

AI-Driven Discovery and Characterisation of Ferroelectric Photovoltaics

Matthew A. H. Walker¹, Cyprien Bone¹, Bradley A. A. Martin¹, Keith T. Butler¹

¹Department of Chemistry, University College London, 20 Gower Street, London, UK

matthew.walker.21@ucl.ac.uk

Materials that combine spontaneous polarisation with strong optical absorption are promising candidates for next-generation photovoltaic (PV) devices capable of exceeding the efficiency limits of conventional thin-film technologies[1]. However, most strongly polar materials are wide-band-gap oxides with poor absorption across the UV-visible spectrum, limiting their practical PV performance. Previous discovery efforts have largely relied on screening existing materials databases, restricting exploration of chemical space to known compounds.

Recent advances in generative machine learning[2-3] offer a route beyond database-driven discovery. Conditioned generative models can propose crystal structures with targeted properties, enabling exploration of previously inaccessible regions of materials space. In parallel, machine-learned interatomic potentials (MLIPs) enable accurate atomistic simulations at a fraction of the computational cost of *ab initio* methods.

Here, we present an end-to-end workflow for the discovery and characterisation of ferroelectric photovoltaic materials. We employ *CrystalLLM- π* [4], a large language model for crystal generation, conditioned to propose materials with polar space groups and high predicted photovoltaic efficiencies. Candidate structures are then evaluated using *MACE-Field*[5], an extension of the equivariant MLIP capable of learning electric field-dependent energy functionals. By relaxing the polarisation, we identify the corresponding high-symmetry, non-polar reference phases, allowing the change in polarisation to be computed along a well-defined branch of the polarisation lattice: an essential requirement for characterisation of ferroelectric materials. We further screen candidates using machine learning and density functional theory (DFT) calculations in high-accuracy, carbon-efficient combinations[6]. Finally, we explore the most promising candidates in greater depth using hybrid DFT.

This work demonstrates how modern machine-learning tools can be combined to both propose and rigorously evaluate novel materials, enabling targeted discovery of ferroelectric photovoltaic compounds beyond existing materials databases.

References

- [1] K.T. Butler, J. M. Frost, A. Walsh, *Energy Environ. Sci.*, 2015, **8**, 838-848.
- [2] L. M. Antunes, K. T. Butler, R. Grau-Crespo, *Nat. Commun.*, 2024, **15**, 10570.
- [3] H. Park, A. Onwuli, A. Walsh, *Nat. Commun.*, 2025, **16**, 4379.
- [4] C. Bone, M. A. H. Walker, K. Leng, L. M. Antunes, R. Grau-Crespo, A. Aligayev, J. Dominguez, K. T. Butler, *arXiv:2511.21299* 2025.
- [5] B. A. A. Martin, A. M. Ganose, V. Kapil, T. Li, K. T. Butler, *arXiv:2508.17870*, 2026.
- [6] M. Walker, K. T. Butler, *Mater. Horiz.* 2026, **13**, 1993-2004.

Figures

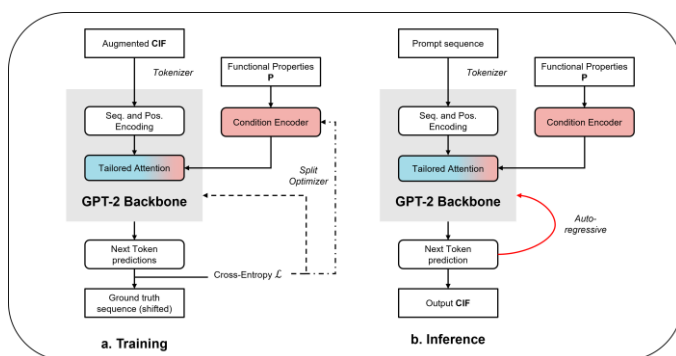


Figure 1. *CrystalLLM- π* architecture from ref. [4]

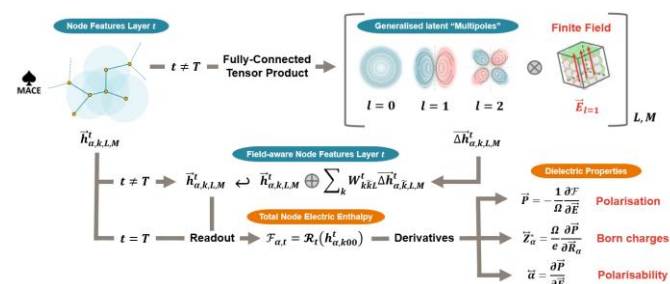


Figure 2. *MACE-Field* architecture from ref. [5]