AI-Accelerated Design and Modeling of Organic Electrochemical Energy Materials: From Redox-Active Molecules to Polymer Electrolytes

Macromolecular Materials and Engineering

Macromolecular Materials and Engineering Open Access

www.mame-journal.de

REVIEW OPEN ACCESS

AI-Accelerated Design and Modeling of Organic Electrochemical Energy Materials: From Redox-Active Molecules to Polymer Electrolytes

Zhan-Yun Zhang[1] | Giannis Savvas[2] | Jichen Li[1] | Yue Zhou[3] | Thanh Trung Le[1] | Daniel Brandell[1] | Rocío Mercado[2] | Chao Zhang[1,4]

[1]Department of Chemistry - Ångström Laboratory, Uppsala University, Uppsala, Sweden | [2]Department of Computer Science & Engineering, Chalmers University of Technology & University of Gothenburg, Gothenburg, Sweden | [3]Department of Information Technology, Uppsala University, Uppsala, Sweden | [4]Wallenberg Initiative Materials Science for Sustainability, Uppsala University, Uppsala, Sweden

Correspondence: Rocío Mercado (rocio.mercado@chalmers.se) | Chao Zhang (chao.zhang@kemi.uu.se)

Received: 13 June 2026 | **Revised:** 14 August 2026 | **Accepted:** 19 August 2026

Keywords: AI | artificial intelligence | batteries | design | electrolytes | generative models | machine learning | MLIPs | modeling | polymers

ABSTRACT

Organic electrochemical energy materials (OEEMs) offer a vast design space for rechargeable batteries, redox-flow batteries, supercapacitors, and mixed ionic–electronic devices, yet their performance depends on tightly coupled thermal, redox, transport, mechanical, morphological, and interfacial properties that can rarely be optimized independently. This review examines how artificial intelligence (AI) is reshaping the computational design and modeling of these materials, tracing the field's shift from small redox-active molecules toward polymeric electrolytes and electrodes. We cover digital representations and data sources, data-driven property prediction, machine-learning interatomic potentials, generative molecular and polymer design, and large-language model-assisted (agentic) workflows. It is written for researchers entering the area from either direction, whether computational chemists curious about AI, or machine learning (ML) researchers new to electrochemical materials. Rather than an exhaustive survey, we give a selective snapshot of fast-moving literature, using representative examples to draw out practical, transferable lessons, and we provide a concise checklist of best practices to help newcomers evaluate and report their own work before publication. We close with our perspective on where the field is heading, together with the persistent challenges of experimental data quality, validation, polymer representation, reproducibility, and modeling integration. The accompany online dataset catalog is maintained as a living community resource that evovles with the rapidly expanding OEEM data landscape.

1 | Introduction

The transition to safe, sustainable, and high-performance electrochemical energy technologies requires materials that can be designed not only for energy density, but also for abundance, processability, mechanical robustness, recyclability, and compatibility with diverse device architectures [1]. Organic electrochemical energy materials (OEEMs) occupy a distinctive position in this landscape: built predominantly from earth-abundant light elements (C, H, N, O, and S) rather than scarce transition metals, they are lightweight, synthetically tunable, and potentially more sustainable to produce and recycle. Small organic molecules serve as electrolyte solvents, redox-active species, additives, and charge-storage units, while polymers are used as solid and gel polymer electrolytes, organic electrodes, ionomers, membranes, binders, and mixed ionic–electronic conductors [2–5]. Their appeal lies further in the breadth of chemical space that can be accessed through molecular substitution, backbone engi-

neering, copolymerization, crosslinking, and supramolecular organization.

At the same time, this chemical flexibility creates a major design problem. Useful organic electrochemical materials must satisfy several coupled requirements: redox stability, solubility, ionic conductivity, ion transference, viscosity, mechanical integrity, thermal stability, processability, and long-term electrochemical compatibility [6]. These properties rarely optimize independently. For example, a solvent or redox-active molecule with favorable redox energetics may suffer from poor solubility or excessive viscosity; a polymer electrolyte with good mechanical strength may display limited segmental mobility and low ionic conductivity; and a mixed conductor may require a delicate balance between electronic structure, ion uptake, swelling, and morphological stability. As a result, purely intuition-driven design and one-property-at-a-time screening become increasingly inefficient as the target materials move from simple molecules to formulations, polymers, and device-relevant condensed phases.

This difficulty coincides with a broader change in how chemical and materials space is explored. Data-driven molecular and materials design has moved from descriptor-based structure–property regression toward representation learning, transferable models, and closed-loop discovery. In chemistry, graph neural networks (GNNs) and message-passing neural networks (MPNNs) established a flexible route for learning molecular representations directly from connectivity and, later, three-dimensional geometry [7–9]. In polymer informatics, platforms, such as Polymer Genome [10] demonstrated how curated datasets and chemically meaningful fingerprints can support rapid prediction of polymer properties, while BigSMILES [11] and related representations addressed the more fundamental problem of encoding macromolecular ensembles rather than single molecules. In atomistic modeling, machine-learning interatomic potentials (MLIPs) have made it possible to approach first-principles accuracy for condensed-phase simulations at a cost comparable to molecular dynamics (MD), opening the door to explicit modeling of liquid electrolytes, redox-active molecules in solution, molecular crystals, and polymer electrolytes [12–14]. In parallel, generative models, active learning, and large language model-based agents are beginning to change the workflow itself: instead of screening a fixed library of materials, models can propose candidates, query simulations or experiments, extract knowledge from literature, and refine the search through feedback [15–17].

Several recent reviews have addressed parts of this landscape from computational or AI perspectives. Molecular modeling of organic redox-active battery materials has been reviewed in detail by Fornari and Silva [18], and computational and data-driven approaches for solid polymer electrolytes have also been summarized by Wang et al. [19]. Broader perspectives can also be found in the literature discussing how machine learning (ML) may transform electrochemical science, battery materials discovery, and generative AI for electrochemical energy storage [20–23]. The present review takes a complementary view by focusing on AI-accelerated design and modeling of OEEMs across small molecules, polymers, and molecular crystals. We connect four layers that are often treated separately: (1) digital representation and data, (2) property prediction, (3) machine-learning-accelerated atomistic simulation, and (4) generative or agentic design workflows. This organization is intended to highlight not only recent successes, but also the methodological bottlenecks that remain when moving from molecular candidates to realistic electrochemical materials.

The review is organized as follows. Section 2 discusses structural representations and data sources for small molecules, molecular crystals, and polymers. Section 3 focuses on data-driven property prediction for redox-active molecules and polymers. Section 4 covers MLIPs and their applications in molecular liquids, redox chemistry, crystalline systems, and polymeric systems. Section 5 discusses molecular and polymer generation, including proper benchmarking of generated candidates. Section 6 examines large language models (LLMs) in closed-loop workflows for extraction, analysis, inverse design, and feedback. Section 7 concludes with best practices and open challenges in (experimental) data quality, validation, polymer representation, reproducibility, and modeling integration.

2 | Structural Representations and Data Source

The predictive performance of ML models relies on three main factors: how molecular or materials structure is encoded, the quality and quantity of training data. Summarized in Figure 1, commonly used representations can be grouped into five classes: string-based representations, graph-based representations, physicochemical descriptors, cheminformatics fingerprints, hand-crafted geometric descriptors. Raw structural representations (string-based and graph-based) are commonly used in end-to-end representation learning, whereas the other three classes of representations are commonly used in feature-based learning, although hybrid approaches are the most common and robust in most cases. In the following section, we discuss examples in these different classes across the three material domains: small molecules, polymers, and molecular crystals. Subsequently, we summarize notable data sources, which can be useful for applications in liquid organic electrolytes in rechargeable batteries, organic electrodes, polymer electrolytes, and organic mixed ion-electron conductors.

2.1 | Structural Representations

2.1.1 | String-Based Representations

String-based representations encode chemical structures as character sequences and are therefore directly compatible with natural-language-processing architectures. For small molecules, the dominant representation remains SMILES (Simplified Molecular Input Line Entry System), which serializes a molecular graph into an ASCII string [29]. Two well-known limitations of SMILES are non-uniqueness and syntactic fragility - minor perturbations frequently yielding chemically invalid structures. Canonicalization algorithms mitigate the first problem, while newer string formats attempt to improve robustness more fundamentally. SELFIES (SELF-Referencing Embedded Strings) [30] guarantees that every valid string corresponds to a chemically valid molecular graph, which is especially advantageous in generative design, while DeepSMILES [31] offers a solution by simplifying ring-closure and branch notations. Crucially, these

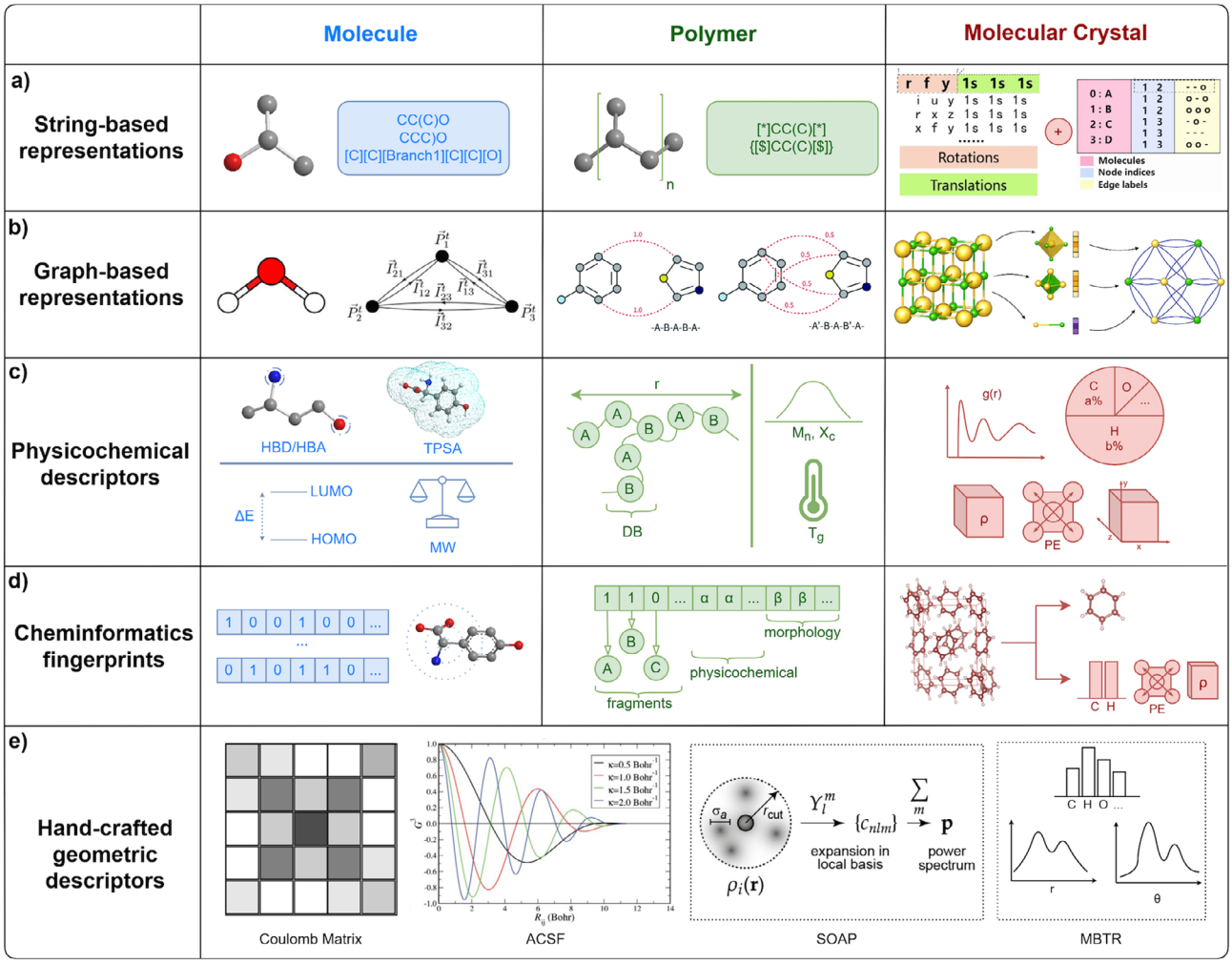

FIGURE 1 | Overview of molecular, polymer, and molecular crystal representations. (a) String-based representations: SMILES, DeepSMILES, and SELFIES for small molecules; PSMILES and BigSMILES for polymers; and SLICES-PLUS for molecular crystals. The SLICES-PLUS illustration is adapted from ref. [24]. Copyright 2025 Elsevier B.V. (b) Graph-based representations: small-molecule graph illustration adapted from ref. [25] under the terms of CC-BY licence (Copyright 2020 Shao et al.) Alternating-copolymer graph illustration adapted from ref. [26], under the terms of CC-BY licence (Copyright 2022 Aldeghi and Coley). Molecular crystal graph illustration adapted from ref. [27] with permission (Copyright 2018 American Physical Society). (c) Physicochemical descriptors: HBD/HBA count, TPSA, HOMO–LUMO gap, and MW for small molecules; chain length, degree of branching, number-average MW, degree of crystallinity, and glass transition temperature (T_g) for polymers; and radial distribution function, chemical composition, density, packing efficiency (PE), and lattice parameters for molecular crystals. (d) Cheminformatics fingerprints: Morgan fingerprints and MACCS keys for small molecules; hierarchical fingerprints for polymers, comprising atomic-fragment occurrence, physicochemical descriptors, and morphological descriptors following the concepts described in ref. [10]; molecular crystal fingerprints comprising component fingerprints and physicochemical descriptors of the bulk material. (e) Hand-crafted geometric descriptors: representative Coulomb-matrix; G^2 radial symmetry function adapted from ref. [28] with permission (Copyright 2011 AIP Publishing LLC); SOAP representation adapted from ref. [13] under the terms of CC-BY licence (Copyright 2021 Deringer et al.); and many-body tensor representation (MBTR).

string-based representations have directly enabled the application of Transformer-based foundational models. Architectures such as MolBERT [32] and ChemBERTa [33] generate highly transferable, context-aware embeddings from small-molecule strings via self-supervised pre-training.

For polymers, standard SMILES is inherently inadequate because it cannot capture the stochasticity or model complex polymer structures, such as branched or ring polymers. BigSMILES [11] addresses this by extending the notation with stochastic bonding descriptors to define precise connectivity patterns within an ensemble. Generative BigSMILES [34] further enhances this representation by incorporating additional statistical information, such as molecular weight (MW) distributions, reactivity ratios, and ensemble instance size. However, the simpler and more common approach is to represent polymers via their repeat-unit SMILES (PSMILES). Similarly, polymer SELFIES (P-SELFIES) [35] introduces specialized tokens to define the connection points. More recent approaches, such as HAPPY (Hierarchically Abstracted rePeat unit of PolYmers) [36], have introduced hierarchical and coarse-grained representations. By abstracting repeat units into chemically meaningful substructures, HAPPY provides

a more compact and information-dense representation while retaining key structural features. Similar to small molecules, these string-based representations for polymers have also enabled polymer-specific Transformer models, such as polyBERT [37], and TransPolymer [38], which learn contextual embeddings from tokenized repeat units.

For crystals, purely molecular strings are generally insufficient because they do not encode packing, symmetry, or periodicity. As indicated in Figure 1, emerging crystal-string representations include SLICES and symmetry/Wyckoff-based sequence representations. SLICES was proposed as a scalable text representation for periodic materials [24], while Wyckoff-position-based encodings have recently been used in generative crystal modeling [39].

2.1.2 | Graph-Based Representations

Graph-based descriptors are a natural representation for molecules and materials because chemical structure is relational, with atoms represented as nodes and covalent bonds or spatial contacts represented as edges. In molecular representation, their importance lies not in introducing molecular graphs themselves, which had long been used in cheminformatics, but in enabling a shift from fixed fingerprints to learned structural descriptors. Early neural graph approaches generalized circular fingerprints and allowed molecular features to be learned directly from connectivity rather than manually enumerated substructures [40, 41]. The message-passing framework later provided a general formulation for neural graph learning by describing molecular representations as iterative updates of atom or bond features over local graph neighborhoods [7].

Geometry-aware graph descriptors developed alongside connectivity-based graph learning, particularly for quantum-chemical, conformer-dependent, and atomistic simulation tasks. DTNN and SchNet introduced distance-based representations for atomistic systems [8, 42], while DimeNet, SphereNet, and GemNet showed that angular and directional information can further improve three-dimensional molecular graph learning [43–45]. This line of work was further advanced by equivariant representations, which encode geometric features in a way that respects rotations, translations, and reflections. PaiNN explored equivariant message passing for tensorial molecular properties [46], while NequIP and MACE systematically introduced equivariant atomistic graph representations to accurate and scalable MLIPs [9, 14] (see more discussion in Section 4.1)

Polymeric systems pose a distinct challenge to conventional graph representations developed for small molecules. A small molecule is naturally represented as a finite, chemically specific atom–bond graph, where the molecular identity is largely captured by explicit connectivity. A polymer chain can also be expanded into an atomistic graph, but this representation is often inefficient and conceptually incomplete: the graph is large, highly patterned, and its size depends on the degree of polymerization. More importantly, the chemically relevant description of a polymer is not only its local atom–bond connectivity, but also how monomer units are arranged along and across chains, including sequence, stereochemistry, branching, end groups, and copolymer architecture. Thus, polymer representation requires moving from a single molecular graph to a hierarchical description that connects monomer chemistry, chain construction, and higher-level architectural and conformational features. Aldeghi and Coley addressed this challenge by representing a polymer ensemble as a weighted directed graph constructed from monomer structures, connection sites, monomer stoichiometries, and probabilistic links between repeat units, with the degree of polymerization included as an additional weighting factor [26]. In a complementary direction, Gurnani et al. developed polyGNN by constructing periodic graphs from polymer repeat units and training a multitask MPNN to learn polymer features directly, reducing the dependence on handcrafted chemostructural fingerprints for large-scale polymer property prediction [47]. Polymer graph descriptors therefore move beyond isolated repeat-unit representations and instead encode composition, architecture and ensemble-level variability (polydispersity).

Crystalline materials pose a different representational problem because the relevant structure is not a finite molecule but an extended periodic network. CGCNN constructed crystal graphs by treating atomic sites as nodes and connecting each site to neighboring atoms in periodically repeated unit cells, so that local coordination environments could be learned directly from crystal structures [27]. MEGNet extended this representation by using separate atom, bond, and global state attributes, allowing local bonding environments and system-level information to be combined within one graph framework [48]. ALIGNN further modified the crystal graph by adding a line graph built from interatomic bonds, which allowed bond-angle information to be passed explicitly in addition to pairwise atomic interactions [49]. More recent crystal graph descriptors have continued this trend by refining how periodic neighbors, geometric relations, and structural symmetry are encoded before learning begins. Nevertheless, it is worth mentioning that many recent GNN architectures mentioned previously (PaiNN, NequIP, MACE) work equally work for both molecules and crystals.

2.1.3 | Physicochemical Descriptors

Physicochemical descriptors encode chemically or physically meaningful quantities derived from molecular structure, experimental measurements, or electronic structure calculations, such as density functional theory (DFT) and ab initio methods.

In cheminformatics and quantitative structure–property/activity relationship (QSPR/QSAR) modeling, molecular descriptors are often classified according to the structural information required for their calculation [50, 51]. Constitutional descriptors depend only on the molecular formula or atom counts; examples include MW, elemental composition, heavy-atom count, heteroatom count, and counts of specific atom types [50, 52]. Topological or graph-based descriptors, however, are calculated from the molecular graph, and therefore encode bonding patterns, functional groups, and connectivity. Examples include hydrogen-bond donor (HBD) and acceptor (HBA) counts, ring counts, aromatic atom counts, rotatable bond counts, topological polar surface area (TPSA), molecular connectivity indices, and topological indices [52]. Geometry-dependent descriptors require one or more molecular conformations and capture spatial properties such as molecular volume, solvent-accessible surface area [50].

Moving from constitutional to geometry-dependent descriptors generally increases structural detail, but also increases dependence on conformer generation, geometry optimization, and therefore, computational cost.

For small molecules, widely used and inexpensive to compute descriptors include MW, TPSA, HBD/HBA counts, aromaticity, ring counts, and rotatable bond counts [52]. Some other descriptors are obtained from quantum-chemical calculations, including dipole moment [53], orbital energies, ionization potential and solvation free energy. These quantities provide information about electronic structure, intermolecular interactions and reactivity, with the trade-off being higher computational cost.

For polymers, physicochemical descriptors must account for both repeat-unit chemistry and bulk material properties. Monomer-level descriptors are direct analogues of small-molecule descriptors, including atom counts, functional group counts, heteroatom content, and aromaticity. Polymer-specific descriptors also encode architectural and chain-level features such as side-chain structure, degree of branching (DB), and copolymer composition [54]. In addition to chemistry-derived descriptors, experimentally measured or computed bulk quantities such as number-average MW (M_n), degree of polymerization (DP), degree of crystallinity (X_c), glass transition temperature (T_g), and melting point (T_m) are frequently used as target properties or auxiliary descriptors in polymer informatics [10]. Thus, polymer descriptors often bridge molecular-scale repeat-unit chemistry with mesoscale chain architecture and macroscopic material properties.

For molecular crystals, physicochemical descriptors further include three-dimensional packing and translational symmetry. Structure-based descriptors extend this representation by incorporating unit-cell volume, density, packing efficiency (PE), coordination number, radial distribution function, local atomic environments, and intermolecular contact statistics [55]. Electronic-structure-derived descriptors, such as band gap, lattice energy, density of states, and charge distribution, are also used to provide additional information relevant to optical, electronic, mechanical, and transport properties. These physics-informed descriptors are particularly important for crystalline systems because properties may depend not only on molecular identity or composition, but also on polymorphism, crystal packing, local coordination, and long-range order [27, 55].

2.1.4 | Cheminformatics Fingerprints

Fingerprint-based representations encode structure as fixed-length binary or count vectors and are widely used with classical ML models because they are fast to compute and perform well on modestly sized datasets. For small molecules, the most common examples are key-based fingerprints such as MACCS keys [56] and circular fingerprints, such as ECFP (Extended-Connectivity Fingerprints)/Morgan fingerprints [57].

For polymers, the application of fingerprint-based representations introduces additional complexity due to their stochastic and hierarchical nature: they lack a single well-defined structure; instead, they are better described as distributions over repeating units, chain lengths, and architectures (e.g., linear, branched, crosslinked). To address this, one common approach is oligomer-based featurization, where short chain segments are explicitly constructed and fingerprinted to approximate local chemical environments [58]. Alternatively, hierarchical and multiscale fingerprints have been developed to encode both local chemistry and global polymer attributes [54]. This approach often concatenates the occurrence of a fixed set of atomic fragments with physicochemical and morphological descriptors, including features such as the shortest topological distance between rings, fraction of atoms that are part of side-chains, and the length of the largest side-chain [10].

For molecular crystals, the limitations of traditional fingerprints are even more pronounced. Molecular fingerprints inherently assume discrete, non-periodic structures and therefore fail to capture periodic boundary conditions, crystal packing motifs, and long-range intermolecular interactions that are critical for crystalline systems [55]. For this case, hand-crafted geometric descriptors and physicochemical descriptors serve as more robust and effective approaches to featurize molecular crystals.

2.1.5 | Hand-Crafted Geometric Descriptors

Despite the popularity of the end-to-end learning with graph-based representations mentioned previously, hand-crafted descriptors remain a useful class of molecular and materials representations with many chemical insights. These descriptors provide compact and interpretable inputs for conventional machine-learning models, while allowing domain knowledge to be built directly into the feature design [55, 59]. Hand-crafted descriptors are not only interpretable but also effective in small-data regimes. Coupled with kernel or linear models, they remain strong baselines and can reach state-of-the-art performance. Some philosophies also influenced modern equivariant GNNs.

Molecular systems provided one of the earliest and most extensively studied settings for hand-crafted geometric descriptors, owing to their finite size, non-periodic structure, and well-defined atomic composition. A direct strategy is to encode all atoms in a molecule into a single representation constructed from nuclear charges and interatomic distances. The Coulomb matrix follows this whole-molecule view by placing nuclear charges and pairwise Coulomb-like interactions into a molecular feature matrix, while sorted, eigenspectrum, and Bag-of-Bonds variants reduce atom-ordering dependence by reordering, diagonalizing, or grouping its entries according to element-pair type [60, 61]. Huang et al. describes molecules at the same molecular scale by implementing force-field-inspired Bonding, Angular, and higher order terms (BA) in their BAML model [62]. Other global descriptors summarize molecular geometry through distributions of internal coordinates. Faber et al. used histograms of distance, angle, and dihedral (HDAD) for molecular representation [63], which Dobbelaere et al. later modified this idea by replacing the fixed histogram features with a learned representation, termed Gaussian Learned (GauL) HDAD [64]. The many-body tensor representation (MBTR) further generalized this distributional strategy by encoding finite or periodic atomistic structures through element-resolved distributions of one-, two-, and three-

body geometric terms [65]. These descriptors move from ordered pairwise matrices toward chemically resolved many-body summaries of molecular geometry, but whole-structure encodings remain sensitive to system size, composition, and the diversity of the structures being compared.

A complementary strategy treats molecular systems as a collection of local chemical environments rather than a single indivisible object. This view reflects the way chemical regularity appears across compound space and the rising of the intermolecular interactions in condensed phase systems. Atom-centered symmetry functions (ACSFs) were a key step in this direction, describing each atomic neighborhood through radial and angular functions of the neighbor distribution, with translational, rotational, and permutational invariance built into the descriptor construction. This formulation provided the input representation for the Behler–Parrinello family of neural-network potentials [28]. Smooth Overlap of Atomic Positions (SOAP) developed the same local view in a density-based form, replacing discrete neighbor lists with smooth atomic densities and comparing environments through rotationally invariant power-spectrum kernels [66, 67]. The Faber-Christensen- Huang-Lilienfeld (FCHL) representation further refined local many-body descriptors for molecular quantum-property learning by combining element-resolved interaction channels with kernel-compatible representations [68]. The atomic cluster expansion placed such local invariant descriptors within asystematically improvable many-body basis framework [69], providing a direct conceptual link to the equivariant GNNs discussed in the following section [55].

Polymers are conventionally specified by their constitutional repeating units, but repeat-unit chemistry alone does not determine polymer properties. Identical constitutional units can correspond to different molecular-weight distributions, sequence distributions, branching architectures, and crystalline or amorphous morphologies. As previously mentioned in the cheminformatics fingerprints section, the hierarchical fingerprint systematizes this repeat-unit description by organizing features across atomic fragments, block-level cheminformatics descriptors, and chain-level shape attributes [10, 70]. For copolymers and architecturally complex polymers, repeat-unit fingerprints are further augmented with monomer fractions, composition-weighted descriptors, dyad and triad statistics, block lengths, tacticity, branching indices, grafting density, and architecture labels for random, alternating, block, graft, star, or network structures [71]. Descriptors derived from explicit polymer configurations then move beyond constitutional and sequence information by summarizing molecular-dynamics or Monte Carlo ensembles through end-to-end distance, radius of gyration, persistence length, backbone-dihedral populations, orientational order parameters, pair-correlation functions, and static structure factors [72]. Persistent-homology descriptors provide a more topological level of representation by encoding components, loops, and voids in amorphous polymer configurations, resolving packing motifs that are not captured by pair-correlation functions alone [73].

Crystalline materials require descriptors that encode atomic structure under periodic boundary conditions and remain invariant to equivalent choices of origin, unit cell, and supercell. Early descriptors extended molecular matrix representations to periodic systems, as in the sine matrix and Ewald-sum matrix, which replace isolated pairwise distances with periodic functions or lattice-summed interactions [74]. These global descriptors introduced periodicity into Coulomb-matrix-like representations, but retained sensitivities to atom indexing, cell choice, and fixed-length vector construction. Local environment descriptors were subsequently adapted to crystals by evaluating neighbor distributions under periodic boundary conditions. Periodic ACSF, SOAP, and MBTR representations describe crystallographic sites through radial, angular, density-based, or many-body distributional features and then aggregate site environments into structure-level descriptors [65, 67]. Such local descriptors capture coordination chemistry and short-range order, but finite cutoffs and pooling can miss distinctions in stacking, long-range ordering, and framework topology. Crystal-specific descriptors therefore also use Voronoi tessellations, coordination fingerprints, pore descriptors, ring statistics, and network-topology features to encode packing geometry and periodic connectivity more directly [75–77].

2.2 | Data Source

Table 1 categorizes notable experimental and computational data sources across different system types and data labels. The selected datasets cover key property classes relevant to molecular modeling, including potential energy surfaces, electronic properties such as HOMO/LUMO energies, electron affinities, and ionization potentials; thermophysical properties such as melting points and glass-transition temperatures; and transport properties such as ionic conductivities and diffusivities. To complement the table, we have created a GitHub repository, https://github.com/Teoroo-CMC/OEEM-data-catalog , which provides continuous updates on publicly accessible datasets.

3 | Data-Driven Property Prediction

3.1 | A Theoretical Minimum on Regression and Classification Methods

Mapping from a structural representation–such as handcrafted descriptors, fingerprints, or learned features–to a target property, thereby establishing a quantitative structure–property relationship and enabling faster screening and design of candidate compound is a supervised learning problem. Depending on the form of the target, supervised learning is divided into classification and regression. Classification refers to problems in which the target belongs to a finite set of discrete categories, such as whether an organic molecule is redox-active or whether a polymer electrolyte meets a prescribed conductivity criterion. Regression refers to problems in which the target is a continuous quantity, such as redox potential, dielectric constant, donor number or glass-transition temperature.

In supervised learning, the dataset is composed of labeled examples (x_i, y_i), where x_i denotes the representation of a molecule or material and y_i denotes the corresponding target property. In practice, the available data are commonly divided into three subsets: a training set, used to fit the model parameters; a

TABLE 1 | Publicly available data sources with properties relevant to electrochemical and battery systems (see the living OEEM-data-catalog repository for more details).

System	Dataset	Method	Size	Key Properties
Small molecule	JCESR [78]	B3LYP/6-31+G(d)	25K	EA/IP
	MPcules [79]	ωB97M-V/def2-TZVPPD	577K	EA/IP, redox potential
	OMEAD [80]	B3LYP/6–311G(d,p)	26K	HOMO/LUMO, EA/IP, redox potential
	ASMS [81]	Exp	2K	aqueous solubility
	Gutmann [82]	Exp	224	acceptor number, donor number
	Schrödinger-VISC [83]	Exp	3K	viscosity
	CIAC-RP [84]	ωB97XD/6-31+(d,p)/SMD	1K	reduction potential
	OMol25 [85]	ωB97M-V/def2-TZVPD	83M	HOMO/LUMO
	Batt-P30K [6]	ωB97X-V/def2-TZVPPD/SMD	30K	HOMO/LUMO, EA/IP
Molecular crystal	OMDB-GAP1 [86]	PBE	12K	E_g
	JARVIS-DFT [87]	PBE+OptB88vdW	37K	E_g
Electrolyte mixture	NUAA-COND [88]	MD	2K	Li diffusion coefficient, ionic conductivity
	CALiSol-23 [89]	Exp	13K	ionic conductivity
	DiffMix [90]	Exp	25K	ionic conductivity
Polymer	VitrimerVAE [91]	MD	8K	T_g
	Chemprop [26]	xTB+B3LYP/DZP	42K	EA/IP
	PoLyInfo [92]	Exp	>200K	T_g, Tensile modulus
	PolymerElectrolyte [93]	Exp	655	T_g, ionic conductivity
	ChemPropPred [94]	Exp	10K	ionic conductivity
	HTP-MD [95, 96]	MD	6K	diffusion coefficients, ionic conductivity, transference number

validation set, used to tune hyperparameters and select among candidate models; and a test set, used only for the final evaluation of predictive performance. This separation (at least between training and validation) is essential to assess whether a model captures generalizable structure–property relationships rather than merely memorizing the training data. When the dataset is small, as is often the case in electrochemical materials studies, k-fold cross-validation (CV) is widely used to obtain a more robust estimate of model performance by repeatedly splitting the data into training and validation folds and averaging the resulting metrics [97].

The purpose of training is to determine a function f that predicts $\hat{\mathbf{y}} = f(\mathbf{x})$ from $\mathbf{x}$ with good generalization beyond the training data. To achieve this goal, the learning algorithm seeks to minimize the so-called loss function $L_\theta(f(\mathbf{x})) + \lambda R(f)$, where first term represents the empirical error and the second term is the regularization penalty. Here, θ is the parameter vector, and the λ is the weight of the regularization $R(f)$. The typical loss function for classification tasks is the cross-entropy loss $L = -\sum_k y_k \log(p_k)$, where where y_k is the target value for class k, usually a one-hot label indicator, and p_k is the predicted probability that the sample belongs to class k. As for the regression tasks, the commonly used loss function is the mean squared error (MSE), i.e., $\frac{1}{N}\sum_i^N (y_i - \hat{y}_i)^2$, where N is the number of data points used in the training. On this basis, a wide range of supervised learning algorithms have been applied to molecular and materials property prediction, including multiple linear regression (MLR), logistic regression (LR), support vector machine (SVM), random forest (RF), gradient-boosting methods (e.g., XGBoost), Gaussian process regression (GPR), and neural network (NN).

3.2 | Applications in Redox-Active Molecules

3.2.1 | Redox Properties

The electrochemical window (ECW), defined as the difference between the oxidation and reduction potentials of a given compound, is a fundamental thermodynamic property. To explore novel electrolyte solvents with wide ECWs, Manna et al. [98] built a series of ML models to predict both oxidation and reduction potentials on a dataset including 308 molecules. They found that the tree-based methods (e.g., RF and XGBoost) achieve higher accuracy than those of other models (e.g., SVM), in which the XGBoost model achieves the lowest mean absolute error (MAE) of 0.37 and 0.25 V for reduction and oxidation potentials, respectively. Du et al. [84] trained a SVM model based on a reduction-potential dataset containing 1,003 samples, achieving a MAE of 0.243 V and a root-mean-square-error (RMSE) of 0.334 V. Clearly, a substantial dataset is essential for making reliable and precise data-driven predictions. To address this challenge, Persson and co-workers [79] constructed an open dataset, MPcules, containing electron affinities (EAs), ionization

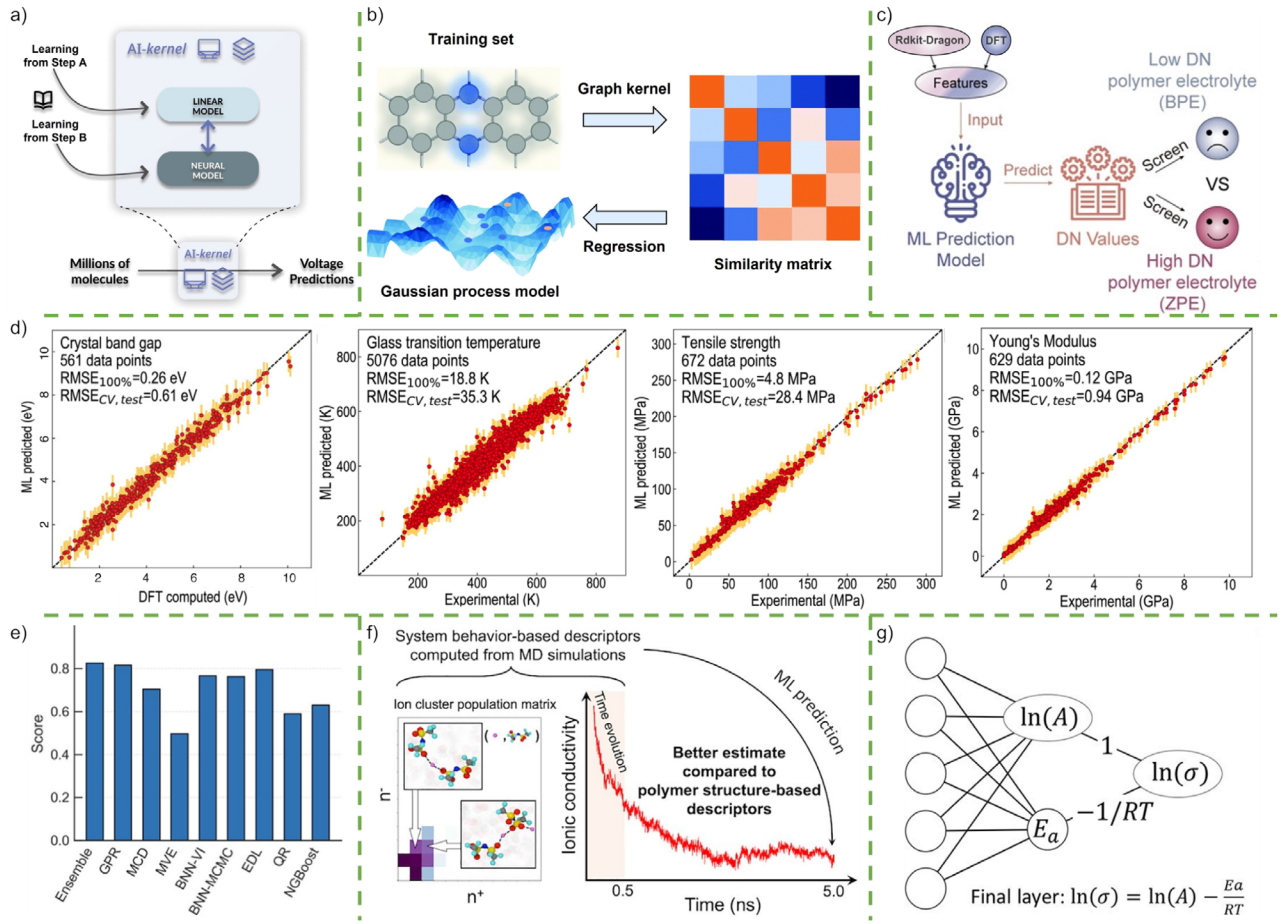

FIGURE 2 | Summary of data-driven methods for property prediction in redox-active molecules. (a) Workflow combining a NN-based reduction-potential predictor (Step B) with a linear model for open-circuit voltage predictor of lithiation reactions (Step A). Adapted from ref. [80] under the terms of the CC-BY license (Copyright 2022 Carvalho et al.). (b) Schematic diagram of the pipeline for constructing a graph-based GPR model for the prediction of aqueous solvation free energy. Adapted from ref. [106] with permission (Copyright 2021 Royal Society of Chemistry). (c) Workflow for screening molecular building units with high donor numbers (DNs) for polymer electrolytes. Adapted from ref. [113] with permission (Copyright 2025 Wiley-VCH GmbH). (d) Predictive performance of GPR models for crystal band gap, glass transition temperature, tensile strength, and Young–s modulus based on the Polymer Genome dataset. Adapted from ref. [119] with permission (Copyright 2020 American Institute of Physics). (e) Predictive R^2 scores of nine ML models for predicting the experimental glass transition temperatures of out-of-distribution polymers. Adapted from ref. [124] with permission (Copyright 2025 American Chemical Society). (f) Schematic illustration of the strategy for predicting equilibrium ion-transport properties in LiTFSI–homopolymer systems using MD-driven descriptors. Adapted from ref. [127] with permission (Copyright 2023 American Chemical Society). (g) Chemistry-informed MPNN architecture incorporating the Arrhenius equation for predicting experimental ionic conductivity of polymer electrolytes. Adapted from ref. [94] under the terms of the CC-BY license (Copyright 2023 Bradford et al.).

potentials (IPs), and redox potentials for more than 1 70 000 molecules. Using around 30 000 electrically neutral records from this dataset, Xu et al. [99] recently developed a GNN-based model that achieved a MAE below 0.06 V for both oxidation and reduction potentials. By further integrating this model with two additional ML models for water solubility and adsorption energies on zinc anode surfaces, a total of 48 candidate molecules were screened out as promising organic electrolyte additives for aqueous zinc-ion batteries. Araujo and co-workers [80] developed a NN-based predictor for reduction potentials using 26 218 data points, achieving an R^2 of 0.88 and an MAE of around 0.1 V. The predictions from this model were subsequently used to estimate the open-circuit voltage of lithiation reactions via a linear model (Figure 2a), which enabled the identification of 459 novel promising organic electrode materials with theoretical energy densities exceeding 1000 Wh/kg from a pool of 20 million molecules.

Accurate determination of redox potentials requires the reorganization energy, which is computationally expensive. Therefore, the difference between the vertical IP and EA is often used as a proxy to the ECW [79]. Although it is not recommended to use the energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) as the proxy to ECW because the former is electronic property and the later is thermodynamic property [100], HOMO and LUMO levels were nevertheless used in the regression tasks for redox-active organic molecules. Süleyman and co-workers [101] developed a linear regression model to map the DFT-calculated LUMO level to the experimental reduction potential of 7 conjugated sulfonamide cathode materials, achieving a RMSE of 0.057 V. Wang et al. [102] constructed a dataset containing calculated HOMO–LUMO gaps and solvation free energies for 1517 quinone molecules, and the XGBoost model trained on it

achieved a MAE of 0.1794 eV for the HOMO–LUMO gap. The model enabled rapid property prediction for 1 00 000 virtual quinone molecules, revealing that introducing $-SO_3H$, $-PO_3H$, and –COOH groups onto 1,2- or 1,4-anthraquinone backbones is a good route to develop aqueous organic redox flow battery (AORFB) materials. To probe the origin of reductive stability and accelerate the design of electrolyte solvents, Chen and co-workers [103] built a RF model to predict the LUMO energies of 1,399 ion-solvent complexes, achieving a MAE of 0.68 eV. They also applied the shapley additive explanations (SHAP) analysis and found that the molecular dipole moment and radius are the most significant descriptors affecting the reductive stability of coordinated solvent molecules.

3.2.2 | Solubility

The solubility of organic compounds is another crucial design parameter, which directly determines the achievable concentration of redox-active species and the energy density of AORFBs [15]. To enable effective high-throughput molecular discovery for AORFBs, Aspuru–Guzik and co-workers [104] proposed a multiple descriptor multiple kernel (MultiDK) based method to predict the molecular intrinsic solubility (i.e., Log *S*) of 1676 compounds. It demonstrated a twenty-fold CV RMSE of 0.61 and a R^2 of 0.91. Importantly, the MultiDK model can be extended to predict the pH-dependent Log *S* by incorporating the *n*-octanol-to-water partition coefficient and the pH-dependent distribution coefficient. Wang and co-workers [105] constructed a comprehensive open-access database, SOMAS, containing approximately 12 000 records of organic molecular solubility in aqueous solution. This resource provides a valuable foundation for developing more reliable ML models for solubility prediction. Xu et al. [99] built a logistic regression model utilizing 263 experimental water solubility records to identify if a molecule is soluble, achieving a *F*1 score of 0.916.

The solvation free energy (SFE) of organic molecules is closely associated with their solubility and can be computed relatively easily using DFT. To achieve accurate prediction of aqueous SFE, Gao et al. [106] proposed a novel molecular graph kernel and trained a GPR model on 588 experimental SFE records, which possesses a MAE of 0.72 kcal/mol and a RMSE of 1.03 kcal/mol on the test set (Figure 2b). In order to estimate the solubility of quinone derivatives, Wang et al. [102] trained a XGBoost model to predict the SFEs of 1517 quinone molecules, achieving a MAE of 0.8629 kcal/mol. Recently, Park et al. [107] built two ML models for predicting DFT-calculated specific energies and aqueous SFEs of over 103,000 organic molecules, which reported a RMSE of 0.056 Wh/kg and 0.175 kcal/mol, respectively. Based on these predictors, they further developed a new framework, SPARKLE, to zero-shot discovery of high-performance and low-cost organic electrode materials.

3.2.3 | Ion-Solvent Interactions

To estimate the ion-solvent interaction strength, the binding energy is a cost-effective property, which is defined as the energy difference between the ion–solvent complex and the isolated ion and solvent. In general, a more negative binding energy indicates a stronger ion–solvent interaction. After calculating the binding energies between five alkali metal ions (Li, Na, K, Rb, and Cs) and 70 solvents, Ishikawa et al. [108] built an exhaustive search with linear regression model, which reported a MAE of 0.127 eV. Based on the regression formula, they found that ionic radius and natural bond orbital (NBO) charge on coordinating O atom are the two most important descriptors. While the binding energy can provide valuable insights, but it is quite difficult to accurately characterize the interaction strength using solely one solvent molecule and one ion [109].

The donor number (DN) proposed by Gutmann is defined as the reaction enthalpy for the 1:1 adduct reaction between antimony pentachloride ($SbCl_5$) and a given solvent at low concentration in the main solvent 1,2-dichloroethane (DCE) [110]. Since the DN value quantifies the Lewis basicity of a given molecule, it can be used as a design parameter for engineering battery solvents [111]. Based on molecular descriptors, Hu et al. [112] presented four ML models to predict the DN values of 211 organic molecules, among which the categorical boosting algorithm showed the best performance, with a MAE of 3.923 kcal/mol and a RMSE of 5.221 kcal/mol. To enable accurate DN prediction, Gao et al. [113] recently developed a stacked ML model on a DN dataset including 210 samples, achieving a test MAE of 3.632 kcal/mol. Furthermore, the model was employed to screen polymer building units with relatively high DN values, facilitating the rational design of polymer electrolytes for batteries (Figure 2c).

3.2.4 | Transport Properties

Besides the redox and thermodynamic properties mentioned before, transport properties play an equally important role in organic electrochemical system. Following the Walden's rule, the electrolyte solution should have a sufficiently low viscosity to ensure fast ion conductivity. Yuan et al. [114] estimated the viscosities of 277 fluorinated 1,2-diethoxyethane solvents using non-equilibrium MD simulations, and built a multilayer perceptron model to predict the simulated viscosities based on features related to fluorination sites, with a RMSE of 0.2 cP. After curating a comprehensive dataset containing more than 4000 viscosity records of small organic molecules, Chew et al. [83] trained a series of molecular descriptor-based ML models and GNN to predict viscosity. Among these approaches, the light gradient-boosting machine (LGBM) model achieved the best performance, with a RMSE of 0.13 in log(cP) unit. Interestingly, the authors also found that incorporating MD–derived descriptors improved predictive accuracy in low-data regimes, whereas no significant improvement was observed when the training set exceeded 1000 samples. Battery electrolytes typically consist of mixed solvents; therefore, predicting the viscosity of multi-component systems is highly relevant. In this regard, Bilodeau et al. [115] proposed a directed MPNN model to predict the viscosity of binary liquid mixtures as a function of composition and temperature. After being trained on 39,077 mixture data points from the literature, the model attains a MAE of 0.043 and RMSE of 0.080 in log(cP) unit.

Unlike the viscosity which can be a solely solvent property, ionic conductivity requires the involvement of ionic species. Using the Δ-learning strategy, Chen et al. [116] developed RF and XGBoost based ML models to predict the difference between COSMO-RS

calculated and experimentally measured electrical conductivity of ionic liquids (ILs). Their findings suggest that the XGBoost model is capable of attaining a MAE of 0.396 ln(S/cm) with a dataset of 2168 data points for 242 distinct IL systems. Shi et al. [88] constructed a simulated ionic conductivity dataset comprising 2,604 electrolyte formulations. Based on this dataset, a RF-based model was developed using component-related features, such as, the mole fractions, ion–molecule binding energies, coordination numbers, and weighted viscosity over all components. The resulting model achieved a MAE of 5.99×10^{-4} S/cm and indicated that the viscosity is the most important feature, which is not a surprise. Recently, the SEED team at Bytedance [117] proposed a unified framework for the design of high-performance liquid electrolyte formulations, integrating a GNN-based predictor built on more than 10 000 experimental ionic conductivity records with a generative diffusion model. Interesting, all high ionic conductivity (>0.01 S/cm) formulation discovered in that work involves ethyl acetate.

3.3 | Applications in Polymers

3.3.1 | Glass Transition Temperature

The glass transition temperature T_g is one of the most important characteristics of polymers, because it governs the onset of segmental mobility that controls both the mechanical response of the polymer matrix and transport-related processes such as ion diffusion and ionic conductivity. Huang et al. [118] proposed a weighted-chained-SMILES that incorporates the compositions of different monomers to represent the polymers. Using this representation, a GNN model built on 5650 homopolymer experimental T_g records achieved a MAE of 24.32 K and a RMSE of 36.01 K, while another GNN model built on 3569 binary copolymer samples achieved a MAE of 13.24 K and a RMSE of 24.45 K. Based on the hierarchical polymer fingerprints in the Polymer Genome project, Ramprasad and co-workers [119] developed a GPR model on a dataset with 5,076 T_g data points, which exhibited a RMSE of 18.8 K (Figure 2d). Antoniuk et al. [120] developed a periodic polymer graph representation to capture the intrinsic periodicity of polymers and employed it to construct MPNN-based predictors for T_g using a dataset containing 7235 records. Compared with the conventional monomer graph representation, the periodic polymer graph representation reduced the predictive RMSE from 39.83 to 37.35 K, suggesting that more accurate polymer representations can lead to improved predictive performance. Dong et al. [121] improved polymer representation by employing an adjacency matrix and a graph autoencoder to capture local and global structural information from monomers. By incorporating the adjacency matrix and graph autoencoder losses into the label loss, they developed the Molecular Structural Regularized Graph Convolutional Network with Reinforcement Learning (MSRGCN-RL), where the weights of different loss components are adaptively optimized via reinforcement learning. Trained on a dataset of 960 polyimides, the model achieved an RMSE of only 10.49 K for T_g prediction. Instead of directly predicting T_g, Zheng et al. [91] developed a Δ-learning strategy to predict the difference in T_g between MD-simulated values and experimental measurements. Using molecular fingerprints of repeating units as input features, a GPR model was trained on a dataset of 8424 vitrimers and achieved a mean MAE of 28.07 K. Subsequently, this model was employed as the predictor in an inverse design framework, enabling the identification of a novel vitrimer with a T_g of 311–317 K. Recently, Kunchapu et al. [122] developed PolyMetriX, a digital polymer informatics workflow integrating curated datasets, feature engineering, polymer-specific data splitting, and robust model construction. On a T_g dataset comprising 7367 polymer entries, the best predictor built within PolyMetriX achieved an MAE of approximately 33 K.

Beyond developing predictors, Li and co-workers [123] systematically benchmarked different combinations of polymer representations, structural features, and ML models for predicting T_g on a dataset with 6923 records. They found that repeat unit-based representations outperform monomer-based ones, due to the inclusion of bond connectivity. Among structural features, the Morgan fingerprints with frequency performed best, followed by SMILES, molecular embeddings, and molecular graphs. Among different ML models, the RF model achieved the best averaged performance on 566 out-of-distribution test polymers with a R^2 of around 0.60 and a MAE of around 53 K. They further evaluated the uncertainties of nine different ML models across four polymer properties, including T_g (Figure 2e), showed that the NN-based ensemble model was leading the overall performance on accuracy and Spearman–s rank correlation coefficient [124]. Meanwhile, Zheng et al. evaluated different combinations of ML models and feature sets for predicting simulated T_g of 8424 vitrimers [125]. They found that LASSO with Mordred descriptors achieved the highest predictive accuracy among single models (RMSE = 15.72 K), while its performance was slightly worse than that of the ensemble model (RMSE = 15.19 K). Recently, Zhang and co-workers [126] showed that incorporating equivariant representations into the neural network reduces the test MAE of T_g prediction on 7166 homopolymers from 31.29 to 24.42 K, compared with the conventional edge-conditioned convolution network.

3.3.2 | Redox Properties

Similar to liquid electrolytes, the solid polymer electrolytes (SPEs) also require a large ECW to be compatible with high-voltage rechargeable batteries. However, accurately computing the redox potentials for large polymer structures is less common. Therefore, ML prediction of redox properties are limited to EA/IP and band gaps. Ramprasad and co-workers [119] employed the GPR models to predict the computational EA and IP of about 400 polymers, which resulting an average CV RMSEs of 0.18 eV for EA and 0.21 eV for IP. To overcome the limitation of small datasets, Aldeghi and Coley [26] constructed a large dataset comprising computational EA/IP values for 42 966 polymers with different chain architectures derived from 691 monomers. Based on it, they further proposed a weighted directed message passing neural network (wD-MPNN) which achieving a RMSE of 0.03 eV for both EA and IP. A GPR model for predicting polymer crystal band gaps (E_g) was also built by Ramprasad and co-workers [119] on a dataset with 561 records, which reported a RMSE of 0.26 eV (Figure 2d). By integrating with two additional GPR models for electron injection barrier (ϕ_e) and T_g, a novel AI-based framework, polyG2G [128], was proposed for designing polymer dielectrics with large E_g, high ϕ_e and also high T_g. To explore heat-resistant polysulfates for electrostatic energy storage, Li et al. [129]

developed two NN models to predict the E_g of 3381 polymers and the T_g of 6906 polymers using the Morgan fingerprints with frequency as input features. These models achieved MAEs of 0.32 eV and 19.19 K for E_g and T_g, respectively.

3.3.3 | Ionic Conductivity

For SPEs, the ionic conductivity is a key design parameter. In this regard, Segalman and co-workers [93] compiled a dataset of polymer Li^+-electrolyte including 655 unique polymer-anion-salt concentration entries and 5225 conductivity data points. Subsequently, several ML models were built for ionic conductivity prediction, among which the RF regressor achieved the best performance with an R^2 greater than 0.7 and a MAE below 0.5 on the log(S/cm) scale. Feature importance analysis indicated that polymer MW, T_g, the existence of electronegative heteroatoms, and anion size are the most influential descriptors. Khajeh et al. [127] proposed a set of features extracted from short-time MD simulations (0.5 ns) on LiTFSI-homopolymer systems, including the distribution of ion cluster sizes, cluster populations, and the time evolution of transport properties. Regarding the prediction of long-time simulated ionic conductivity (5 ns), their benchmark results on a dataset with 6024 records reveal that the introduction of MD features leads to a reduction of the MAE from 0.120 to 0.092 on the log(S/cm) scale (Figure 2f). Similar improvements were also observed in the prediction of Li diffusivity, $TFSI^-$ diffusivity, polymer diffusivity, and transference number, with the most significant improvement seen for the latter, where the MAE decreased from 0.158 to 0.108.

Using the ML-based predictors for ionic conductivity, one can attempt to screen SPEs database. Oyaizu and co-workers [130] constructed a large database ($N \approx 10^4$) of Li^+-conducting polymer electrolytes, including composition ratios, temperature, and the corresponding ionic conductivity data. Using this database, a GNN model was developed by integrating features of polymers, salts, and additives to predict ionic conductivity, reported an MAE lower than 1 on the log(S/cm) scale. The model was then used to screen new polymer electrolytes, leading to the identification of two polyphenylene sulfide-based electrolytes with experimental conductivities of approximately 10^{-3} S/cm at room temperature. On a similar note, Shao–Horn and co-workers [94] developed a state-of-the-art chemistry-informed MPNN architecture, in which the readout layers are designed in accordance with the Arrhenius equation (Figure 2g). After trained on dataset including 7133 experimental ionic conductivity data points, the model reported a MAE of 1.00 ± 0.03 on the log(S/cm) scale. Using ML models, they found that polyethylene oxide (PEO) and polytrimethylene carbonate (PTMC) show almost opposite anion-dependence on the ionic conductivity.

3.3.4 | Mechanical Properties

When designing novel SPEs, the mechanical properties (e.g., tensile strength and elastic modulus) of polymers play a critical role and polymers should be strong enough to tolerate stress variations from environmental and electrochemical conditions. Based on Polymer Genome fingerprints, Ramprasad and co-workers [119] trained two GPR models to predict experimental tensile strength and Youngs modulus of around 600 polymers, with average CV RMSEs of 4.75 and 120 MPa, respectively (Figure 2d). Recently, Li and co-workers [131] developed a multi-task NN by incorporating seven additional properties as auxiliary tasks on a dataset containing 2468 polymers, achieving validation MAE values of 8.59 MPa for tensile strength and 300 MPa for Young–s modulus. The obtained models were subsequently employed for high-throughput screening of polymers with desired properties, and their predictions were further validated by MD simulations.

Data-driven property prediction is increasingly enabling rapid screening and discovery of new electrochemical energy materials. However, predictive performance remains strongly dependent on the target property and on the quality, quantity, and coverage of the available data. Predictions of molecular electronic and redox-related properties, such as HOMO/LUMO energies, EA/IP, and redox potentials, are becoming relatively mature, whereas transport properties such as ionic conductivity and thermophysical or solution properties such as T_g and solubility remain considerably more challenging.

4 | ML-Accelerated Simulation and ML Interatomic Potentials

4.1 | Machine-Learning Interatomic Potentials in a Nutshell

MLIPs are data-driven surrogate models of the Born–Oppenheimer potential energy surface (PES), trained on electronic-structure calculations rather than prescribed through fixed analytical forms [12, 13]. Within this framework, the construction of an interatomic potential can be recast as the approximation of the map from nuclear coordinates to the ground-state energy, i.e., a high-dimensional function-learning problem defined on atomic configurations. A key physical simplification is provided by the nearsightedness of electronic matter, which implies that local energetic contributions depend predominantly on finite chemical environments and therefore permits the total energy to be approximated as a sum of local atomic terms subject to translational, rotational, and permutational symmetries [132, 133]. Under these conditions, flexible nonlinear regressors, including neural networks and kernel methods, provide a universal approximation to the representation of the PES. Their practical advantage is that the expensive electronic-structure calculations are incurred only during reference-data generation, whereas deployment requires only the evaluation of local environments and a surrogate model prediction, which leads to highly favorable, often near-linear, scaling with system size. MLIPs are therefore very attractive for simulating condensed phase systems over extended time-scales because it resides at a sweet-spot between chemical accuracy and computational efficiency.

In the 1990s, neural networks had already been explored for fitting potential-energy surfaces, motivated by their universal approximation capability for nonlinear functions; however, these early efforts were largely restricted to low-dimensional or fixed-

topology problems and did not yet provide a general treatment of variable-sized, permutation-invariant atomic environments [134]. The breakthrough came with the Behler–Parrinello neural network (BPNN), which decomposed the total energy into a sum of atomic contributions evaluated by local environments and represented each neighborhood by symmetry-preserving atom-centered descriptors, later formalized as ACSFs [28, 133]. This construction decoupled the number of model parameters from the total number of atoms, while enforcing translational, rotational, and permutational invariance, and thereby transformed neural-network potentials from system-size-fixed PES fits into a practical framework for large atomistic systems of arbitrary size and composition. The impact was rapidly demonstrated in realistic applications in liquid water, solid-liquid interfaces, and high pressure science [135–138].

Despite of the huge success of BPNN, its capability still depended on carefully handcrafted local descriptors, whose design became increasingly cumbersome as chemical complexity grew [55]. This limitation motivated a shift toward the end-to-end GNNs, which treat atoms as nodes and bonds as edges, in which atomic features are iteratively refined through neighbor-dependent message passing instead of being encoded by manually designed descriptors [7, 41]. SchNet represents a first implementation of this idea for atomistic simulation of periodic systems by learning directly from nuclear charges and continuous three-dimensional coordinates through continuous-filter convolutions, thereby removing the need for manually constructed local descriptors [8]. Its strong performance on molecular and materials property and force-learning benchmarks showed that learned geometric representations could match descriptor-based models while offering substantially greater flexibility and extensibility across chemical space. This inspires many follow-up works on applying graph, convolutional or MPNNs for modeling molecules and materials, such as PhysNet [139], CGCNN [27], PiNet [25] etc.

In these MPNNs that we have mentioned as well as the popular DeepMD code [140, 141], the internal representation were based on rotationally invariant scalar features. NequIP addressed this limitation by introducing E(3)-equivariant convolutions over tensor contractions in a systematic manner, so that intermediate representations transform covariantly under rotations, reflections, and translations and can therefore retain directional information and higher-order many-body correlations [9]. This advance translated directly into a substantial practical gain. In its original study, NequIP showed a remarkable data efficiency across chemically and structurally diverse systems, including small molecules, water in different phases, an amorphous solid, a reaction at a solid–gas interface, and a lithium superionic conductor, and in several cases outperformed earlier MLIPs with up to three orders of magnitude fewer training structures. The benefits of including equivariant features have also been confirmed by PaiNN [46], NewtonNet [142], TensorNet [143], PiNet2 [144] etc.

As model capacity, training compute, and data diversity increased, graph-based models for chemistry and materials began to exhibit the same broad empirical trend seen in other foundation-model settings, namely that performance and generalization can improve systematically with scale [145, 146]. In atomistic simulation, this observation suggested that sufficiently expressive MLIPs, when pretrained on large and diverse datasets, could learn reusable representations of local chemical environments that support transfer across domains, stronger out-of-distribution robustness, and efficient downstream adaptation. This led to the emergence of the atomistic foundation-model paradigm, for which MACE-MP-0 provided an early and influential example [147]. Pretrained on a large public materials dataset, it showed that a general-purpose model could be deployed across chemically and structurally diverse problems, as evinced also in DPA-2 [148], UMA [149], sevenNet-Omni [150] etc. Interested readers are welcome to follow the leaderboard, such as MatBench [151] on the foundation atomistic model, where new models continuously show up. It is worth noting that knowledge distillation can be applied to foundation atomistic models to the best efficiency and accuracy tradeoff with lightweight bespoke models [152, 153].

4.2 | Applications of MLIPs

4.2.1 | Applications of MLIPs in Molecular Liquids and Crystals

Small-molecule in condensed-phase systems (liquid and molecular crystals) are ideal application domain for MLIPs. [154] In molecular liquids, solvation structure and transport are governed by many-body polarization, local coordination, and collective solvent organization, whereas conventional force fields usually compress these effects into fixed partial charges and pairwise nonbonded parameters. In reactive molecular systems, chemical reaction and oxidation/reduction move the system between distinct electronic states, but classical force fields typically assume a fixed charge distribution and a single parameterized potential-energy surface, making it difficult to treat reaction free energies, redox potentials, and redox-coupled solvation in a consistent manner [155]. In molecular crystals, weak intermolecular interactions give rise to small but decisive free-energy differences between competing packings, yet empirical force fields frequently lack the accuracy needed to balance dispersion, electrostatics, hydrogen bonding, and anharmonic lattice motion at the few kJ mol^{-1} scale.

The first regime of application is on the description of many-body intermolecular interactions in molecular liquids, where solvation, local coordination, and ion transport are jointly controlled by environment-dependent polarization and packing. In carbonate electrolytes, classical fixed-charge force fields often fail to simultaneously reproduce Li^+ coordination environments, solvent exchange kinetics, viscosity, and ionic diffusivity. Csányi and co-workers [156] constructed an Gaussian approximated potential (GAP) potential with their SOAP descriptor for ethylene carbonate (EC) + ethyl methyl carbonate (EMC) mixtures, as illustrated in Figure 3a. They used the model to compute composition-dependent solvation structures and coordination motifs that are difficult to capture with fixed classical force field parameters. The Schrödinger team [157] used a high-dimensional neural-network potential and quantum-mechanical (QM) cluster calculations (~10K) to simulate liquid electrolytes and evaluate viscosity and ion diffusion, demonstrating that transport observables can be brought closer to QM accuracy while retaining long MD sampling. Nevertheless, a recent work from Svahn et al. [158]

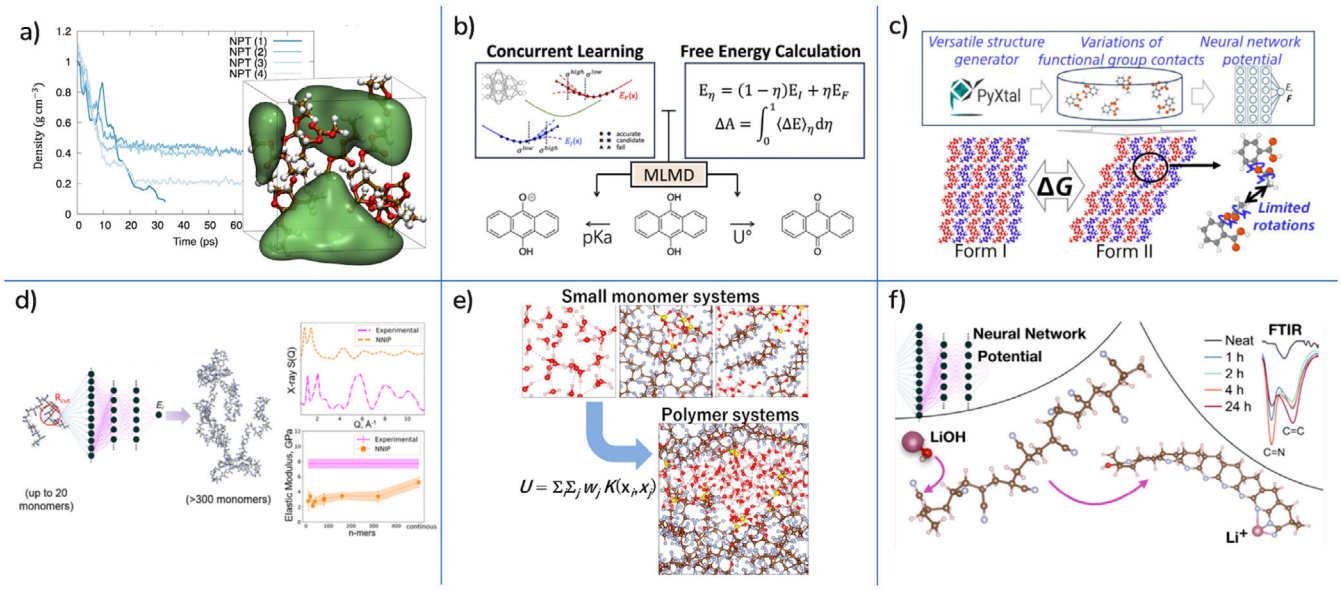

FIGURE 3 | Representative applications of MLIPs for small-molecule in condensed phase (panels (a)–(c)) and polymer (panels (d) – (f)) simulations. (a) MLIP simulations of ethylene carbonate (EC) + ethyl methyl carbonate (EMC) mixtures where bubble forms with unstable interatomic potential under NPT ensemble;reproduced from ref. [156] under the terms of CC-BY licence (Copyright 2023 Magdau et al.). (b) MLIP-based explicit-solvent sampling enables redox potentials and acidity constants to be computed via free energy perturbation. Reproduced with permission from ref. [166] (Copyright 2024 American Chemical Society). (c) Finite-temperature free-energy evaluation of competing molecular-crystal packing in different aspirin polymorphs with MLIPs. The illustration is reproduced from ref. [169] under the terms of CC-BY licence (Copyright 2024 Hattori). (d) The structure and mechanical property investigation of bulk polyacrylonitrile (PAN) with MLIPs [172]. Reproduced with permission (Copyright 2024 the American Chemical Society). (e) The hydrated perfluorinated-ionomer study with MLIPs that resolves morphology-, hydration-, and coordination-coupled proton transport. The illustration is reproduced with permission from ref. [176] (Copyright 2023 American Chemical Society). (f) The reactive PAN study that demonstrates MLIP simulations of electron-coupled Li^+ transport and ring formation. Reproduced from ref. [179] under the terms of CC-BY licence (Copyright 2026 Chahal-Crockett). Panels (a), (c), and (f) are reproduced under the applicable Creative Commons licenses, and panels (b), (d), and (e) are reproduced with permission from the American Chemical Society. Copyright remains with the respective copyright holders.

raised the question on the stability of MLIP trained only on the QM cluster data for modeling liquid electrolytes, as compared to its counterpart using periodic DFT data.

More recent works have shifted the emphasis from single-formulation accuracy to transferable electrolyte-property prediction across chemical space. Gong et al. [159] performed massive number of QM cluster calculations (~720 K) and introduced the BAMBOO MLIP framework and evaluated density, viscosity, ionic conductivity, atomic charge distributions, and solvation-structure fractions across different solvent–salt combinations, showing that MLIP-based MD can be used as a predictive tool for liquid-electrolyte development. Cheng and co-workers [160] further developed a domain-oriented universal machine-learning potential for battery electrolytes, by sampling randomly composed datasets over broad solvent and salt spaces, enabling the computation of density, solvation structures, ionic conductivity, operating-temperature windows, and Li^+ coordination lifetimes across diverse formulations. These studies mark a shift from using MLIPs to reproduce isolated electrolyte observables toward using them to screen and rationalize liquid-electrolyte formulations through coupled structural, dynamical, and transport properties.

The second regime of application lies in reactive molecular systems, where the limitation of classical force fields is more fundamental: chemical reaction (e.g., proton activity), electron attachment, and oxidation/reduction move the system between distinct electronic states and therefore change both the relevant potential-energy surface and the solvent response. Shao et al. applied the BPNN to study the change of ionic conductivity as a function of salt concentration and temperature in NaOH (sol) system where the OH^- transfer involves the interchange between chemical bond and hydrogen bond, i.e., the so-called Grotthuss mechanism [161]. Their study revealed a surprising finding that the ion-ion correlation does not manifest in highly concentrated electrolyte because of a high viscosity [161, 162]. Lan et al. [163] used a BPNN trained at correlated wave-function accuracy (MP2) to simulate the hydrated electron in bulk water, recovering the cavity structure, localization dynamics after electron injection, diffusion mechanism, and vibrational spectroscopy of the solvated electron—a system for which conventional empirical force fields lack a natural atomistic representation. Jinnouchi, Karsai, and Kresse [164] combined GAP with thermodynamic integration to compute finite-temperature redox potentials for aqueous Fe^{3+}/Fe^{2+}, Cu^{2+}/Cu^+, and Ag^{2+}/Ag^+ redox couples, capturing the large coordination changes that accompany reduction and achieving quantitative agreement with experiment. A similar strategy to combine MLIP and free energy calculations have been also applied to the pK_a calculations and the aqueous redox flow battery systems, as shown by Cheng group [165, 166] (Figure 3b). Kocer et al. [167] further targeted redox chemistry in solution using fourth-generation machine-learning potentials, i.e., charge-transfer augmented BPNN, showing for aqueous

Fe^{2+}/Fe^{3+} systems that oxidation states can be assigned consistently with the global ionic composition and that electron-transfer processes between ferrous and ferric ions can be described within MD.

The third regime of application is about molecular crystals, where weak intermolecular interactions control thermodynamic observables through small but decisive free-energy differences between competing packings. Classical force fields can often generate plausible crystal structures, but they frequently fail to rank polymorphs reliably because dispersion, hydrogen bonding, electrostatics, and anharmonic lattice motion are balanced on a narrow energy scale. Firaha et al. [168] highlighted the severity of this problem in real-world crystal-form selection, showing that pharmaceutically relevant hydrate and anhydrate stability requires solid–solid free-energy differences with standard errors on the order of 1–2 kJ mol^{-1}. Hattori and Zhu [169] addressed this challenge directly for aspirin by using a machine-learned potential to evaluate finite-temperature free-energy differences between competing polymorphs, as summerized in Figure 3c, revisiting a long-debated stability problem beyond static lattice-energy comparisons or empirical force-field rankings. Butler et al. [170] used active learning from organic crystal-structure-prediction landscapes to build MLIPs that rerank large and diverse candidate crystal structures at near-DFT accuracy, directly addressing the computational bottleneck created by tens to hundreds of thousands of trial structures in organic crystal structure prediction (CSP). Michaelides and coworkers [171] further extended this direction to finite-temperature molecular-crystal modeling by generating MLIPs for the X23 molecular-crystal dataset and computing sublimation enthalpies while accounting for anharmonicity and nuclear quantum effects, achieving sub-chemical accuracy with respect to experiment.

4.2.2 | Applications of MLIPs in Polymers

Polymers present a distinctive challenge for atomistic simulation because their key observables emerge from physics spanning ångström-scale monomer chemistry to nanometer-scale chain conformations, morphologies, and phase-separated domains, and from femtosecond bond vibrations to slow segmental and glassy relaxation. Cases have already merged on the development of bespoke MLIP models for polymeric systems. Chahal et al. [172] investigated polyacrylonitrile (PAN), highlighted in Figure 3d, a polar polymer in which dipole–dipole interactions, hydrogen bonding, and stereochemical disorder strongly influence molecular ordering. By training DeepMD on small-scale DFTMD data for oligomers and applying them to larger polymer configurations, they reproduced amorphous bulk PAN structure against experimental X-ray structure factors and connected molecular orientation to density and elastic modulus. A similar workflow has been applied to high-performance engineering polymers of polyether ether ketone (PEEK) [173], where MLIPs have been used to compare crystalline, amorphous, and single-chain forms and to analyze phonon thermal transport across chain and interchain directions. Beside using DeepMD, moment tensor potential (MTP) has also been used to investigate the heat trasnport in crystalline-polymer of polyethylene and polythiophene by Reicht et al. [174], which shows how chain-direction phonons, anharmonicity, and real-space versus reciprocal-space analyses affect the computed thermal conductivity. An equivariant GNN has also been explored for developing polymer specific MLIPs (SimPoly) and showed promising results on predicting density and glass transition temperature [175].

Beyond conformational sampling with MLIPs that bears a great similarity to the examples on simulating liquid electrolytes mentioned in the previous section, MLIPs have also been applied to reactive polymeric systems. Jinnouchi et al. [176] investigated dry and hydrated perfluorinated ionomers using an actively learned MLIP with GAP, shown schematically in Figure 3e, reproducing heterogeneous hydrophilic/hydrophobic domains, water and proton diffusion coefficients, and short Grotthuss chains under high hydration. Along the same direction, Hänseroth et al. [177] studied commercial anion-exchange membranes, where fine-tuned MACE-MP-0 with new DFT lables captured hydration-dependent hydroxide mobility, the transition from isolated water clusters to connected hydrogen-bond networks, and experimental trends in diffusion and activation energies. In addition to the proton migration in polymer electrolyte, where MLIPs are clearly indispensable, the coupled ion-electron motion and their effects on the chemical reaction is also well-suited for MLIP applications [4, 178]. In this regard, Chahal–Crockett et al. [179] applied DeepMD trained on nonequilibrium PAN configurations to study nucleophile-initiated cyclization and electron-coupled Li^+ transport, as illustrated in Figure 3f, resolving reaction free energies and pathways in the sequential ring formation. Mori et al. [180] further extended this direction to reactive polymerization and curing, demonstrating that an atomistic foundation model can describe polymer growth, polycondensation, and interfacial epoxy cross-linking within a unified framework without constructing a separate reactive force field for each polymer chemistry [180].

Before leaving this section, it is worth to mention practical and recent alternatives to MLIP when it comes to the materials modeling of polymers [175]. Due to the significantly larger size of polymer systems and much longer time-scale required for capturing the segmental dynamics, efforts have been made to reparameterized classical force field parameters with ML techniques or combine the short-ranged MLIPs with classical force fields for accuracy [181]. These approaches are particularly useful when no chemical reaction is involved in the polymeric systems and a high-throughput screening for identifying the lead compound (rather than mechanistic study) is the goal.

Machine-learning interatomic potentials (MLIPs) have evolved from descriptor-based models to graph neural networks, equivariant architectures, and pretrained atomistic foundation models. Their application to molecular liquids is becoming increasingly mature, particularly for predicting structural, dynamical, and transport properties, whereas molecular crystals and polymeric systems remain comparatively underexplored. Major frontiers include redox-active processes, reaction-coupled ion transport, and the simulation of OEEMs under realistic electrochemical conditions, where changes in electronic state, charge transfer, and applied potential (field) must be treated consistently.

5 | Molecular and Polymer Generation

5.1 | Foundations of Generative AI

Generative AI (GenAI) refers to a class of deep learning models designed to produce novel samples by drawing from a learned approximation of the data distribution. Formally, given a training dataset with observations x drawn from an unknown true data distribution $p_{\text{data}}(x)$, a generative model learns a parameterized distribution $p_\theta(x)$ such that samples drawn from $p_\theta(x)$ closely resemble the training data. GenAI has achieved remarkable success across modalities, such as text [182] and image [183] generation, and is increasingly applied to generate novel molecular [184] and polymer [185] candidates. More recently, these approaches have been extended to the design of organic electrochemical materials [20, 21, 186].

When applied to molecular and polymer design, generative models can operate in three modes: unconditional [187], conditional [188], and goal-directed [189] generation. In unconditional generation, the model learns the underlying data distribution $p_\theta(x)$ and draws samples from it without any constraints. For example, a model trained on a dataset of organic molecules might generate chemically reasonable structures resembling those in the training set, without targeting any specific chemical property. Conditional generation instead introduces a target property y and models the conditional distribution $p_\theta(x|y)$, enabling the model to produce samples that satisfy predefined constraints. Lastly, goal-directed approaches treat generation as an explicit optimization task, in which a scoring function $f(x)$ guides the search toward desired regions of chemical space. The scoring function typically evaluates molecular or polymer properties such as the HOMO-LUMO gap, ionic conductivity, or redox potential. In this way, generative models can iteratively propose and refine candidates tailored to a specific application. Conditional and goal-directed approaches are also sometimes collectively referred to as *de novo* or *inverse* design.

To enable the generation of novel candidate structures, various architectures have been proposed in the literature [184, 190]. These can be broadly grouped by their training paradigms, and the field has clearly evolved over time. Early efforts relied on variational autoencoders (VAEs) [191], which learn a continuous latent space, and generative adversarial networks (GANs) [192], which rely on adversarial training between a generator and a discriminator. Both paradigms have been increasingly superseded by diffusion models (DMs) [193], which now dominate much of the recent literature, offering more stable training and higher-quality samples. DMs learn to generate data through a two-stage process, namely the forward and reverse processes. The forward process $q(x_t|x_{t-1})$ gradually corrupts data by adding noise over a sequence of time steps, while the reverse process $p_\theta(x_{t-1}|x_t)$ trains a neural network to iteratively denoise and recover the original data point. In the standard denoising diffusion probabilistic model (DDPM) formulation [194], models are trained to minimize the error between the predicted and the actual noise added at each time step. A closely related and rapidly growing direction is flow matching (FM) [195], which learns a continuous-time velocity field that transports a simple prior distribution (e.g., Gaussian noise) to the target data distribution. Unlike DMs, which require a fixed forward noising schedule and learn to reverse it, FM directly regresses an optimal transport path between distributions, typically yielding faster sampling with fewer integration steps. Although FM has not yet been applied to organic electrochemical materials, its combination of sampling efficiency and compatibility with geometric equivariance makes it a natural candidate for this domain.

Orthogonally, models may also differ in their *generation strategy*. VAEs and GANs typically produce the entire structure in a single pass, while DMs and FM models generate through global iterative refinement over many denoising or integration steps. In contrast, autoregressive models [196] generate structures sequentially, one token, atom, or bond at a time, conditioning each step on all previously generated elements. Formally, the joint distribution is decomposed into a product of conditional distributions, $p(x) = \prod_i p(x_i|x_1, \ldots, x_{i-1})$. Autoregressive generation is thus a *strategy* rather than an architecture family in its own right, and can be implemented with various neural network backbones, including recurrent neural networks (RNNs) [189], GNNs [197], and Transformer decoders [198]; the latter underpin many recent LLM-based molecular generators, which are discussed separately in Section 6.

In a complementary way, reinforcement learning (RL) provides a goal-directed optimization layer [199] that operates on top of an underlying generative architecture, treating the generator as a policy network whose outputs are steered toward desirable structures by reward signals. In such a setting, a generative model acts as a policy network π_θ that constructs candidate structures [189] and receives a reward $r(x)$ based on the quality of the generated output. The policy is optimized to maximize the expected reward $\mathbb{E}[r(x)]$ over the distribution of generated outputs, typically via policy gradient methods [200]. Central to this framework is the concept of the *oracle*: an evaluation function that returns a property score for a given structure. In molecular design, the oracle could in principle be a DFT calculation, an MD simulation, or a wet-lab experiment; however, the iterative nature of RL requires thousands to millions of evaluations per optimization run, making such high-fidelity oracles impractical. In practice, the oracle is therefore approximated by a *surrogate model*—typically a neural network property predictor trained on available data (as discussed in Section 3)—that is fast enough to be called within the RL loop at each policy update step. This surrogate-as-oracle setup is what makes RL-driven molecular generation computationally tractable. In constrained or budget-limited settings, the quality of a generative framework is often assessed by how well it can identify high-performing structures with a fixed number of oracle calls [201]. A key advantage of RL over purely conditional generative approaches is its ability to optimize external objectives via cheap surrogate evaluations, making it particularly well-suited to closed-loop materials discovery pipelines.

5.2 | Applications of Generative AI

Generative models have gained traction in the design of organic electrochemical materials for next-generation batteries. The application scope includes liquid and polymer electrolytes for Li-ion batteries, as well as redox-active molecules for flow batteries (Figure 4). It is worth noting that, compared to data-

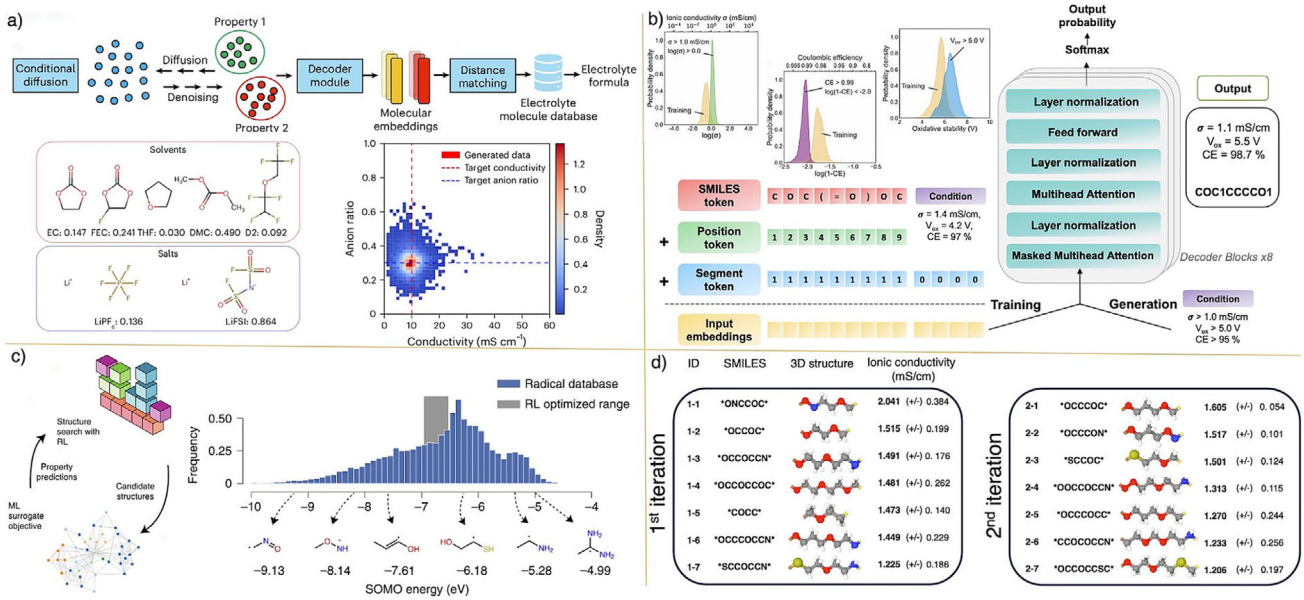

FIGURE 4 | Generative AI application summary for organic electrochemical materials covering examples for liquid electrolytes, redox-active molecules, and polymer electrolytes. (a) Diffusion model for generating multi-solvent and multi-salt mixtures for liquid electrolytes [117], adapted with permission (Copyright 2026 Springer Nature group). (b) Transformers decoder for conditional electrolyte formulation design [202] adapted under the terms of CC-BY license (Copyright 2026 Kim et al.). (c) Multi-objective RL framework for the discovery of redox active molecules [203], adapted under the terms of CC-BY license (Copyright 2022 Sowndarya et al.). (d) Iterative materials discovery framework for polymer electrolyte discovery [204], adapted under CC-BY license (Coyprright 2025 Khajeh et al.).

driven property prediction (Section 3) or MLIPs (Section 4), the application of generative models to organic electrochemical materials remains a relatively young area with a smaller body of published work. Nevertheless, the studies reviewed below illustrate how different generative paradigms are being adapted to this domain and highlight the diversity of design targets already addressed.

5.2.1 | Applications in Small Molecules

Regarding small molecules, Yang et al. [117] developed a conditional permutation-invariant diffusion model, trained on 30.000 synthetic multi-solvent and multi-salt electrolyte formulations labeled by a surrogate neural network. The generated output embedding is matched to entries in an existing database, producing novel molar ratios with optimized anion ratios and ionic conductivity. Experimental validation of 11 generated multi-solvent formulations showed that 10 exhibited high ionic conductivities, ranging from 11.99 mS/cm to 18.77 mS/cm, significantly higher than the mean conductivity of 5.57 mS/cm measured for random experimental formulation, with Raman spectroscopy confirming high anion ratios and one formulation achieving a Coulombic efficiency above 95 %. On a similar note, Kim et al. [202] developed ElectrolyteGPT, a Transformer decoder model trained on a curated library of 1 million electrolyte-relevant molecules. The study introduced a line notation, which represents a complete electrolyte formulation, including multiple salts, solvents, molar ratios, and concentrations, as a single text string. This framework enables multi-property conditional generation, allowing the model to optimize up to 10 parameters simultaneously. The researchers experimentally validated the model's results by synthesizing two novel solvents and a multi-component formulation that achieved a coulombic efficiency of 99.4 %. In a related effort, Yoon et al. [205] used a VAE with a property predictor in the latent space to guide the generation of electrolyte additives with optimized HOMO-LUMO levels, which are strongly related to the solid electrolyte interphase (SEI) of the battery. DFT calculations on approximately 1.000 newly generated candidates validated the latent-space property predictor, yielding a mean absolute error of only 0.05 eV for HOMO and 0.07 eV for LUMO relative to the DFT ground truth. In a complementary direction, Pritom et al. [206] combined an unconditional GAN with an MPNN property predictor and DFT validation in a multi-stage screening funnel. The GAN, trained on the GDB-11 dataset, first generated 30.000 chemically valid candidates without any property conditioning; the MPNN then filtered this pool to 50 candidates based on predicted enthalpy of formation and HOMO-LUMO thresholds, and DFT calculations were used as a final high-fidelity screen. This procedure yielded 26 promising molecules, mainly from the nitrile class, with oxidation potentials of 55.6 to 6.9 V versus Li/Li^+. Taken together, these four studies illustrate how distinct generative modeling strategies can be adapted to closely related electrolyte design problems.

Moving beyond Li-ion batteries, Sowndarya et al. [203] applied a reinforcement learning agent to perform goal-directed optimization of radical scaffolds for aqueous redox flow batteries. The scalar reward jointly balanced the reduction and oxidation potentials, cell voltage, radical stability, and synthesizability, providing a representative example of multi-objective, goal-directed design. Out of 3.8 million evaluated structures, DFT validation of the 1,078 top candidates showed that 80.5 % met the redox requirements and, independently, 41.9 % passed the stability threshold, while further retrosynthesis analysis identified 32 radicals with practical synthetic routes. To our knowledge, this

remains one of the few examples of GenAI-driven design for redox flow batteries, suggesting room for further work in this direction.

5.2.2 | Applications in Polymers

GenAI for polymers is still in an early, developing stage, yet some initial efforts have been made to model polymer electrolytes for Li-ion batteries. Yang et al. [17] compared a Transformer-based decoder with two diffusion-based models for the de novo design of polymer electrolytes. The models were trained on a subset of the HTP-MD dataset [95, 96], which contains 6.024 polymer electrolytes, to generate monomer repeat units represented as PSMILES, both unconditionally and conditionally on ionic conductivity. Across the standard validity, uniqueness, and novelty metrics, minGPT achieved the highest mean score of 0.773, compared to 0.767 for diffusion-LM and 0.736 for 1D diffusion. To validate the conditional model, the authors generated 100 000 candidates and screened the top 46 via MD simulations. They identified 17 novel candidates with ionic conductivities exceeding the highest value in the training set (5.07×10^{-4} S/cm), with the top candidate reaching 1.13×10^{-3} S/cm, more than doubling the best training-set conductivity. Building on this framework, Khajeh et al. [204] introduced a closed-loop discovery pipeline that integrated the conditional minGPT model, MD-based computational evaluation, and a feedback mechanism for iterative refinement. Across two discovery iterations, 100 candidates were validated via MD, and the mean ionic conductivity of generated polymers rose from 0.06 mS/cm in the seed training set to 0.85 mS/cm by the second iteration. In total, the framework identified 14 distinct polymer repeating units that surpassed the ionic conductivity of the industry benchmark, PEO.

5.3 | Benchmarking and Evaluating Generative Models

To assess the quality of generated molecules and polymers, several metrics are commonly employed [207]. At a basic level, validity measures whether generated samples correspond to chemically valid structures, uniqueness quantifies the proportion of non-duplicate samples, novelty measures the fraction of generated structures not present in the training set, and diversity captures the structural variety among the generated samples, typically computed as the average pairwise Tanimoto distance over molecular or polymer fingerprints. Beyond these, distribution-based metrics such as the Fréchet ChemNet Distance [208] quantify how closely the distribution of generated molecules resembles that of the training set. Finally, synthetic accessibility scores [209] provide an estimate of how easily a generated structure can be synthesized in practice. While such scores are reasonably reliable for small molecules, their applicability to polymers remains limited due to the complexity of polymerization processes.

For conditional and goal-directed tasks, it is equally important to evaluate whether the generated structures exhibit the desired property values. The success rate measures the fraction of generated candidates satisfying a predefined property threshold, providing a direct indicator of how reliably the model generates on-target structures. However, a high success rate alone is insufficient: A model that achieves it by repeatedly proposing variations of a single high-scoring scaffold offers limited utility. The number of modes captures the diversity of solutions by counting distinct high-scoring regions of chemical space explored by the model. Rediscovery of known high-performing materials provides an additional diagnostic, testing whether a well-trained model can recover established benchmarks from first principles. Finally, sample efficiency, quantified by the oracle burden, i.e., the number of scoring function calls required to identify N on-target structures, is particularly critical when each oracle evaluation corresponds to an expensive DFT calculation, MD simulation, or wet-lab experiment.

Transformer decoders and diffusion models have produced some of the most compelling results in generative design of OEEMs to date, including experimentally validated electrolyte formulations. A major bottleneck, however, remains the scarcity of high-quality labeled electrochemical data, such that many generative workflows still rely heavily on computational or surrogate oracles for candidate screening and optimization. Flow matching, goal-directed reinforcement learning for electrolyte design, and more property- and budget-aware benchmarking represent promising directions for the next stage of the field.

6 | LLMs in Closed-Loop Agentic Workflows for OEEMs

Previous sections have reviewed data-driven property prediction and generative models for chemical space exploration. A related emerging direction is to place such models inside LLM-based agentic systems, where users can define targets, configure generation tasks, run external tools, and interpret outputs through natural language interaction [210]. For example, ChatINVENT integrates the REINVENT molecular generation framework into a conversational LLM agent, allowing reward functions, molecular representations, generation runs, and result interpretation to be handled through a natural-language interface [211]. Although ChatINVENT has been demonstrated mainly in drug discovery, it illustrates a transferable workflow pattern in which generative models are wrapped in language-based interfaces to configure objectives, run design tasks, and interpret results.

This section adopts this workflow-level perspective for OEEMs; see Figure 5 for an overview. Rather than organizing the discussion by model type, we focus on how LLMs can function within agentic research workflows that connect material representations, unstructured literature, analysis, validation, and iterative feedback.

6.1 | Language Interfaces for Molecules, Formulations, and Polymers

LLM-based workflows begin with the conversion of molecules, formulations, and materials descriptions into language-ready representations. LLMs do not directly access the material itself. They first receive a serialized and tokenized input. The mapping from molecules and materials to language is therefore not a

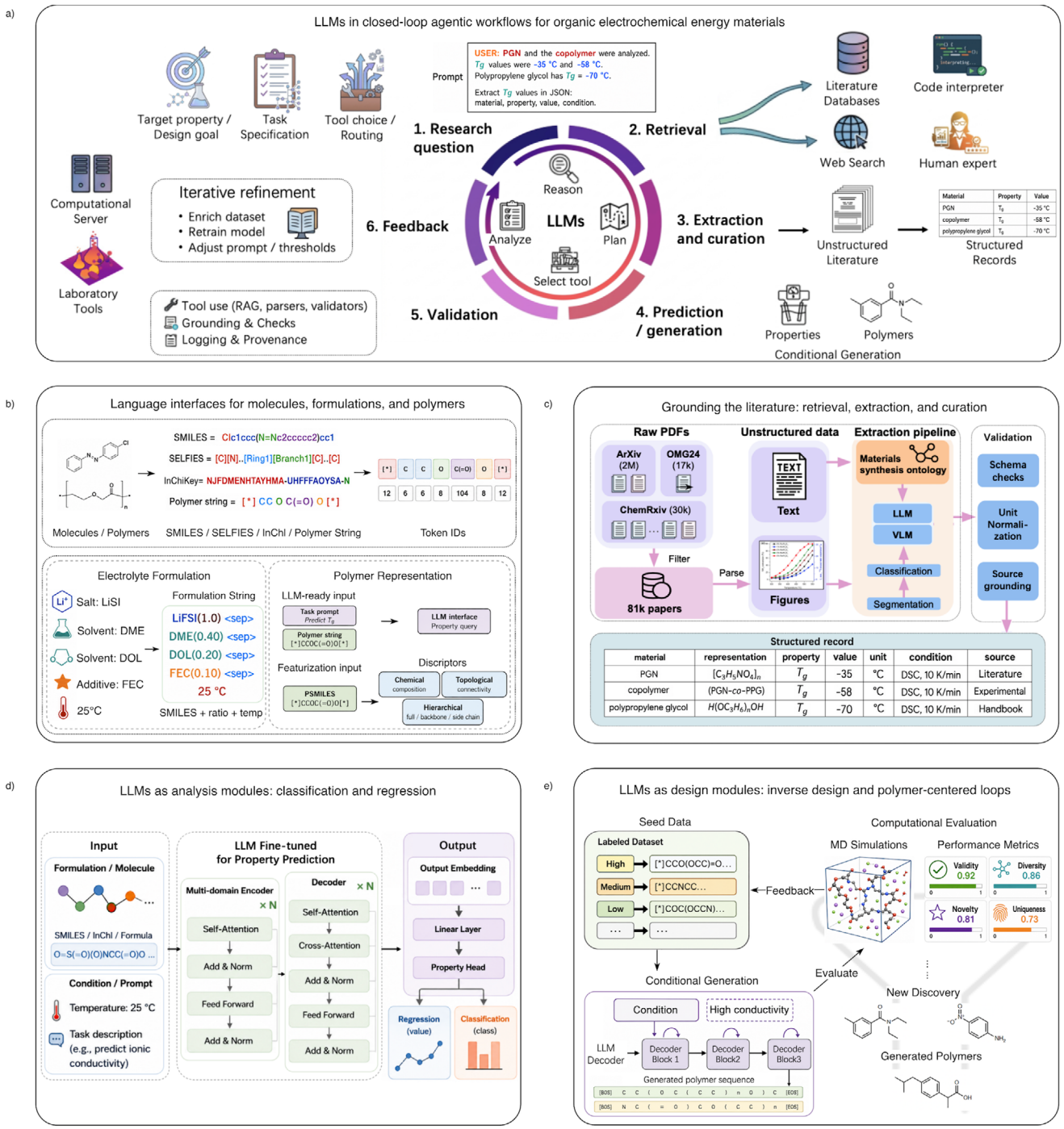

FIGURE 5 | LLMs in closed-loop agentic workflows for OEEMs. (a) Overview of the closed-loop in silico agentic workflow and feedback closure. Some graphical elements are adapted from [214] under the terms of CC-BY license (Copyright 2024 Bran et al.). (b) Language interfaces for molecules, formulations, and polymers. Some graphical elements are adapted from [212] under the terms of CC-BY license (Copyright 2024 Jablonka et al.). (c) Grounding the literature: retrieval, extraction, and curation. Adapted from Lederbauer et al. [215] under the terms of CC-BY license (Copyright 2025 Lederbauer et al.); validation steps added by the authors. (d) LLMs as analysis modules: classification and regression. (e) LLMs as design modules: inverse design and polymer-centered loops. Some graphical elements are adapted from [212] under the terms of CC-BY license (Copyright 2024 Jablonka et al.).

secondary implementation detail, but a prerequisite for the workflow to function.

For small molecules, this interface can take several forms, including IUPAC names, SMILES, SELFIES, and InChI. These representations compress molecular structure into line-based formats that can be embedded into prompts, in-context examples, or fine-tuning datasets. Jablonka et al. showed that chemistry and materials tasks can be reformulated in a language-interfaced way, covering classification, regression, and inverse design across molecules, materials, and reactions [212]. Their results also suggest that useful performance can be obtained across different chemical representations, provided that the representation can be consistently integrated into a language-based task formulation.

Beyond single-molecule inputs, many practical materials problems require formulation-level representations that combine component identities, compositions, and experimental conditions. Language-based interfaces can encode this information within the same prompt. Such representations are compatible with LLM and foundation-model pipelines, as they carry conditional and relational information in a form that is close to natural language.

The need for robust language-ready representations becomes particularly clear for polymers, where digital representation is itself a major challenge. PolyNC addresses this by combining natural language prompts with chemical language in a unified text-to-text framework, where task instructions are paired directly with polymer SMILES. It also reports that tokenization choices affect performance for polymer SMILES, with character-level tokenization outperforming group-wise tokenization in the tasks evaluated in that study [213]. This finding may not generalize across all datasets or polymer language models, but it shows that representation choices can affect model performance. PolyMetriX further emphasizes that polymer informatics also requires standardized and hierarchical digital representations, including full-polymer, backbone, and side-chain levels, to support robust downstream learning and reproducible workflows [122]. Once materials are converted into language-ready records, LLMs can more effectively support both upstream curation and downstream prediction or design.

6.2 | Grounding the Literature: Retrieval, Extraction, and Curation

LLMs can help transform unstructured literature and experimental reports into structured, grounded, and reusable records. For OEEMs, relevant knowledge is rarely presented in a single clean format. Instead, it is often distributed across the main text, supplementary information, tables, figure captions, and experimental sections. The practical challenge is therefore not only to locate relevant papers, but also to convert dispersed scientific information into records that can support downstream analysis and design.

A first step in this process is retrieval. In an LLM-based workflow, retrieval often begins with a research question, target property, or material class, and proceeds through query-guided document selection and passage-level localization. At this stage, the goal is to identify the specific textual and visual evidence that contains synthesis details, material identifiers, test conditions, and reported outcomes. This step narrows a large corpus into a smaller set of relevant and interpretable evidence units.

Once relevant passages have been located, LLMs can support extraction and curation. Here, the task is to identify material names, structural representations, formulations, synthesis procedures, test conditions, property values, units, and provenance, and to convert them into schema-consistent records. A useful schema may include fields such as material, representation, property, value, unit, condition, and source. The task therefore goes beyond summarization. It becomes a process of building machine-readable scientific records that can be reused across datasets, benchmarks, and closed-loop design pipelines.

Because these records are intended for downstream prediction or generation, extraction alone is usually not sufficient. They also need to be grounded and validated. This includes source linking, unit normalization, range checks, schema validation, and consistency checks across text, tables, and figures. Such safeguards are important because hallucinated or weakly grounded records can propagate errors into later stages of the workflow. ChemCrow provides a general chemistry example of this principle, showing how tool-grounded agents can constrain LLM outputs with external chemistry tools and knowledge sources [214]. For OEEM workflows, the same logic would need to be implemented through domain-specific databases, formulation constraints, and validation rules.

LeMat-Synth provides a relevant example of this workflow by combining LLMs and vision-language models (VLMs) to extract synthesis procedures and performance data from text and figures, and to organize the outputs into ontology-structured, machine-readable records [215]. MaCBench complements this perspective by evaluating data extraction, experimental understanding, and results interpretation in chemistry and materials science, while also highlighting current limitations of multimodal scientific assistants [216]. Once curated and validated records are available, they can feed into the predictive and generative modules discussed in the following subsections.

Recently, Wang et al. introduced Chat-RFB, a question-answering framework for redox flow batteries that combines LLMs with knowledge graphs [217]. Using a corpus of 5353 articles, they generated a structured database with 164232 nodes and 853939 relationships. The system outperformed general-purpose LLMs on redox-flow-battery questions (> 90% accuracy) and reduced hallucinations compared with native models, such as GPT-4o and DeepSeek-v3.

6.3 | LLMs as Analysis Modules: Classification and Regression

After knowledge has been curated into structured records, LLMs can support classification and regression tasks that turn language-ready inputs into property estimates, candidate scores, or rankings for later screening and validation.

Jablonka et al. provide a representative example for small molecules, where GPT-3 was fine-tuned to answer chemistry questions in natural language across classification and regression tasks [212]. Their work shows that molecular property prediction can be reformulated as a question–answer or structured completion problem, using representations such as IUPAC names, SMILES, or SELFIES. This formulation is especially relevant in low-data settings, where the fine-tuned model was often competitive with, or even better than, dedicated baselines. Related results have also been reported for organic molecules by Xie et al., who showed that fine-tuned GPT-3 can identify chemically meaningful patterns from SMILES and perform robust coarse-grained molecular property classification [218].

Formulation-based electrolyte systems provide a complementary case, where the input is no longer a single structure but a composition-dependent mixture under external conditions. Zohair et al. addressed this setting by fine-tuning a chemical foundation model on 13 666 ionic conductivity measurements curated from the lithium-ion battery literature. Their SMI-TED-IC model represents each formulation as a string that combines the canonical SMILES of each constituent with its composition, and then concatenates this representation with temperature as an additional condition [219]. This case is particularly relevant for the present review because it shows how LLM-style or foundation-model-style interfaces can handle mixed-format inputs, where molecular identity, formulation ratio, and experimental context need to be considered together. Here, regression helps rank and screen candidate formulations before more resource-intensive validation steps. Classification and regression can therefore serve as useful analytical components in an agentic materials workflow. They help triage candidates, prioritize experiments, and provide structured signals for subsequent validation or design. For example, PolyNC combines natural language prompts with polymer SMILES in a unified framework that supports both regression and classification [213].

6.4 | LLMs as Design Modules: Inverse Design and Polymer-Centered Loops

Beyond analysis, LLMs can also support inverse design by proposing candidates under target constraints and linking generation to scoring, filtering, simulation, database lookup, and iterative refinement. Polymer electrolytes are a natural focus for this discussion because they require sequence-like representations, property-conditioned generation, and downstream validation, making them a useful setting for examining how language-based interfaces can support materials design [17].

A recent ChemRxiv preprint on PolySea illustrates this emerging direction in polymer informatics. It reports a domain-specific LLM framework with capabilities in property prediction, inverse design, and knowledge extraction. In the inverse design setting, PolySea is described as generating polymer structures under target property constraints, followed by evaluation through surrogate prediction and external validation. The generated polymers are not treated as final answers. Instead, they are further assessed with a GNN surrogate, Polymer Genome, and density functional theory calculations, and promising candidates are retained for subsequent analysis [220]. This example suggests how generation, evaluation, and iteration could be connected in a polymer-centered design loop.

These design loops also rely on the polymer representations and data infrastructure discussed above. Candidate generation is only useful when generated structures can be represented, screened, and validated in a reproducible workflow. For polymer electrolytes, this means that structure encoding, target-property specification, and downstream validation need to remain connected throughout the design process.

Uni-Electrolyte provides a related battery-oriented example. It is an artificial intelligence platform for electrolyte molecule design that integrates candidate generation with property-based screening, retrosynthesis planning, and reaction-network analysis [221]. Although this platform is not polymer-specific, it is still useful as a complementary case because it shows that small-molecule electrolyte design can also be organized as a multi-stage loop rather than a one-shot search process. Across these examples, LLM-enabled design becomes more useful when it is coupled to downstream scoring, simulation, database checks, and iterative feedback, rather than treated as a standalone generation step.

6.5 | Closing the Loop: Feedback and Iterative Refinement

For LLM-based agentic workflows to remain useful across repeated discovery cycles, evaluation should function not only as assessment, but also as a feedback mechanism. In a closed-loop setting, the feedback can guide later curation, prompt revision, model adaptation, candidate filtering, and validation.

ChemBench, MaCBench, and LeMat-GenBench provide useful reference points for this purpose. ChemBench offers a systematic framework for evaluating chemical knowledge and reasoning, and is designed to support both standard and tool-augmented systems [222]. MaCBench extends this perspective to multimodal chemistry and materials tasks that are closer to realistic workflows, including data extraction, experimental understanding, and results interpretation [216]. LeMat-GenBench evaluates generative outputs through a multi-dimensional metric suite including validity, stability, novelty, uniqueness, diversity, and efficiency [223]. Although it is focused on crystalline materials, it still illustrates an important point for design workflows: downstream feedback is often multi-objective and cannot be reduced to a single score.

Such signals become useful when they can be fed back into the system. In practice, this may involve enriching datasets with newly validated records, retraining or fine-tuning models on corrected examples, and adjusting prompts, schemas, thresholds, or tool-routing strategies. This logic is already visible in tool-augmented scientific agents. In ChemCrow, the agent iteratively observes tool responses and adapts its next action accordingly, and can revise synthesis procedures on the basis of platform validation feedback until a valid executable procedure is obtained [214]. At the data and workflow level, infrastructures such as LeMat-Synth and PolyMetriX suggest how this feedback can be stabilized through structured records, extensible schemas, curated datasets, and reproducible pipelines that support later rounds of prediction or design [122, 215].

Practical limitations also become more informative in this setting. Hallucinations, unit mismatches, representation ambiguity, incomplete context, and benchmark-to-practice gaps do not only reduce reliability; they also indicate where the workflow needs correction. ChemBench shows that strong models can still be overconfident and limited on some knowledge-intensive tasks [222]. MaCBench similarly shows that multimodal systems can perform well on perception-heavy tasks while still struggling with spatial reasoning, cross-modal synthesis, and multi-step inference [216]. In a feedback-oriented workflow, such failure modes can guide the next intervention, for example by improving

validation rules, refining prompts, integrating better databases, or retraining models on more representative examples.

LLMs are increasingly being used to connect individual components of materials discovery workflows, transforming dispersed knowledge into evidence that can support prediction, simulation, and materials design. Current studies largely showcase individual components or partial loops. However, fully integrated OEEM-specific agentic workflows that connect representation, prediction, generation, simulation, experimental validation, and iterative feedback have yet to be convincingly demonstrated.

7 | Where We are and Where to go

Taken together, the examples surveyed in this review show that AI is no longer a peripheral tool for OEEMs, but an emerging design layer connecting representations, data, property prediction, atomistic simulation, generative models, and LLM-assisted workflows. Yet the field will advance only if methodological progress is matched by rigorous validation, transparent reporting, reusable data and code, and closer alignment between computational targets and experimentally measurable quantities. We therefore close by outlining practical best practices, identifying key missing pieces, and emphasizing the need to narrow the experiment–modeling gap.

7.1 | Starting With the Best Practices

To help researchers entering the field evaluate and report their own work before publication, we have made a practical checklist, extending the best-practice recommendations of Artrith et al. [224]:

1. **Baseline models**: Simple baselines, such as linear or fingerprint-based models, should be established before deploying more complex architectures. If a deep model does not outperform a fingerprint-based ridge regression or random forest model, the added complexity is unwarranted—particularly on the small datasets typical of this field.
2. **Internal validation**: Reliable estimates of generalization depend on careful validation. Unbiased data splitting (e.g., stratified, scaffold-based, or polymer-specific splits) helps prevent data leakage [122], and CV strategies such as five-fold CV should be used to obtain statistically meaningful performance estimates [97].
3. **External validation**: Model predictions, whether of properties or of generated structures, should be treated as hypotheses until corroborated by higher-fidelity calculations, MD simulations, or wet lab experiments.
4. **Sample efficiency over surface metrics**: For generative models, resist over-optimizing easy distributional metrics such as validity, uniqueness, or novelty ("V.U.N."), which are necessary but weak indicators of practical utility. The metric that matters most is *sample efficiency*: how many high-performing candidates a model generates per oracle call [201] or labeled data point seen. Because each oracle call ultimately corresponds to a costly DFT calculation, MD simulation, or wet lab experiment, sample efficiency maps directly onto the real expense of discovering a new material. Generative methods should thus be benchmarked under a fixed evaluation budget, reporting the number (or quality) of candidates meeting the design target per fixed number of oracle calls.
5. **Machine-learning interatomic potentials**: Where possible, build on foundation MLIPs (e.g., via the Matbench Discovery leaderboard[1]) rather than training bespoke potentials from scratch. Validate them within the chemical and thermodynamic space of interest to avoid extrapolation, and use distillation or fine-tuning of foundation models to balance accuracy and efficiency [152].
6. **Grounded and auditable LLM workflows**: LLM-based agents warrant particular caution when their outputs feed later stages of a discovery pipeline, since a hallucinated value, missing experimental condition, or unsupported design suggestion can be propagated downstream as if it were reliable data. Each LLM-assisted step should produce structured outputs that remain traceable to a source document or an external domain tool; extracted records should be validated before entering any dataset, and generated candidates should be treated as hypotheses until independently verified [210, 214].
7. **Reproducibility**: Reproducible and sustainable machine-learning research in materials discovery, both within individual research groups and across the broader community, requires open data and code. Datasets, preprocessing workflows, trained model weights, and training and evaluation scripts should be made freely available upon publication through open repositories.

7.2 | Looking for the Missing Points

The best practices above describe how to do AI-driven design well; we close by highlighting four computational gaps and one in the experiment-modeling coupling, in our view, most limit progress currently.

The first is *polymer representation*. Despite the proliferation of polymer-aware formats surveyed in Section 2, no representation yet captures the stochastic, hierarchical, and conformational nature of real polymer structures in a form that is both learnable and widely adopted. Most studies still collapse a polymer to a single repeat unit, discarding the sequence, dispersity, tacticity, branching, end-group, and morphological information that often governs properties from T_g to ionic conductivity. We need representations that are explicitly hierarchical—linking monomer chemistry, chain architecture, and ensemble descriptors—include synthesis and processing information, and are standardized, so that a given polymer maps to the same representation across datasets and models [122]. Progress here would propagate to every downstream task in this review.

A second gap is *integration*. The methods surveyed here—property predictors (Section 3), MLIPs (Section 4), and generative models (Section 5)—are still largely built and benchmarked as

siloed components, whereas the agentic workflows of Section 6 point toward closed loops that retrieve knowledge, propose candidates, validate them by simulation or experiment, and refine the search through feedback. Realizing this for OEEMs takes more than wrapping existing models in a language interface: it needs interoperable interfaces between representation, prediction, simulation, and generation; verifiable feedback and human alignment; and uncertainty-aware policies for deciding which expensive oracle to call next. The building blocks now exist as individual models, but end-to-end demonstrations on battery-relevant systems remain scarce.

A third gap, underlying the first two, is *data*. As the compilation in Table 1 shows, experimental datasets for organic electrochemical materials are small, heterogeneous, and scattered across incompatible formats, limiting both training and fair comparison. Lowering the barriers to data sharing and creating FAIR[2], machine-readable records with standardized units and metadata, shared benchmarks, and routine release of datasets, splits, and model weights, would advance the field as much as, if not more than, any new architecture. Experience from more mature data-sharing efforts in chemistry suggests that the decisive factors are consistent schemas, low-friction deposition, and community norms or mandates, rather than the underlying technology [225]. The extraction pipelines of Section 6 offer one route to turn dispersed published results into reusable structured records [215].

A fourth gap is *evaluation* itself. Generative models are still judged largely by distributional proxies (e.g., V.U.N.) that are easily inflated and only weakly predictive of practical value. What ultimately matters is sample efficiency: the number of high-quality candidates surfaced per oracle call [201], since each call corresponds to a costly DFT calculation, MD simulation, or experiment and therefore sets the real cost of discovery. Benchmark suites such as ChemBench [222] and LeMat-GenBench [223] are beginning to standardize multi-objective evaluation, but the community still lacks budget-aware benchmarks that report how many candidates meet a design target under a fixed number of oracle calls. Closing the loop between shared data, the right metrics, and open benchmarks so that progress is measured by candidates discovered per experiment rather than by distributional scores is, in our view, one of the clearest gaps warranting research investment in the next phase of AI-accelerated design. [225]

7.3 | Closing the Experiment–Modeling Gap

The last but not least hurdle for AI-based exploration of organic electrochemical materials – or for all in silico methods exploring these chemistries – is the lack of reliable experimental data on similar systems, and systematic approaches to experimentally generate large sets of high-quality data. There are several dimensions to this complexity. First, the energy field is an applied area, and the systems explored experimentally are often device oriented and target materials performance, rather than distinct physical and chemical properties. Therefore, high-quality data are scare, and with limited reproducibility and uniformity. Second, energy devices are complex, often comprising many different interacting components, both in terms of electrochemically active and inactive materials. This renders it profoundly difficult to obtain data, which are straight-forwardly comparable to those predicted computationally, and which are often based on more simplistic chemical systems. Third, there are profound uncertainties in the experimental measurement data which is clearly related to unsharp methodology, less robust instrumentation and difficulties in interpretations. Both thermal (e.g., T_g) and electrochemical responses (e.g., redox potentials) are often generated with large errors and poor reproducibility, and depend on scan rate, instrumentation, setups, etc., and therefore difficult to compare with computationally generated results. For example, there is a shortage of adequate reference electrode systems, which can generate reliable voltages for the electroactive components in these systems. Moreover, many methods (e.g., cyclic voltammetry) generally function reliably in solution chemistry setups, but not in setups with are adequate for many of the targeted devices, such as batteries. Fourth, the employed organic materials are generally far from uniform. They are semi-crystalline, contain different types of defects, have varying speciation and coordination chemistry, etc., which means that they in principle should not generate a uniform electrochemical response. This is related to the very nature of organic compounds in electrochemical systems, where different reactions and processes co-exist and are difficult to separate in the data. These factors all affect the possibility to do proper benchmarking, and thereby to ultimately perform a robust assessment of the accuracy of these AI-based methods.

There is thus a need to, parallel to the rapid progress of AI-based methods, develop novel experimental methodologies for exploration of organic energy materials. There are today many inherent shortcomings. For example, the standard measurement techniques of the ECW are today based on voltametric methods, where electrode substrates, side-reaction and the employed solution chemistry will play significant roles and influence data quality. [226] Ionic conductivity and transference [227], in turn, rely heavily on data interpretation from less sharp electrochemical methods, such as impedance spectroscopy and polarization. [228] The very different results for ionic mobility generated by nuclear magnetic resonance (NMR) techniques and electrochemical methods – naturally depending on that they are performed under either equilibrium or non-equilibrium conditions, respectively – highlights the difficulty of comparisons with computationally obtained results. [229] Here, a better understanding of the methodologies need to go hand-in-hand with development of more advanced methods that are now becoming available, often utilizing large-scale infrastructure, that are increasing the resolution in both space and time. [230] Moreover, exploration of useful and robust model systems – not necessarily the best performing materials – for polymer electrolytes and redox-active organic compounds will provide more reliable datasets and interpretations.

Author Contributions

Zhan-Yun Zhang: data curation, investigation, methodology, project administration, visualization, writing – original draft preparation. **Giannis Savvas**: data curation, investigation, methodology, visualization, writing – original draft preparation. **Jichen Li**: investigation, methodology, visualization, writing – original draft preparation. **Yue Zhou**: investigation, methodology, visualization, writing – original draft prepara-

tion. **Thanh Trung Le**: investigation, methodology, visualization, writing – original draft preparation. **Daniel Brandell**: supervision, resources, writing – original draft preparation, writing – review & editing. **Rocío Mercado**: supervision, methodology, funding acquisition, writing – original draft preparation, writing – review & editing. **Chao Zhang**: supervision, methodology, funding acquisition, project administration, writing – original draft preparation, writing – review & editing.

Acknowledgments

This work was partially supported by the Wallenberg Initiative Materials Science for Sustainability (WISE) funded by the Knut and Alice Wallenberg Foundation and the Swedish Energy Agency (Project no. P2025-03176). GS and RM acknowledge funding support from the Wallenberg AI, Autonomous Systems, and Software Program (WASP), supported by the Knut and Alice Wallenberg Foundation. The authors declare no competing interests.

Conflicts of Interest

The authors declare no conflicts of interests.

Data Availability Statement

The data that support the findings of this study are openly available in GitHub at https://github.com/Teoroo-CMC/OEEM-data-catalog.

Endnotes

[1] https://matbench-discovery.materialsproject.org

[2] Findable, Accessible, Interoperable, Reusable

References

1. D. Larcher and J.-M. Tarascon, "Towards Greener and More Sustainable Batteries for Electrical Energy Storage," *Nature Chemistry* 7, no. 1 (January 2015): 19–29, https://doi.org/10.1038/nchem.2085.

2. J. Kim, Y. Kim, J. Yoo, G. Kwon, Y. Ko, and K. Kang, "Organic Batteries for a Greener Rechargeable World," *Nature Reviews Materials* 8, no. 1 (September 2022): 54–70, https://doi.org/10.1038/s41578-022-00478-1.

3. M. Li, R. P. Hicks, Z. Chen, et al., "Electrolytes in Organic Batteries," *Chemical Reviews* 123, no. 4 (February 2023): 1712–1773, https://doi.org/10.1021/acs.chemrev.2c00374.

4. B. D. Paulsen, K. Tybrandt, E. Stavrinidou, and J. Rivnay, "Organic Mixed Ionic–Electronic Conductors," *Nature Materials* 19, no. 1 (January 2020): 13–26, https://doi.org/10.1038/s41563-019-0435-z.

5. S. T. Keene, V. Gueskine, M. Berggren, G. G. Malliaras, K. Tybrandt, and I. Zozoulenko, "Exploiting Mixed Conducting Polymers in Organic and Bioelectronic Devices," *Physical Chemistry Chemical Physics* 24, no. 32 (2022): 19144–19163, https://doi.org/10.1039/D2CP02595G.

6. Z.-Y. Zhang, R. Mercado, T. T. Le, and C. Zhang, "Unlocking the Chemical Space for Rechargeable Batteries with a Generative Solvent Design System," *ACS Nano* 20, no. 29 (July 2026): 20714–20729, https://doi.org/10.1021/acsnano.6c06255.

7. J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural Message Passing for Quantum Chemistry," In *International Conference on Machine Learning* (PMLR, 2017), 1263–1272.

8. K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, "SchNet – A Deep Learning Architecture for Molecules and Materials," *Journal of Chemical Physics* 148, no. 24 (March 2018): 241722, https://doi.org/10.1063/1.5019779.

9. S. Batzner, A. Musaelian, L. Sun, et al., "E(3)-Equivariant Graph Neural Networks for Data-Efficient and Accurate Interatomic Potentials," *Nature Communications* 13, no. 1 (May 2022): 2453, https://doi.org/10.1038/s41467-022-29939-5.

10. C. Kim, A. Chandrasekaran, T. D. Huan, D. Das, and R. Ramprasad, "Polymer Genome: A Data-Powered Polymer Informatics Platform for Property Predictions," *Journal of Physical Chemistry C* 122, no. 31 (2018): 17575–17585, https://doi.org/10.1021/acs.jpcc.8b02913.

11. T.-S. Lin, C. W. Coley, H. Mochigase, et al., "BigSMILES: A Structurally-Based Line Notation for Describing Macromolecules," *ACS Central Science* 5, no. 9 (2019): 1523–1531, https://doi.org/10.1021/acscentsci.9b00476.

12. J. Behler, "Four Generations of High-Dimensional Neural Network Potentials," *Chemical Reviews* 121, no. 16 (August 2021): 10037–10072, https://doi.org/10.1021/acs.chemrev.0c00868.

13. V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, "Gaussian Process Regression for Materials and Molecules," *Chemical Reviews* 121, no. 16 (August 2021): 10073–10141, https://doi.org/10.1021/acs.chemrev.1c00022.

14. I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csanyi, "MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields," *Advances in Neural Information Processing Systems* 35 (December 2022): 11423–11436.

15. J. Noh, H. A. Doan, H. Job, et al., "An Integrated High-Throughput Robotic Platform and Active Learning Approach for Accelerated Discovery of Optimal Electrolyte Formulations," *Nature Communications* 15, no. 1 (March 2024): 2757, https://doi.org/10.1038/s41467-024-47070-5.

16. E. Jin, A. C. Nica, M. Galkin, et al., "OXtal: An All-Atom Diffusion Model for Organic Crystal Structure Prediction," *arXiv* (2025): arXiv:2512.06987 [cs].

17. Z. Yang, W. Ye, X. Lei, D. Schweigert, H.-K. Kwon, and A. Khajeh, "De novo design of polymer electrolytes using GPT-based and diffusion-based generative models," *npj Computational Materials* 10, no. 1 (2024): 296, https://doi.org/10.1038/s41524-024-01470-9.

18. R. P. Fornari and P. De Silva, "Molecular Modeling of Organic Redox-Active Battery Materials," *WIREs Computational Molecular Science* 11, no. 2 (March 2021): e1495, https://doi.org/10.1002/wcms.1495.

19. K. Wang, H. Shi, T. Li, et al., "Computational and Data-Driven Modelling of Solid Polymer Electrolytes," *Digital Discovery* 2, no. 6 (2023): 1660–1682, https://doi.org/10.1039/d3dd00078h.

20. A. Mistry, A. A. Franco, S. J. Cooper, S. A. Roberts, and V. Viswanathan, "How Machine Learning Will Revolutionize Electrochemical Sciences," *ACS Energy Letters* (March 2021): 1422–1431, https://doi.org/10.1021/acsenergylett.1c00194.

21. S. Li and F. You, "GenAI for Scientific Discovery in Electrochemical Energy Storage: State-of-the-Art and Perspectives From Nano- and Micro-Scale," *Small* 20, no. 50 (December 2024): 2406153, https://doi.org/10.1002/smll.202406153.

22. Y.-C. Gao, X. Chen, Y.-H. Yuan, et al., "Accelerating Battery Innovation: AI-Powered Molecular Discovery," *Chemical Society Reviews* 54, no. 21 (2025): 9630–9684, https://doi.org/10.1039/D5CS00053J.

23. T. Lombardo, M. Duquesnoy, H. El-Bouysidy, et al., "Artificial Intelligence Applied to Battery Research: Hype or Reality?," *Chemical Reviews* 122, no. 12 (June 2022): 10899–10969, https://doi.org/10.1021/acs.chemrev.1c00108.

24. B. Wang, Z. Xu, Z. Han, et al., "SLICES-PLUS: A Crystal Representation Leveraging Spatial Symmetry," *Materials & Design* 253 (May 2025): 113856, https://doi.org/10.1016/j.matdes.2025.113856.

25. Y. Shao, M. Hellström, P. D. Mitev, L. Knijff, and C. Zhang, "PiNN: A Python Library for Building Atomic Neural Networks of Molecules and Materials," *Journal of Chemical Information and Modeling* 60, no. 3 (March 2020): 1184–1193, https://doi.org/10.1021/acs.jcim.9b00994.

26. M. Aldeghi and C. W. Coley, "A Graph Representation of Molecular Ensembles for Polymer Property Prediction," *Chemical Science* 13, no. 35 (2022): 10486–10498, https://doi.org/10.1039/d2sc02839e.

27. T. Xie and J. C. Grossman, "Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material

Properties," *Physical Review Letters* 120, no. 14 (April 2018): 145301, https://doi.org/10.1103/PhysRevLett.120.145301.

28. J. Behler, "Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials," *The Journal of Chemical Physics* 134, no. 7 (February 2011): 074106, https://doi.org/10.1063/1.3553717.

29. D. Weininger, "SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules," *Journal of Chemical Information and Computer Sciences* 28, no. 1 (February 1988): 31–36, https://doi.org/10.1021/ci00057a005.

30. M. Krenn, F. Häse, A. Nigam, P. Friederich, and A. Aspuru-Guzik, "Self-Referencing Embedded Strings (SELFIES): A 100% Robust Molecular String Representation," *Machine Learning: Science and Technology* 1, no. 4 (October 2020): 045024, https://doi.org/10.1088/2632-2153/aba947.

31. N. O'Boyle and A. Dalke, *DeepSMILES: An Adaptation of SMILES for Use in Machine-Learning of Chemical Structures*, (ChemRxiv, 2018), no. 0919, https://doi.org/10.26434/chemrxiv.7097960.v1.

32. J. Li and X. Jiang, "Mol-BERT: An Effective Molecular Representation with BERT for Molecular Property Prediction," *Wireless Communications and Mobile Computing* 2021, no. 1 (2021): 7181815, https://doi.org/10.1155/2021/7181815.

33. R. Singh, A. A. Barsainyan, R. Irfan, et al., "ChemBERTa-3: An Open-Source Training Framework for Chemical Foundation Models," *Digital Discovery* 5, no. 2 (February 2026): 662–685, https://doi.org/10.1039/D5DD00348B.

34. L. Schneider, D. Walsh, B. Olsen, and J. D. Pablo, "Generative BigSMILES: An Extension for Polymer Informatics, Computer Simulations & ML/AI," *Digital Discovery* 3, no. 1 (January 2024): 51–61, https://doi.org/10.1039/D3DD00147D.

35. H. Sahu, W. Xiong, A. Savit, S. S. Shukla, and R. Ramprasad, "POLYT5: An Encoder-Decoder Foundation Chemical Language Model for Generative Polymer Design," *npj Artificial Intelligence* 2, no. 1 (March 2026): 30, https://doi.org/10.1038/s44387-026-00087-1.

36. J. Ahn, G. P. Irianti, Y. Choe, and S.-M. Hur, "Enhancing Deep Learning Predictive Models with HAPPY (Hierarchically Abstracted rePeat Unit of PolYmers) Representation," *npj Computational Materials* 10, no. 1 (May 2024): 110, https://doi.org/10.1038/s41524-024-01293-8.

37. C. Kuenneth and R. Ramprasad, "polyBERT: A Chemical Language Model to Enable Fully Machine-Driven Ultrafast Polymer Informatics," *Nature Communications* 14, no. 1 (July 2023): 4099, https://doi.org/10.1038/s41467-023-39868-6.

38. C. Xu, Y. Wang, and A. Barati Farimani, "TransPolymer: A Transformer-Based Language Model for Polymer Property Predictions," *npj Computational Materials* 9, no. 1 (April 2023): 64, https://doi.org/10.1038/s41524-023-01016-5.

39. N. Kazeev, W. Nong, I. Romanov, et al., "Wyckoff Transformer: Generation of Symmetric Crystals," arXiv (2025): arXiv:2503.02407.

40. D. Duvenaud, D. Maclaurin, J. Aguilera-Iparraguirre, et al., "Convolutional Networks on Graphs for Learning Molecular Fingerprints," *arXiv* (2015): arXiv:1509.09292.

41. S. Kearnes, K. McCloskey, M. Berndl, V. Pande, and P. Riley, "Molecular Graph Convolutions: Moving Beyond Fingerprints," *Journal of Computer-Aided Molecular Design* 30, no. 8 (August 2016): 595–608, https://doi.org/10.1007/s10822-016-9938-8.

42. K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, "Quantum-Chemical Insights From Deep Tensor Neural Networks," *Nature Communications* 8, no. 1 (January 2017): 13890, https://doi.org/10.1038/ncomms13890.

43. J. Gasteiger, J. Groß, and S. Günnemann, "Directional Message Passing for Molecular Graphs," *arXiv* (2022): arXiv:2003.03123.

44. Y. Liu, L. Wang, M. Liu, X. Zhang, B. Oztekin, and S. Ji, "Spherical Message Passing for 3D Graph Networks," *arXiv* (2021): arXiv:2102.05013.

45. J. Gasteiger, F. Becker, and S. Günnemann, "GemNet: Universal Directional Graph Neural Networks for Molecules," *arXiv* (2021): arXiv:2106.08903.

46. K. Schütt, O. Unke, and M. Gastegger, "Equivariant Message Passing for the Prediction of Tensorial Properties and Molecular Spectra," in *Proceedings of the 38th International Conference on Machine Learning* (PMLR, 2021), 9377–9388.

47. R. Gurnani, C. Kuenneth, A. Toland, and R. Ramprasad, "Polymer Informatics at Scale with Multitask Graph Neural Networks," *Chemistry of Materials* 35, no. 4 (February 2023): 1560–1567, https://doi.org/10.1021/acs.chemmater.2c02991.

48. C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong, "Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals," *Chemistry of Materials* 31, no. 9 (May 2019): 3564–3572, https://doi.org/10.1021/acs.chemmater.9b01294.

49. K. Choudhary and B. DeCost, "Atomistic Line Graph Neural Network for Improved Materials Property Predictions," *npj Computational Materials* 7, no. 1 (November 2021): 185, https://doi.org/10.1038/s41524-021-00650-1.

50. R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors* (Wiley, 2000).

51. H. Moriwaki, Y.-S. Tian, N. Kawashita, and T. Takagi, "Mordred: A Molecular Descriptor Calculator," *Journal of Cheminformatics* 10, no. 1 (February 2018): 4, https://doi.org/10.1186/s13321-018-0258-y.

52. G. Landrum, P. Tosco, B. Kelley, et al., *rdkit/rdkit: 2026_03_1 (Q1 2026) Release*,' (Zenodo, 2026).

53. K. Muuga, L. Knijff, and C. Zhang, "Molecular Electrostatic Potentials From Machine Learning Models for Dipole and Quadrupole Predictions," *AI for Science* 2, no. 2 (June 2026): 025002, https://doi.org/10.1088/3050-287X/ae531a.

54. R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, "Machine Learning in Materials Informatics: Recent Applications and Prospects," *npj Computational Materials* 3, no. 1 (December 2017): 54, https://doi.org/10.1038/s41524-017-0056-5.

55. F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, and M. Ceriotti, "Physics-Inspired Structural Representations for Molecules and Materials," *Chemical Reviews* 121, no. 16 (August 2021): 9759–9815, https://doi.org/10.1021/acs.chemrev.1c00021.

56. J. L. Durant, B. A. Leland, D. R. Henry, and J. G. Nourse, "Reoptimization of MDL Keys for Use in Drug Discovery," *Journal of Chemical Information and Computer Sciences* 42, no. 6 (2002): 1273–1280, https://doi.org/10.1021/ci010132r.

57. D. Rogers and M. Hahn, "Extended-Connectivity Fingerprints," *Journal of Chemical Information and Modeling* 50, no. 5 (May 2010): 742–754, https://doi.org/10.1021/ci100050t.

58. L. Tao, G. Chen, and Y. Li, "Machine Learning Discovery of High-Temperature Polymers," *Patterns* 2, no. 4 (April 2021): 100225, https://doi.org/10.1016/j.patter.2021.100225.

59. M. F. Langer, A. Goeßmann, and M. Rupp, "Representations of Molecules and Materials for Interpolation of Quantum-Mechanical Simulations via Machine Learning," *npj Computational Materials* 8, no. 1 (March 2022): 41, https://doi.org/10.1038/s41524-022-00721-x.

60. M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, "Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning," *Physical Review Letters* 108, no. 5 (January 2012): 058301, https://doi.org/10.1103/PhysRevLett.108.058301.

61. K. Hansen, F. Biegler, R. Ramakrishnan, et al., "Machine Learning Predictions of Molecular Properties: Accurate Many-Body Potentials and Nonlocality in Chemical Space," *Journal of Physical Chemistry Letters* 6, no. 12 (June 2015): 2326–2331, https://doi.org/10.1021/acs.jpclett.5b00831.

62. B. Huang and O. A. Von Lilienfeld, "Communication: Understanding Molecular Representations in Machine Learning: The Role of Uniqueness

and Target Similarity," *Journal of Chemical Physics* 145, no. 16 (October 2016): 161102, https://doi.org/10.1063/1.4964627.

63. F. A. Faber, L. Hutchison, B. Huang, et al., "Prediction Errors of Molecular Machine Learning Models Lower than Hybrid DFT Error," *Journal of Chemical Theory and Computation* 13, no. 11 (November 2017): 5255–5264, https://doi.org/10.1021/acs.jctc.7b00577.

64. M. R. Dobbelaere, P. P. Plehiers, R. Van de Vijver, C. V. Stevens, and K. M. Van Geem, "Learning Molecular Representations for Thermochemistry Prediction of Cyclic Hydrocarbons and Oxygenates," *The Journal of Physical Chemistry A* 125, no. 23 (June 2021): 5166–5179, https://doi.org/10.1021/acs.jpca.1c01956.

65. H. Huo and M. Rupp, "Unified Representation of Molecules and Crystals for Machine Learning," *Machine Learning: Science and Technology* 3, no. 4 (December 2022): 045017, https://doi.org/10.1088/2632-2153/aca005.

66. A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, "Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons," *Physical Review Letters* 104, no. 13 (April 2010): 136403, https://doi.org/10.1103/PhysRevLett.104.136403.

67. A. P. Bartók, R. Kondor, and G. Csányi, "On Representing Chemical Environments," *Physical Review B* 87, no. 18 (May 2013): 184115, https://doi.org/10.1103/PhysRevB.87.184115.

68. A. S. Christensen, L. A. Bratholm, F. A. Faber, and O. Anatole von Lilienfeld, "FCHL Revisited: Faster and More Accurate Quantum Machine Learning," *Journal of Chemical Physics* 152, no. 4 (January 2020): 044107, https://doi.org/10.1063/1.5126701.

69. R. Drautz, "Atomic Cluster Expansion for Accurate and Transferable Interatomic Potentials," *Physical Review B* 99, no. 1 (January 2019): 014104, https://doi.org/10.1103/PhysRevB.99.014104.

70. H. Doan Tran, C. Kim, L. Chen, et al., "Machine-Learning Predictions of Polymer Properties with Polymer Genome," *Journal of Applied Physics* 128, no. 17 (November 2020): 171104, https://doi.org/10.1063/5.0023759.

71. R. A. Patel, C. H. Borca, and M. A. Webb, "Featurization Strategies for Polymer Sequence or Composition Design by Machine Learning," *Molecular Systems Design & Engineering* 7, no. 6 (2022): 661–676, https://doi.org/10.1039/d1me00160d.

72. R. Ma, D. Huang, T. Zhang, and T. Luo, "Determining Influential Descriptors for Polymer Chain Conformation Based on Empirical Force Fields and Molecular Dynamics Simulations," *Chemical Physics Letters* 704 (July 2018): 49–54, https://doi.org/10.1016/j.cplett.2018.05.035.

73. Y. Shimizu, T. Kurokawa, H. Arai, and H. Washizu, "Higher-Order Structure of Polymer Melt Described by Persistent Homology," *Scientific Reports* 11, no. 1 (January 2021): 2274, https://doi.org/10.1038/s41598-021-80975-5.

74. F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, "Crystal Structure Representations for Machine Learning Models of Formation Energies," *International Journal of Quantum Chemistry* 115, no. 16 (2015): 1094–1101, https://doi.org/10.1002/qua.24917.

75. L. Ward, R. Liu, A. Krishna, et al., "Including Crystal Structure Attributes in Machine Learning Models of Formation Energies via Voronoi Tessellations," *Physical Review B* 96, no. 2 (July 2017): 024104, https://doi.org/10.1103/PhysRevB.96.024104.

76. Y. Lee, S. D. Barthel, P. Dłotko, S. M. Moosavi, K. Hess, and B. Smit, "Quantifying Similarity of Pore Geometry in Nanoporous Materials," *Nature Communications* 8, no. 1 (May 2017): 15396, https://doi.org/10.1038/ncomms15396.

77. S. M. Moosavi, A. Nandy, K. M. Jablonka, et al., "Understanding the Diversity of the Metal-Organic Framework Ecosystem," *Nature Communications* 11, no. 1 (August 2020): 4068, https://doi.org/10.1038/s41467-020-17755-8.

78. X. Qu, A. Jain, N. N. Rajput, et al., "The Electrolyte Genome Project: A Big Data Approach in Battery Materials Discovery," *Computational Materials Science* 103 (June 2015): 56–67, https://doi.org/10.1016/j.commatsci.2015.02.050.

79. E. W. C. Spotte-Smith, O. A. Cohen, S. M. Blau, et al., "A Database of Molecular Properties Integrated in the Materials Project," *Digital Discovery* 2, no. 6 (2023): 1862–1882, https://doi.org/10.1039/D3DD00153A.

80. R. P. Carvalho, C. F. Marchiori, D. Brandell, and C. M. Araujo, "Artificial Intelligence-Driven In-Silico Discovery of Novel Organic Lithium-Ion Battery Cathodes," *Energy Storage Materials* 44 (2022): 313–325, https://doi.org/10.1016/j.ensm.2021.10.029.

81. J. Wang, T. Hou, and X. Xu, "Aqueous Solubility Prediction Based on Weighted Atom Type Counts and Solvent Accessible Surface Areas," *Journal of Chemical Information and Modeling* 49, no. 3 (March 2009): 571–581, https://doi.org/10.1021/ci800406y.

82. V. Gutmann, *The Donor-Acceptor Approach to Molecular Interactions* (Plenum Press, 1978).

83. A. K. Chew, M. Sender, Z. Kaplan, et al., "Advancing Material Property Prediction: Using Physics-Informed Machine Learning Models for Viscosity," *Journal of Cheminformatics* 16, no. 1 (March 2024): 31, https://doi.org/10.1186/s13321-024-00820-5.

84. J. Du, J. Guo, Q. Sun, et al., "Data-Diven Discovery of Carbonyl Organic Electrode Molecules: Machine Learning and Experiment," *Journal of Materials Chemistry A* 12, no. 20 (2024): 12034–12042, https://doi.org/10.1039/D4TA00136B.

85. D. S. Levine, M. Shuaibi, E. W. C. Spotte-Smith, et al., "The Open Molecules 2025 (OMol25) Dataset, Evaluations, and Models," *arXiv* (2025): arXiv:2505.08762.

86. B. Olsthoorn, R. M. Geilhufe, S. S. Borysov, and A. V. Balatsky, "Band Gap Prediction for Large Organic Crystal Structures with Machine Learning," *Advanced Quantum Technologies* 2, no. 7-8 (August 2019): 1900023, https://doi.org/10.1002/qute.201900023.

87. K. Choudhary, K. F. Garrity, A. C. E. Reid, et al., "The Joint Automated Repository for Various Integrated Simulations (JARVIS) for Data-Driven Materials Design," *npj Computational Materials* 6, no. 1 (November 2020): 173, https://doi.org/10.1038/s41524-020-00440-1.

88. J. Shi, X. Feng, Z. Xu, and X. Zhang, "Molecular Dynamics Simulation and Artificial Intelligence-Driven Development of New Lithium Electrolytes with High Ionic Conductivity and Understanding Ion Transport," *Journal of Physical Chemistry Letters* 16 (2025): 11777–11786, https://doi.org/10.1021/acs.jpclett.5c02681.

89. P. De Blasio, J. Elsborg, T. Vegge, E. Flores, and A. Bhowmik, "CALiSol-23: Experimental electrolyte conductivity data for various Li-salts and solvent combinations," *Scientific Data* 11, no. 1 (July 2024): 750, https://doi.org/10.1038/s41597-024-03575-8.

90. S. Zhu, B. Ramsundar, E. Annevelink, et al., "Differentiable Modeling and Optimization of Non-Aqueous Li-Based Battery Electrolyte Solutions Using Geometric Deep Learning," *Nature Communications* 15, no. 1 (October 2024): 8649, https://doi.org/10.1038/s41467-024-51653-7.

91. Y. Zheng, P. Thakolkaran, A. K. Biswal, et al., "AI-Guided Inverse Design and Discovery of Recyclable Vitrimeric Polymers," *Advanced Science* 12 (2024): e2411385, https://doi.org/10.1002/advs.202411385.

92. M. Ishii, T. Ito, H. Sado, and I. Kuwajima, "NIMS Polymer Database PoLyInfo (I): An Overarching View of Half a Million Data Points," *Science and Technology of Advanced Materials: Methods* 4, no. 1 (December 2024): 2354649, https://doi.org/10.1080/27660400.2024.2354649.

93. N. S. Schauser, G. A. Kliegle, P. Cooke, R. A. Segalman, and R. Seshadri, "Database Creation, Visualization, and Statistical Learning for Polymer Li+-Electrolyte Design," *Chemistry of Materials* 33, no. 13 (2021): 4863–4876, https://doi.org/10.1021/acs.chemmater.0c04767.

94. G. Bradford, J. Lopez, J. Ruza, et al., "Chemistry-Informed Machine Learning for Polymer Electrolyte Discovery," *ACS Central Science* 9, no. 2 (2023): 206–216, https://doi.org/10.1021/acscentsci.2c01123.

95. T. Xie, A. France-Lanord, Y. Wang, et al., "Accelerating Amorphous Polymer Electrolyte Screening by Learning to Reduce Errors in Molecular Dynamics-Simulated Properties," *Nature Communications* 13, no. 1 (June 2022): 3415, https://doi.org/10.1038/s41467-022-30994-1.

96. T. Xie, H.-K. Kwon, D. Schweigert, et al., "A Cloud Platform for Sharing and Automated Analysis of Raw Data From High-Throughput Polymer MD Simulations," *APL Machine Learning* 1, no. 4 (2023): 046108, https://doi.org/10.1063/5.0160937.

97. J. R. Ash, C. Wognum, R. Rodríguez-Pérez, et al., "Practically Significant Method Comparison Protocols for Machine Learning in Small Molecule Drug Discovery," *Journal of Chemical Information and Modeling* 65, no. 18 (September 2025): 9398–9411, https://doi.org/10.1021/acs.jcim.5c01609.

98. S. Manna, S. S. Manna, and B. Pathak, "Integrated Supervised and Unsupervised Machine Learning Approach to Map the Electrochemical Windows Over 4500 Solvents for Battery Applications," *ACS Applied Materials & Interfaces* 16, no. 32 (August 2024): 42138–42152, https://doi.org/10.1021/acsami.4c06243.

99. G. Xu, Y. Li, J. Li, et al., "Toward Stable Zinc Anode: An AI-Assisted High-Throughput Screening of Electrolyte Additives for Aqueous Zinc-Ion Battery," *Angewandte Chemie International Edition* 64, no. 39 (September 2025): e202511389, https://doi.org/10.1002/anie.202511389.

100. P. Peljo and H. H. Girault, "Electrochemical potential window of battery electrolytes: the HOMO–LUMO misconception," *Energy & Environmental Science* 11, no. 9 (2018): 2306–2309, https://doi.org/10.1039/C8EE01286E.

101. X. Zhou, C. Xu, X. Guo, et al., "Computational and Machine Learning-Assisted Discovery and Experimental Validation of Conjugated Sulfonamide Cathodes for Lithium-Ion Batteries," *Advanced Energy Materials* 15, no. 25 (July 2025): 2401658, https://doi.org/10.1002/aenm.202401658.

102. F. Wang, J. Li, Z. Liu, T. Qiu, J. Wu, and D. Lu, "Computational Design of Quinone Electrolytes for Redox Flow Batteries Using High-Throughput Machine Learning and Theoretical Calculations," *Frontiers in Chemical Engineering* 4 (January 2023): 1086412, https://doi.org/10.3389/fceng.2022.1086412.

103. Y.-C. Gao, N. Yao, X. Chen, L. Yu, R. Zhang, and Q. Zhang, "Data-Driven Insight into the Reductive Stability of Ion–Solvent Complexes in Lithium Battery Electrolytes," *Journal of the American Chemical Society* 145, no. 43 (November 2023): 23764–23770, https://doi.org/10.1021/jacs.3c08346.

104. S. Kim, A. Jinich, and A. Aspuru-Guzik, "MultiDK: A Multiple Descriptor Multiple Kernel Approach for Molecular Discovery and Its Application to Organic Flow Battery Electrolytes," *Journal of Chemical Information and Modeling* 57, no. 4 (April 2017): 657–668, https://doi.org/10.1021/acs.jcim.6b00332.

105. P. Gao, A. Andersen, J. Sepulveda, et al., "SOMAS: A Platform for Data-Driven Material Discovery in Redox Flow Battery Development," *Scientific Data* 9, no. 1 (December 2022): 740, https://doi.org/10.1038/s41597-022-01814-4.

106. P. Gao, X. Yang, Y.-H. Tang, et al., "Graphical Gaussian Process Regression Model for Aqueous Solvation Free Energy Prediction of Organic Molecules in Redox Flow Batteries," *Physical Chemistry Chemical Physics* 23, no. 43 (2021): 24892–24904, https://doi.org/10.1039/D1CP04475C.

107. J. Park, F. Sorourifar, M. R. Muthyala, et al., "Zero-Shot Discovery of High-Performance, Low-Cost Organic Battery Materials Using Machine Learning," *Journal of the American Chemical Society* 146, no. 45 (November 2024): 31230–31239, https://doi.org/10.1021/jacs.4c11663.

108. A. Ishikawa, K. Sodeyama, Y. Igarashi, T. Nakayama, Y. Tateyama, and M. Okada, "Machine Learning Prediction of Coordination Energies for Alkali Group Elements in Battery Electrolyte Solvents," *Physical Chemistry Chemical Physics* 21, no. 48 (2019): 26399–26405, https://doi.org/10.1039/C9CP03679B.

109. Y. Chen, Z. Yu, P. Rudnicki, et al., "Steric Effect Tuned Ion Solvation Enabling Stable Cycling of High-Voltage Lithium Metal Battery," *Journal of the American Chemical Society* 143, no. 44 (November 2021): 18703–18713, https://doi.org/10.1021/jacs.1c09006.

110. V. Gutmann, "Empirical Parameters for Donor and Acceptor Properties of Solvents," *Electrochimica Acta* 21, no. 9 (September 1976): 661–670, https://doi.org/10.1016/0013-4686(76)85034-7.

111. J. Xu, J. Zhang, T. P. Pollard, et al., "Electrolyte Design for Li-Ion Batteries Under Extreme Operating Conditions," *Nature* 614, no. 7949 (February 2023): 694–700, https://doi.org/10.1038/s41586-022-05627-8.

112. H. Hu, Y. Shan, Q. Zhao, J. Wang, L. Wu, and W. Liu, "The Prediction of Donor Number and Acceptor Number of Electrolyte Solvent Molecules Based on Machine Learning," *Journal of Energy Chemistry* 98 (November 2024): 374–382, https://doi.org/10.1016/j.jechem.2024.06.050.

113. Y. Gao, S. Qi, M. Li, et al., "Machine Learning Assisted Prediction of Donor Numbers: Guiding Rational Fabrication of Polymer Electrolytes for Lithium-Ion Batteries," *Angewandte Chemie International Edition* 64, no. 2 (January 2025): e202411437, https://doi.org/10.1002/anie.202411437.

114. X. Yuan, X. Chen, Y. Zhou, Z. Yu, and X. Kong, "Design Principles of Fluoroether Solvents for Lithium Metal Battery Electrolytes Unveiled by Extensive Molecular Simulation and Machine Learning," *Journal of Energy Chemistry* 102 (March 2025): 52–62, https://doi.org/10.1016/j.jechem.2024.10.021.

115. C. Bilodeau, A. Kazakov, S. Mukhopadhyay, et al., "Machine Learning for Predicting the Viscosity of Binary Liquid Mixtures," *Chemical Engineering Journal* 464 (May 2023): 142454, https://doi.org/10.1016/j.cej.2023.142454.

116. Z. Chen, J. Chen, Y. Qiu, et al., "Prediction of Electrical Conductivity of Ionic Liquids: From COSMO-RS Derived QSPR Evaluation to Boosting Machine Learning," *ACS Sustainable Chemistry & Engineering* 12, no. 17 (April 2024): 6648–6658, https://doi.org/10.1021/acssuschemeng.4c00307.

117. Z. Yang, Y. Wu, X. Han, et al., "A Unified Predictive and Generative Solution for Liquid Electrolyte Formulation," *Nature Machine Intelligence* 8, no. 2 (January 2026): 186–196, https://doi.org/10.1038/s42256-025-01173-w.

118. Q. Huang, Z. Chen, Z. Lin, W. Li, W. Yu, and L. Zhu, "Enhancing Copolymer Property Prediction through the Weighted-Chained-SMILES Machine Learning Framework," *ACS Applied Polymer Materials* 6, no. 7 (2024): 3666–3675, https://doi.org/10.1021/acsapm.3c02715.

119. H. D. Tran, C. Kim, L. Chen, et al., "Machine-Learning Predictions of Polymer Properties with Polymer Genome," *Journal of Applied Physics* 128, no. 17 (2020): 171104, https://doi.org/10.1063/5.0023759.

120. E. R. Antoniuk, P. Li, B. Kailkhura, and A. M. Hiszpanski, "Representing Polymers as Periodic Graphs with Learned Descriptors for Accurate Polymer Property Predictions," *Journal of Chemical Information and Modeling* 62, no. 22 (2022): 5435–5445, https://doi.org/10.1021/acs.jcim.2c00875.

121. C. Dong, D. Li, and J. Liu, "Glass Transition Temperature Prediction of Polymers via Graph Reinforcement Learning," *Langmuir* 40, no. 35 (September 2024): 18568–18580, https://doi.org/10.1021/acs.langmuir.4c01906.

122. S. Kunchapu and K. M. Jablonka, "PolyMetriX: An Ecosystem for Digital Polymer Chemistry," *npj Computational Materials* 11, no. 1 (2025): 312, https://doi.org/10.1038/s41524-025-01823-y.

123. L. Tao, V. Varshney, and Y. Li, "Benchmarking Machine Learning Models for Polymer Informatics: An Example of Glass Transition Temperature," *Journal of Chemical Information and Modeling* 61, no. 11 (2021): 5395–5413, https://doi.org/10.1021/acs.jcim.1c01031.

124. H. Tang, T. Yue, and Y. Li, "Assessing Uncertainty in Machine Learning for Polymer Property Prediction: A Benchmark Study," *Journal of Chemical Information and Modeling* (2025), https://doi.org/10.1021/acs.jcim.5c00550.

125. Y. Zheng, A. K. Biswal, Y. Guo, et al., "Toward Sustainable Polymer Design: A Molecular Dynamics-Informed Machine Learning Approach for Vitrimers," *Digital Discovery* 4, no. 9 (2025): 2559–2569, https://doi.org/10.1039/d5dd00239g.

126. Z. Long, H. Lu, and Z. Zhang, "Large-Scale Glass-Transition Temperature Prediction with an Equivariant Neural Network for Screening Polymers," *ACS Omega* 9, no. 5 (February 2024): 5452–5462, https://doi.org/10.1021/acsomega.3c06843.

127. A. Khajeh, D. Schweigert, S. B. Torrisi, L. Hung, B. D. Storey, and H.-K. Kwon, "Early Prediction of Ion Transport Properties in Solid Polymer Electrolytes Using Machine Learning and System Behavior-Based Descriptors of Molecular Dynamics Simulations," *Macromolecules* 56, no. 13 (2023): 4787–4799, https://doi.org/10.1021/acs.macromol.3c00416.

128. R. Gurnani, D. Kamal, H. Tran, et al., "polyG2G: A Novel Machine Learning Algorithm Applied to the Generative Design of Polymer Dielectrics," *Chemistry of Materials* 33, no. 17 (2021): 7008–7016, https://doi.org/10.1021/acs.chemmater.1c02061.

129. H. Li, H. Zheng, T. Yue, et al., "Machine Learning-Accelerated Discovery of Heat-Resistant Polysulfates for Electrostatic Energy Storage," *Nature Energy* 10 (2024): 90–100, https://doi.org/10.1038/s41560-024-01670-z.

130. K. Hatakeyama-Sato, T. Tezuka, M. Umeki, and K. Oyaizu, "AI-Assisted Exploration of Superionic Glass-Type Li+ Conductors with Aromatic Structures," *Journal of the American Chemical Society* 142, no. 7 (2020): 3301–3305, https://doi.org/10.1021/jacs.9b11442.

131. T. Yue, J. He, L. Tao, and Y. Li, "High-Throughput Screening and Prediction of High Modulus of Resilience Polymers Using Explainable Machine Learning," *Journal of Chemical Theory and Computation* 19, no. 14 (July 2023): 4641–4653, https://doi.org/10.1021/acs.jctc.3c00131.

132. E. Prodan and W. Kohn, "Nearsightedness of Electronic Matter," *Proceedings of the National Academy of Sciences* 102, no. 33 (August 2005): 11635–11638, https://doi.org/10.1073/pnas.0505436102.

133. J. Behler and M. Parrinello, "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces," *Physical Review Letters* 98, no. 14 (April 2007): 146401, https://doi.org/10.1103/PhysRevLett.98.146401.

134. T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, "Neural Network Models of Potential Energy Surfaces," *Journal of Chemical Physics* 103, no. 10 (September 1995): 4129–4137, https://doi.org/10.1063/1.469597.

135. T. Morawietz, A. Singraber, C. Dellago, and J. Behler, "How van der Waals Interactions Determine the Unique Properties of Water," *Proceedings of the National Academy of Sciences* 113, no. 30 (July 2016): 8368–8373, https://doi.org/10.1073/pnas.1602375113.

136. V. Quaranta, M. Hellström, and J. Behler, "Proton-Transfer Mechanisms at the Water–ZnO Interface: The Role of Presolvation," *Journal of Physical Chemistry Letters* 8, no. 7 (April 2017): 1476–1483, https://doi.org/10.1021/acs.jpclett.7b00358.

137. C. Schran, F. L. Thiemann, P. Rowe, E. A. Müller, O. Marsalek, and A. Michaelides, "Machine Learning Potentials for Complex Aqueous Systems Made Simple," *Proceedings of the National Academy of Sciences* 118, no. 38 (September 2021): e2110077118, https://doi.org/10.1073/pnas.2110077118.

138. B. Cheng, G. Mazzola, C. J. Pickard, and M. Ceriotti, "Evidence for Supercritical Behaviour of High-Pressure Liquid Hydrogen," *Nature* 585, no. 7824 (September 2020): 217–220, https://doi.org/10.1038/s41586-020-2677-y.

139. O. T. Unke and M. Meuwly, "PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges," *Journal of Chemical Theory and Computation* 15, no. 6 (June 2019): 3678–3693, https://doi.org/10.1021/acs.jctc.9b00181.

140. L. Zhang, J. Han, H. Wang, R. Car, and W. E, "Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics," *Physical Review Letters* 120, no. 14 (April 2018): 143001, https://doi.org/10.1103/PhysRevLett.120.143001.

141. H. Wang, L. Zhang, J. Han, and W. E, "DeePMD-kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics," *Computer Physics Communications* 228 (July 2018): 178–184, https://doi.org/10.1016/j.cpc.2018.03.016.

142. M. Haghighatlari, J. Li, X. Guan, et al., "NewtonNet: A Newtonian Message Passing Network for Deep Learning of Interatomic Potentials and Forces," *Digital Discovery* 1, no. 3 (June 2022): 333–343, https://doi.org/10.1039/D2DD00008C.

143. G. Simeon and G. De Fabritiis, "TensorNet: Cartesian Tensor Representations for Efficient Learning of Molecular Potentials," in *Advances in Neural Information Processing Systems*, eds. A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Vol 36 (Curran Associates, Inc., 2023), 37334–37353.

144. J. Li, L. Knijff, Z.-Y. Zhang, L. Andersson, and C. Zhang, "PiNN: Equivariant Neural Network Suite for Modeling Electrochemical Systems," *Journal of Chemical Theory and Computation* 21, no. 3 (February 2025): 1382–1395, https://doi.org/10.1021/acs.jctc.4c01570.

145. A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, "Scaling Deep Learning for Materials Discovery," *Nature* 624, no. 7990 (December 2023): 80–85, https://doi.org/10.1038/s41586-023-06735-9.

146. E. C.-Y. Yuan, Y. Liu, J. Chen, et al., "Foundation Models for Atomistic Simulation of Chemistry and Materials," *Nature Reviews Chemistry* 10, no. 3 (March 2026): 212–230, https://doi.org/10.1038/s41570-025-00793-5.

147. I. Batatia, P. Benner, Y. Chiang, et al., "A Foundation Model for Atomistic Materials Chemistry," *The Journal of chemical physics* 163 (2025): 3372267, https://doi.org/10.1063/5.0297006.

148. D. Zhang, X. Liu, X. Zhang, et al., "DPA-2: A Large Atomic Model as a Multi-Task Learner," *npj Computational Materials* 10, no. 1 (December 2024): 293, https://doi.org/10.1038/s41524-024-01493-2.

149. B. M. Wood, M. Dzamba, X. Fu, et al., "UMA: A Family of Universal Models for Atoms," *Advances in Neural Information Processing Systems* 38 (2026): 129391–129427.

150. J. Kim, J. You, Y. Park, et al., "Optimizing Cross-Domain Transfer for Universal Machine Learning Interatomic Potentials," *Nature Communications* 17 (2025): 3432.

151. A. Dunn, Q. Wang, A. Ganose, D. Dopp, and A. Jain, "Benchmarking Materials Property Prediction Methods: The Matbench Test set and Automatminer Reference Algorithm," *npj Computational Materials* 6, no. 1 (September 2020): 138, https://doi.org/10.1038/s41524-020-00406-3.

152. A. Sudhama, Z.-Y. Zhang, and C. Zhang, "Foundation Model Distillation with PiNNAcLe: Application to the ZnCl2 Aqueous Electrolyte Solution," *Electrochimica Acta* 572 (October 2026): 149136, https://doi.org/10.1016/j.electacta.2026.149136.

153. J. L. A. Gardner, D. F. T. D. Toit, C. B. Mahmoud, et al., "Distillation of Atomistic Foundation Models Across Architectures and Chemical Domains," *arXiv* (2025): arXiv:2506.10956.

154. Y. Shao, L. Knijff, F. M. Dietrich, K. Hermansson, and C. Zhang, "Modelling Bulk Electrolytes and Electrolyte Interfaces with Atomistic Machine Learning," *Batteries & Supercaps* 4, no. 4 (April 2021): 585–595, https://doi.org/10.1002/batt.202000262.

155. T. P. Senftle, S. Hong, M. M. Islam, et al., "The ReaxFF Reactive Force-Field: Development, Applications and Future Directions," *npj Computational Materials* 2, no. 1 (March 2016): 15011, https://doi.org/10.1038/npjcompumats.2015.11.

156. I.-B. Magdău, D. J. Arismendi-Arrieta, H. E. Smith, C. P. Grey, K. Hermansson, and G. Csányi, "Machine Learning Force Fields for Molecular Liquids: Ethylene Carbonate/Ethyl Methyl Carbonate Binary Solvent," *npj Computational Materials* 9, no. 1 (August 2023): 146, https://doi.org/10.1038/s41524-023-01100-w.

157. S. Dajnowicz, G. Agarwal, J. M. Stevenson, et al., "High-Dimensional Neural Network Potential for Liquid Electrolyte Simulations," *Journal of Physical Chemistry B* 126, no. 33 (August 2022): 6271–6280, https://doi.org/10.1021/acs.jpcb.2c03746.

158. V. Svahn, I.-B. Magdău, S. P. Niblett, G. Csányi, K. Hermansson, and J. Kullgren, "Limitations of Cluster-Trained MLIPs for Liquid Density and Diffusivity," *Journal of Chemical Theory and Computation* 22, no. 7 (April 2026): 3660–3671, https://doi.org/10.1021/acs.jctc.5c02043.

159. S. Gong, Y. Zhang, Z. Mu, et al., "A Predictive Machine Learning Force-Field Framework for Liquid Electrolyte Development," *Nature Machine Intelligence* 7, no. 4 (April 2025): 543–552, https://doi.org/10.1038/s42256-025-01009-7.

160. F. Wang, Y.-H. Tang, Z.-B. Ma, Y.-C. Jin, and J. Cheng, "Domain-Oriented Universal Machine Learning Potential Enables Fast Exploration of Chemical Space of Battery Electrolytes," *Nature Communications* 17, no. 1 (December 2025): 1226, https://doi.org/10.1038/s41467-025-67982-0.

161. Y. Shao, M. Hellström, A. Yllö, et al., "Temperature Effects on the Ionic Conductivity in Concentrated Alkaline Electrolyte Solutions," *Physical Chemistry Chemical Physics* 22, no. 19 (May 2020): 10426–10430, https://doi.org/10.1039/C9CP06479F.

162. Y. Shao, K. Shigenobu, M. Watanabe, and C. Zhang, "Role of Viscosity in Deviations From the Nernst–Einstein Relation," *The Journal of Physical Chemistry B* 124, no. 23 (June 2020): 4774–4780, https://doi.org/10.1021/acs.jpcb.0c02544.

163. J. Lan, V. Kapil, P. Gasparotto, M. Ceriotti, M. Iannuzzi, and V. V. Rybkin, "Simulating the Ghost: Quantum Dynamics of the Solvated Electron," *Nature Communications* 12, no. 1 (February 2021): 766, https://doi.org/10.1038/s41467-021-20914-0.

164. R. Jinnouchi, F. Karsai, and G. Kresse, "Machine Learning-Aided First-Principles Calculations of Redox Potentials," *npj Computational Materials* 10, no. 1 (May 2024): 107, https://doi.org/10.1038/s41524-024-01295-6.

165. F. Wang and J. Cheng, "Automated Workflow for Computation of Redox Potentials, Acidity Constants, and Solvation Free Energies Accelerated by Machine Learning," *The Journal of Chemical Physics* 157, no. 2 (July 2022): 024103, https://doi.org/10.1063/5.0098330.

166. F. Wang, Z. Ma, and J. Cheng, "Accelerating Computation of Acidity Constants and Redox Potentials for Aqueous Organic Redox Flow Batteries by Machine Learning Potential-Based Molecular Dynamics," *Journal of the American Chemical Society* 146, no. 21 (May 2024): 14566–14575, https://doi.org/10.1021/jacs.4c01221.

167. E. Kocer, R. E. Haouari, C. Dellago, and J. Behler, "Machine Learning Potentials for Redox Chemistry in Solution," *arXiv* (2024): arXiv:2410.03299.

168. D. Firaha, Y. M. Liu, J. van de Streek, et al., "Predicting crystal form stability under real-world conditions," *Nature* 623, no. 7986 (November 2023): 324–328, https://doi.org/10.1038/s41586-023-06587-3.

169. S. Hattori and Q. Zhu, "Revisiting Aspirin Polymorphic Stability Using a Machine Learning Potential," *ACS Omega* 9, no. 34 (August 2024): 36589–36599, https://doi.org/10.1021/acsomega.4c04782.

170. P. W. V. Butler, R. Hafizi, and G. M. Day, "Machine-Learned Potentials by Active Learning From Organic Crystal Structure Prediction Landscapes," *Journal of Physical Chemistry A* 128, no. 5 (February 2024): 945–957, https://doi.org/10.1021/acs.jpca.3c07129.

171. F. D. Pia, B. X. Shi, V. Kapil, A. Zen, D. Alfè, and A. Michaelides, "Accurate and Efficient Machine Learning Interatomic Potentials for Finite Temperature Modelling of Molecular Crystals," *Chemical Science* 16, no. 25 (2025): 11419–11433, https://doi.org/10.1039/D5SC01325A.

172. R. Chahal, M. D. Toomey, L. T. Kearney, et al., "Deep-Learning Interatomic Potential Connects Molecular Structural Ordering to the Macroscale Properties of Polyacrylonitrile," *ACS Applied Materials and Interfaces* 16, no. 28 (July 2024): 36878–36891, https://doi.org/10.1021/acsami.4c04491.

173. H. Cui, W. Hua, L. Cao, Y. Jin, and Y. Wang, "Deep-Neural-Network Molecular Dynamics Investigation of Phonon Thermal Transport in Polyether Ether Ketone," *Computational Materials Science* 249 (February 2025): 113641, https://doi.org/10.1016/j.commatsci.2024.113641.

174. L. Reicht, L. Legenstein, S. Wieser, and E. Zojer, "Analysing Heat Transport in Crystalline Polymers in Real and Reciprocal Space," *npj Computational Materials* 12, no. 1 (February 2026): 129, https://doi.org/10.1038/s41524-026-01988-0.

175. G. N. C. Simm, J. Hélie, H. Schulz, et al., "SimPoly: Simulation of Polymers with Machine Learning Force Fields Derived From First Principles," *arXiv* (2025): _arXiv: 2510.13696.

176. R. Jinnouchi, S. Minami, F. Karsai, C. Verdi, and G. Kresse, "Proton Transport in Perfluorinated Ionomer Simulated by Machine-Learned Interatomic Potential," *Journal of Physical Chemistry Letters* 14, no. 14 (April 2023): 3581–3588, https://doi.org/10.1021/acs.jpclett.3c00293.

177. J. Hänseroth, M. N. Qaisrani, M. Moradi, K. Skadell, and C. Dreßler, "Revealing Hydroxide Ion Transport Mechanisms in Commercial Anion-Exchange Membranes at Nano-Scale From Machine-learned Interatomic Potential Simulations," *arXiv* (2026): arXiv:2603.13705.

178. S. T. Keene, V. Gueskine, M. Berggren, G. G. Malliaras, K. Tybrandt, and I. Zozoulenko, "Exploiting Mixed Conducting Polymers in Organic and Bioelectronic Devices," *Physical Chemistry Chemical Physics* 24, no. 32 (August 2022): 19144–19163, https://doi.org/10.1039/D2CP02595G.

179. R. Chahal-Crockett, M. D. Toomey, L. T. Kearney, et al., "Concerted Electron-Ion Transport by Polyacrylonitrile Elucidated with Reactive Deep Learning Potentials," *Journal of the American Chemical Society* 148 (June 2026): 23457, https://doi.org/10.1021/jacs.6c05078.

180. H. Mori, S. Tonogai, Y. Miyazaki, A. Hayashi, and M. Takayanagi, "Ready-to-Use Polymerization Simulations Combining Universal Machine Learning Interatomic Potential with Time-Dependent Bond Boosting for Polymer and Interface Design," *Journal of Physical Chemistry B* 130, no. 9 (March 2026): 2710–2720, https://doi.org/10.1021/acs.jpcb.5c08192.

181. C. Tu, X. Li, J. Chen, B. Sun, and K. Yu, "Enhancing Thermal Conductivity Computation of Polymers via Machine Learning Techniques," *Journal of Physical Chemistry B* 129, no. 33 (August 2025): 8593–8602, https://doi.org/10.1021/acs.jpcb.5c03656.

182. A. Singh, A. Fry, A. Perelman, et al., "OpenAI GPT-5 System Card," *arXiv* (2026): arXiv:2601.03267.

183. P. Esser, S. Kulal, A. Blattmann, et al., "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis," *arXiv* (2024): arXiv:2403.03206.

184. Y. Sun, L. Chen, Z. Jing, et al., "Generative AI for the Design of Molecules: Advances and Challenges," *Journal of Chemical Information and Modeling* 65, no. 23 (December 2025): 12668–12690, https://doi.org/10.1021/acs.jcim.5c02234.

185. T. Yue, L. Tao, V. Varshney, and Y. Li, "Benchmarking Study of Deep Generative Models for Inverse Polymer Design," *Digital Discovery* 4 (2025): 910–926, https://doi.org/10.1039/d4dd00395k.

186. D. Rajagopal, A. Cierpka, B. Nestler, and A. Koeppe, "Generative Deep Learning for Advanced Battery Materials," *Batteries & Supercaps* 9, no. 2 (February 2026): e202500494, https://doi.org/10.1002/batt.202500494.

187. M. J. Kusner, B. Paige, and J. M. Hernández-Lobato, "Grammar Variational Autoencoder," In *International Conference on Machine Learning*, (PMLR, 2017).

188. S. Kang and K. Cho, "Conditional Molecular Design with Deep Generative Models," *Journal of Chemical Information and Modeling* 59, no. 1 (January 2019): 43–52, https://doi.org/10.1021/acs.jcim.8b00263.

189. M. Olivecrona, T. Blaschke, O. Engkvist, and H. Chen, "Molecular De-novo Design Through Deep Reinforcement Learning," *Journal of Cheminformatics* 9, no. 1 (December 2017): 48, https://doi.org/10.1186/s13321-017-0235-x.

190. L. Gao, S. Song, J. Lin, Y. Xu, L. Wang, and L. Du, "AI-Assisted Design of Advanced Polymeric Materials: Challenges and Solutions," *Advanced Materials* 38, no. 7 (February 2026): e16857, https://doi.org/10.1002/adma.202516857.

191. D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," *arXiv* (2013): arXiv:1312.6114.

192. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative Adversarial Networks," *Communications of the ACM* 63, no. 11 (2020), 139–144, https://doi.org/10.1145/3422622.

193. M. Xu, L. Yu, Y. Song, C. Shi, S. Ermon, and J. Tang, "GeoDiff: A Geometric Diffusion Model for Molecular Conformation Generation," arXiv (2022): arXiv:2203.02923.

194. J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," *Advances in Neural Information Processing Systems* 33 (2020): 6840–6851.

195. Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le, "Flow Matching for Generative Modeling," *arXiv* (2023): arXiv:2210.02747.

196. V. Bagal, R. Aggarwal, P. K. Vinod, and U. D. Priyakumar, "MolGPT: Molecular Generation Using a Transformer-Decoder Model," *Journal of Chemical Information and Modeling* 62, no. 9 (May 2022): 2064–2076, https://doi.org/10.1021/acs.jcim.1c00600.

197. R. Mercado, T. Rastemo, E. Lindelöf, et al., "Graph networks for molecular design," *Machine Learning: Science and Technology* 2, no. 2 (June 2021): 025023, https://doi.org/10.1088/2632-2153/abcf91.

198. H. H. Loeffler, J. He, A. Tibo, et al., "Reinvent 4: Modern AI-Driven Generative Molecule Design," *Journal of Cheminformatics* 16, no. 1 (February 2024): 20, https://doi.org/10.1186/s13321-024-00812-5.

199. R. S. Sutton and A. Barto, *Reinforcement learning: an introduction* (The MIT Press, 2020).

200. A. Agarwal, S. M. Kakade, J. D. Lee, and G. Mahajan, "On the Theory of Policy Gradient Methods: Optimality, Approximation, and Distribution Shift," *arXiv* (2020, October): arXiv:1908.00261, http://arxiv.org/abs/1908.00261.

201. W. Gao, T. Fu, J. Sun, and C. W. Coley, "Sample Efficiency Matters: A Benchmark for Practical Molecular Optimization," *arXiv* (2022, October): arXiv:2206.12411.

202. J. Kim, K.-H. Wang, R. Kumar, P. Ma, and C. V. Amanchukwu, "Generative Electrolyte Solvent and Formulation Discovery," *JACS Au* 6, no. 4 (April 2026): 2288–2302, https://doi.org/10.1021/jacsau.5c01628.

203. S. S. S. V., J. N. Law, C. E. Tripp, et al., "Multi-Objective Goal-Directed Optimization of de Novo Stable Organic Radicals for Aqueous Redox Flow Batteries," *Nature Machine Intelligence* 4, no. 8 (August 2022): 720–730, https://doi.org/10.1038/s42256-022-00506-3.

204. A. Khajeh, X. Lei, W. Ye, et al., "A Materials Discovery Framework Based on Conditional Generative Models Applied to the Design of Polymer Electrolytes," *Digital Discovery* 4, no. 1 (2025): 11–20, https://doi.org/10.1039/D4DD00293H.

205. D. Yoon, J. Lee, and S. Lee, "Deep Learning-Driven Molecular Generation and Electrochemical Property Prediction for Optimal Electrolyte Additive Design," *Applied Sciences* 15, no. 7 (March 2025): 3640, https://doi.org/10.3390/app15073640.

206. R. Pritom and M. M. Islam, "Generative AI-Driven Discovery of Next-Generation Electrolytes for Alkali Metal Batteries," *Journal of Chemical Information and Modeling* 66, no. 7 (April 2026): 3747–3758, https://doi.org/10.1021/acs.jcim.6c00135.

207. D. Polykovskiy, A. Zhebrak, B. Sanchez-Lengeling, et al., "Molecular Sets (MOSES): A Benchmarking Platform for Molecular Generation Models," *Frontiers in Pharmacology* 11 (December 2020): 565644, https://doi.org/10.3389/fphar.2020.565644.

208. K. Preuer, P. Renz, T. Unterthiner, S. Hochreiter, and G. Klambauer, "Fréchet ChemNet Distance: A Metric for Generative Models for Molecules in Drug Discovery," *Journal of Chemical Information and Modeling* 58, no. 9 (September 2018): 1736–1741, https://doi.org/10.1021/acs.jcim.8b00234.

209. P. Ertl and A. Schuffenhauer, "Estimation of Synthetic Accessibility Score of Drug-Like Molecules Based on Molecular Complexity and Fragment Contributions," *Journal of Cheminformatics* 1 (June 2009): 8, https://doi.org/10.1186/1758-2946-1-8.

210. D. A. Boiko, R. MacKnight, B. Kline, and G. Gomes, "Autonomous Chemical Research with Large Language Models," *Nature* 624, no. 7992 (December 2023): 570–578, https://doi.org/10.1038/s41586-023-06792-0.

211. J. He, H. Lai, L. Saigiridharan, et al., "Democratising Real-World Drug Discovery Through Agentic AI," *Drug Discovery Today* 31, no. 2 (March 2026): 104605, https://doi.org/10.1016/j.drudis.2026.104605.

212. K. M. Jablonka, P. Schwaller, A. Ortega-Guerrero, and B. Smit, "Leveraging Large Language Models for Predictive Chemistry," *Nature Machine Intelligence* 6, no. 2 (February 2024): 161–169, https://doi.org/10.1038/s42256-023-00788-1.

213. H. Qiu, L. Liu, X. Qiu, X. Dai, X. Ji, and Z.-Y. Sun, "PolyNC: a Natural and Chemical Language Model for the Prediction of Unified Polymer Properties," *Chemical Science* 15, no. 2 (2023): 534–544, https://doi.org/10.1039/d3sc05079c.

214. A. M. Bran, S. Cox, O. Schilter, C. Baldassari, A. D. White, and P. Schwaller, "Augmenting Large Language Models with Chemistry Tools," *Nature Machine Intelligence* 6, no. 5 (May 2024): 525–535, https://doi.org/10.1038/s42256-024-00832-8.

215. M. Lederbauer, S. Betala, X. Li, et al., "LeMat-Synth: A Multi-Modal Toolbox to Curate Broad Synthesis Procedure Databases From Scientific Literature," *arXiv* (2025): arXiv:2510.26824 [cs].

216. N. Alampara, M. Schilling-Wilhelmi, M. Ríos-García, et al., "Probing the Limitations of Multimodal Language Models for Chemistry and Materials Research," *Nature Computational Science* 5 (2025): 952–961, https://doi.org/10.1038/s43588-025-00836-3.

217. H.-T. Wang, X. Bai, Z. Zheng, et al., "Chat-RFB: A Flow Battery Chat System Leveraging Knowledge Graphs and Large language Models," *Digital Discovery* 5, no. 3 (March 2026): 1401–1410, https://doi.org/10.1039/D5DD00494B.

218. Z. Xie, X. Evangelopoulos, M. H. Omar, A. Troisi, A. I. Cooper, and L. Chen, "Fine-Tuning GPT-3 for Machine Learning Electronic and Functional Properties of Organic Molecules," *Chemical Science* 15, no. 2 (January 2024): 500–510, https://doi.org/10.1039/D3SC04610A.

219. M. Zohair, V. Sharma, E. A. Soares, et al., "Chemical Foundation Model-Guided Design of High Ionic Conductivity Electrolyte Formulations," *npj Computational Materials* 11, no. 1 (August 2025): 283, https://doi.org/10.1038/s41524-025-01774-4.

220. H. Qiu, J. Zhao, E. Jing, et al., *Introducing PolySea: An LLM-Based Polymer Smart Evolution Agent*, (Cabridge University Press, 2025), no. 0417, https://doi.org/10.26434/chemrxiv-2025-zw65g.

221. X. Chen, M. Liu, S. Yin, Y.-C. Gao, N. Yao, and Q. Zhang, "Uni-Electrolyte: An Artificial Intelligence Platform for Designing Electrolyte Molecules for Rechargeable Batteries," *Angewandte Chemie International Edition* 64, no. 30 (2025): e202503105, https://doi.org/10.1002/anie.202503105.

222. A. Mirza, N. Alampara, S. Kunchapu, et al., "A Framework for Evaluating the Chemical Knowledge and Reasoning Abilities of Large Language Models Against the Expertise of Chemists," *Nature Chemistry* 17, no. 7 (July 2025): 1027–1034, https://doi.org/10.1038/s41557-025-01815-x.

223. S. Betala, S. P. Gleason, A. Ramlaoui, et al., "LeMat-GenBench: A Unified Evaluation Framework for Crystal Generative Models," *arXiv* (2026): arXiv:2512.04562 [cs].

224. N. Artrith, K. T. Butler, F.-X. Coudert, et al., "Best Practices in Machine Learning for Chemistry," *Nature Chemistry* 13, no. 6 (June 2021): 505–508, https://doi.org/10.1038/s41557-021-00716-z.

225. R. Mercado, S. M. Kearnes, and C. W. Coley, "Data Sharing in Chemistry: Lessons Learned and a Case for Mandating Structured Reaction Data," *Journal of Chemical Information and Modeling* 63, no. 14 (July 2023): 4253–4265, https://doi.org/10.1021/acs.jcim.3c00607.

226. G. Hernández, I. L. Johansson, A. Mathew, C. Sångeland, D. Brandell, and J. Mindemark, "Going Beyond Sweep Voltammetry: Alternative Approaches in Search of the Elusive Electrochemical Stability of Polymer Electrolytes," *Journal of The Electrochemical Society* 168, no. 10 (October 2021): 100523, https://doi.org/10.1149/1945-7111/ac2d8b.

227. Y. Shao, H. Gudla, J. Mindemark, D. Brandell, and C. Zhang, "Ion Transport in Polymer Electrolytes: Building New Bridges between Experiment and Molecular Simulation," *Accounts of Chemical Research* 57, no. 8 (April 2024): 1123–1134, https://doi.org/10.1021/acs.accounts.3c00791.

228. D. T. Hickson and N. Balsara, "Toward Establishing Uniqueness of Experimentally Determined Transference Numbers," *Journal of The Electrochemical Society* 172, no. 4 (April 2025): 040528, https://doi.org/10.1149/1945-7111/adca04.

229. K. Elbouazzaoui, A. Mahun, V. Shabikova, et al., "Enabling High-Voltage Polymer-Based Solid-State Batteries Through Reinforcements with $LiAlO_2$ Fillers," *Advanced Energy Materials* 15, no. 26 (July 2025): 2405249, https://doi.org/10.1002/aenm.202405249.

230. X. Wu, J. Billaud, I. Jerjen, et al., "Operando Visualization of Morphological Dynamics in All-Solid-State Batteries," *Advanced Energy Materials* 9, no. 34 (September 2019): 1901547, https://doi.org/10.1002/aenm.201901547.