forms, such as axial, planar, and helical chirality, are largely unaddressed, representing a significant gap in current methodologies.

---

\*Corresponding author.

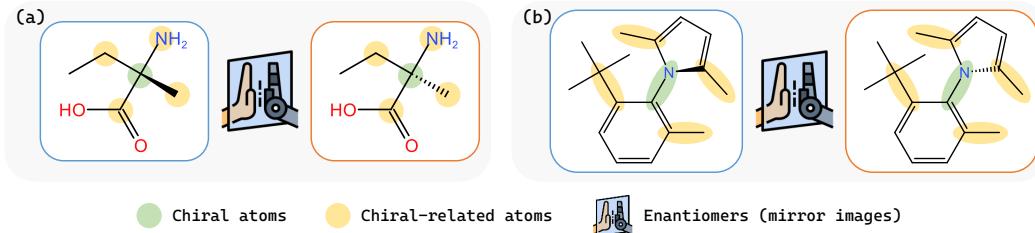


Figure 1: Examples of central chirality (a) and axial chirality (b). The configuration of the chiral atoms (in green) is determined by the spatial arrangement of chiral-related atoms (in orange).

Among the various forms of molecular chirality, central chirality and axial chirality represent the two most prevalent categories. **Central chirality** arises when a tetrahedral atom (typically carbon) is bonded to four distinct substituents. Such an atom constitutes a *chiral center*, and its stereochemical configuration (R or S) (Cahn et al., 1966; Favre & Powell, 2013) is determined by the spatial arrangement of its substituents, as illustrated in Figure 1(a). **Axial chirality** typically results from restricted rotation around a chemical axis, often a bond connecting two aromatic or aliphatic groups. Here, chirality (Ra/M or Sa/P) is determined by the relative arrangement of substituent groups around the stereogenic axis, as shown in Figure 1(b).

Existing computational approaches encounter fundamental limitations in capturing chirality. E(3)-invariant architectures that rely exclusively on pairwise distances or bond angles, such as SchNet (Schütt et al., 2017) and DimeNet (Gasteiger et al., 2020b;a), are inherently incapable of distinguishing enantiomers since these geometric quantities remain unchanged under reflection. While SE(3)-invariant models, including SphereNet (Liu et al., 2022), ChiRo (Adams et al., 2022), and ChiGNN (Yan et al., 2025), address this limitation through directional information and equivariant features, they face critical challenges. Most notably, molecular graphs are typically treated as homogeneous structures where chiral atoms are not explicitly differentiated, causing stereochemical signals to be diluted during global aggregation. Recent approaches like ChiGNN attempt to incorporate chirality through permutation-resolution strategies, but they still fail to extract expressive, localized descriptors that faithfully characterize chiral centers or stereogenic axes. These constraints highlight the urgent need for a unified framework capable of systematically modeling diverse chirality types while explicitly encoding stereochemical features that govern molecular behavior.

To address the above limitations, we propose the **ChiDeK** model (**Chiral Determinant Kernels**), which learns molecular chirality based on the determinant of the chirality matrix and applies cross-attention between chiral and non-chiral atoms to explicitly capture stereogenic information. The introduced chiral determinant kernels embed the SE(3)-invariant chirality matrix (Shi et al., 2026) for each chiral atom, capturing stereogenic information for both central and axial configurations. A cross-attention mechanism enhanced by pair representations then propagates the influence of chiral atoms (queries) with chiral-related and non-chiral atoms (keys and values) into the global molecular representation, enhancing stereochemical sensitivity. Our main contributions are:

- We introduce ChiDeK, a unified architecture that systematically encodes both central and axial chirality, and evaluate it comprehensively on multiple chirality-aware prediction tasks.
- We design the chiral determinant kernel for chiral atoms that embeds the SE(3)-invariant chirality matrix, effectively capturing stereogenic features for chirality.
- We construct and release a dataset of axially chiral molecules for the prediction of electronic circular dichroism (ECD) and optical rotation (OR), providing a benchmark for this underexplored stereogenic type.
- Across multiple tasks, including R/S classification, enantiomer ranking, ECD prediction, and OR prediction, ChiDeK demonstrates substantial improvements over state-of-the-art baselines, particularly for the axial chirality task.

## 2 RELATED WORK

**Learning 3D Representations of Molecules.** Recent advances in molecular machine learning have heavily leveraged 3D geometric message passing and attention architectures, where interatomic interactions are modeled via internal coordinates such as distances, angles, and torsions. SchNet (Schütt et al., 2017), a pioneering architecture for predicting quantum mechanical properties, parameterizes messages using radial basis expansions of pairwise distances. DimeNet (Gasteiger et al., 2020b) and DimeNet++ (Gasteiger et al., 2020a) extend this framework by incorporating angular features through spherical Bessel functions, while DeepH-E3 (Gong et al., 2023) learns Hamiltonians from density functional theory (DFT) calculations. BOTNet (Batatia et al., 2025) further simplifies neural equivariant interatomic potentials (NequIP) (Batzner et al., 2022), a high-performing equivariant message passing framework. These models are designed to be  $E(3)$ -invariant, which are robust to translations, rotations, and reflections. While this invariance is desirable for many property prediction tasks, it enforces reflection-symmetry, rendering such models unable to distinguish enantiomers that differ only by mirror inversion (Satorras et al., 2021; Dumitrescu et al., 2025).

To remedy this shortcoming,  $SE(3)$ -invariant architectures have been developed, which maintain sensitivity to reflections. Tensor Field Networks (Thomas et al., 2018) and  $SE(3)$ -Transformers (Fuchs et al., 2020) achieve equivariance through the use of spherical harmonics and irreducible representations. Path-MPNN (Flam-Shepherd et al., 2021), GemNet (Klicpera et al., 2021), and SphereNet (Liu et al., 2022) further enhance 3D GNNs by embedding angular and torsional information. More recently, MoleculeSED (Liu et al., 2023a) introduces reflection-sensitive stochastic differential equations for molecular modeling.

**Explicit Representation Learning of Molecular Chirality.** A number of studies explicitly incorporate stereochemical information into molecular representations. Early efforts introduce hand-crafted descriptors such as the chirality code (Aires-de Sousa et al., 2004), physicochemical atomic stereo (PAS) descriptors (Zhang & Aires-de Sousa, 2006), relative chirality indices (RCI) (Natarajan et al., 2007), and chiral cliffs (Schneider et al., 2018). More recently, MAP4C (Orsi & Reymond, 2024) extends molecular fingerprints to capture stereochemical tags directly. Linear notations such as SMILES (Weininger, 1988), SELFIES (Krenn et al., 2022), and fragSMILES (Mastrolorito et al., 2025) encode molecules as sequences that can include chirality tags, enabling models like Transformers (Vaswani et al., 2017) to process stereochemical information (Yoshikai et al., 2024; Yang et al., 2025). However, these tags serve only as symbolic indicators without explicit 3D structural context, making the resulting representations less informative for capturing stereogenic geometry. Graph neural networks have also been adapted to represent chirality more explicitly. Tetra-DMPNN (Pattanaik et al., 2020) modifies message passing in 2D GNNs with asymmetric updates to encode tetrahedral chirality. SPMS (Xu et al., 2021) introduces spherical projection descriptors with convolutional architectures for stereostructure modeling. ChiRo (Adams et al., 2022) learns 3D representations invariant to bond rotations yet sensitive to stereoisomers. MolKGNN (Liu et al., 2023c) leverages the tetrahedron volume calculation to capture the neighbor ordering for chirality. CFFN (Du et al., 2023) combines 2D graph topology and 3D geometry to build stereochemistry-aware molecular representations. GCPNet (Morehead & Cheng, 2024) develops geometric representations tailored for biomolecular graphs. ChiENN (Gaiński et al., 2023) and ChiGNN (Yan et al., 2025) explicitly consider atom permutations related to chirality to improve stereochemical discrimination. ChiralCat (Peng et al., 2025) integrates 3D molecular structures with LLM-generated textual descriptions to perform chiral type classification.

Despite these advances, the vast majority of prior work focuses primarily on central chirality. Other stereogenic forms, such as axial chirality, remain largely underexplored, limiting current models’ ability to generalize across the full spectrum of stereochemistry. Although methods such as SPMS are theoretically capable of representing additional forms of chirality, they have not been systematically evaluated in these settings, leaving their practical effectiveness uncertain.

## 3 PRELIMINARIES

### 3.1 CHIRAL AND CHIRAL-RELATED ATOMS

To construct a meaningful molecular representation that explicitly encodes chirality, we distinguish between *chiral atoms* and their corresponding *chiral-related atoms* (Figure 1). Given a molecule  $z$

with  $M$  atoms, where the index set of atoms is  $\mathcal{I} = \{1, \dots, M\}$ , we define three disjoint subsets :

$$\mathcal{I}_c, \quad \mathcal{I}_r = \bigcup_{i \in \mathcal{I}_c} \{j : j \in \mathcal{N}(i)\} \setminus \mathcal{I}_c, \quad \mathcal{I}_n = \mathcal{I} \setminus (\mathcal{I}_c \cup \mathcal{I}_r), \quad (1)$$

where  $\mathcal{I}_c$ ,  $\mathcal{I}_r$ , and  $\mathcal{I}_n \subset \mathcal{I}$  are the index sets of chiral, chiral-related, and non-chiral atoms, respectively, and  $\mathcal{N}(i)$  denotes the index set of substituents around a chiral atom  $i$ .

For central chirality, a chiral atom  $i \in \mathcal{I}_c$  typically has four different substituent groups. The corresponding four neighboring atoms define its chiral-related set  $\mathcal{N}(i)$ . For axial chirality, chirality arises along a stereogenic axis (e.g., in biaryls), where steric hindrance prevents free rotation. In this case, the substituent groups are partitioned into four distinct sets around the axis, and we select the nearest atom from each group as the chiral-related atoms.

### 3.2 CHIRALITY MATRIX

To explicitly encode stereogenic information, we adopt the chirality matrix introduced in ChiralFinder (Shi et al., 2026). Formally, for a chiral atom  $i \in \mathcal{I}_c$  with ordered substituents  $\mathcal{N}(i) = (r_{i,1}, r_{i,2}, r_{i,3}, r_{i,4})$ , the *chirality matrix* is defined as

$$\mathbf{M}_C(i) := ((\mathbf{x}_{r_{i,1}} - \mathbf{x}_i) \ (\mathbf{x}_{r_{i,2}} - \mathbf{x}_i) \ (\mathbf{x}_{r_{i,4}} - \mathbf{x}_{r_{i,3}}))^\top, \quad (2)$$

where  $\mathbf{x}_i \in \mathbb{R}^3$  denotes the 3D position of the chiral atom  $i$ , and  $\mathbf{x}_{r_{i,j}}$  denotes the position of its  $j$ -th substituent. The corresponding *chirality product* is then defined as

$$P_C(i) := ((\mathbf{x}_{r_{i,1}} - \mathbf{x}_i) \times (\mathbf{x}_{r_{i,2}} - \mathbf{x}_i)) \cdot (\mathbf{x}_{r_{i,4}} - \mathbf{x}_{r_{i,3}}) = \det(\mathbf{M}_C(i)), \quad (3)$$

where  $\times$  denotes the vector cross product,  $\cdot$  denotes the dot product, and  $\det(\cdot)$  is the determinant of a given matrix. Appendix D.1 provides details on how  $\mathcal{N}(i)$  is determined.

**Proposition 3.1.** *Given a chiral atom  $i$ , the chirality product  $P_C(i)$  is invariant under rigid-body translation and rotation, and changes sign under reflection:*

$$\det(\mathbf{M}_C(i)) = \det(R_1 \mathbf{M}_C(i)), \quad \det(R_2 \mathbf{M}_C(i)) = -\det(\mathbf{M}_C(i)), \quad (4)$$

where  $R_1 \in \text{SE}(3)$  denotes any rigid-body motion with rotation  $r \in \text{SO}(3)$  and translation  $t \in \mathbb{R}^3$ , and  $R_2 \in O^-(3)$  denotes any reflection (orthogonal transformation with determinant  $-1$ ).

**Lemma 3.1.** *Let  $(\mathbf{z}_1, \mathbf{z}_2)$  be a pair of enantiomers differing only in the stereochemical configuration of atom  $i$ . Then*

$$\text{configuration}(i) = \begin{cases} R, & \det(\mathbf{M}_C(i)) > 0, \\ S, & \det(\mathbf{M}_C(i)) < 0. \end{cases} \quad (5)$$

Proof of Lemma 3.1 is provided in Appendix F.1.

## 4 METHOD

### 4.1 CHIDEK FRAMEWORK

The proposed ChiDeK framework, illustrated in Figure 2(a), is designed to capture stereochemical information by explicitly modeling chiral atoms through chiral determinant kernels. Given a molecule  $\mathbf{z} = (\mathbf{X}, \mathbf{H})$ , where  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_M) \in \mathbb{R}^{3M}$  specifies the 3D coordinates of the  $M$  atoms, and  $\mathbf{H} = (\mathbf{h}_1, \dots, \mathbf{h}_M) \in \mathbb{R}^{dM}$  encodes atom-level features, we construct a molecular graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where the node set  $\mathcal{V}$  is partitioned into chiral atoms, chiral-related atoms, and non-chiral atoms, and the edge set  $\mathcal{E}$  incorporates geometric distances between atoms.

The forward pass of ChiDeK proceeds in three stages: (1) a *chiral encoder* computes chiral-sensitive embeddings for chiral atoms via chiral determinant kernels, while three separate multi-layer projectors compute chiral-invariant embeddings for chiral, chiral-related, and non-chiral atoms; (2) the multi-layer *chiral cross-attention* mechanism refines embeddings of chiral atoms (queries) by attending to heterogeneous neighbors (keys/values), guided by distance-aware Gaussian Kernel with Pair Type (GKPT) biases; (3) the resulting chiral-aware representations are aggregated and passed through task-specific heads (e.g., chirality classification, spectrum regression).

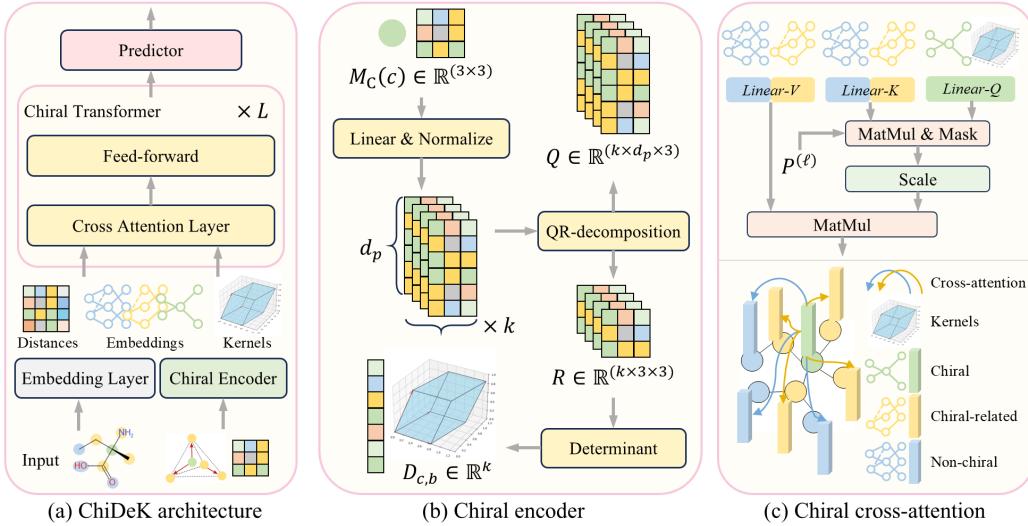


Figure 2: Overview of the ChiDeK architecture (a). It consists of a chiral encoder (b), a chiral transformer incorporating  $L$  cross attention layers (c), and a predictor for predicting chiral properties.

## 4.2 CHIRAL ENCODER

A key requirement for learning molecular chirality is to design features that are sensitive to reflections but invariant to global translations and rotations. The classical determinant of a  $3 \times 3$  chirality matrix  $M_C(i)$  provides exactly such a property, as it encodes the signed volume spanned by three bond vectors around a chiral atom. To generalize this construction into a high-dimensional latent space, we introduce the chiral encoder based on chiral determinant kernels.

**Chiral Determinant Kernel.** To explicitly represent chiral atoms with stereogenic information, we introduce the chiral determinant kernel, a determinant-based kernel that encodes the oriented volume spanned by triplets of projected atomic vectors with QR decomposition.

Let  $M_C \in \mathbb{R}^{B \times 3 \times 3}$  denote the batch of chirality matrices and  $\mathbf{W} \in \mathbb{R}^{k \times d_p \times 3}$  the learnable weights, where  $k$  is the number of kernels and  $d_p$  is the projection dimension. After broadcasting to  $M'_C \in \mathbb{R}^{B \times k \times 3 \times 3}$  and  $\mathbf{W}' \in \mathbb{R}^{B \times k \times d_p \times 3}$ , we compute

$$\mathbf{O}_{b,k} = \mathbf{W}'_k M'_{C,b}, \quad \mathbf{O} \in \mathbb{R}^{B \times k \times d_p \times 3}, \quad (6)$$

where  $M'_{C,b} \in \mathbb{R}^{3 \times 3}$  is the  $b$ -th chirality matrix,  $\mathbf{W}'_k \in \mathbb{R}^{d_p \times 3}$  is the  $k$ -th kernel, and multiplication is standard matrix multiplication along the last two dimensions. Then, we reshape the output  $\mathbf{O}$  from  $\mathbb{R}^{B \times k \times d_p \times 3}$  to  $\mathbf{O}' \in \mathbb{R}^{(BK) \times d_p \times 3}$ . Applying layer normalization along the  $d_p$  dimension yields

$$\tilde{\mathbf{O}} = \text{LayerNorm}(\mathbf{O}') \in \mathbb{R}^{(BK) \times d_p \times 3}. \quad (7)$$

Next, we perform a QR decomposition for each slice:

$$\tilde{\mathbf{O}}_{(bk)} = \mathbf{Q}_{(bk)} \mathbf{R}_{(bk)}, \quad \mathbf{Q}_{(bk)}^\top \mathbf{Q}_{(bk)} = \mathbf{I}_3, \quad \mathbf{Q}_{(bk)} \in \mathbb{R}^{d_p \times 3}, \quad \mathbf{R}_{(bk)} \in \mathbb{R}^{3 \times 3}, \quad (8)$$

for all  $b \in \{1, \dots, B\}$  and  $k \in \{1, \dots, k\}$ . We then compute the determinant of each factor:

$$\mathbf{D}_{c,b,k} = \det(\mathbf{R}_{(bk)}), \quad \mathbf{D}_c \in \mathbb{R}^{B \times k}. \quad (9)$$

Thus, each chiral atom is embedded into a  $k$ -dimensional representation that generalizes the determinant, provides differentiability (Roberts & Roberts, 2020), and explicitly encodes stereochemistry. Then the initial hidden representations of chiral, chiral-related, and non-chiral atoms are

$$\mathbf{H}_c = \mathbf{D}_c + f_c(\mathbf{H}_{\mathcal{I}_c}), \quad \mathbf{H}_r = f_r(\mathbf{H}_{\mathcal{I}_r}), \quad \mathbf{H}_n = f_n(\mathbf{H}_{\mathcal{I}_n}), \quad (10)$$

where  $f_c, f_r, f_n$  are embedding layers for atomic features. We finally prepend a learnable global chiral token before chiral atoms to enhance chiral representation.

### 4.3 CHIRAL TRANSFORMER

**Gaussian kernel with pair type (GKPT).** Distances between atoms provide crucial stereochemical cues. To encode them, we adopt GKPT (Shuaibi et al., 2021; Zhou et al., 2023). Formally, for input distance  $x$  and pair type  $e$ , the GKPT is defined as

$$\text{GKPT}((x, e), \mu, \sigma) = \mathcal{G}(\mathbf{E}_1(e) \cdot x + \mathbf{E}_2(e), \mu, \sigma), \quad (11)$$

where  $\mathcal{G}(x', \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x' - \mu}{\sigma}\right)^2\right)$ ,  $\mathbf{E}_1, \mathbf{E}_2 \in \mathbb{R}^{N_e \times k}$  are learnable embeddings,  $N_e$  is the number of pair types, and  $\mu, \sigma \in \mathbb{R}^k$  parameterize the Gaussian kernels. Here, we distinguish between two types of edges: chiral atoms with chiral-related atoms and with non-chiral atoms. The GKPT module is applied to geometric distance  $x_{ij}$  with edge type  $e_{ij}$ , between atom  $i \in \mathcal{I}_c$  and  $j \in \mathcal{I}_r \cup \mathcal{I}_n$ , producing the bias term

$$p_{ij}^{(0)} = \text{GKPT}((x_{ij}, e_{ij}), \mu, \sigma) \mathbf{w}_p, \quad (12)$$

with learnable projection  $\mathbf{w}_p$ . The resulting pairwise bias matrix is

$$\mathbf{P}^{(0)} = [p_{ij}^{(0)}]_{1 \leq i \leq |\mathcal{I}_c|, 1 \leq j \leq |\mathcal{I}_r| + |\mathcal{I}_n|} \in \mathbb{R}^{|\mathcal{I}_c| \times (|\mathcal{I}_r| + |\mathcal{I}_n|)}. \quad (13)$$

This pairwise representation encodes chemically relevant interactions, informing the subsequent attention computation and enabling the model to learn complex spatial dependencies.

**Chiral Cross-attention.** We introduce a cross-attention mechanism enhanced by pairwise representations, where chiral atoms serve as queries, while chiral-related and non-chiral atoms provide keys and values. Let  $\mathbf{H}_c$ ,  $\mathbf{H}_r$ , and  $\mathbf{H}_n$  denote the hidden representations of chiral atoms  $\mathcal{I}_c$ , chiral-related atoms  $\mathcal{I}_r$ , and non-chiral atoms  $\mathcal{I}_n$ , respectively. At each layer  $\ell$  (totally  $L$  layers), the linear projections for keys and values are separated according to atom types and defined as

$$\mathbf{K}_r^{(\ell)} = \mathbf{H}_r W_{K_r}^{(\ell)\top}, \quad \mathbf{V}_r^{(\ell)} = \mathbf{H}_r W_{V_r}^{(\ell)\top}, \quad \mathbf{K}_n^{(\ell)} = \mathbf{H}_n W_{K_n}^{(\ell)\top}, \quad \mathbf{V}_n^{(\ell)} = \mathbf{H}_n W_{V_n}^{(\ell)\top}, \quad (14)$$

where  $W_{K_r}^{(\ell)}, W_{V_r}^{(\ell)}, W_{K_n}^{(\ell)}$ , and  $W_{V_n}^{(\ell)} \in \mathbb{R}^{k \times k}$  are learnable projection matrices. Let  $\mathbf{H}_c^{(\ell)}$  denotes the output at layer  $\ell$ , the resulting queries, keys, and values are

$$\mathbf{Q}^{(\ell)} = \mathbf{H}_c^{(\ell-1)} W_Q^{(\ell)\top}, \quad \mathbf{K}^{(\ell)} = [\mathbf{K}_r^{(\ell)}, \mathbf{K}_n^{(\ell)}], \quad \mathbf{V}^{(\ell)} = [\mathbf{V}_r^{(\ell)}, \mathbf{V}_n^{(\ell)}], \quad (15)$$

where  $W_Q^{(\ell)} \in \mathbb{R}^{k \times k}$  is the learnable query projection,  $\mathbf{H}_c^{(0)} = \mathbf{H}_c$ , and  $[\cdot, \cdot]$  denotes concatenation.

At each layer  $\ell$ , we maintain a learnable pairwise bias  $p_{ij}^{(\ell)}$  between atom  $i$  (chiral) and atom  $j$  (chiral-related or non-chiral), updated via the query-key interaction. See Appendix C for details.

**Predictor.** We utilize a lightweight predictor to predict chirality properties. It consists of two linear layers with GELU activation Hendrycks & Gimpel (2016). This simple yet effective architecture maps the final global representation to the target prediction space.

### 4.4 THEORETICAL ANALYSIS FOR CHIRAL ENCODER

The chirality product obtained from the chirality matrix (Shi et al., 2026) exhibits inherent sensitivity to stereochemical variation. We demonstrate that a chiral determinant kernel maintains this desirable characteristic.

**Lemma 4.1.** *Let  $\mathbf{M}_C(i) \in \mathbb{R}^{3 \times 3}$  be the chirality matrix of atom  $i$ , and let  $\mathbf{W} \in \mathbb{R}^{d_p \times 3}$  be a full column-rank projection. Define QR decomposition of  $\mathbf{O} = \mathbf{W} \mathbf{M}_C(i) \in \mathbb{R}^{d_p \times 3}$  as*

$$\mathbf{O} = \mathbf{Q}\mathbf{R}, \quad \mathbf{Q}^\top \mathbf{Q} = \mathbf{I}_3, \quad \mathbf{Q} \in \mathbb{R}^{d_p \times 3}, \quad \mathbf{R} \in \mathbb{R}^{3 \times 3}. \quad (16)$$

*Then the generalized chirality product  $P_{\mathbf{W}}(i) = \det(\mathbf{R})$  satisfies*

$$P_{\mathbf{W}}(i) = \alpha(\mathbf{W}) \cdot P_C(i), \quad (17)$$

*where  $P_C(i)$  is the chirality product, and  $\alpha(\mathbf{W}) > 0$  is a scaling factor depending only on  $\mathbf{W}$  but not on the atom coordinates.  $P_{\mathbf{W}}(i)$  satisfies Proposition 3.1 and Lemma 3.1.*

The proof of Lemma 4.1 is provided in Appendix F.2.

**Discussion on rank deficiency.** If  $\mathbf{W}$  is not full column-rank, then  $\mathbf{O}$  becomes rank-deficient and the determinant  $\det(\mathbf{R})$  degenerates to 0, resulting in the loss of chirality information. While the magnitude of  $\det(\mathbf{R})$  depends on  $\alpha(\mathbf{W})$ , the sign remains a reliable indicator of chiral configurations as long as  $\mathbf{W}$  has full column rank. Hence, ensuring rank sufficiency of  $\mathbf{W}$  is critical for stability and discriminative power. To guarantee full column rank of  $\mathbf{W}$ , we adopt two strategies in practice:

**Weight regularization.** We introduce a penalty term on the weight as part of the loss function:

$$\mathcal{L}_{\text{reg}} = \|\mathbf{W}^\top \mathbf{W} - \mathbf{I}_3\|^2, \quad (18)$$

encouraging  $\mathbf{W}$  to preserve rank during training.

**Auxiliary QR on weights.** Alternatively, we apply QR decomposition directly on  $\mathbf{W}$  before projection, ensuring that projected neighbor vectors remain linearly independent.

## 5 EXPERIMENTS

### 5.1 MAIN EXPERIMENTS

**Datasets.** We evaluate our method on both centrally and axially chiral molecules using multiple benchmarks, as shown in Appendix Table 4. R/S classification and enantiomer ranking datasets for central chirality are from ChIRo (Adams et al., 2022). ECD prediction for central chirality is from CMCDS (Li et al., 2025). ECD and OR prediction for axial chirality are constructed by ourselves, named ACMP (axial chiral molecular properties). See Appendix A for details about dataset processing, construction, and splits.

**Baselines.** We benchmark ChiDeK against both an E(3)-invariant baseline, DimeNet++ (Gasteiger et al., 2020a), and several SE(3)-invariant baselines, including SphereNet (Liu et al., 2022), ChIRo (Adams et al., 2022), ECDFormer (Li et al., 2025), and ChiGNN (Yan et al., 2025). We also consider the 2D chiral GNN Tetra-DMPNN with permutation (p) and concatenation (c) variants (Pattanaik et al., 2020), and the stereostructure descriptor-based SPMS Xu et al. (2021).

**Implementation Details.** All experiments are implemented in PyTorch (Paszke et al., 2019) and conducted on NVIDIA RTX 3090 GPUs. See Appendix D for more training details.

#### 5.1.1 EXPERIMENTAL RESULTS FOR CENTRAL CHIRALITY

Table 1: R/S classification, enantiomer ranking, and ECD prediction results for central chirality. **Best** and second-best are marked.

Method	R/S (%)	Ranking (%)	Position	Number	Symbol (%)
	Acc $\uparrow$	Acc $\uparrow$	RMSE $\downarrow$	RMSE $\downarrow$	Acc $\uparrow$
DimeNet++	$65.7 \pm 2.9$	$58.4 \pm 0.2$	<u><math>2.12 \pm 0.19</math></u>	$1.04 \pm 0.15$	$50.8 \pm 0.1$
Tetra-DMPNN (c)	<u><math>99.7 \pm 0.1</math></u>	$70.1 \pm 0.5$	$2.38 \pm 0.12$	$1.25 \pm 0.09$	$50.0 \pm 0.1$
Tetra-DMPNN (p)	<u><math>99.7 \pm 0.1</math></u>	$67.6 \pm 0.6$	$2.36 \pm 0.16$	$1.28 \pm 0.08$	$50.0 \pm 0.1$
SphereNet	$98.2 \pm 0.2$	$68.6 \pm 0.3$	$2.36 \pm 0.15$	<u><math>1.02 \pm 0.11</math></u>	$51.9 \pm 0.3$
ChIRo	$98.5 \pm 0.2$	<u><math>72.0 \pm 0.5</math></u>	$2.67 \pm 0.13$	$1.22 \pm 0.13$	$51.0 \pm 0.4$
ECDFormer	$92.3 \pm 1.2$	$58.6 \pm 0.3$	<b><math>2.02 \pm 0.10</math></b>	<b><math>1.01 \pm 0.09</math></b>	$50.3 \pm 0.3$
ChiGNN	$84.5 \pm 0.9$	$59.6 \pm 0.4$	$2.39 \pm 0.18$	$1.24 \pm 0.10$	$49.9 \pm 0.5$
SPMS	$81.4 \pm 0.7$	$60.4 \pm 0.4$	$2.58 \pm 0.17$	$1.22 \pm 0.12$	$50.9 \pm 0.3$
ChiDeK (Ours)	<b><math>99.8 \pm 0.1</math></b>	<b><math>72.8 \pm 0.2</math></b>	$2.20 \pm 0.14$	$1.18 \pm 0.09$	<b><math>53.3 \pm 0.6</math></b>

**R/S Classification.** Table 1 reports the results. ChiDeK achieves the best performance close to 100%. This clearly demonstrates the effectiveness of our chirality-aware design. In contrast, the E(3)-invariant DimeNet++ completely fails on this task, which is as expected. We present a visualization of cross attention weights in Appendix G. Note that, as emphasized in prior work (Adams

et al., 2022), this R/S classification task is necessary but not sufficient for meaningful chiral learning. Nonetheless, strong performance on this benchmark is an essential prerequisite for chirality-sensitive applications.

**Enantiomer Ranking.** Table 1 demonstrates that ChiDeK achieves optimal performance, surpassing the second-best method ChIRo by 0.8%. The modest performance gap between ChiDeK and ChIRo can be attributed to the limited structural diversity of the dataset, as all molecules contain only a single chiral center.

**ECD Spectrum Prediction.** Since the number and positions of peaks remain invariant across enantiomers while the symbols of peak intensities are inverted, accurately predicting peak heights represents the most critical aspect of ECD spectrum modeling. A model that successfully recovers peak counts and positions but fails to distinguish peak height symbols is ineffective and impractical. Consequently, we require a model that jointly predicts all three components, providing comprehensive ECD spectrum prediction. As shown in Table 1, ChiDeK achieves optimal performance in predicting peak height symbols while maintaining comparable accuracy for peak positions and counts.

However, the sign accuracy remains only slightly above 50% across all models. We hypothesize that this is largely due to *dataset inconsistencies*: the original CMCDS dataset provides RDKit-generated (Landrum, 2025) conformers, whereas the ECD labels are computed from optimized geometries. This mismatch causes poor performance. This underscores the inherent difficulty of reliable ECD spectrum modeling and suggests that consistent datasets are essential for more robust evaluation.

### 5.1.2 EXPERIMENTAL RESULTS FOR AXIAL CHIRALITY

Table 2: OR and ECD prediction results for axial chirality. **Best** and second-best are marked.

Method	Rotation (%)		Position	Number	Symbol (%)
	Acc $\uparrow$	RMSE $\downarrow$	RMSE $\downarrow$	RMSE $\downarrow$	Acc $\uparrow$
DimeNet++	$50.0 \pm 0.0$	$3.62 \pm 0.17$	$1.13 \pm 0.08$	$50.0 \pm 0.1$	
Tetra-DMPNN (c)	$50.0 \pm 0.1$	$3.15 \pm 0.12$	$1.20 \pm 0.15$	$52.8 \pm 0.2$	
Tetra-DMPNN (p)	$50.0 \pm 0.1$	$3.16 \pm 0.10$	$1.05 \pm 0.12$	$52.8 \pm 0.2$	
SphereNet	$52.5 \pm 0.2$	$3.36 \pm 0.22$	$1.08 \pm 0.09$	$52.4 \pm 0.4$	
ChIRo	$50.0 \pm 0.1$	$3.78 \pm 0.18$	$1.11 \pm 0.11$	$51.1 \pm 0.3$	
ECDFormer	$53.5 \pm 0.3$	$3.89 \pm 0.25$	$1.16 \pm 0.14$	$51.5 \pm 0.5$	
ChiGNN	$50.2 \pm 0.1$	$2.89 \pm 0.12$	$1.06 \pm 0.13$	$50.8 \pm 0.8$	
SPMS	$65.0 \pm 0.6$	$3.69 \pm 0.21$	$1.16 \pm 0.15$	$60.4 \pm 0.3$	
ChiDeK (Ours)	<b><math>69.2 \pm 0.5</math></b>	$3.24 \pm 0.14$	<b><math>1.05 \pm 0.12</math></b>	<b><math>71.2 \pm 0.6</math></b>	

**OR Prediction.** Table 2 shows that ChiDeK delivers the strongest performance, exceeding the second-best method by a noticeable 4.2% margin. Notably, all baseline models, except SPMS, achieve accuracy below 55%, indicating their inability to capture axial chirality due to their neglect of stereogenic axes. SPMS, which incorporates spherical stereostructural projections, can partially encode such information. These findings underscore the inherent complexity of the task and highlight the limitations of previous models in accurately representing complex chirality.

**ECD Spectrum Prediction.** Table 2 reports the results on the axial chirality ECD prediction task. ChiDeK achieves the highest accuracy in predicting the symbols of peak heights, exceeding the second-best model by a substantial margin of 10.8%. It also delivers competitive performance in predicting peak positions and numbers. Similar to OR prediction, SPMS exhibits limited performance, while all other baseline models achieve less than 55% accuracy in symbol prediction. For fine-grained analysis, we present performance breakdown across seven axial-chirality subtypes in Appendix B.1. Additionally, we evaluate general molecular property prediction tasks in Appendix B.2 to confirm that our model maintains full expressive power, where chirality is less critical.

Figure 3 presents a representative prediction comparison. ChiDeK successfully assigns opposite peak-height symbols to opposite stereochemical configurations, thereby correctly distinguishing axial enantiomers. In contrast, ChiGNN and Tetra-DMPNN (permute) produce identical peak-height

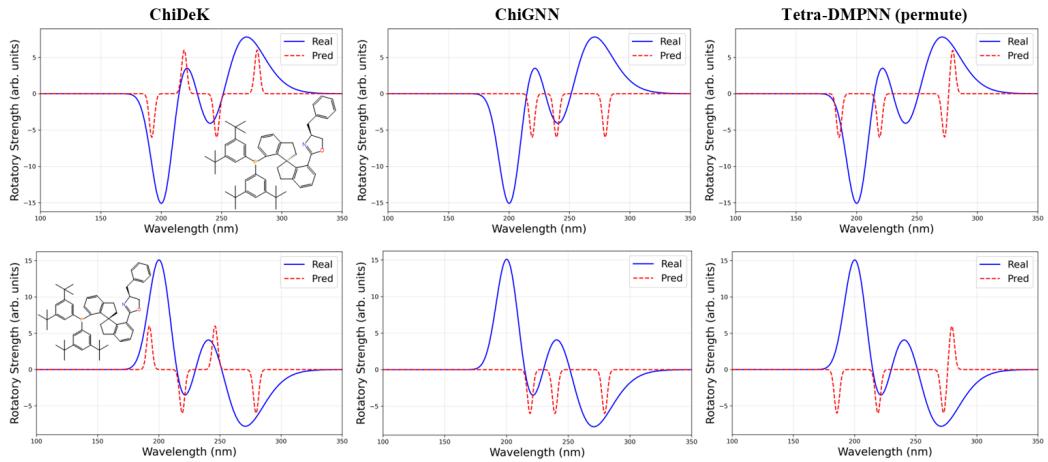


Figure 3: An axial ECD prediction example of a pair of enantiomers. Each row shows the predictions for one configuration across models, with the two rows corresponding to opposite configurations.

symbols across enantiomers, failing to capture stereochemical differences. These results highlight the robustness and stereochemical sensitivity of ChiDeK in handling axial chirality. See Appendix E for more explanations and prediction examples.

**Rotation Analysis of Axial Chirality Representations.** To investigate how molecular representations generated by ChiDeK vary under rotation along the chiral axis, we systematically vary the torsion angle, with details provided in Appendix E.3. Figure 4 illustrates two representative examples. In both cases, the 18 conformers are partitioned into two opposite configurations (each comprising 9 conformers), consistent with the expected stereochemical symmetry. The trajectory reflects a transition from one configuration to its opposite and then back to the original, while the cosine similarity plots show high similarity within the same configuration and vice versa.

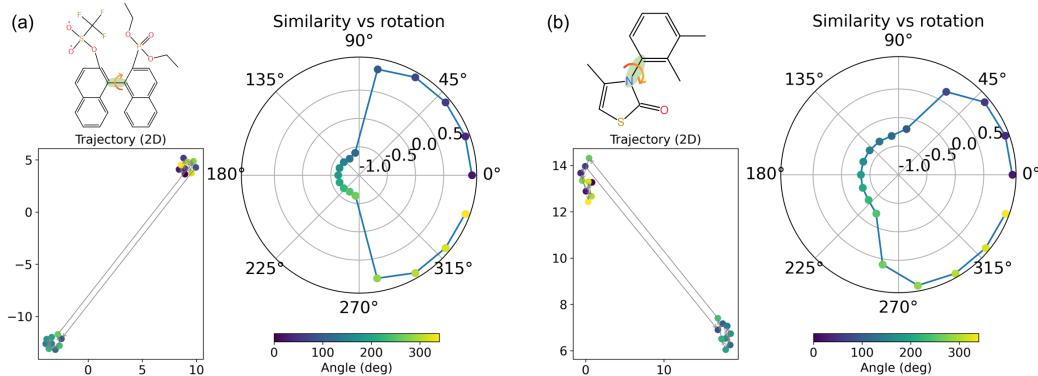


Figure 4: Visualization of how ChiDeK representations change under rotations along the chiral axis for two molecules. Each point in the trajectory plot obtained by UMAP (McInnes et al., 2018) represents the learned embedding at a given rotation angle, while each point in the polar plot denotes the cosine similarity between the embedding and the reference at degree 0.

## 5.2 ABLATION STUDIES

We perform ablation studies on the axial ECD prediction task, using the accuracy of peak height symbols as the primary evaluation metric. We examine strategies to address rank deficiency (QR and Reg denote QR decomposition and regularization, respectively), types of chiral encoders, and atom separation strategies (Sep.), as shown in Table 3. The ablation of missing or mislabeled chiral atoms

is presented in Appendix B.3. The comparison with ChiralFinder-enhanced model is presented in Appendix B.4.

Table 3: Ablation on model components.

Model Components			Acc (%) $\uparrow$
Chiral Encoder	Rank strategy	Separate	
Linear w/o. $M_C$	None	✓	$51.2 \pm 1.2$
Linear w. $M_C$	None	✓	$68.7 \pm 0.6$
Kernel-based	QR	✗	$69.8 \pm 0.4$
Kernel-based	Reg	✓	$71.0 \pm 0.5$
Kernel-based	None	✓	$68.3 \pm 0.6$
Kernel-based	QR	✓	$71.2 \pm 0.6$

**(i) Strategies to Address Rank Deficiency.** Table 3 compares three approaches for handling rank deficiency: regularization (Reg), QR decomposition (QR), and no intervention (None). When handling rank deficiency, both Reg and QR outperform None, achieving similar performance.

**(ii) Choice of Chirality Encoder.** We examine alternative chiral encoder designs in Table 3. A linear encoder without access to  $M_C$  performs poorly (close to random guessing), indicating that it cannot capture chirality. In contrast, our kernel-based encoder yields a higher accuracy, confirming its ability to encode stereogenic features effectively.

**(iii) Separating Chiral Atom Types.** We test the effect of explicitly distinguishing chiral-related atoms ( $\mathcal{I}_r$ ) from non-chiral atoms ( $\mathcal{I}_n$ ). Models that separate chiral-related atoms and non-chiral atoms outperform those that treat them jointly, highlighting the importance of modeling their distinct roles within stereochemical environments.

## 6 CONCLUSION

We present ChiDeK (**Chiral Determinant Kernels**), a unified framework for learning chiral molecular representations. By embedding the  $SE(3)$ -invariant chirality matrix through chiral determinant kernels and employing cross-attention between chiral and non-chiral atoms, ChiDeK explicitly captures both central and axial stereogenic features. Across multiple benchmarks, ChiDeK substantially outperforms baselines, particularly for axially chiral molecules. We further contribute a benchmark dataset for axial chirality, encompassing both ECD and OR prediction tasks. This dataset provides a foundation for future research in stereochemistry-aware machine learning. In future work, we plan to extend our unified stereochemical representation to additional forms of chirality, such as planar and helical chirality, and evaluate its utility in broader downstream applications, including docking-score prediction for chiral ligand-protein interactions and enantioselectivity prediction in asymmetric catalysis.

## ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China (No. 2023YFC2811500) and the National Natural Science Foundation of China (No. 62272300).

## REPRODUCIBILITY STATEMENT

Code and data are available at <https://github.com/Meteor-han/ChiDeK>.

## REFERENCES

Keir Adams, Lagnajit Pattanaik, and Connor W. Coley. Learning 3d representations of molecular chirality with invariance to bond rotations. In *International Conference on Learning Representations*, 2022. URL <https://openreview.net/forum?id=hm2tNDdgaFK>.

João Aires-de Sousa, Johann Gasteiger, Ivan Gutman, and Dušica Vidović. Chirality codes and molecular structure. *Journal of Chemical Information and Computer Sciences*, 44(3):831–836, 2004.

Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. Gfn2-xtb—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *Journal of Chemical Theory and Computation*, 15(3):1652–1671, 2019. doi: 10.1021/acs.jctc.8b01176. PMID: 30741547.

Ilyes Batatia, Simon Batzner, Dávid Péter Kovács, Albert Musaelian, Gregor N. C. Simm, Ralf Drautz, Christoph Ortner, Boris Kozinsky, and Gábor Csányi. The design space of  $e(3)$ -equivariant atom-centred interatomic potentials. *Nature Machine Intelligence*, 7(1):56–67, 2025. URL <https://doi.org/10.1038/s42256-024-00956-x>.

Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P. Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E. Smidt, and Boris Kozinsky.  $E(3)$ -equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13, 2022. doi: 10.1038/s41467-022-29939-5.

Robert S Cahn, Christopher Ingold, and Vladimir Prelog. Specification of molecular chirality. *Angewandte Chemie International Edition in English*, 5(4):385–415, 1966.

Jessica Ceramella, Domenico Iacopetta, Angelica Franchini, Michele De Luca, Carmela Saturnino, Inmaculada Andreu, Maria Stefania Sinicropi, and Alessia Catalano. A look at the importance of chirality in drug activity: Some significative examples. *Applied Sciences*, 2022.

Wenjie Du, Xiaoting Yang, Di Wu, FenFen Ma, Baicheng Zhang, Chaochao Bao, Yaoyuan Huo, Jun Jiang, Xin Chen, and Yang Wang. Fusing 2d and 3d molecular graphs as unambiguous molecular descriptors for conformational and chiral stereoisomers. *Briefings in bioinformatics*, 24(1), 2023.

Alexandru Dumitrescu, Dani Korpela, Markus Heinonen, Yogesh Verma, Valerii Iakovlev, Vikas Garg, and Harri Lähdesmäki.  $E(3)$ -equivariant models cannot learn chirality: Field-based molecular generation. In *The Thirteenth International Conference on Learning Representations*, 2025. URL <https://openreview.net/forum?id=mXHTifc1Fn>.

Henri A. Favre and Warren H. Powell (eds.). *Nomenclature of Organic Chemistry – IUPAC Recommendations and Preferred Names 2013*. Royal Society of Chemistry, 2013. ISBN 978-1-84973-306-9. doi: 10.1039/9781849733069.

Daniel Flam-Shepherd, Tony C Wu, Pascal Friederich, and Alan Aspuru-Guzik. Neural message passing on high order paths. *Machine Learning: Science and Technology*, 2(4):045009, 2021. doi: 10.1088/2632-2153/abf5b8.

M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, and et al. Gaussian 09, revision a.02, 2016a. Gaussian Inc. Wallingford CT.

M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, and et al. Gaussian 16, revision b.01, 2016b. Gaussian Inc. Wallingford CT.

Fabian B. Fuchs, Daniel E. Worrall, Volker Fischer, and Max Welling.  $Se(3)$ -transformers: 3d roto-translation equivariant attention networks. In *Advances in Neural Information Processing Systems 34 (NeurIPS)*, 2020.

Piotr Gaiński, Michał Koziarski, Jacek Tabor, and Marek Śmieja. ChiENN: Embracing molecular chirality with graph neural networks. In *Machine Learning and Knowledge Discovery in Databases: Research Track*, volume 14171 of *Lecture Notes in Computer Science*, pp. 37–54. Springer, Cham, 2023. doi: 10.1007/978-3-031-43418-1\_3.

Johannes Gasteiger, Shankari Giri, Johannes T. Margraf, and Stephan Günnemann. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. In *Machine Learning for Molecules Workshop, NeurIPS*, 2020a.

Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. In *International Conference on Learning Representations (ICLR)*, 2020b.

A. Gaucherand, E. Yen-Pon, A. Domain, and et al. Enantioselective synthesis of molecules with multiple stereogenic elements. *Chemical Society Reviews*, 2024.

Xiaoxun Gong, He Li, Nianlong Zou, Runzhang Xu, Wenhui Duan, and Yong Xu. General framework for e(3)-equivariant neural network representation of density functional theory hamiltonian. *Nature Communications*, 14(1):2848, 2023.

Yan-Mei He, Yuan-Zheng Cheng, Yingdong Duan, Yan-Dong Zhang, Qing-Hua Fan, Shu-Li You, Sanzhong Luo, Shou-Fei Zhu, Xue-Feng Fu, and Qi-Lin Zhou. Recent progress of asymmetric catalysis from a chinese perspective. *CCS Chemistry*, 5(12):2685–2716, 2023.

D. Hendrycks and K. Gimpel. Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*, 2016.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014. Preprint at <https://arxiv.org/abs/1412.6980>.

Johannes Klicpera, Florian Becker, and Stephan Günnemann. Gemnet: Universal directional graph neural networks for molecules. In A. Beygelzimer, Y. Dauphin, P. Liang, and J. Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems*, 2021. URL [https://openreview.net/forum?id=HS\\_sOaxS9K-](https://openreview.net/forum?id=HS_sOaxS9K-).

Mario Krenn, Qianxiang Ai, Senja Barthel, Nessa Carson, Angelo Frei, Nathan C Frey, Pascal Friederich, Théophile Gaudin, Alberto Alexander Gayle, Kevin Maik Jablonka, et al. Selfies and the future of molecular string representations. *Patterns*, 3(10), 2022.

G. Landrum. Rdkit: open-source cheminformatics software, 2025. URL <http://www.rdkit.org>. (accessed Feb 11, 2025).

Hao Li, Da Long, Li Yuan, Yonghong Tian, Xinchang Wang, and Fanyang Mo. Decoupled peak property learning for efficient and interpretable electronic circular dichroism spectrum prediction. *Nature Computational Science*, pp. 1–11, 2025. doi: 10.1038/s43588-024-00757-7.

Shengchao Liu, Weitao Du, Zhiming Ma, Hongyu Guo, and Jian Tang. A group symmetric stochastic differential equation model for molecule multi-modal pretraining. In *Proceedings of the 40th International Conference on Machine Learning*, ICML’23. JMLR.org, 2023a.

Yaoran Liu, Zilong Wu, Daniel W Armstrong, Herman Wolosker, and Yuebing Zheng. Detection and analysis of chiral molecules as disease biomarkers. *Nature Reviews Chemistry*, 7(5):355–373, 2023b.

Yi Liu, Limei Wang, Meng Liu, Yuchao Lin, Xuan Zhang, Bora Oztekin, and Shuiwang Ji. Spherical message passing for 3d molecular graphs. In *International Conference on Learning Representations*, 2022. URL <https://openreview.net/forum?id=givsRXs0t9r>.

Yunchao (Lance) Liu, Yu Wang, Oanh Vu, Rocco Moretti, Bobby Bodenheimer, Jens Meiler, and Tyler Derr. Interpretable chirality-aware graph neural network for quantitative structure activity relationship modeling in drug discovery. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(12):14356–14364, Jun. 2023c. doi: 10.1609/aaai.v37i12.26679.

Fabrizio Mastrolorito, Fulvio Ciriaco, Orazio Nicolotti, and Francesca Grisoni. Enhancing deep chemical reaction prediction with advanced chirality and fragment representation. *Chemical Communications*, 2025. doi: 10.1039/D5CC02641E.

Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. Umap: Uniform manifold approximation and projection. *J. Open Source Softw.*, 3:861, 2018.

Rebecca U McVicker and Niamh M O’Boyle. Chirality of new drug approvals (2013–2022): Trends and perspectives. *Journal of Medicinal Chemistry*, 2024.

Alex Morehead and Jianlin Cheng. Geometry-complete perceptron networks for 3d molecular graphs. *Bioinformatics*, 40(2):btae087, 2024. URL <https://doi.org/10.1093/bioinformatics/btae087>.

G. P. Moss. Basic terminology of stereochemistry (iupac recommendations 1996). *Pure and Applied Chemistry*, 68(12):2193–2222, 1996. doi: doi:10.1351/pac199668122193. URL <https://doi.org/10.1351/pac199668122193>.

Ramanathan Natarajan, Subhash C Basak, and Terrence S Neumann. Novel approach for the numerical characterization of molecular chirality. *Journal of Chemical Information and Modeling*, 47(3):771–775, 2007.

Markus Orsi and Jean-Louis Reymond. One chiral fingerprint to find them all. *Journal of Chemicalinformatics*, 16(1):53, 2024. ISSN 1758-2946. doi: 10.1186/s13321-024-00849-6. URL <https://doi.org/10.1186/s13321-024-00849-6>.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, , and et al. Pytorch: An imperative style, high-performance deep learning library. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. Curran Associates Inc., Red Hook, NY, USA, 2019.

Lagnajit Pattanaik, Octavian-Eugen Ganea, Ian Coley, Klavs F. Jensen, William H. Green, and Connor W. Coley. Message passing networks for molecules with tetrahedral chirality. *ArXiv*, abs/2012.00094, 2020.

Paola Peluso and Bezhan Chankvetadze. Recognition in the domain of molecular chirality: from noncovalent interactions to separation of enantiomers. *Chemical Reviews*, 122(16):13235–13400, 2022.

Yichuan Peng, Gufeng Yu, Runhan Shi, Letian Chen, Xi Wang, Wenjie Du, Xiaohong Huo, and Yang Yang. Chiralcat: Molecular chirality classification with enhanced spatial representation using learnable queries. *Artificial Intelligence Chemistry*, 3(2):100091, 2025. ISSN 2949-7477. doi: 10.1016/j.aichem.2025.100091.

Denisa Roberts and Lucas Roberts. Qr and lq decomposition matrix backpropagation algorithms for square, wide, and deep matrices and their software implementation. *ArXiv*, abs/2009.10071, 2020.

Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139, pp. 9323–9332. PMLR, 2021. URL <https://proceedings.mlr.press/v139/satorras21a.html>.

Nadine Schneider, Richard A. Lewis, Nikolas Fechner, and Peter Ertl. Chiral cliffs: investigating the influence of chirality on binding affinity. *ChemMedChem*, 13, 2018.

Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. In *NIPS*, 2017.

Qian Shen, Lu Wang, Hui Zhou, Huidi Jiang, Lushan Yu, and Su Zeng. Stereoselective binding of chiral drugs to plasma proteins. *Acta Pharmacologica Sinica*, 34:998 – 1006, 2013.

Runhan Shi, Chi Zhang, Gufeng Yu, Xiaohong Huo, and Yang Yang. A unifying geometric framework for computational representation of stereoisomers based on mixed product. *Cell Reports Physical Science*, 7(1):103065, 2026. doi: 10.1016/j.xcrp.2025.103065.

M. Shuaibi, A. Kolluru, A. Das, and et al. Rotation invariant graph neural networks using spin convolutions. *arXiv preprint arXiv:2106.09575*, 2021.

Nathaniel Thomas, Tess E. Smidt, Steven M. Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick F. Riley. Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds. *ArXiv*, abs/1802.08219, 2018.

Ashish Vaswani, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Neural Information Processing Systems*, 2017.

David Weininger. Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences*, 28(1):31–36, Feb 1988. ISSN 0095-2338. doi: 10.1021/ci00057a005. URL <https://doi.org/10.1021/ci00057a005>.

Zhenqin Wu, Bharath Ramsundar, Evan N. Feinberg, Joseph Gomes, Caleb Geniesse, Aneesh S. Pappu, Karl Leswing, and Vijay Pande. Moleculenet: a benchmark for molecular machine learning. *Chem. Sci.*, 9:513–530, 2018. doi: 10.1039/C7SC02664A.

Li-Cheng Xu, Xin Li, Miao-Jiong Tang, Luo-Tian Yuan, Jia-Yu Zheng, Shuo-Qing Zhang, and Xin Hong. A molecular stereostructure descriptor based on spherical projection. *Synlett*, 32(18):1837–1842, 2021. doi: 10.1055/s-0040-1705977.

Jiaxin Yan, Haiyuan Wang, Wensheng Yang, Xiaonan Ma, Yajing Sun, and Wenping Hu. Chignn: Interpretable algorithm framework of molecular chiral knowledge-embedding and stereosensitive property prediction. *Journal of Chemical Information and Modeling*, 65(7):3239–3247, 2025. doi: 10.1021/acs.jcim.4c02259.

Yifei Yang, Runhan Shi, Zuchao Li, Shu Jiang, Bao-Liang Lu, Qibin Zhao, Yang Yang, and Hai Zhao. Batgpt-chem: A foundation large model for chemical engineering. *Research*, 8:0827, 2025. doi: 10.34133/research.0827.

Yasuhiro Yoshikai, Tadahaya Mizuno, Shumpei Nemoto, and Hiroyuki Kusuhara. Difficulty in chirality recognition for transformer architectures learning chemical structures from string representations. *Nature Communications*, 15(1):1197, 2024.

Qing-You Zhang and Joao Aires-de Sousa. Physicochemical stereodescriptors of atomic chiral centers. *Journal of Chemical Information and Modeling*, 46(6):2278–2287, 2006.

Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, and Guolin Ke. Uni-mol: A universal 3d molecular representation learning framework. In *The Eleventh International Conference on Learning Representations*, 2023.

## A DATASETS

Table 4: Summary of downstream datasets.

Dataset	Chirality	Task	Size	Split
ChIRo (R/S)	Central	R/S classification	466K conformers / 78K mols	70/15/15
ChIRo (Ranking)	Central	Enantiomer ranking	335K conformers / 69K mols	70/15/15
CMCDS	Central	ECD prediction	22,182 conformers / mols	80/10/10
ACMP	Axial	ECD and OR prediction	1,192 conformers / mols	80/10/10

### A.1 CENTRAL CHIRALITY

We adopt datasets from ChIRo (Adams et al., 2022): (i) R/S classification: the task is to classify the configuration (R or S) of a tetrahedral chiral center. The dataset is a subset of PubChem3D, comprising about 466K conformers of 78K enantiomers, each with exactly one tetrahedral chiral center. The same 70/15/15 splits are used. (ii) Enantiomer ranking: the task is to predict docking score rankings, evaluating whether the predicted relative ordering between enantiomers matches the ground truth. This dataset includes 335K conformers of 69K enantiomers (34.5K pairs), with enantiomers kept in the same data partition. We sample one conformer for each enantiomer in the batch during training. The same 70/15/15 splits are used.

In addition, we employ the CMCDS dataset for ECD prediction (Li et al., 2025), formulated as a multi-task, multi-label problem with three sub-tasks: predicting (1) the number of spectral peaks, (2) their positions, and (3) the sign of peak heights. CMCDS consists of ECD spectra for 22,190 chiral molecules (11,095 pairs), generated by large-scale theoretical calculations using Gaussian16 B.01 (Frisch et al., 2016b). After excluding four molecules without chiral centers, the dataset contains 22,182 conformers (11,091 pairs). We use random 80/10/10 splits while ensuring enantiomers are within the same partition.

### A.2 AXIAL CHIRALITY

We introduce a new benchmark, ACMP, for electronic circular dichroism (ECD) and optical rotation (OR) prediction. Specifically, we curate 650 axially chiral molecules from ChiralFinder (Shi et al., 2026) and generate initial 3D conformations using RDKit (Landrum, 2025). To refine these structures, a preliminary conformational optimization is first carried out using GFN2-xTB (Bannwarth et al., 2019), which is a semiempirical quantum mechanical method. Following this, the molecular geometries are fully optimized using the Gaussian 09 package (Frisch et al., 2016a), employing the B3LYP functional and the 6-31G(d) basis set.

For ECD, to obtain the spectral data, time-dependent DFT (TD-DFT) calculations are performed to compute the first 30 electronic excited states. These calculations use the long-range corrected CAM-B3LYP functional and the 6-311+G(d,p) basis set, yielding the excitation wavelengths and their corresponding rotatory strengths. ECD spectra curves are generated with Gaussian broadening (FWHM = 2/3), and an example is presented in Figure 5. We uniformly sample spectra at 1 nm intervals to produce sequence-form ECD representations. The labeling process follows the same procedure as CMCDS: peak positions are uniformly discretized into 20 classes, while peak heights are encoded using their sign (positive or negative). The number of peaks ranges from 0 to 6, and its distribution is illustrated in Figure 6. For OR, we compute the optical rotation at 589.3 nm using CAM-B3LYP/6-31G\*\*. The resulting scalar rotation value is then mapped to a binary classification label, forming a two-class prediction task. All Gaussian calculations are conducted using 400 GB of memory with 192 CPU cores. To obtain enantiomeric counterparts, we generate opposite configurations by reflecting atomic coordinates across the z-axis. We finally get a dataset of 1,192 enantiomers (596 pairs). For evaluation, we adopt random 80/10/10 train/validation/test splits, ensuring that enantiomer pairs are kept within the same partition.

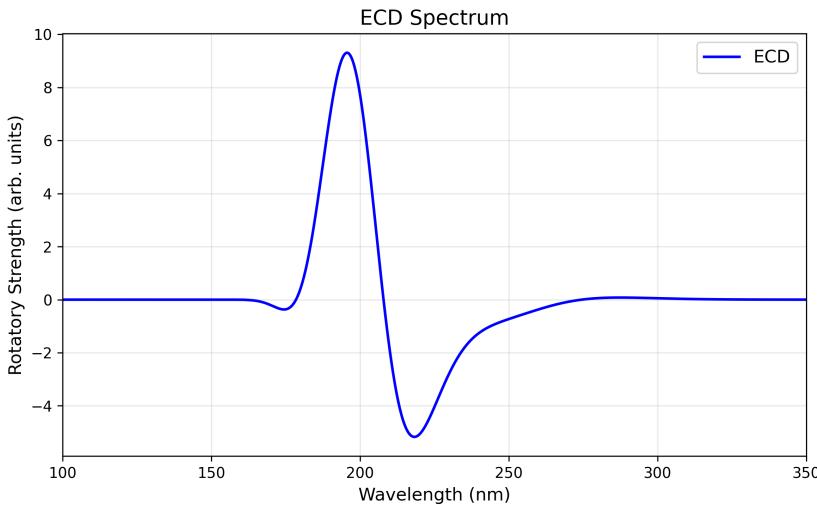


Figure 5: The ECD curve example for an axial chiral molecule.

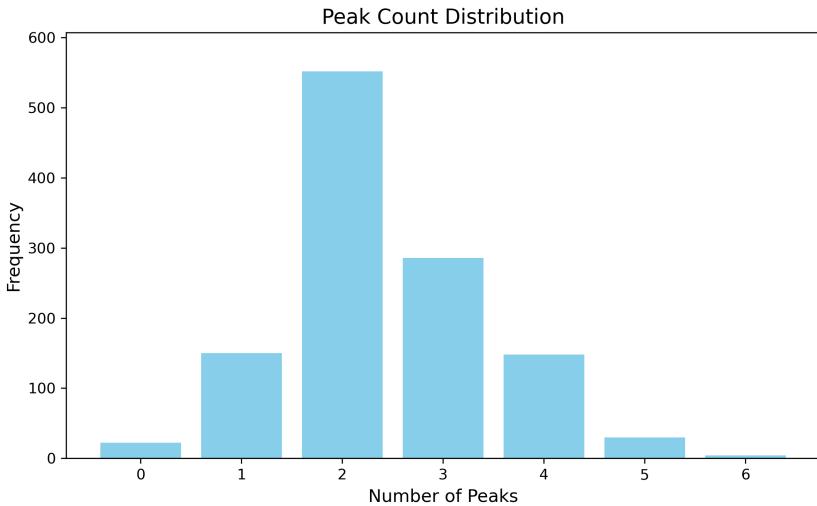


Figure 6: Distribution of ECD peak numbers in ACMP dataset.

## B ADDITIONAL EXPERIMENTS

### B.1 FINE-GRAINED PERFORMANCE ON AXIAL-CHIRALITY SUBTYPES

To further investigate ChiDeK’s performance across structural diversity, we provide a subtype-level breakdown on the test set, covering all seven axial-chirality subtypes. Table 5 reports detailed results for axial ECD-height prediction.

We observe that ChiDeK achieves strong performance on most subtypes ( $\geq 70\%$ ), while Heterobiaryl (C–N) and Heterobiaryl (C–B) show lower performance ( $\leq 60\%$ ), highlighting potential directions for future improvement.

### B.2 EVALUATION ON GENERAL MOLECULAR PROPERTY PREDICTION

To assess the broader applicability of ChiDeK beyond chirality-sensitive tasks, we evaluate the model on several moderate-sized MoleculeNet datasets Wu et al. (2018) where chirality is less

Table 5: Performance breakdown of seven subtypes for the axial ECD-height task.

Subtype	# Enantiomers	# Height symbols	Acc (%) $\uparrow$
Chiral atom pair	10	26	$71.5 \pm 1.8$
Allene-like	2	4	$80.0 \pm 9.9$
Spiral atom and chain	4	8	$90.0 \pm 5.0$
Biaryl	70	194	$72.9 \pm 0.5$
Heterobiaryl (C-N)	12	20	$57.0 \pm 4.0$
Heterobiaryl (C-B)	6	16	$60.0 \pm 3.1$
Nonbiaryl	16	46	$70.0 \pm 1.6$
Total/Average	120	314	$71.2 \pm 0.6$

critical, including FreeSolv, BACE, BBBP, ClinTox, and SIDER. These benchmarks test general molecular property prediction rather than chiral discrimination.

As summarized in Table 6, ChiDeK achieves performance comparable to standard GNNs without pretraining (D-MPNN, Attentive FP). While it does not surpass highly optimized SOTA models on these general tasks, the results indicate that ChiDeK introduces no harmful inductive bias and retains broad representational capacity for generic molecular property prediction.

Table 6: Results on less chiral-related MoleculeNet datasets.

Metric	ROC-AUC (%) $\uparrow$				RMSE $\downarrow$
	BBBP	BACE	ClinTox	SIDER	
Datasets	2039	1513	1478	1427	642
# Molecules					
# Tasks	1	1	2	27	1
<i>W/o. pre-training</i>					
D-MPNN	$71.0 \pm 0.3$	$80.9 \pm 0.6$	$90.6 \pm 0.6$	$57.0 \pm 0.7$	$2.082 \pm 0.082$
Attentive FP	$64.3 \pm 1.8$	$78.4 \pm 0.0$	$84.7 \pm 0.3$	$60.6 \pm 3.2$	$2.073 \pm 0.183$
ChiDeK (Ours)	$68.9 \pm 0.8$	$78.1 \pm 0.9$	$79.2 \pm 1.1$	$58.3 \pm 0.8$	$2.412 \pm 0.109$
<i>W. pre-training</i>					
PretrainGNN	$68.7 \pm 1.3$	$84.5 \pm 0.7$	$72.6 \pm 1.5$	$62.7 \pm 0.8$	$2.764 \pm 0.002$
Uni-Mol	$72.9 \pm 0.6$	$85.7 \pm 0.2$	$91.9 \pm 1.8$	$65.9 \pm 1.3$	$1.480 \pm 0.048$

### B.3 ROBUSTNESS TO NOISY OR INCOMPLETE CHIRALITY ANNOTATIONS

To evaluate ChiDeK’s robustness to imperfect chirality annotations, we conduct experiments on both central and axial chirality.

**Central chirality (R/S classification).** We systematically introduce noise by removing or mislabeling (1:1) chiral atoms in 2.5%, 5%, 7.5%, and 10% of molecules. For mislabeling, chiral centers are replaced with non-chiral atoms; for missing centers, the labels are simply removed. Table 7 shows a smooth and gradual degradation rather than a sharp collapse. With moderate levels of noise, the model maintains strong accuracy, demonstrating that the cross-attention architecture provides meaningful tolerance to incomplete or partially incorrect stereochemical inputs. These results confirm that ChiDeK is robust to moderate annotation errors, while also showing that reliable chiral identification naturally leads to stronger performance.

**Axial chirality (ECD-height prediction).** We evaluate the effect of removing ground-truth axial labels (with noise) on ECD-height prediction performance. In this setting, the average coverage (percentage of identified chiral atoms covering true chiral atoms) is 0.924, and the average IoU (intersection over union) between identified and true chiral atoms is 0.645. Similar to central chiral-

Table 7: Results of central R/S classification with different ratios of noise.

Ratio	W/o. noise	2.5%	5%	7.5%	10%
ChiDeK	99.8 $\pm$ 0.1	99.1 $\pm$ 0.0	98.2 $\pm$ 0.0	96.9 $\pm$ 0.1	96.1 $\pm$ 0.2

ity, removing ground-truth guidance reduces performance as shown in Table 8, demonstrating that reliable chiral identification is essential for optimal model performance.

Table 8: Results of axial ECD-height prediction with or without noise.

Ratio	W/o. noise	W. noise
ChiDeK	71.2 $\pm$ 0.6	65.8 $\pm$ 0.4

#### B.4 COMPARISON WITH CHIRALFINDER-ENHANCED MODEL

We integrate the ChiralFinder-derived chirality matrix into ChIRo by concatenating the matrix features with the atomic embeddings (using zero vectors for non-chiral atoms) and evaluate this augmented model on the same axial-chirality benchmarks.

As shown in Table 9, the inclusion of the chirality matrix yields a clear improvement over the original ChIRo. Nevertheless, ChiDeK continues to outperform both variants. This result indicates that while the chirality matrix is beneficial, ChiDeK provides stronger representational capacity for modeling axial chirality.

Table 9: Comparison with ChIRo enhanced by the chirality matrix from ChiralFinder.

Method	Rotation (%)	Symbol (%)
ChIRo	50.0 $\pm$ 0.1	51.1 $\pm$ 0.3
ChIRo + Chirality matrix	62.1 $\pm$ 0.3	61.8 $\pm$ 1.0
ChiDeK	69.2 $\pm$ 0.5	71.2 $\pm$ 0.6

## C CHIRAL CROSS-ATTENTION

The learnable pairwise bias  $p_{ij}^{(\ell)}$  is updated by:

$$p_{ij}^{(\ell+1)} = \frac{\mathbf{Q}_i^{(\ell)} \left( \mathbf{K}_j^{(\ell)} \right)^\top}{\sqrt{k}} + p_{ij}^{(\ell)}, \quad (19)$$

where  $k$  is the hidden dimension. The attention outputs are then computed as

$$\text{Attention} \left( \mathbf{Q}_i^{(\ell)}, \mathbf{K}^{(\ell)}, \mathbf{V}^{(\ell)} \right) = \text{softmax} \left( \frac{\mathbf{Q}_i^{(\ell)} \left( \mathbf{K}^{(\ell)} \right)^\top}{\sqrt{k}} + \mathbf{p}_i^{(\ell-1)} \right) \mathbf{V}^{(\ell)}. \quad (20)$$

Finally, the outputs of  $L$  chiral cross-attention layers are aggregated by average pooling to get

$$\mathbf{H}_{\text{global}} = \frac{1}{|\mathcal{I}_c|} \sum_{i \in \mathcal{I}_c} \mathbf{H}_{c,i}^{(L)}. \quad (21)$$

This global representation serves as the input for downstream chirality property prediction.

## D IMPLEMENTATION DETAILS

Whenever possible, baseline results are taken directly from the original publications; otherwise, we reproduce them using publicly available code. All results are reported over five test folds or five random runs.

### D.1 INITIALIZATION OF ATOMS

**Identification of Chiral and Chiral-related Atoms.** For centrally chiral molecules, we use RDKit (Landrum, 2025) to detect chiral centers and their corresponding neighbors as chiral-related atoms. For axially chiral molecules, the chiral axis is labeled by chemists, while chiral-related atoms are determined using ChiralFinder (Shi et al., 2026). If ChiralFinder fails, we fall back to selecting up to four nearest neighbors as chiral-related atoms.

An example for ordering chiral-related atoms  $\mathcal{N}(i)$  for atom  $i$  is presented in Figure 7. After identifying the chiral atom (in green), suppose the indices of its chiral-related atoms (in orange) are A, B, C, and D, respectively. We use RDKit to calculate the CIP value for each atom (1, 4, 2, 5, respectively) and sort the indices in ascending order of their CIP priorities.

**Initial Atom Features.** Each atom is initialized with a 52-dimensional feature vector, consistent with ChIRo (Adams et al., 2022), including atom type, degree, formal charge, number of hydrogens, and hybridization state. Chiral atoms are additionally associated with their chirality matrix, while non-chiral and chiral-related atoms share the same feature design. The embedding layers are implemented as multi-layer perceptrons.

### D.2 HYPERPARAMETER OPTIMIZATION

We tune hyperparameters via grid search (Table 10). Training epochs are set to 10 for R/S configuration classification, 200 for enantiomer ranking, and 50 for both central and axial ECD prediction. We use Adam optimizer (Kingma & Ba, 2014) and adopt a cosine learning rate schedule with a minimum learning rate of  $0.1 \times$  the initial value. As for the enantiomer ranking task, we follow ChIRo (Adams et al., 2022) and add the margin ranking loss.

Table 10: Hyperparameter search space for ChiDeK.

Hyperparameter	Search Space
# Chiral Transformer Layers	{4, 8, 12, 16}
# Chiral Transformer Heads	{2, 4, 8}
Hidden Dimension	{64, 128, 256, 512}
Projection Dimension	{32, 64, 128, 256, 512}
Rank Strategy	{QR, Reg}
Learning Rate	{1e-4, 5e-4}
Batch Size	{16, 32, 64, 128, 256}
Weight of MarginRankingLoss	{0.1, 0.5, 1.0}

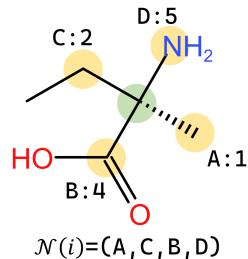


Figure 7: Ordering chiral-related atoms.

## E ANALYSIS OF AXIAL ECD PREDICTION

### E.1 BASELINES FAIL TO LEARNING AXIAL CHIRALITY

Axial chirality presents unique challenges for ECD prediction. Our results demonstrate that ChiDeK achieves strong performance in this setting, primarily due to its explicit encoding of chirality through

the proposed chirality determinant kernels. By directly modeling stereochemically relevant features, ChiDeK is capable of distinguishing axial enantiomers and accurately predicting peak height symbols, a property that is essential for practical chirality-sensitive spectrum prediction.

By contrast, all baseline methods except SPMS perform poorly on axially chiral molecules, with each exhibiting specific limitations for distinct reasons, as analyzed below:

- **DimeNet++**: As an  $E(3)$ -invariant model that relies only on distances and angles, DimeNet++ is fundamentally incapable of distinguishing mirror-related structures. Its poor performance is therefore expected.
- **Tetra-DMPNN**: This model augments 2D GNNs with a central-chirality readout, but has no mechanism to capture axial chirality. As a result, it fails on this task.
- **SphereNet** and **ChiRo**: In principle, these models incorporate torsional information and should have some sensitivity to stereochemistry. However, axial chirality involves subtle global spatial arrangements that are not well captured by their architectures, leading to limited performance.
- **ECDFFormer**: While this model introduces additional chiral encoding, it is designed for central chirality and lacks the necessary generalization to axial cases.
- **ChiGNN**: This method explicitly resolves atom permutations around tetrahedral centers, making it effective for central chirality. However, it does not generalize to axial stereogenic axes and thus completely fails in this setting.

## E.2 ECD PREDICTION EXAMPLES

Figure 8 presents two additional representative prediction comparisons. ChiDeK successfully assigns opposite peak-height symbols to opposite stereochemical configurations, while ChiGNN and Tetra-DMPNN (permutation variant) produce identical peak-height symbols across enantiomers, failing to capture stereochemical differences.

## E.3 DETAILS FOR ROTATING CHIRAL AXIS

We generate 18 conformers by rotating the torsion angle in 20-degree increments over 360 degrees. Each conformer corresponds to distinct coordinates, yielding different chirality matrices and thus different representations. We extract the embeddings before the predictor, apply UMAP (McInnes et al., 2018) for visualization, and record the trajectory as the torsion angle progresses from 0 to 340 degrees. Additionally, we calculate the cosine similarity between each embedding and the reference conformer at degree 0, and visualize the results using polar plots.

## F PROOF DETAILS

### F.1 CHIRALITY MATRIX

*Proof of Lemma 3.1.* Let  $(z_1, z_2)$  be a pair of enantiomers that differ only in the stereochemical configuration of atom  $i$ . By definition of enantiomers there exists an isometry  $T$  of  $\mathbb{R}^3$  that maps the atomic coordinates of  $z_1$  to those of  $z_2$  and which is a composition of a rigid-body motion and a reflection; equivalently, one may write  $T = R_2 \circ R_1$  with  $R_1 \in \text{SE}(3)$  and  $R_2 \in O^-(3)$  (an orthogonal transformation with determinant  $-1$ ). In particular, the local coordinate frame used to form the chirality matrix  $M_C(i)$  for  $z_2$  is obtained from that for  $z_1$  by applying the linear part of  $T$ , which has determinant  $-1$ .

Since  $i$  is a chiral atom, its substituent vectors (used to build  $M_C(i)$ ) are non-coplanar, hence  $\det(M_C(i)) \neq 0$ . The sign of the determinant is well-defined.

Applying Proposition 3.1, first note that the rigid-body component  $R_1 \in \text{SE}(3)$  leaves the determinant invariant:

$$\det(M_C^{(z_1)}(i)) = \det(R_1 M_C^{(z_1)}(i)). \quad (22)$$

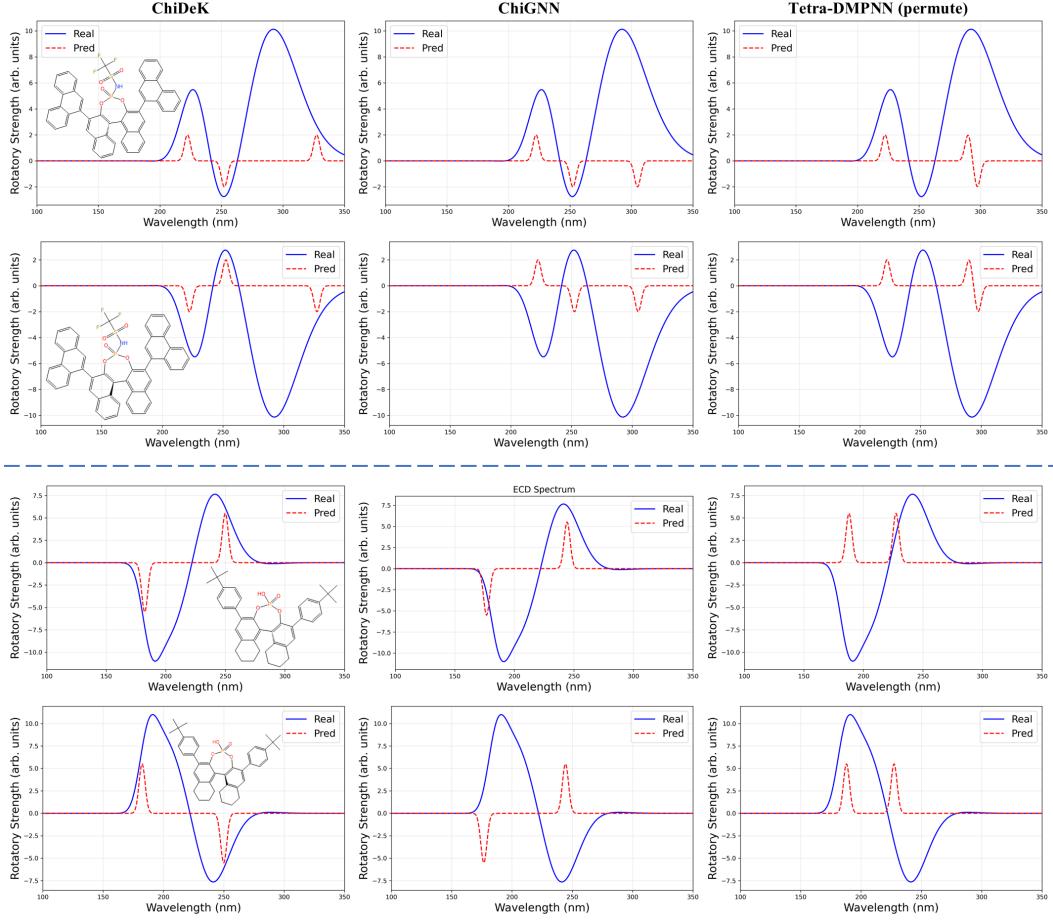


Figure 8: Axial ECD prediction examples of two pairs of enantiomers.

Then the reflection  $R_2 \in O^-(3)$  flips the sign of the determinant:

$$\det(R_2 R_1 M_C^{(z_1)}(i)) = -\det(R_1 M_C^{(z_1)}(i)) = -\det(M_C^{(z_1)}(i)). \quad (23)$$

$R_2 R_1 M_C^{(z_1)}(i)$  is exactly the chirality matrix for atom  $i$  in the enantiomer  $z_2$ , i.e.,

$$M_C^{(z_2)}(i) = R_2 R_1 M_C^{(z_1)}(i). \quad (24)$$

Therefore,

$$\det(M_C^{(z_2)}(i)) = -\det(M_C^{(z_1)}(i)). \quad (25)$$

It follows that the determinants for the two enantiomers have opposite signs. By the sign convention in the lemma (positive determinant  $\mapsto$  ‘R’, negative determinant  $\mapsto$  ‘S’), the stereochemical configuration of  $i$  is ‘R’ when  $\det(M_C(i)) > 0$  and ‘S’ when  $\det(M_C(i)) < 0$ .  $\square$

Figure 9 illustrates how reflection flips the sign of the volume (determinant), yielding an opposite configuration.

## F.2 CHIRAL DETERMINANT KERNEL

*Proof of Lemma 4.1.* Let  $M = M_C(i) \in \mathbb{R}^{3 \times 3}$  and  $W \in \mathbb{R}^{d_p \times 3}$  be full column rank (so the rank is 3). Define  $O := WM \in \mathbb{R}^{d_p \times 3}$  and perform QR decomposition:

$$O = QR, \quad Q^\top Q = I_3, \quad Q \in \mathbb{R}^{d_p \times 3}, \quad R \in \mathbb{R}^{3 \times 3} \text{ upper-triangular.} \quad (26)$$

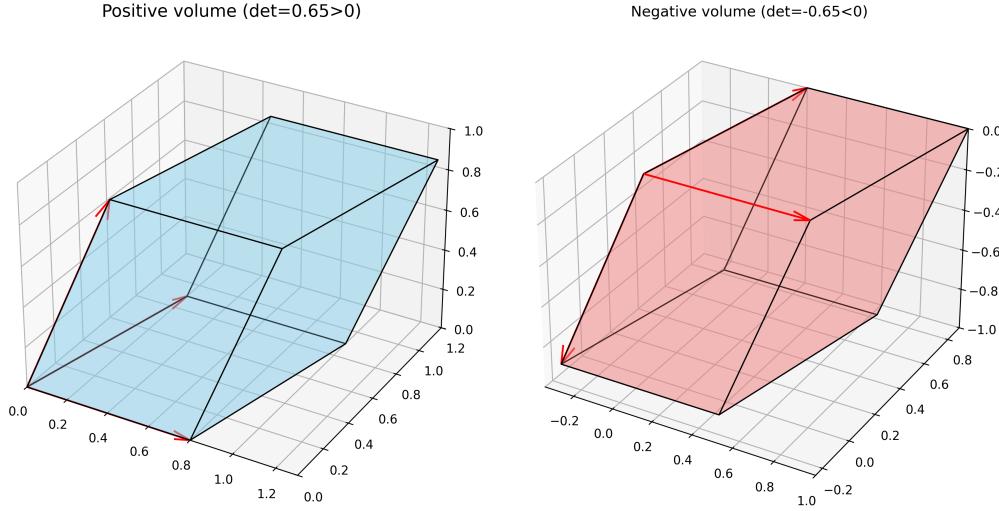


Figure 9: An example of reflection that flips the sign of volume.

Unlike the conventional QR convention, here we allow the diagonal entries of  $\mathbf{R}$  to be negative, as is the case in numerical packages such as PyTorch. Consequently, the sign of  $\det(\mathbf{R})$  is not predetermined but remains well defined.

Consider the Gram matrix

$$\mathbf{G} := \mathbf{W}^\top \mathbf{W}. \quad (27)$$

Since  $\mathbf{W}$  has full column rank,  $\mathbf{G} \succ 0$  and  $\det(\mathbf{G}) > 0$ . We have

$$\mathbf{R}^\top \mathbf{R} = \mathbf{O}^\top \mathbf{O} = (\mathbf{W}\mathbf{M})^\top (\mathbf{W}\mathbf{M}) = \mathbf{M}^\top \mathbf{W}^\top \mathbf{W}\mathbf{M} = \mathbf{M}^\top \mathbf{G}\mathbf{M}. \quad (28)$$

Taking determinants gives

$$\det(\mathbf{R})^2 = \det(\mathbf{M}^\top \mathbf{G}\mathbf{M}) = \det(\mathbf{M})^2 \det(\mathbf{G}). \quad (29)$$

Hence

$$|\det(\mathbf{R})| = |\det(\mathbf{M})| \cdot \sqrt{\det(\mathbf{G})}. \quad (30)$$

To recover the sign, observe that  $\det(\mathbf{O}) = \det(\mathbf{Q}) \det(\mathbf{R})$ . Since  $\mathbf{O} = \mathbf{W}\mathbf{M}$ , this implies

$$\det(\mathbf{R}) = \frac{\det(\mathbf{W}\mathbf{M})}{\det(\mathbf{Q})}. \quad (31)$$

Here  $\det(\mathbf{Q}) \in \{\pm 1\}$  because  $\mathbf{Q}$  has orthonormal columns. Substituting gives

$$\det(\mathbf{R}) = \frac{\det(\mathbf{W}) \det(\mathbf{M})}{\det(\mathbf{Q})}. \quad (32)$$

Combining with equation 30, the factor depending on  $\mathbf{W}$  is

$$\alpha(\mathbf{W}) := \frac{|\det(\mathbf{W})|}{|\det(\mathbf{Q})|} \sqrt{\det((\mathbf{W}^\top \mathbf{W})/(\mathbf{W}^\top \mathbf{W}))}. \quad (33)$$

Since  $\det(\mathbf{Q}) = \pm 1$ , this ambiguity only flips the sign in a way consistent with  $\det(\mathbf{M})$ . Therefore, we can absorb it into the proportionality constant and write

$$\det(\mathbf{R}) = \sqrt{\det(\mathbf{G})} \det(\mathbf{M}) = \alpha(\mathbf{W}) P_{\mathbf{C}}(i), \quad (34)$$

with  $\alpha(\mathbf{W}) = \sqrt{\det(\mathbf{W}^\top \mathbf{W})} > 0$  independent of  $\mathbf{M}$ .

Finally, the invariance properties follow directly: if  $\mathbf{M} \mapsto r\mathbf{M}$  for  $r \in \text{SO}(3)$ , then  $\det(\mathbf{M})$  is unchanged, hence  $P_{\mathbf{W}}(i)$  is invariant. If  $\mathbf{M} \mapsto s\mathbf{M}$  for  $s \in O^-(3)$  with  $\det(s) = -1$ , then  $\det(\mathbf{M})$  flips sign, hence  $P_{\mathbf{W}}(i)$  flips sign as well. Thus  $P_{\mathbf{W}}(i)$  inherits Proposition 3.1 and Lemma 3.1.  $\square$

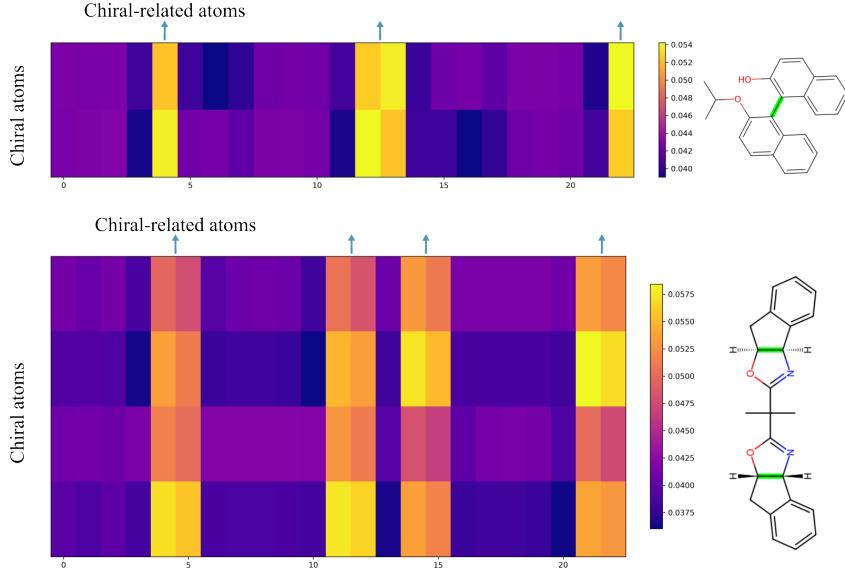


Figure 10: Two examples of cross attention weights. The chiral axis for each molecule is highlighted in green.

## G VISUALIZATION OF THE CHIRAL TRANSFORMER

To better understand how ChiDeK leverages chiral information, we visualize the cross-attention weights between chiral atoms (used as queries) and all other atoms (used as keys and values) on the R/S classification task. Figure 10 illustrates two representative test examples from the axial ECD prediction task, depicting the head-averaged cross-attention weights in the final layer of the chiral transformer. In both cases, we observe that attention weights assigned to chiral-related atoms are higher than those assigned to non-chiral atoms. This demonstrates that ChiDeK successfully prioritizes stereochemically relevant atoms when integrating local chiral information into the global molecular representation. Conversely, non-chiral atoms receive comparatively lower weights, suggesting that the model effectively distinguishes their less critical role in stereospecific interactions.