REM3DI: Learning smooth, chiral 3D molecular representations from equivariant atomistic foundation models

Steffen Wedig STEFFEN@DUNIA.AI
Department of Physics, University of Cambridge, UK
Dunia Innovations, Berlin, Germany
Rokas Elijošius RE344@CAM.AC.UK
Department of Engineering, University of Cambridge, UK
Christoph Schran CS2121@CAM.AC.UK
Department of Physics, University of Cambridge, UK
Lars Leon Schaaf LLS34@CAM.AC.UK
Department of Physics, University of Cambridge, UK

Editors: List of editors' names

Abstract

Accurate molecular descriptors are central to property prediction and similarity screening in computational drug discovery, but prevailing approaches rely on two-dimensional graphs that underuse 3D geometry. We introduce REM3DI, a framework that builds 3D molecular representations by leveraging latent atomic features from equivariant machine-learned interatomic potentials (MLIPs). The local atom centred MLIP features are pooled with permutation-invariant attention aggregation to obtain molecule-level embeddings. To add chiral sensitivity, REM3DI constructs pseudo-scalar channels, which change sign under mirror reflection. Given the scarcity of experimental training data, we pretrain the model via self-supervised denoising. Across small-molecule benchmarks and dedicated chirality tests, REM3DI achieves competitive performance with strong 2D baselines and consistently distinguishes enantiomers on tasks where stereochemistry matters. Importantly, REM3DI sidesteps limitations of 2D cheminformatics descriptors which encode the presence of molecular substructures. REM3DI provides a unified route from physics-based atomic embeddings to versatile, chirality-aware molecular representations for property prediction and virtual screening.

Keywords: Equivariance, pseudo-scalar, ML force fields, molecular representations, drug discovery, atomistic foundation models

1. Introduction

Modern machine-learning interatomic potentials (MLIPs) trained on quantum-mechanical energies and forces learn rich, physics-grounded latent representations of atomic environments. These models are typically used to drive atomistic simulations (e.g., molecular dynamics) to estimate observables such as density, solubility, or protein–ligand binding affinities, often requiring 10^6–10^8 model evaluations over multiple days to obtain statistically significant calculation of observables. In this work we ask a simple question: can the latent features of state-of-the-art atomistic foundation models Batzner et al. (2022); Anderson et al. (2019); Batatia et al. (2024); Kovács et al. (2025); Wood et al. (2025); Plé et al. (2025); Neumann et al. (2024); Bronstein et al. (2021); Rhodes et al. (2025); Fu

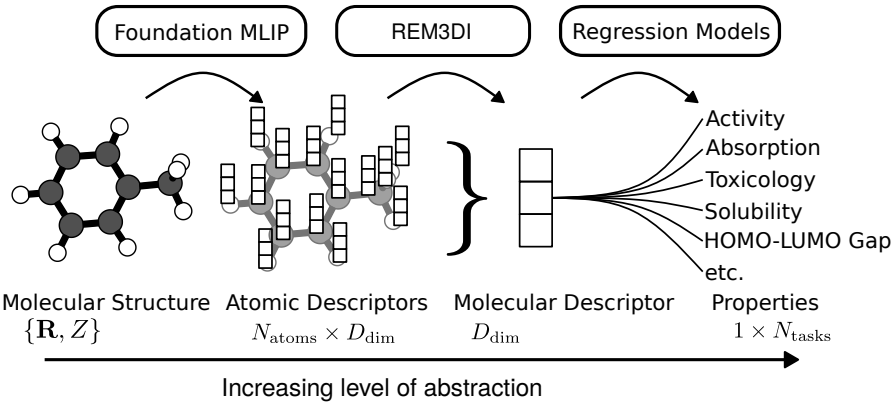

Figure 1: **REM3DI architecture.** Leveraging atom centred features from atomistic foundation models to construct a fixed dimension molecular descriptors for property prediction task.

et al. (2025) be repurposed for direct quantitative structure–property prediction (QSPR) on drug-like molecules, thereby sidestepping costly free-energy calculations Cournia et al. (2017)?

Classical cheminformatics pipelines in drug discovery often rely on fixed molecular descriptors (e.g., ECFP Rogers and Hahn (2010)) that encode hand-crafted patterns or substructures in the 2D molecular graph. Here, we extract invariant per-atom embeddings from pretrained MLIPs and aggregate them into a fixed-size, 3D molecular representation suitable for downstream tasks such as property prediction and virtual screening. In summary, we:

1. Propose a compact, physics-informed, 3D-aware, and transferable molecular descriptor REM3DI that outperforms ECFP fingerprints on pharmaceutical property prediction tasks.

2. Introduce a principled construction of chirality-sensitive features (pseudoscalar channels) from equivariant MLIP embeddings to model stereochemistry of molecules.

3. Introduce a pre-training method for our architecture that exploits large unlabelled datasets to allow for interpretable descriptors of complex chemistry. The lack of requirement for bonding networks for example allows for the treatment of transition metal complexes.

2. REM3DI architecture

Overview An overview of the REM3DI (REpresentation learning for Molecules using 3D Information) architecture is shown in Figure 2a, showing how 3D molecular structures are converted to a fixed length molecular descriptor. Given atomic species and 3D coordinates, a foundation MLIP produces per-atom invariant and equivariant features. The equivariant MLIP features used to create chiral embeddings, which encode stereochemical information of each atoms environment. A structure-encoding block builds a geometric tensor G that can encode either (i) 2D graph topology via featurised random-walk probabilities or directly (ii)

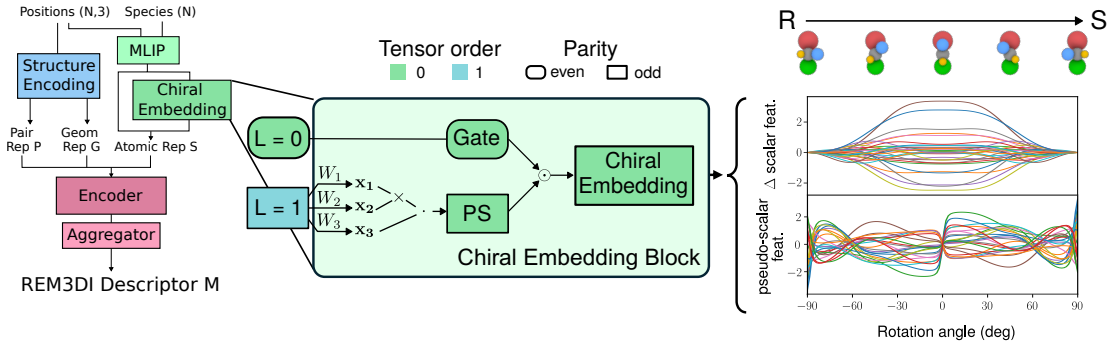

Figure 2: **Constructing pseudo-scalars from equivariant features to capture molecular chirality.** Showing the overall REM3DI architecture (a), focusing on the chiral embedding block. A plot of the chiral atom level embeddings of a molecule as it smoothly transforms to its mirror image, showing the even and odd properties of the relative scalar and pseudo-scalar features (b). Final performance on chiral prediction tasks (c).

3D geometry via radial basis function embeddings of the pairwise distances. The encoder then iteratively refines a pair P and atomic representation S. The encoder outputs are subsequently aggregated into a fixed-size molecular descriptor that supports downstream tasks such as property prediction and similarity screening.

Constructing pseudo-scalars from equivariant features We propose a chiral embedding that encodes stereochemical information explicitly from equivariant latent features. The MLIP scalar features are invariant to inversion and therefore cannot distinguish two enantiomers (mirror images of molecules). By contrast, many pharmacologically relevant properties are not inversion-symmetric, since they depend on interactions with chiral environments such as proteins and solvents. Augmenting invariant descriptors with chiral-sensitive terms (e.g. pseudoscalars that change sign under inversion) allows the model to resolve mirror configurations and predict enantiomer-specific properties.

We construct pseudo-scalars, which are invariant with respect to rotation but change sign upon reflection using the MLIPs equivariant features. Firstly, we construct three tensors a, b, c using learnable linear mixing of node features across the channel dimensions,

$$a_{i,k} = \sum_{k'} W^{(a)}_{k,k'} \tilde{h}_{i,k'}, \quad b_{i,k} = \sum_{k'} W^{(b)}_{k,k'} \tilde{h}_{i,k'}, \quad c_{i,k} = \sum_{k'} W^{(c)}_{k,k'} \tilde{h}_{i,k'}, \tag{1}$$

where k and k' are channel dimensions, h_i are equivariant node features on atom i and W are weights that maintain the equivariant of h as implemented in the `e3nn` library Geiger and Smidt (2022). Taking e3nn tensor products between these three tensors with spherical order $L > 1$ leads to paths that result in odd-scalars. In the case of an L=1 tensor (vector) this is just the vector triple product $(a \times b) \cdot c$, which is known to result in chiral features. The cross product $(a \times b)$ results in an even vector (ie a pseudo vector) which when dot producted with an odd vector (polar/standard vector) results in an odd scalar. Similar paths exists for higher order tensors ($L > 1$) resulting in multiple paths to pseudo scalars. To construct

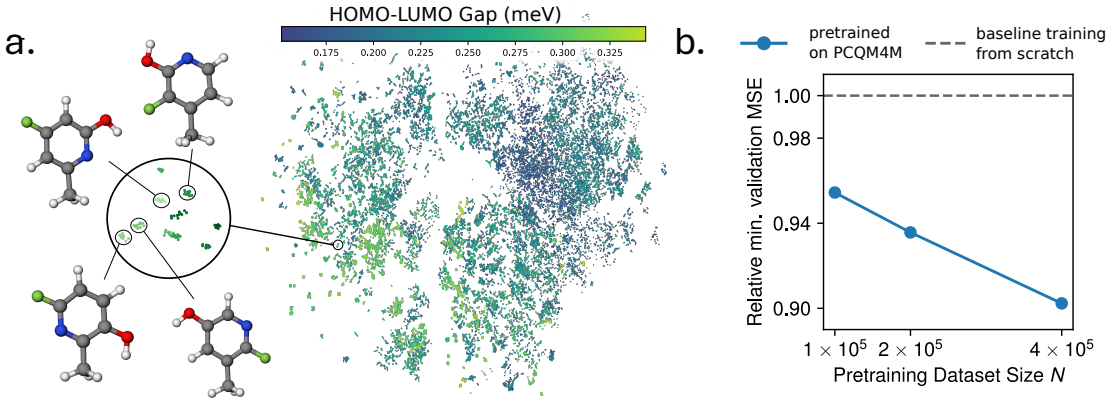

Figure 3: **Denoising pretraining.** Showing a UMAP (a) of the resulting pre-trained molecular representations, with insets showing that chemically similar molecules are clustered together. Performance on HOMO-LUMO gap regression tasks on a 10k subsplit of QM9 (b) as a function of pre-training dataset size.

pseudo scalars one needs a minimum of three tensors and two equivariant tensor products (see proof in Appendix A.4). As shown in Figure 2 (left), an environment-aware gate learns to amplify pseudo-scalar channels, and a subsequent chiral embedding (i.e. linear/identity or bias-free MLP with odd activation function) outputs the chiral embedding.

Figure 2 (right) shows how the chiral embeddings and normal scalar features vary as a molecule smoothly transforms between its mirror image through the rotation of a functional group. As visible the scalar features are identical for their mirror images ($\theta = -90, 90$). Conversely the pseudo features are exactly negative each other. Furthermore, as the molecule transitions through an a-chiral conformation at $\theta = 0$, the pseudo scalar features are strictly zero, showing the correct encoding of the underlying symmetry.

We test REM3DI on chiral retention time dataset Xu et al. (2023) with and without chiral embedding. Without the chiral embedding, REMEDI cannot distinguish mirror images and so predicts identical retention times, leading to high errors. As outlined in Appendix C.4, the errors decrease from 0.458 to 0.237 in log normalised retention times with the chiral embedding. Post workshop submission we will test our method on a recently released Pfizer dataset Elijošius et al. (2025a) to be able to compare to state of the art approaches.

Denoising pretraining of REM3DI Experimental labels for pharmaceutically relevant molecules are scarce, while unlabeled molecules are abundant (e.g., ZINC, PubChem, ChEMBL) as SMILES convertible to graphs; 3D conformers can be generated at scale via MLIP-driven relaxations Irwin et al. (2012); Kim et al. (2024); Mendez et al. (2018); Cohen et al. (2025). We exploit this by pretraining REM3DI with a denoising objective: the encoder produces a molecular descriptor that a decoder uses to reconstruct intentionally corrupted atomic embeddings via descriptor-to-atom cross-attention. Denoising is a standard self-supervised paradigm Ho et al. (2020); Song et al. (2020); prior molecular work perturbs Cartesian coordinates Liao et al. (2024); Liu et al. (2022); Zaidi et al. (2022);

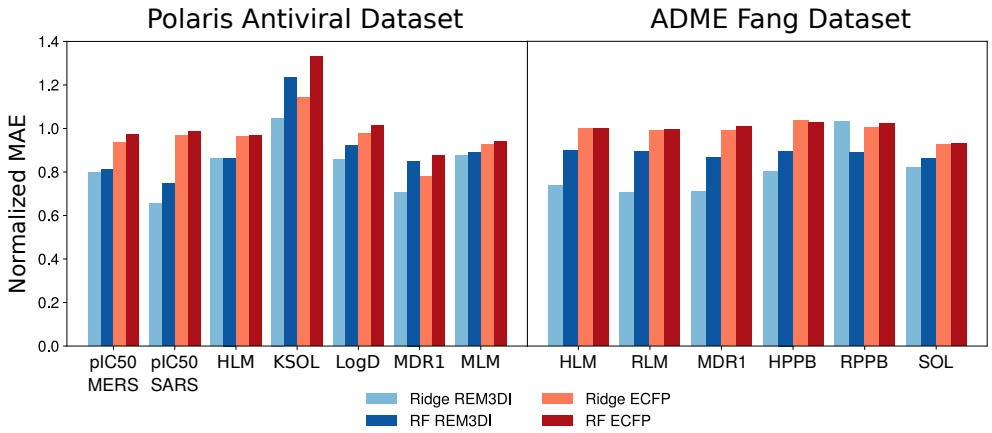

Figure 4: **Pretrained REM3DI descriptors outperform ECFP fingerprints on drug property prediction tasks.** We trained ridge regression and random forest (RF) model on both REM3DI (blue) and ECFP descriptors (red). REM3DI Ridge based regressors perform best in 11 out of 12 tasks.

Wijaya et al. (2024), which can yield unphysical, high-energy structures Feng et al. (2023); Liu et al. (2025). Instead, we inject isotropic Gaussian noise at a fixed level directly in the MACE descriptor space, preserving chemical validity while encouraging robustness to local perturbations. Pretraining on large unlabeled sets (e.g., PCQM4M Nakata and Maeda (2023)) yields rich, interpretable representations: on QM9, a UMAP of our descriptors forms clusters that share functional groups but differ in their arrangement around aromatic rings; conformers from short (200,fs) MD map to nearby regions (Fig. 3).

3. Results: Property prediction in drug discovery

We evaluate pretrained REM3DI on pharmacological property prediction against ECFP Rogers and Hahn (2010). Benchmarks are two polarishub datasets, the Polaris antiviral set MacDermott-Opeskin et al. (2025) and the ADME Fang set Fang et al. (2023), covering 12 ADMET tasks with only hundreds of labels Wijaya et al. (2024). We pretrain on 200k GEOM Drugs molecules Nikitin et al. (2025); Axelrod and Gómez-Bombarelli (2022), compute REM3DI descriptors and ECFPs, and fit task-specific ridge regression Hilt and Seegrist (1977) and random forests Breiman (2001) with hyperparameters tuned via repeated k-fold cross-validation Ash et al. (2024). Across all tasks, REM3DI outperforms ECFP; ridge is typically the strongest regressor. Further experimental results on transition metal complexes and the effect of positional encoding are show in in Appendix C. These are particularly challenging examples for traditional hand-built descriptors, which often assume a fixed covalent bonding pattern which don't apply to transition metal complexes.

4. Conclusion

REM3DI is a physics-informed, smooth, conformation-aware descriptor that aggregates equivariant equivariant MLIP features into a fixed-size molecular embedding. It combines (i) a chiral embedding that forms pseudoscalars from MLIP features, (ii) positional encod-

ings of molecular structure, and (iii) denoising pretraining on large unlabeled data. On pharmacological benchmarks, REM3DI consistently outperforms ECFP while remaining broadly applicable wherever the underlying MLIP is valid (including transition-metal compounds). Small, noisy datasets still limit absolute accuracy, but denoising pretraining yields structured "maps of chemistry" that improve data efficiency.

Outlook We observe strong initial performance and scaling with data set size. To push towards SOTA performance on larger datasets, we suggest various improvements which will be investigated before the workshop.

- **Scale and efficiency.** Increase pre-training data, parameters, and train time; reduce encoder cost with linear/FlashAttention, gradient checkpointing, and fused kernels for near-linear memory and higher throughput. Sweep descriptor width (a 256-D REM3DI already competes with 2048-bit ECFP).
- **Unified pretraining with MLIPs.** Co-pretrain REM3DI and an equivariant MLIP end-to-end with multi-task losses (energy/force consistency + descriptor denoising) to improve sample efficiency and reduce reliance on a single foundation model.
- **Broader domains & evaluations.** Extend to charged/spin systems and periodic materials using MLIPs with electrostatics and PBCs; evaluate retrieval and MD-trajectory embeddings. Emphasize OOD splits, conformer/chirality robustness, and provide reproducible releases (code, configs, checkpoints).

References

Brandon Anderson, Truong Son Hy, and Risi Kondor. Cormorant: Covariant molecular neural networks. *Advances in neural information processing systems*, 32, 2019.

Jeremy R. Ash, Cas Wognum, Raquel Rodríguez-Pérez, Matteo Aldeghi, Alan C. Cheng, Djork-Arné Clevert, Ola Engkvist, Cheng Fang, Daniel J. Price, Jacqueline M. Hughes-Oliver, and W. Patrick Walters. Practically significant method comparison protocols for machine learning in small molecule drug discovery. November 2024. doi: 10.26434/chemrxiv-2024-6dbwv-v2.

Simon Axelrod and Rafael Gómez-Bombarelli. Geom, energy-annotated molecular conformations for property prediction and molecular generation. *Scientific Data*, 9(1), April 2022. ISSN 2052-4463. doi: 10.1038/s41597-022-01288-4.

Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N. Kinch, R. Dustin Schaeffer, Claudia Millán, Hahnbeom Park, Carson Adams, Caleb R. Glassman, Andy DeGiovanni, Jose H. Pereira, Andria V. Rodrigues, Alberdina A. van Dijk, Ana C. Ebrecht, Diederik J. Opperman, Theo Sagmeister, Christoph Buhlheller, Tea Pavkov-Keller, Manoj K. Rathinaswamy, Udit Dalwadi, Calvin K. Yip, John E. Burke, K. Christopher Garcia, Nick V. Grishin, Paul D. Adams, Randy J. Read, and David Baker. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876, August 2021. ISSN 1095-9203. doi: 10.1126/science.abj8754.

David Balcells and Bastian Bjerkem Skjelstad. tmqm dataset—quantum geometries and properties of 86k transition metal complexes. *Journal of Chemical Information and Modeling*, 60(12):6135–6146, November 2020. ISSN 1549-960X. doi: 10.1021/acs.jcim.0c01041.

Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, Matthew Avaylon, William J. Baldwin, Fabian Berger, Noam Bernstein, Arghya Bhowmik, Samuel M. Blau, Vlad Cărare, James P. Darby, Sandip De, Flaviano Della Pia, Volker L. Deringer, Rokas Elijošius, Zakariya El-Machachi, Fabio Falcioni, Edvin Fako, Andrea C. Ferrari, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Clare P. Grey, Petr Grigorev, Shuang Han, Will Handley, Hendrik H. Heenen, Kersti Hermansson, Christian Holm, Jad Jaafar, Stephan Hofmann, Konstantin S. Jakob, Hyunwook Jung, Venkat Kapil, Aaron D. Kaplan, Nima Karimitari, James R. Kermode, Namu Kroupa, Jolla Kullgren, Matthew C. Kuner, Domantas Kuryla, Guoda Liepuoniute, Johannes T. Margraf, Ioan-Bogdan Magdău, Angelos Michaelides, J. Harry Moore, Aakash A. Naik, Samuel P. Niblett, Sam Walton Norwood, Niamh O'Neill, Christoph Ortner, Kristin A. Persson, Karsten Reuter, Andrew S. Rosen, Lars L. Schaaf, Christoph Schran, Benjamin X. Shi, Eric Sivonxay, Tamás K. Stenczel, Viktor Svahn, Christopher Sutton, Thomas D. Swinburne, Jules Tilly, Cas van der Oord, Eszter Varga-Umbrich, Tejs Vegge, Martin Vondrák, Yangshuai Wang, William C. Witt, Fabian Zills, and Gábor Csányi. A foundation model for atomistic materials chemistry, 2024.

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(1):2453, 2022.

Jan C. Brammer, Gerd Blanke, Claudia Kellner, Alexander Hoffmann, Sonja Herres-Pawlis, and Ulrich Schatzschneider. Tucan: A molecular identifier and descriptor applicable to the whole periodic table from hydrogen to oganesson. *Journal of Cheminformatics*, 14 (1), September 2022. ISSN 1758-2946. doi: 10.1186/s13321-022-00640-5.

Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001. ISSN 0885-6125. doi: 10.1023/a:1010933404324.

Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.

Orion Cohen, Janosh Riebesell, Rhys Goodall, Adeesh Kolluru, Stefano Falletta, Joseph Krause, Jorge Colindres, Gerbrand Ceder, and Abhijeet S. Gangan. Torchsim: An efficient atomistic simulation engine in pytorch, 2025.

Zoe Cournia, Bryce Allen, and Woody Sherman. Relative binding free energy calculations in drug discovery: Recent advances and practical considerations. *Journal of Chemical Information and Modeling*, 57(12):2911–2937, December 2017. ISSN 1549-960X. doi: 10.1021/acs.jcim.7b00564.

Tri Dao, Daniel Y. Fu, Stefano Ermon, Atri Rudra, and Christopher Ré. Flashattention: Fast and memory-efficient exact attention with io-awareness, 2022.

Rokas Elijošius, Emma King-Smith, Felix A. Faber, Louise Bernier, Simon Berritt, William P. Farrell, Xinjun Hou, Jacquelyn L. Klug-McLeod, Jason Mustakis, Neal W. Sach, Qingyi Yang, Roger M. Howard, and Alpha A. Lee. Predictive design of crystallographic chiral separation. *Nature Communications*, 16(1), August 2025a. ISSN 2041-1723. doi: 10.1038/s41467-025-62825-4.

Rokas Elijošius, Fabian Zills, Ilyes Batatia, Sam Walton Norwood, Dávid Péter Kovács, Christian Holm, and Gábor Csányi. Zero shot molecular generation via similarity kernels. *Nature Communications*, 16(1), July 2025b. ISSN 2041-1723. doi: 10.1038/s41467-025-60963-3.

Cheng Fang, Ye Wang, Richard Grater, Sudarshan Kapadnis, Cheryl Black, Patrick Trapa, and Simone Sciabola. Prospective validation of machine learning algorithms for absorption, distribution, metabolism, and excretion prediction: An industrial perspective. *Journal of Chemical Information and Modeling*, 63(11):3263–3274, May 2023. ISSN 1549-960X. doi: 10.1021/acs.jcim.3c00160.

Shikun Feng, Yuyan Ni, Yanyan Lan, Zhi-Ming Ma, and Wei-Ying Ma. Fractional denoising for 3d molecular pre-training. July 2023. doi: 10.48550/ARXIV.2307.10683.

J. Thorben Frank, Stefan Chmiela, Klaus-Robert Müller, and Oliver T. Unke. Euclidean fast attention: Machine learning global atomic representations at linear cost, 2024.

Xiang Fu, Brandon M. Wood, Luis Barroso-Luque, Daniel S. Levine, Meng Gao, Misko Dzamba, and C. Lawrence Zitnick. Learning smooth and expressive interatomic potentials for physical property prediction, 2025.

Mario Geiger and Tess Smidt. e3nn: Euclidean neural networks, 2022.

Stephen R Heller, Alan McNaught, Igor Pletnev, Stephen Stein, and Dmitrii Tchekhovskoi. Inchi, the iupac international chemical identifier. *Journal of Cheminformatics*, 7(1), May 2015. ISSN 1758-2946. doi: 10.1186/s13321-015-0068-4.

Donald E. Hilt and Donald W. Seegrist. *Ridge, a computer program for calculating ridge regression estimates /*. Dept. of Agriculture, Forest Service, Northeastern Forest Experiment Station,, 1977. doi: 10.5962/bhl.title.68934.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf.

John J. Irwin, Teague Sterling, Michael M. Mysinger, Erin S. Bolstad, and Ryan G. Coleman. Zinc: A free tool to discover chemistry for biology. *Journal of Chemical Information and Modeling*, 52(7):1757–1768, June 2012. ISSN 1549-960X. doi: 10.1021/ci3001277.

Sunghwan Kim, Jie Chen, Tiejun Cheng, Asta Gindulyte, Jia He, Siqian He, Qingliang Li, Benjamin A Shoemaker, Paul A Thiessen, Bo Yu, Leonid Zaslavsky, Jian Zhang, and Evan E Bolton. Pubchem 2025 update. *Nucleic Acids Research*, 53(D1):D1516–D1525, November 2024. ISSN 1362-4962. doi: 10.1093/nar/gkae1059.

Dávid Péter Kovács, J. Harry Moore, Nicholas J. Browning, Ilyes Batatia, Joshua T. Horton, Yixuan Pu, Venkat Kapil, William C. Witt, Ioan-Bogdan Magdău, Daniel J. Cole, and Gábor Csányi. Mace-off: Short-range transferable machine learning force fields for organic molecules. *Journal of the American Chemical Society*, 147(21):17598–17611, May 2025. ISSN 1520-5126. doi: 10.1021/jacs.4c07099.

Yi-Lun Liao, Tess Smidt, Muhammed Shuaibi, and Abhishek Das. Generalizing denoising to non-equilibrium structures improves equivariant force fields, 2024.

Shengchao Liu, Hongyu Guo, and Jian Tang. Molecular geometry pretraining with se(3)-invariant denoising distance matching. June 2022. doi: 10.48550/ARXIV.2206.13602.

Yurou Liu, Jiahao Chen, Rui Jiao, Jiangmeng Li, Wenbing Huang, and Bing Su. DenoiseVAE: Learning molecule-adaptive noise distributions for denoising-based 3d molecular pre-training. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=ym7pr83XQr.

Hugo MacDermott-Opeskin, Jenke Scheen, Cas Wognum, Joshua T Horton, Devany West, Alexander Matthew Payne, Maria A Castellanos, Sean Colby, Edward Griffen, David Cousins, Jessica Stacey, Lauren Reid, Jasmin Cara Aschenbrenner, Daren Fearon, Blake Balcomb, Peter Marples, Charles W.E. Tomlinson, Ryan Lithgo, Max Winokan, Haim Barr, Noa Lahav, Michael Lavi, Shirley Duberstein, Galit Cohen, Gwendolyn Fate, Bruce Lefker, Ralph Robinson, Tamas Szommer, Nick Lynch, Mallory Tollefson, Cynthia Xu, Jonny Hsu, Julien St-Laurent, Honore Etsmoberg, Lu Zhu, Andrew Quirke, Mohamed Iliyas Abdul Haleem, Irfan Alibay, Gunjan Baid, Benjamin Birnbaum, Kevin Bishop, Hugo Bohorquez, Ashmita Bose, C. J. Brown, Jackson Burns, Lianjin Cai, Ruel Cedeno, Vladimir Chupakhin, Finlay Clark, Daniel Cole, Carles Corbi-Verge, Muhammad Danial, Alec Davi, Wim Dehaen, Niklas Piet Doering, Alexis Dougha, Bryce Eakin, Anatol Ehrlich, Rokas Elijosius, Jozef Fülöp, Anthony Gitter, Yaowen Gu, Teresa Head-Gordon, Ellena Jiang, Benjamin Kaminow, Sina Khosravi, Asma Feriel Khoualdi, Eelke Bart Lenselink, Zhirong Liu, Yue Liu, Sijie Liu, Yizhou Ma, Patrick Maher, Imke Mayer, Antonia Mey, Floriane Montanari, Taoyu Niu, Ryusei Ogino, Ashok Palaniappan, Xiaolin Pan, Auro Patnaik, Long-Hung Pham, Luis Pinto, Justin Purnomo, Alexander Rich, Lars Schaaf, Christoph Schran, Satya Pratik Srivastava, Kunyang Sun, Zhaoxi Sun, Valerij Talagayev, Balamurugan Thirukonda Subramanian Balakrishnan, Alexandre Tkatchenko, Wojtek Treyde, Austin Tripp, Nopsinth Vithayapalert, Yingze Wang, Azmine Toushik Wasi, Steffen Wedig, Bofei Xu, Weijun Zhou, Frank von Delft, Alpha Lee, Karla Kirkegaard, Peter Sjö, James Fraser, and John D. Chodera. A computational community blind challenge on pan-coronavirus drug discovery data. July 2025. doi: 10.26434/chemrxiv-2025-zd9mr.

Eric Meggers. Targeting proteins with metal complexes. *Chemical Communications*, (9): 1001, 2009. ISSN 1364-548X. doi: 10.1039/b813568a.

David Mendez, Anna Gaulton, A Patrícia Bento, Jon Chambers, Marleen De Veij, Eloy Félix, María Paula Magariños, Juan F Mosquera, Prudence Mutowo, Michał Nowotka, María Gordillo-Marañón, Fiona Hunter, Laura Junco, Grace Mugumbate, Milagros Rodriguez-Lopez, Francis Atkinson, Nicolas Bosc, Chris J Radoux, Aldo Segura-Cabrera, Anne Hersey, and Andrew R Leach. Chembl: towards direct deposition of bioassay data. *Nucleic Acids Research*, 47(D1):D930–D940, November 2018. ISSN 1362-4962. doi: 10.1093/nar/gky1075.

Amil Merchant, Simon Batzner, Samuel S. Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624(7990): 80–85, November 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06735-9.

Katja Dralle Mjos and Chris Orvig. Metallodrugs in medicinal inorganic chemistry. *Chemical Reviews*, 114(8):4540–4563, January 2014. ISSN 1520-6890. doi: 10.1021/cr400460s.

Maho Nakata and Toshiyuki Maeda. Pubchemqc b3lyp/6-31g*//pm6 dataset: the electronic structures of 86 million molecules using b3lyp/6-31g* calculations, 2023.

Aditya Nandy, Chenru Duan, Michael G. Taylor, Fang Liu, Adam H. Steeves, and Heather J. Kulik. Computational discovery of transition-metal complexes: From high-throughput screening to machine learning. *Chemical Reviews*, 121(16):9927–10000, July 2021. ISSN 1520-6890. doi: 10.1021/acs.chemrev.1c00347.

Mark Neumann, James Gin, Benjamin Rhodes, Steven Bennett, Zhiyi Li, Hitarth Choubisa, Arthur Hussey, and Jonathan Godwin. Orb: A fast, scalable neural network potential. October 2024. doi: 10.48550/ARXIV.2410.22570.

Filipp Nikitin, Ian Dunn, David Ryan Koes, and Olexandr Isayev. Geom-drugs revisited: Toward more chemically accurate benchmarks for 3d molecule generation, 2025.

Thomas Plé, Olivier Adjoua, Anouar Benali, Evgeny Posenitskiy, Corentin Villot, Louis Lagardère, and Jean-Philip Piquemal. A foundation model for accurate atomistic simulations in drug design. May 2025. doi: 10.26434/chemrxiv-2025-f1hgn-v3.

Maria H. Rasmussen, Magnus Strandgaard, Julius Seumer, Laura K. Hemmingsen, Angelo Frei, David Balcells, and Jan H. Jensen. Smiles all around: structure to smiles conversion for transition metal complexes. *Journal of Cheminformatics*, 17(1), April 2025. ISSN 1758-2946. doi: 10.1186/s13321-025-01008-1.

Benjamin Rhodes, Sander Vandenhaute, Vaidotas Šimkus, James Gin, Jonathan Godwin, Tim Duignan, and Mark Neumann. Orb-v3: atomistic simulation at scale, 2025.

David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50(5):742–754, April 2010. ISSN 1549-960X. doi: 10.1021/ci100050t.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations, 2020.

Hannes Stärk, Dominique Beaini, Gabriele Corso, Prudencio Tossou, Christian Dallago, Stephan Günnemann, and Pietro Liò. 3d infomax improves gnns for molecular property prediction. 2021. doi: 10.48550/ARXIV.2110.04126.

Kevin Tirta Wijaya, Minghao Guo, Michael Sun, Hans-Peter Seidel, Wojciech Matusik, and Vahid Babaei. Two-stage pretraining for molecular property prediction in the wild, 2024.

Brandon M. Wood, Misko Dzamba, Xiang Fu, Meng Gao, Muhammed Shuaibi, Luis Barroso-Luque, Kareem Abdelmaqsoud, Vahe Gharakhanyan, John R. Kitchin, Daniel S. Levine, Kyle Michel, Anuroop Sriram, Taco Cohen, Abhishek Das, Ammar Rizvi, Sushree Jagriti Sahoo, Zachary W. Ulissi, and C. Lawrence Zitnick. Uma: A family of universal models for atoms, 2025.

Hao Xu, Jinglong Lin, Dongxiao Zhang, and Fanyang Mo. Retention time prediction for chromatographic enantioseparation by quantile geometry-enhanced graph neural network. *Nature Communications*, 14(1), May 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-38853-3.

Sheheryar Zaidi, Michael Schaarschmidt, James Martens, Hyunjik Kim, Yee Whye Teh, Alvaro Sanchez-Gonzalez, Peter Battaglia, Razvan Pascanu, and Jonathan Godwin. Pre-training via denoising for molecular property prediction, 2022.

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. May 2022. doi: 10.26434/chemrxiv-2022-jjm0j.

Appendix A. Pseudo-scalar construction

Below we outline the required background for the pseudo scalar construction in the chiral embedding block. Firstly we define a spherical tensor expressed in the spherical harmonic basis and then outline the tensor products needed to construct pseudo scalars which change sign upon inversion but remain invariant to rotations.

A.1. Spherical tensors and their transformation law

A spherical tensor of rank l in a spherical harmonic basis has $2l+1$ components indexed by m

$$T^{(l)} = \{T^{(l)}_m\}^l_{m=-l}, \tag{2}$$

that transform under a rotation $R \in SO(3)$ as

$$(R \cdot T)^{(l)}_m = \sum_{m'=-l}^{l} D^{(l)}_{mm'}(R)\, T^{(l)}_{m'}, \tag{3}$$

where $D^{(l)}(R)$ is the Wigner D-matrix, i.e. the $(2l+1)$-dimensional irreducible representation of $SO(3)$.

A.2. Parity and the orthogonal group

To extend to the full orthogonal group $O(3)$ we must specify the action of inversion

$$I: \ \boldsymbol{r} \mapsto -\boldsymbol{r}. \tag{4}$$

This action is encoded by a parity label $\pi \in \{+, -\}$:

$$I \cdot T^{(l,\pi)}_m = \pi\, T^{(l,\pi)}_m. \tag{5}$$

For tensors constructed from ordinary proper tensors the parity is give by

$$\pi = (-1)^l, \tag{6}$$

such that eg. scalars (T^+_0) are even and vectors (T^-_1) are odd. We define pseudo tensors as those where the parity is $\pi = (-1)^{l+1}$,. As such chiral features are an example of a pseudo-scalar T^-_0.

A.3. Clebsch–Gordan decomposition in words

Given two spherical tensors $T^{(l_1,\pi_1)}$ and $U^{(l_2,\pi_2)}$, their tensor product decomposes into irreducible spherical tensors of various ranks:

$$T^{\pi_1}_{l_1} \otimes U^{\pi_2}_{l_2} \ \cong \bigoplus_{l_3=|l_1-l_2|}^{l_1+l_2} V^{\pi_1\pi_2}_{l_3}. \tag{7}$$

where the possible ranks l_3 range from $|l_1 - l_2|$ up to $l_1 + l_2$, in integer steps also called the triangle rule. The parity is the product $\pi_1\pi_2$.

A.4. Constructing a pseudoscalar

Suppose we want a pseudoscalar, V_0^-. By the triangle rule, rank $l_3 = 0$ can only appear if $l_1 = l_2$. The parity of the result is $\pi_1\pi_2$. To obtain $\pi = -$, we need opposite parities. If only one parity per l is available in the primitive feature set (e.g. only the "normal" proper tensors with $\pi = (-1)^l$), then this is impossible. Hence we must firstly construct higher order $l_1, l_2 \geq 1$ pseudo tensors from models containing only polar tensor representation. We couple a tensor $T^{(l,\pi)}$ with a polar tensor $U^{(l_2,-)}$

$$(l, \pi) \otimes (l_2, -) \supset (l, -\pi), \tag{8}$$

which produces a new tensor of the same rank l but flipped parity. Couple this tensor with with another tensor $W^{(l,\pi)}$ of the same order l to produce a scalar $l = 0$:

$$(l, -\pi) \otimes (l, \pi) \supset (0, -). \tag{9}$$

An example of this for a set of 3 tensors of order $l_1 = l_2 = l_3 = 1$ is a simple vector triple product $t \cdot (u \times w)$.

Appendix B. Related Work

2D cheminformatics descriptors. Cheminformatics fingerprints remain strong baselines for molecular property prediction and similarity screenings. Extended Connectivity Fingerprints (ECFP) iteratively hash atom neighborhoods on the molecular graph and fold identifiers into a fixed-length bit vectorRogers and Hahn (2010). The ECFP4 (radius 2) is especially common and are used with a variety of regression models, traditionally random forests Breiman (2001). These descriptors are fast and interpretable, but do not consider 3D geometry. Further, as they rely on the definition of a molecular graph, they particularly struggle with systems with non-covalent bonds such as transition metal complexes Rasmussen et al. (2025).

Machine-learned interatomic potentials (MLIPs) E(3)-equivariant message-passing MLIPs approximate potential-energy surfaces with near-DFT accuracy across broad chemical space. Architecturally, they (i) build local, learned atomic descriptors via message passing within cutoffs and (ii) read out energies with a permutation-invariant head. Forces come from gradients of the energy wrt. the atomic positions to ensure energy conservation. Internally, node features are decomposed into spherical tensors (scalars/vectors/higher-order tensors) that transform with rotation according to the wigner D matrices, the irreducible representations of the SO(3) group (see Appendix A for more details. These learned atomic embeddings encode strong physical priors (symmetry, locality, smoothness) and provide rich, transferable descriptions of local environments and can be reused beyond energy/force prediction, e.g.. in generative models Elijošius et al. (2025b); Baek et al. (2021).

Atomistic foundation models Recently, MLIPs have been trained on large quantum mechanical datasets of energies and forces Merchant et al. (2023); Batatia et al. (2024); Kovács et al. (2025); Wood et al. (2025); Plé et al. (2025); Neumann et al. (2024); Rhodes et al. (2025); Fu et al. (2025). These models generalise to very diverse chemistries without

the need for expensive quantum mechanical calculations and training from scratch. We leverage latent node features from these pre-trained models as atom centred representations.

Learned 3D molecular representations. Recent molecular representation learning architectures incorporate geometry explicitly. 3D Infomax jointly learns from a 2D molecular graph and a 3D positional graph by maximizing mutual information between the two views Stärk et al. (2021). UniMol Zhou et al. (2022) injects pairwise distance encodings to bias attention toward spatially proximate atoms and is pretrained on $\sim$ 209M conformers via masked-atom prediction and 3D position denoising before fine-tuning on property tasks, but still relies on manually selected atom features. We use a more expressive local description, transfer the atomic representation from atomistic foundation models and introduce a novel chiral embedding.

Appendix C. Additional Experiments

This section outlines additional experiments, including validation of various positional encoding approaches, as-well as performance on transition metal complexes.

C.1. Positional encoding

Without positional encoding only local geometry feeds into the molecular descriptor through the MLIP graph neural network. To ensure the relative position of the highly expressive node features can be captured by the architecture, we extend the model by injecting positional information via a pair representation into the attention calculation, which is conditioned on the location and interaction of neighbours. The positional information is contained in the pair representation P which biases the attention calculation in the update of the atomic representation S.

$$\Delta S = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}} + \text{MLP}(P)\right)V. \tag{10}$$

Some positional information is already contained in MLIP descriptors which cover short-range physics. Including the positional encodings to bias the attention reflects both local and long-range effects (conjugation, electrostatics, and π–π stacking) because REM3DI uses all-pairs global attention. This also allows the model to learn distance-dependent weights for atomic neighbourhoods prior to aggregation into the molecular descriptor. The downside of calculating global pairwise distances is its $\mathcal{O}(N^2)$ scaling in compute time, where N is the number of atoms in the system. This is mitigated by the fact that the time complexity of our architecture is already $\mathcal{O}(N^2)$ due to the self-attention calculation, so computing pairwise distances does not negatively impact scaling. We discuss how to reduce the scaling of the self-attention and pairwise distance calculation in the outlook by drawing inspiration from linear scaling attention architectures Frank et al. (2024); Dao et al. (2022).

C.2. Interpretable descriptors for transition metal complexes

REM3DI is designed to operate on a wide variety of atomistic systems without relying on hand-crafted, domain-specific features. This is particularly useful for transition-metal

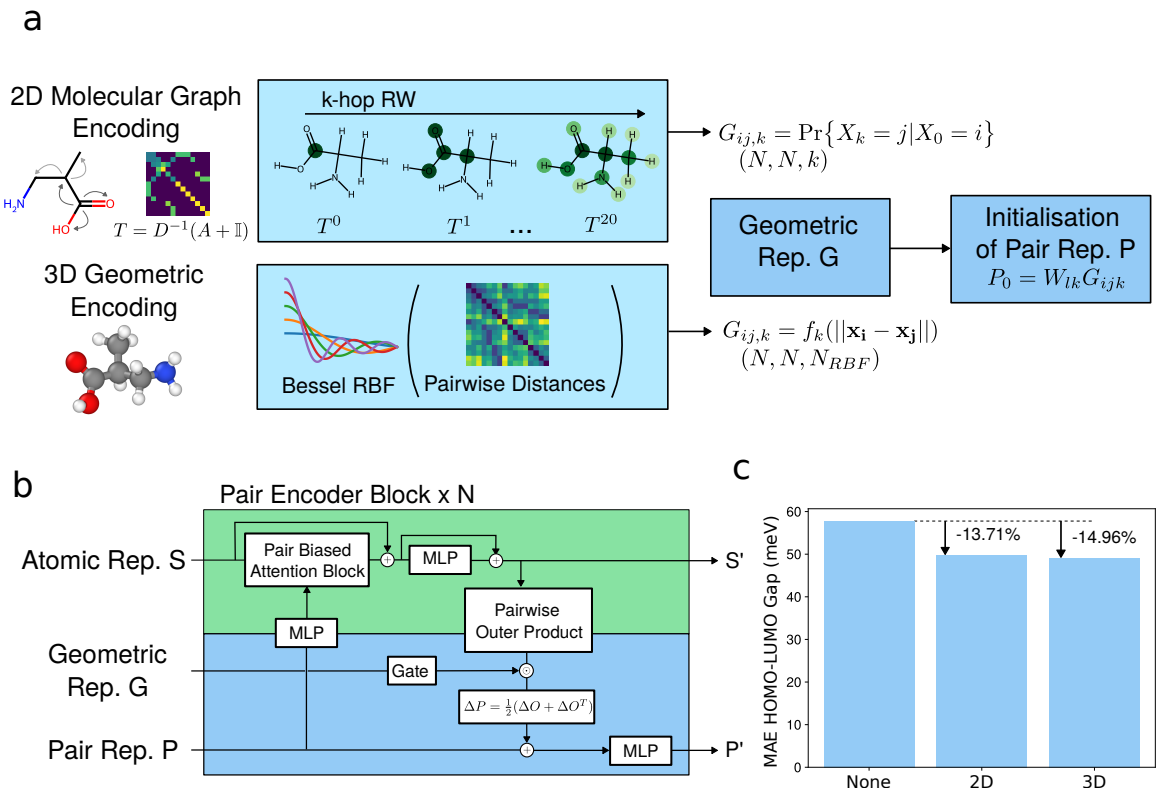

Figure 5: **Positional encoding in encoder.** Showing the construction of the geometric representation G utilising the 2D molecular graph or the 3D distance matrix (a) and its use in the Pair Encoder Block (b). Improved performance with positional encodings both from the 2D graph structure and 3D distance matrix (c).

complexes (TMCs), where chemical diversity is enormous: metals from different rows (3d/4d/5d), a wide range of oxidation and spin states, many ligand classes, and multiple coordination geometries. Hand-built descriptors often assume a fixed covalent bonding pattern. In TMCs, the metal–ligand interaction is typically coordinate (dative) bonding, frequently involves multi-center π interactions, and the very notion of a unique "bond order" is ambiguous Brammer et al. (2022). As a result, classical cheminformatics descriptors that depend on well-defined covalent graphs struggle or require restrictive assumptions (e.g. a small set of metals/ligands and a few interatomic distances chosen a priori, or removing all dative bonds as in InChi) Nandy et al. (2021); Heller et al. (2015); Rasmussen et al. (2025). The lack of general cheminformatics descriptors hinders the further exploitation of transition metal compounds as pharmaceutical agents, particularly as anti-cancer drugs Meggers (2009); Mjos and Orvig (2014). REM3DI captures important chemical trends in transition metal complexes in the tmQM dataset Balcells and Skjelstad (2020). The UMAP projection in figure 6 show that clusters based on the transition metal centre species, the ligand chemistry, or the geometry form. Specific clusters, such as the Rh/Ir/Re/Ru half-sandwich complexes show that the model can also capture complex interactions between

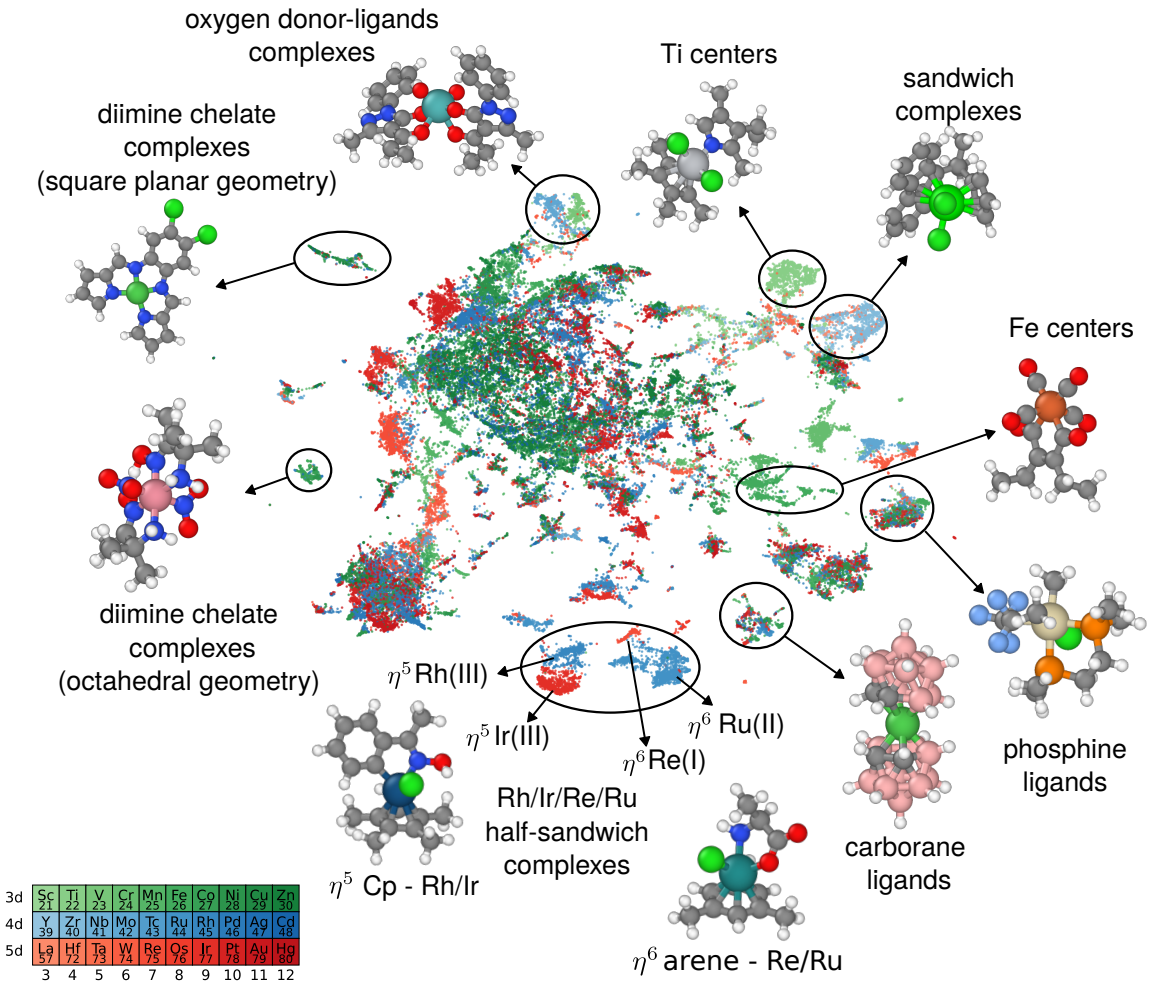

Figure 6: **Transition metal complexes.** UMAP of the molecular descriptors after denoising pre-training without ground-truth labels. Insets highlight the clustering of chemically similar compounds, that group TMCs by geometry, ligand chemistry, and metal centre identity.

different important structural features of the transition metal complexes. This shows that REM3DI is transferable to any system that can be modelled with the underlying MLIP, without having to hand pick relevant descriptors for systems under investigation.

C.3. Quantum Mechanical dataset

We present regression results on the QM9 dataset in Table 1. The performance difference between the two positional encodings is small, with the 3D positions marginally outperforming the 2D encodings. The multitask model shows a 10.8% drop in performance, which shows that here training on multiple tasks in does not lead to performance gains through exploiting shared information between different tasks, but rather to task interference.

Table 1: Performance of REM3DI benchmarked against GNNs on QM9. REM3DI shows reasonable but not competetive performance on response/thermodynamic properties.

	SchNet	E(n)-GNN	DimeNet++	SphereNet	PaiNN	TorchMD-NET	Pre-trained			
GNS-TAT+NN	REM3DI									
+ 3D	REM3DI									
+ 2D	REM3DI									
Multitask										
$\Delta\varepsilon$ (meV)	63.0	48.0	32.6	32.3	45.7	36.1	22.0	47.3	47.8	52.4
μ (D)	0.033	0.029	0.030	0.027	0.012	0.011	0.016			0.0620
α (a_0^3)	0.235	0.071	0.043	0.047	0.045	0.059	0.040			0.132
ZPVE (meV)	1.700	1.550	1.210	1.120	1.280	1.840	1.018			9.90
c_V mol^{-1} K^{-1}	0.033	0.031	0.023	0.022	0.024	0.026	0.020			0.0675

C.4. Chiral dataset

Please see the extended chiral plots in Figure 7 showing the errors of various models with and without chiral embedding.

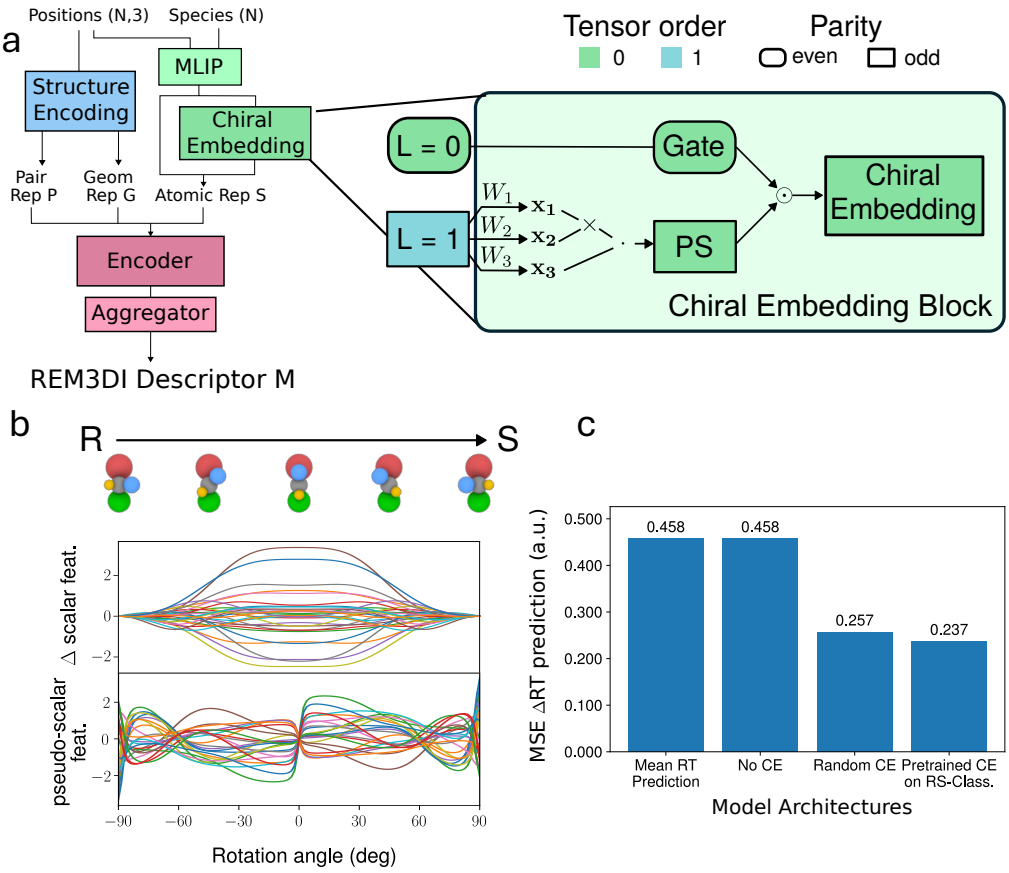

Figure 7: **Constructing pseudo-scalars from equivariant features to capture molecular chirality.** Showing the overall REM3DI architecture (a), focusing on the chiral embedding block. A plot of the chiral atom level embeddings of a molecule as it smoothly transforms to its mirror image, showing the even and odd properties of the relative scalar and pseudo-scalar features (b). Final performance on chiral prediction tasks (c).