Discovery of Sustainable Refrigerants through Physics-Informed RL Fine-Tuning of Sequence Models

Adrien Goldszal [1 2] **Diego Calanzone** [2 3] **Vincent Taboga** [2 3] **Pierre-Luc Bacon** [2 3]
[1]*Ecole Polytechnique, Palaiseau, France* [2]*Mila Québec AI Institute, Montréal, Canada* [3]*Université de Montréal*. Correspondence to: Adrien Goldszal adrien.goldszal@mila.quebec, Diego Calanzone diego.calanzone@mila.quebec.

1. Introduction

The search for new refrigerants balances competing thermodynamic, environmental, and safety objectives. Following the Montreal Protocol's CFC phaseout, hydrofluorocarbons (HFCs) became dominant but now face their own phase down under the Kigali Amendment due to high global warming potential (GWP) [1, 2]. Emerging hydrofluoroolefins (HFOs) remain concerning as most are classified as per- and polyfluoroalkyl substances (PFAS) with unknown environmental impacts [3, 4], while natural alternatives (CO_2, propane) face high-pressure or flammability constraints.

High-throughput screening faces severe constraints in this domain. McLinden et al. [5] demonstrated that once thermodynamic performance, safety, and environmental criteria are enforced, only a handful of candidates remain from existing chemical libraries. The scarcity of measured thermodynamic data, as only 300 refrigerants with known properties exist, limits purely data-driven approaches [6].

We present `Refgen` [1], which embeds physics-based inductive biases into molecular generation: (1) we develop state-of-the-art physics-grounded property predictors trained independently from supervised datasets and physics models (Equations of state (EOS), NASA polynomials, group-contribution methods) to compute key thermodynamic and chemical properties of molecules and (2) we build an RL pipeline to optimize a reward signal balancing the different objectives. This reward is obtained from the property predictors, to guide the LLM towards the generation of optimal refrigerant candidates in SMILES format.

2. Methods

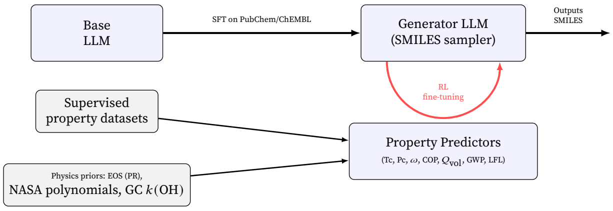

Fig. 1: The `Refgen` framework combines supervised property prediction with physics models (bottom) and RL-guided molecular generation (top).

Physics-Informed Property Prediction Rather than learning complex thermodynamic relationships directly from scarce data, we adopt a two-stage approach: (1) train predictors on measurable base properties for which annotations are more readily available, then (2) use established physics models to compute derived properties.

For refrigerant thermodynamic performance in the vapor compression cycle (VCC), we predict critical temperature (T_c), critical pressure (P_c), and acentric factor (ω), which parameterize the Peng-Robinson equation of state [7]. The EOS enables VCC simulation at any working operating conditions, yielding coefficient of performance (COP) and volumetric cooling capacity (Q_{vol}):

$$\mathrm{COP} = \frac{h_1 - h_4}{h_2 - h_1}, \quad Q_{vol} = \frac{h_1 - h_4}{v_1} \tag{1}$$

where enthalpies (h_i) and specific volume (v_1) are computed via departure functions with NASA polynomial corrections for ideal gas behavior at key parts in the cycle. For environmental impact, we predict radiative efficiency (RE) and estimate atmospheric lifetime (τ) through group-contribution methods for OH radical reaction rates [8]. These combine via IPCC formulas and constants ($\mathrm{AGWP_X}(100) = \mathrm{RE_X} \times \tau_\mathrm{X} \times (1 - e^{-100/\tau_\mathrm{X}})$) [9] to yield GWP_{100}. Lower flammability limit (LFL) is predicted directly.

We fine-tune SMI-TED [10], a transformer encoder pre-trained on molecular representations, for each target property. SMI-TED processes SMILES strings (textual molecular representations) and learns dense embeddings that capture chemical structure. We apply SMILES augmentation: generating multiple valid representations per molecule through different graph traversals, to improve generalization on small datasets. End-to-end finetuning with mean absolute error loss adapts the model to each prediction task, avoiding handcrafted molecular descriptors while leveraging pretrained chemical knowledge.

Generative Model Training We fine-tune Llama 3.2 1B [11] in two stages. First, we supervise finetune (SFT) on 37M filtered SMILES from PubChem, ChEMBL, and SureChEMBL [12, 13, 14] to teach valid unconditional molecular generation. Second, we use Group Relative Policy Optimization (GRPO) [15] to guide the generation toward target property ranges through a weighted multi-objective reward:

$$R_{\text{total}}(\mathbf{x}) = R_{diversity}(\mathbf{x}) \cdot \sum_{i=1}^{k} w_i R_i(\mathbf{x}) \tag{2}$$

---

[1] Code is available at https://github.com/ddidacus/refgen
Full paper with details: `https://arxiv.org/pdf/2509.19588`

where R_i scores predicted properties against acceptance criteria (e.g., $320K < T_c < 420K$, COP > 5, GWP < 10) based on the literature [5]. Weights found by hyperparameter search emphasize thermodynamic performance (COP/Q_{vol}/T_c: 80% total) over environmental metrics (GWP/LFL: 10%), reflecting relative optimization difficulty. The diversity term $R_{diversity}$ penalizes repeated structures based on Tanimoto similarity, preventing mode collapse during optimization.

3. Results

Property predictor validation SMI-TED fine-tuning achieves accurate predictions across all target properties. For COP prediction, we validate against CoolProp's reference EOS on 82 refrigerants, obtaining mean absolute error of 0.25 (median 0.12). Saturation dome reconstruction shows enthalpy RMSE of 17.5 J (median, N=110). For GWP, group-contribution methods for OH reaction rates achieve >60% predictions within factor-2 error on EPA validation data. These results demonstrate reliable coverage of the refrigerant chemical space despite small training datasets (937-4,940 samples per property). More details in Appendix A.

Guided molecular generation RL fine-tuning dramatically shifts the generated distribution toward target property ranges (Table 1). The base model produces large, complex molecules (median 22 atoms, T_c = 841K) with poor refrigerant properties. After GRPO, the model generates compact fluorinated compounds (9 atoms, T_c = 403K) with thermodynamically viable performance (COP=7.99, Q_{vol} = 1.18 MJ/m^3).

Table 1: Property distributions shift after RL fine-tuning (average over 1k generated molecules)

Model	T_c (K)	Atoms	GWP	LFL	COP	Q_{vol}
Base (SFT)	841	22	0.04	0.07	8.66	0.002
RLFT	**403**	**9**	**13**	**0.61**	**7.99**	**1.18**

We apply constraints based on McLinden et al (COP > 5, 320K < T_c < 420K, LFL > 0.1 kg/m^3, excluding unstable and toxic groups `=CF2` and `-OF`) as well as GWP < 10, which is below all existing regulations. Generated molecules reveal a tradeoff between `COP` and Q_{vol} depending on `Tc`, Figure 2. This behaviour, observed in the reference work [5] confirms the importance of our physics-grounded approach, which ensures thermodynamically sound predictions.

Figure 3 compares performance to R-410A (COP=7.39, Q_{vol}=6.61 MJ/m^3), a benchmark refrigerant blend considered to have some of the best thermodynamic properties in the vapor compression cycle. Generated molecules achieve competitive thermodynamic tradeoffs while meeting environmental constraints that R-410A fails (GWP=2088).

Finally, we screen out per- and polyfluorinated substances (`PFAS`) according to the OECD definition [3] for potential environmental risk of toxicity. We observe ~20% of generated molecules to be `non-PFAS`, despite no constraint having been added.

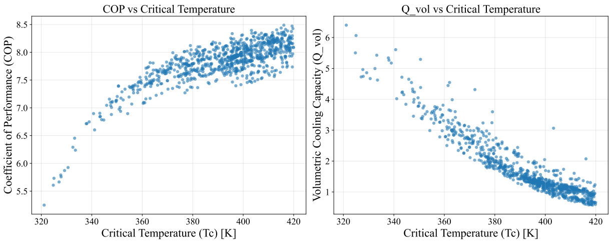

Fig. 2: Comparative plots for `COP` to `Tc` and Q_{vol} to `Tc` for compounds generated with `Refgen`.

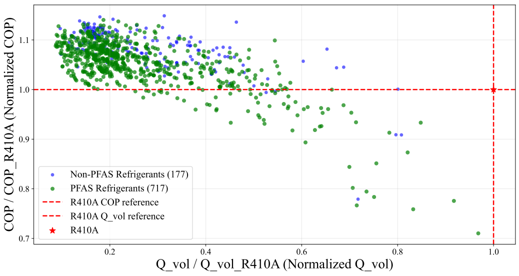

Fig. 3: Generated molecules achieve competitive COP/Q_{vol} tradeoffs vs. R-410A benchmark (red star). PFAS (blue) and non-PFAS (green) candidates both satisfy thermodynamic and environmental constraints.

Table 2: Top non-PFAS candidates with balanced properties

SMILES	**COP**	T_c (K)	**GWP**	**LFL**	Q_{vol}
`C(F)N(F)N(F)F`	6.71	338	3.95	0.52	5.34
`N(F)C(F)N(F)F`	6.72	338	0.25	0.54	5.26

4. Discussion & Conclusion

Our physics-informed generative pipeline rediscovers established refrigerant classes (HFCs, HFOs) under appropriate constraints while proposing novel compounds achieving competitive performance with dramatically lower environmental impact (GWP 0.25-3.95 vs. current refrigerants' 1400-2000). The physics-grounded architecture ensures thermodynamically sound predictions validated against reference equations of state.

Current and future work is focused on integrating additional constraints and predictors for synthesizability, chemical stability and toxicity (workplace exposure standards) which are key for practical compound viability. Experimentally, collaboration with chemists and HVAC engineers is essential to validate predicted properties and assess true efficacy in vapor compression systems. Observed candidate saturation (894 unique molecules from 1M samples) also motivates investigating better exploration strategies in the RL finetuning step to emphasize variety.

References

[1] United Nations Environment Programme Ozone Secretariat. Montreal Protocol on Substances that Deplete the Ozone Layer. `https://ozone.unep.org/treaties/montreal-protocol`, 1987. Accessed: 2025-08-15.

[2] Amendment to the montreal protocol on substances that deplete the ozone layer (kigali amendment). Adopted at the Twenty-Eighth Meeting of the Parties to the Montreal Protocol, Kigali, 10–15 October 2016, October 2016. United Nations Treaty Series, C.N.872.2016.TREATIES-XXVII.2.f; entry into force 1 January 2019 (with certain provisions entering into force 1 January 2033).

[3] OECD. Reconciling terminology of the universe of per- and polyfluoroalkyl substances: Recommendations and practical guidance. Technical Report 61, Organisation for Economic Co-operation and Development (OECD), Paris, 2021.

[4] Juliane Glüge, Katharina Breuer, Armin Hafner, Christian Vering, Dirk Müller, Ian T. Cousins, Rainer Lohmann, Gretta Goldenman, and Martin Scheringer. Finding non-fluorinated alternatives to fluorinated gases used as refrigerants. *Environmental Science: Processes & Impacts*, 26(11):1955–1974, 2024. Published online 2024 Oct 3.

[5] Mark O. McLinden, J. Steven Brown, Riccardo Brignoli, Andrei F. Kazakov, and Piotr A. Domanski. Limited options for low-global-warming-potential refrigerants. *Nature Communications*, 8:14476, 2017.

[6] Andrei Kazakov, Mark O. McLinden, and Michael Frenkel. Computational design of new refrigerant fluids based on environmental, safety, and thermodynamic characteristics. *Industrial & Engineering Chemistry Research*, 51(38):12534–12542, 2012.

[7] Ding-Yu Peng and Donald B. Robinson. A new two-constant equation of state. *Industrial & Engineering Chemistry Fundamentals*, 15(1):59–64, 1976.

[8] Eric S. C. Kwok and Roger Atkinson. Estimation of hydroxyl radical reaction rate constants for gas-phase organic compounds using a structure–reactivity relationship: An update. *Atmospheric Environment*, 29(14):1685–1695, 1995.

[9] Vladimir Zieger, Thibaut Lecompte, Simon Guihéneuf, Yann Guevel, Manuel Bazzana, Thomas Gasser, and Yue He. Climate change metrics: Ipcc ar6 updates, discussions and dynamic assessment applications. *EGUsphere*, 2025. Preprint. Discussion started: 3 February 2025.

[10] Eduardo Soares, Emilio Vital Brazil, Victor Yukio Shirasuna, Renato Cerqueira, Dmitry Zubarev, and Kristin Schmidt. Smi-ted: A large-scale foundation model for materials and chemistry. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2025. Published as a conference paper at ICLR 2025.

[11] Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, and et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024. Version 3, released November 23, 2024.

[12] Sunghwan Kim, Jian Chen, Tiejun Cheng, Asta Gindulyte, Jia He, Sherry He, Qingliang Li, Benjamin A. Shoemaker, Paul A. Thiessen, Bohu Yu, Len Zaslavsky, Jian Zhang, and Evan E. Bolton. Pubchem in 2021: new data content and improved web interfaces. *Nucleic Acids Research*, 49(D1):D1388–D1395, 2021. Accessed: 2025-07-25.

[13] Anna Gaulton, Anne Hersey, Michal Nowotka, A. Patrícia Bento, Jon Chambers, David Mendez, Prudence Mutowo, Fraser Atkinson, Louisa J. Bellis, Élodie Cibrián-Uhalte, Mark Davies, Neil Dedman, Anders Karlsson, Marta P. Magariños, John P. Overington, George Papadatos, Ilenia Smit, and Andrew R. Leach. ChEMBL: a large-scale bioactivity database for drug discovery, 2017. Accessed: 2025-07-25.

[14] George Papadatos, Mark Davies, Nathan Dedman, Jon Chambers, Anna Gaulton, James Siddle, Richard Koks, Sean A. Irvine, Joe Pettersson, Nicko Goncharoff, Anne Hersey, and John P. Overington. SureChEMBL: a large-scale, chemically annotated patent document database. *Nucleic Acids Research*, 44(D1):D1220–D1228, 2016. Accessed: 2025-07-25.

[15] Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, Y.K. Li, Y. Wu, and Daya Guo. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. arXiv preprint arXiv:2405.20094, 2024. Available at `https://github.com/deepseek-ai/DeepSeek-Math`.

[16] Caleb Bell, Yoel Rene Cortes-Pena, and Contributors. Chemicals: Chemical properties component of chemical engineering design library (chedl). `https://github.com/CalebBell/chemicals`, 2016–2024. Accessed: January 16, 2026.

[17] David S. Farina., Sai Krishna Sirumalla, Emily J. Mazeau, and Richard H. West. Extensive high-accuracy thermochemistry and group additivity values for halocarbon combustion modeling. *Industrial & Engineering Chemistry Research*, 60(41):15492–15501, oct 2021.

[18] Balaganesh Muthiah, Shih-Cheng Li, and Yi-Pei Li. Developing machine learning models for accurate prediction of radiative efficiency of greenhouse gases. *Journal of the Taiwan Institute of Chemical Engineers*, 151:105123, 2023.

[19] Max R. McGillen, William P. L. Carter, Abdelwahid Mellouki, John J. Orlando, Bénédicte Picquet-Varrault, and Timothy J. Wallington. Database for the kinetics of the gas-phase atmospheric reactions of organic compounds. *Earth System Science Data*, 12:1203–1216, 2020.

[20] Mathilda Maury, Didier Mathieu, and Johan Jacquemin. General estimation method for lower flammability limits of organic compounds: The simpler the better. *Fuel*, 345:127753, 2023.

[21] Ian Bell and contributors. Coolprop: Open-source thermodynamic and transport properties database. `https://pypi.org/project/coolprop/`, 2025. Version 6.8.0, released April 21, 2025. Available on PyPI. Maintained by Ian Bell and contributors.

Appendix A. Property Predictor Validation

Datasets & Preprocessing We compile datasets with annotated properties for SMILES structures: `Tc`, `Pc` and `w` are extracted from [16]; NASA polynomials for the enthalpy and entropy from [17]; radiative efficiencies from [18]; `k(OH)` reaction constants from [19]; `LFL` values are compiled from [20] and [16].(1) Smiles canonicalization is not performed to avoid loosing data from multiple data entry points which could have relevant information. (2) Smiles augmentation is performed, according to preliminary results showing its benefits to model generalization. For each molecule in the training set, four additional SMILES traversals are generated. (3) Dataset cleaning discards metals and metalloids and single elemental atoms, and ions. (4) we separate train/validation and test splits (70/15/15) by Bemis-Murcko scaffolds, using the RdKit library.

1.1 Thermodynamic Property Prediction

Table A1 shows prediction performance for critical properties on both held-out test sets and external CoolProp validation. External validation against CoolProp [21], an industry-standard library with reference-quality equations of state for 122 refrigerants, demonstrates maintained accuracy despite distribution shift.

Table A1: Critical property prediction on test sets and CoolProp external validation

Dataset	Property	N	R^2	MAE	RMSE
Test Set	T_c	8,079	0.975	12.5 K	21.8 K
	P_c	8,322	0.876	149 kPa	568 kPa
	ω	4,940	0.875	0.040	0.087
CoolProp (External)	T_c	106	0.926	20.7 K	37.6 K
	P_c	107	0.769	502 kPa	1452 kPa
	ω	107	0.636	0.057	0.125

1.1.1 Vapor Compression Cycle Simulation

We verify accurate simulation of complete refrigeration cycles through saturation dome reconstruction quality and COP prediction (T_{evap} = 10°C and T_{cond} = 40°C) against CoolProp's reference calculations and EOS.

Table A2: COP prediction accuracy at standard conditions (T_{evap} = 10°C and T_{cond} = 40°C)

Metric	N	Mean	Median
Dome enthalpy RMSE (J)	110	30.9	17.5
COP absolute error	82	0.252	0.12

Figure A1 illustrates saturation dome reconstruction for two representative refrigerants. The pressure-enthalpy (P-H) diagrams show the Peng-Robinson + NASA polynomial predictions closely match CoolProp's reference multiparameter EOS.

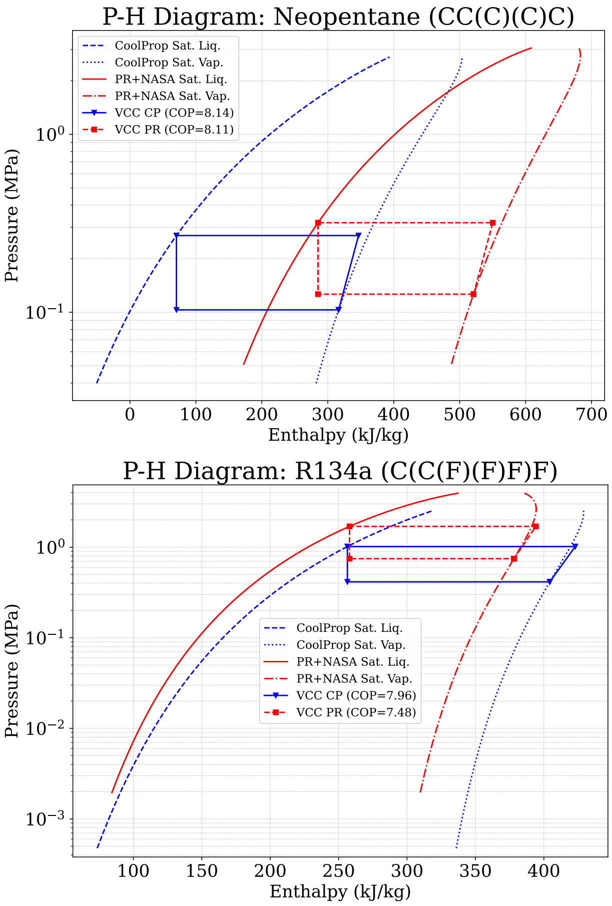

Fig. A1: Comparison between ground-truth saturation dome from CoolProp (blue) and predicted (red) for molecules Neopentane (left) and R134a (right). The COP is computed for condenser and evaporator temperatures at 10°C and 40°C respectively. Dome translation shifts are due to varying reference points in the enthalpy and they are irrelevant for the computation of COP or Q_{vol}.

1.2 Environmental and Safety Properties

Global Warming Potential Direct SMI-TED prediction of radiative efficiency achieves $R^2 = 0.91$ (MAE = 0.016 W m^{-2} kg^{-1}) on the test set of 14,380 molecules. This property combines with atmospheric lifetime approximated through reaction with reaction with the OH radical ($\tau_X = \frac{1}{k_{OH}[OH]}$) to yield GWP via IPCC formulas. For OH radical reaction rates, we implement group-contribution methods based on Kwok & Atkinson [8] and EPA AOPWIN. Validation against 72k PubChem molecules shows 61% of predictions within factor-2 error, demonstrating reliable generalization for GWP estimation ($R^2 = 0.535$, 65% within factor-5 on 220 test molecules with measured GWP).

Flammability. Lower flammability limit (LFL) prediction achieves $R^2 = 0.83$ (MAE = 0.16 kg/m^3) on 171 test molecules, enabling safety constraint enforcement during generation.