

Supporting Information

(Acknowledgments on Data Formatting)

The initial data formatting for this supporting information was performed using Liner AI. Subsequent formatting and structuring of the data for each force field was carried out using ChatGPT version 4.1.

Question

Can AI conduct research? We are planning to participate in a competition to solve this question. I think AI has both strengths and weaknesses, and there may be specific types of papers where AI can be particularly useful. What kind of paper would be better and for what reasons? Also, for the direction of the topic, what would be a good subject in the field of chemistry, and what topics are currently emerging? Please recommend.

Answer

1. What type of paper is good?

I recommend the review article format.

- **Reason 1: Ability to process large amounts of data**

AI excels at quickly collecting, classifying, and organizing thousands of papers. Since review articles involve comparing various research results and extracting patterns, AI's data processing and analytical abilities are best showcased here.

- **Reason 2: Ability to analyze research trends**

AI is useful not only for simple summarization but also for highlighting future research directions or unresolved questions in a field. This fits perfectly with the purpose of a review article.

- **Reason 3: Can also reveal weaknesses**

Experimental papers require "creative ideas" or "field experience," but review articles focus on systematic organization and meta-analysis, allowing AI to avoid its weaknesses.

2. Direction of the topic (Chemistry field)

- Computational Chemistry Review: AI-based drug candidate prediction, machine learning in catalyst design, acceleration and optimization of molecular dynamics (MD) simulations
- Materials Chemistry: AI-driven battery material discovery, new material design for solar cells/fuel cells
- Environmental Chemistry: CO₂ conversion/reduction catalyst research, development of electrochemical catalysts for hydrogen production (HER, OER)

3. Recommended topic keywords

- "The Meeting of AI and Computational Chemistry: A New Paradigm for Catalyst and Drug Development"

- “AI-based Computational Chemistry Review: Toward the Future of Materials, Energy, and Environment”
- “Can AI Replace Experiments? Possibilities and Limitations Seen Through a Computational Chemistry Review”

Summary: The review article format can best utilize AI’s strengths, and the topic direction should focus on the latest research reviews combining AI with computational chemistry (especially drugs, catalysts, materials). The message is that AI is valuable not as a “replacement” for research, but as a “tool for acceleration and expansion.”

Would you like me to list candidate topics for each field of chemistry, or would you prefer a specific paper topic that can be used directly for the competition?

Prompt 2: Selection of Computational Chemistry Review Paper Topic and FF Recommendation

Question

What topics in computational chemistry are worth writing a review article on, i.e., what topics are currently valuable for review articles?

Answer

Saying a review article is “valuable” means (1) it is up-to-date, (2) there is still a lack of comprehensive reviews, and (3) it is a field that researchers will need to refer to in the future. In computational chemistry, topics that meet these criteria include the following.

- 1. AI and Machine Learning-Based Force Field Development and Applications**
 - There are active attempts to correct or newly train existing force fields (CHARMM, AMBER, OPLS, ReaxFF, DeepMD, AMOEBA, ANT, PCFF, etc.) using AI/ML.
 - There is still a lack of reviews that organize the differences in application between materials simulation (batteries, catalysts) and biomolecules (proteins, drugs).
 - “AI-driven Force Fields: Applications from Biomolecules to Materials”
- 2. Computational Chemistry-Based Catalyst Design (Especially Electrochemical Reactions)**
 - CO₂ electroreduction (CO₂RR), oxygen evolution reaction (OER), hydrogen evolution reaction (HER), nitrogen reduction reaction (NRR) are hot topics.
 - There is great value in reviews because there is an explosion of research on active site design, reaction pathway prediction using DFT calculations, and high-speed screening using AI.

- “Computational Insights into Electrocatalysis: AI-assisted Trends and Perspectives”

3. AI-accelerated Molecular Dynamics (MD) and Simulation Acceleration

- Long time-scale protein folding, membrane protein simulations, and complex interfacial reactions were difficult with conventional MD, but AI/ML-based acceleration techniques are now being actively researched.
- There is still a lack of comprehensive reviews, so the value of such a paper is high.
- “Machine Learning in Molecular Dynamics: Accelerating Biomolecular and Material Simulations”

4. Innovation in Drug Discovery Using Computational Chemistry

- The fusion of AI and computational chemistry for virtual screening, molecular docking, and drug-target interaction prediction is very active.
- Especially, reviews connecting AlphaFold structure prediction to computational chemistry-based drug design are still in the early stages.
- “From Structure Prediction to Drug Discovery: AI in Computational Chemistry”

2. Direction of the topic (Chemistry field) Currently, the hottest trend in chemistry is the fusion of computational chemistry and AI. For example: ● Computational Chemistry Review

- AI-based drug candidate prediction (drug discovery)
- Machine learning in catalyst design
- Acceleration and optimization of molecular dynamics (MD) simulations
- Materials Chemistry
- AI-driven battery material discovery (lithium/sodium battery, hydrogen storage, etc.)
- New material design for solar cells/fuel cells
- Environmental Chemistry
- CO₂ conversion/reduction catalyst research (electrocatalysis)
- Development of electrochemical catalysts for hydrogen production (HER, OER)

3. Recommended topic keywords ● “The Meeting of AI and Computational Chemistry: A New Paradigm for Catalyst and Drug Development”

● “AI-based Computational Chemistry Review: Toward the Future of Materials, Energy, and Environment”

● “Can AI Replace Experiments? Possibilities and Limitations Seen Through a Computational Chemistry Review”

?? Summary: ● Paper format: Review article (maximizing AI’s strengths)

● Topic direction: Centered on computational chemistry (especially drugs, catalysts, materials) → Latest research reviews combined with AI

● Message: AI is valuable as a “tool for acceleration and expansion” rather than a “replacement” for research.

Would you like me to list candidate topics for various fields of chemistry as I am now, or would you prefer a specific paper topic that can be immediately used for the competition?

Reason for choosing force field

Prompt 2: Selection of Computational Chemistry Review Paper Topic and FF Recommendation

Question

What topics in computational chemistry are worth writing a review article on, i.e., what topics are currently valuable for review articles?

Answer

Saying a review article is “valuable” means (1) it is up-to-date, (2) there is still a lack of comprehensive reviews, and (3) it is a field that researchers will need to refer to in the future. In computational chemistry, topics that meet these criteria include the following.

1. AI and Machine Learning-Based Force Field Development and Applications

- There are active attempts to correct or newly train existing force fields (CHARMM, AMBER, OPLS, ReaxFF, DeepMD, AMOEBA, ANT, PCFF, etc.) using AI/ML.
- There is still a lack of reviews that organize the differences in application between materials simulation (batteries, catalysts) and biomolecules (proteins, drugs).
- “AI-driven Force Fields: Applications from Biomolecules to Materials”

2. Computational Chemistry-Based Catalyst Design (Especially Electrochemical Reactions)

- CO₂ electroreduction (CO₂RR), oxygen evolution reaction (OER), hydrogen evolution reaction (HER), nitrogen reduction reaction (NRR) are hot topics.
- There is great value in reviews because there is an explosion of research on active site design, reaction pathway prediction using DFT calculations, and high-speed screening using AI.
- “Computational Insights into Electrocatalysis: AI-assisted Trends and Perspectives”

3. AI-accelerated Molecular Dynamics (MD) and Simulation Acceleration

- Long time-scale protein folding, membrane protein simulations, and complex interfacial reactions were difficult with conventional MD, but AI/ML-based acceleration techniques are now being actively researched.
- There is still a lack of comprehensive reviews, so the value of such a paper is high.
- “Machine Learning in Molecular Dynamics: Accelerating Biomolecular and Material Simulations”

4. Innovation in Drug Discovery Using Computational Chemistry

- The fusion of AI and computational chemistry for virtual screening, molecular docking, and drug-target interaction prediction is very active.

- Especially, reviews connecting AlphaFold structure prediction to computational chemistry-based drug design are still in the early stages.
- “From Structure Prediction to Drug Discovery: AI in Computational Chemistry”

FF Recommendation: Top 10 Commonly Used Force Fields

Question

In computational chemistry, MD simulations are frequently used. Can you tell me the 10 most commonly used force fields for this purpose?

Answer

- CHARMM: Organic molecules, biomolecules –widely used across organic/biological systems.
- ANT: Various systems –a relatively new force field, considered for its applicability to various systems.
- AMOEBA: Systems with polarity/induced dipoles –considers polarity and polarizability. Suitable for precise calculations.
- ReaxFF: Reactive systems –can simulate chemical reactions (bond formation/breaking). Essential for reactive systems.
- DeepMD: Materials, inorganic substances –deep learning-based force field. Aims for DFT-level accuracy.
- PCFF (Polymer Consistent FF): Polymers, materials –suitable for polymers and complex solid systems.
- UFF (Universal FF): Entire periodic table –broad elemental coverage. For initial optimization and fast screening.
- GROMOS: Proteins, lipids –widely used in GROMACS. Optimized for biomolecular simulations.

These 10 are essential force fields for handling molecular-level interactions, and it is important to choose appropriately depending on the purpose and system characteristics. If needed, I can also explain the formula structure or parameter features of each force field in detail.

Criteria for Writing the Introduction: Points to Note When Writing a Review Paper Introduction

Question

I'm going to write a review paper with CHARMM, AMBER, OPLS, ANT, AMOEBA, ReaxFF, DeepMD, and PCFF (Polymer Consistent FF). What should I be careful about in the intro part?

Answer

The introduction of a review paper is critical, because it sets the stage not just for summarizing potentials like CHARMM, AMBER, OPLS, ANT, AMOEBA, ReaxFF, DeepMD, PCFF, but also for convincing the reader why your review matters. Here are some things you should be careful about when drafting it:

(a) Balance Between Breadth and Depth

- Since you're covering both classical biomolecular force fields (CHARMM, AMBER, OPLS, ANT), polarizable/multipole force fields (AMOEBA), reactive force fields (ReaxFF), machine learning potentials (DeepMD), and polymer-specific force fields (PCFF), the intro should not deep dive too early.
- Instead, frame them in categories:
 - Traditional biomolecular force fields (CHARMM, AMBER, OPLS, ANT)
 - Advanced polarizable/quantum-informed force fields (AMOEBA)
 - Reactive potentials (ReaxFF)
 - ML-driven approaches (DeepMD)
 - Polymer-specific models (PCFF)
- This prevents overwhelming the reader right at the beginning.

(b) Historical Context

- Start with the history of molecular dynamics simulations: from early Lennard-Jones and hard-sphere models → rise of biomolecular force fields → expansion to reactive/ML approaches.
- Emphasize how force fields evolved to address specific challenges (e.g., fixed charges → polarizability, non-reactive → reactive, hand-parameterized → data-driven).
- This gives the reader a sense of progression and why your review is timely.

(c) Define "Potential / Force Field" Clearly

- Many non-experts confuse "force field" with just a formula.
- In the intro, explicitly state:
 - Functional form (bonded + nonbonded terms)
 - Parameters (empirical, ab initio, ML-trained)

- Domain of applicability (proteins, polymers, chemical reactions, materials)
- That clarity ensures readers won't misinterpret your later comparisons.
- Each force field has strengths and weaknesses (e.g., CHARMM vs AMBER for proteins, ReaxFF vs DeepMD for chemistry, PCFF vs OPLS for polymers).
- Point out that users often face difficulty choosing the right potential, especially with emerging ML-based ones.
- State explicitly that the goal of your review is to guide the selection of potentials under specific conditions.

(d) **Scope & Exclusions**

- Which force fields you will discuss in detail (your list).
- Which ones you will not cover (e.g., niche ones like GROMOS, COMPASS, etc., unless relevant).
- This sets boundaries so reviewers don't criticize you for missing something.

This gives the reader a sense of progression and why your review is timely.

3. Define "Potential / Force Field" Clearly ● Many non-experts confuse "force field" with just a formula.

● In the intro, explicitly state:

○ Functional form (bonded + nonbonded terms) ○ Parameters (empirical, ab initio, ML-trained) ○ Domain of applicability (proteins, polymers, chemical reactions, materials) That clarity ensures readers won't misinterpret your later comparisons.

● Each force field has strengths and weaknesses (e.g., CHARMM vs AMBER for proteins, ReaxFF vs DeepMD for chemistry, PCFF vs OPLS for polymers). ● Point out that users often face difficulty choosing the right potential, especially with emerging ML-based ones. ● State explicitly that the goal of your review is to guide the selection of potentials under specific conditions.

5. Scope Exclusions

○ Which force fields you will discuss in detail (your list).

○ Which ones you will not cover (e.g., niche ones like GROMOS, COMPASS, etc., unless relevant).

This sets boundaries so reviewers don't criticize you for missing something.

6. Applications as Motivation

● Mention real-world importance: drug design, energy materials, catalysis, nanotechnology. ● This makes your review relevant across disciplines.

● Be careful to cite landmark papers in the intro:

○ DeepMD's Zhang et al. papers

○ Original PCFF papers

● Also cite recent reviews/meta-analyses, to show your paper adds new synthesis rather than repeats existing work.

● Don't overload the intro with equations ? keep those for later sections. ● Don't claim one force field is "the best" in the intro ? keep neutrality until the discussion. ● Don't assume the reader knows acronyms (spell them out the first time). ● Avoid redundancy: instead of describing each FF in detail in the intro, motivate why reviewing them together makes sense.

? In summary:

<https://getliner.com/ko/search/s/26944841/t/87830100?msg-entry-type=main>

Review Planning and Introduction Drafting

Planning Steps

- (a) Outlined the history of molecular dynamics and its importance.
- (b) Drafted the definition of potentials and planned a historical overview of potentials.
- (c) Addressed the Lennard-Jones potential.
- (d) Expanded the outline to include specific sections for each potential (CHARMM, AMBER, etc.), focusing on theoretical foundations, strengths, and applications.
- (e) Planned the comparison section, categorizing potentials, and considering hybrid methods.
- (f) Refined the paper's structure based on detailed requirements, solidified the introduction's components and overall layout.
- (g) Sketched out sections dedicated to each potential, ensuring detailed summaries with appropriate citations.
- (h) Finalized the introduction to include history, definitions, and the Lennard-Jones potential, with equations and sentence-level citations.
- (i) Wrote the scope of the paper to align with requirements.
- (j) Drafted the structure for each potential, aiming for comprehensive understanding and comparative analysis.
- (k) Ensured all summaries adhere to the structure: Theoretical Foundation, Parameterization Strategy, Unique Features, Strengths Weaknesses, and Primary Scientific Applications.
- (l) Developed the historical context, including early simulations and the evolution of computational resources.
- (m) Drafted the definition of potential energy in MD simulations and the Lennard-Jones potential and related equations.
- (n) Integrated discussions on early MD simulations and the importance of potentials in MD simulations.
- (o) Organized the paper's scope, focusing on detailed summaries for each potential.

- (p) Outlined the review paper's scope and objectives, and refined the rationale for reference selection.
- (q) Integrated all essential components, including a historical overview, definitions, and equations.
- (r) Finalized the scope and objectives while maintaining sentence-level citations.
- (s) Drafted the introduction, encompassing the history of MD simulations, defining potentials, and detailing the Lennard-Jones potential and its implications.
- (t) Outlined objectives and reference selection rationales.
- (u) Detailed parameterization strategies for CHARMM and the development of CGenFF for drug-like molecules.
- (v) Documented unique features within CHARMM, highlighting its extensive parameter library.
- (w) Refined the introduction and equations, ensuring sentence-level citations before continuing with detailed summaries.
- (x) Detailed parameterization strategies for AMBER, including quantum chemical calculations and fitting to experimental data.
- (y) Wrote details for partial atomic charges from quantum mechanical electrostatic potential fitting and bonded parameters.
- (z) Focused on parameterization strategies for AMBER, discussing derivation of partial atomic charges and validation of bonded parameters.
- (o) Summarized CHARMM and AMBER, detailed OPLS frameworks, and refined summary outlines.
- (o) Began drafting sections for ANT, AMOEBA, ReaxFF, DeepMD, and PCFF, focusing on theoretical foundations, parameterization strategies, unique features, strengths, weaknesses, and applications.
- (o) Outlined the introduction, including history, definitions, and LJ potential sections, and covered scope and objectives.
- (o) Crafted sections for ANT, AMOEBA, ReaxFF, DeepMD, and PCFF, outlining core principles and parameterization methods, and analyzing strengths, weaknesses, and application areas.

The Evolution of Molecular Dynamics Simulations Molecular dynamics (MD) simulations have emerged as a cornerstone in computational science, tracing their origins back to foundational work in the late 1950s. The pioneering efforts of Alder and Wainwright in 1957 marked a critical milestone, as they employed a hard-sphere model to investigate particle collisions, laying the groundwork for the field. This initial research demonstrated the nascent capability of MD methods to simulate real atomic interactions, moving beyond purely theoretical models. Subsequently, in the mid-1960s, Rahman significantly advanced the field by extending MD to more realistic continuous potentials, notably simulating liquid argon [Berne, 2002](#). This progression allowed for the study of increasingly complex molecular systems, moving from simple monoatomic liquids to diatomic and triatomic molecules throughout the 1960s and 1970s [Berne, 2002](#).

. ?A landmark achievement was the first protein simulation in 1976, which utilized an empirical energy function based on physical principles, thereby enabling the modeling of biological macromolecules and initiating a shift towards vital applications in biophysics and materials scienceZhu Lopes, 2012

. ?The ongoing evolution of computational methods, including the development of force fields, advanced integration schemes, and the introduction of multiple time scale algorithms, has been pivotal in expanding the scope of MD simulations to larger systems and longer timescalesBerne, 2002

Perez, 2011

. ?The exponential increase in computing power has transformed simulations from mere picoseconds to microseconds and beyond, encompassing systems with millions of atomsLoeer Winn, 2012

. ?Recent advancements have also integrated quantum mechanical approaches like ab initio molecular dynamics, providing highly accurate descriptions of electronic interactionsCollins et al., 2005

. ?Furthermore, techniques such as accelerated MD methods address challenges related to infrequent-event dynamics, significantly enhancing sampling efficiency and broadening the scope of scientific inquiryHamelberg et al., 2004

Perez, 2011

. ?The continuous refinement of algorithms and computational infrastructure, including parallel computing and GPU acceleration, has made MD simulations widely accessible and indispensable across various scientific disciplinesChen et al., 2021

.

Defining the Potential Energy Function in Molecular Dynamics

The concept of "potential" in molecular dynamics simulations refers to the potential energy function that describes the energy of a molecular system as a consequence of its atomic coordinates. This function is crucial as it encapsulates all interactions among atoms and molecules, including both bonded interactions?such as those defining bond lengths, angles, and torsional rotations?and non-bonded interactions, which include van der Waals forces and electrostatic interactions. ?The potential energy function fundamentally governs the dynamics of the system by determining the forces acting on each atom through its gradient, thereby dictating the trajectories of the atoms over timeHoover, 1986

. ?Mathematically, potentials in MD simulations are formulated within classical mechanics, typically Newtonian frameworks, and can take various functional forms depending on the system being studiedHoover, 1986

Semenov et al., 2023

. ?For instance, a widely used pairwise potential, the Lennard-Jones potential, is mathematically expressed as

??

(

??

)

=

4

??

$V(r)=4\epsilon$, where

??

r is the interatomic distance,

??

ϵ represents the depth of the potential well, and

??

σ denotes the finite distance at which the interparticle potential becomes zero Allen et al., 2006

Semenov et al., 2023

. More intricate potentials, such as many-body potentials, account for interactions involving multiple atoms simultaneously, offering a more accurate representation of physical phenomena in covalently bonded or metallic materials Malerba, 2006

Fan et al., 2015

Wood Thompson, 2017

. The physical significance of the potential is profound, as it directly determines the molecular interactions and the overall behavior of the system simulated by MD Allen et al., 1989

. It provides a quantitative measure of system stability and allows for the monitoring of energy fluctuations, which can signal significant molecular events or conformational changes Gunsteren, 1990

. These potential energy functions are often collectively referred to as force fields, forming the empirical or semi-empirical basis for accurate and computationally feasible simulations across diverse scientific domains Gonzalez, 2011

.

Historical Trajectory of Molecular Dynamics Potentials

The historical evolution of potential functions in molecular dynamics reflects a continuous drive towards greater accuracy and realism in simulating molecular systems Rodger et al., 1992

. This journey began with simple pairwise interaction models, notably the Lennard-Jones (LJ) potential Schwerdtfeger Wales, 2024

Berne, 2002

. The LJ potential, a pairwise additive model, was instrumental in describing van der Waals interactions through an analytical formula that combined both attractive and repulsive terms Monticelli Tieleman, 2012

. Its inherent simplicity and effectiveness established it as a cornerstone for early MD studies, serving as both a practical modeling tool and a fundamental conceptual framework for understanding intermolecular forces Allen et al., 2006

Vanommeslaeghe Jr, 2015

. However, as computational capabilities expanded and theoretical understanding deepened, the limitations of these simple pairwise potentials became evident, particularly in capturing the complex many-body interactions prevalent in condensed phases, intricate biomolecules, and advanced materials. This recognition spurred the development of more sophisticated many-body potential functions, which moved beyond simple pairwise sums to integrate effects such as molecular geometry, electronic polarization, and detailed electronic structure more accuratelyGuenza et al., 2018

. These advanced potentials significantly enhanced the realism of simulated molecular phenomena by providing a better description of system energetics and dynamics.

?In recent years, the field has witnessed a paradigm shift with the advent of machine learning (ML) techniques for constructing interatomic potentialsMartinBarrios NavasConyedo, 2024

Hou et al., 2024

. These ML-based potentials, particularly deep neural network potentials, are trained on extensive databases of quantum mechanical calculationsBalyakin et al., 2020

. This data-driven approach allows them to capture the complex potential energy surface with near first-principles accuracy while maintaining a much-improved computational efficiency compared to direct quantum mechanical calculationsZhang et al., 2018

Chen et al., 2021

. Examples like neural network potentials have emerged to approximate intricate potential energy landscapes without relying on rigid, predefined functional forms, thereby improving transferability and enabling simulations of larger and more complex systems?ugec et al., 2024

. This new generation of potentials represents a significant frontier in molecular dynamics, effectively bridging the long-standing gap between accuracy and scalabilityHou et al., 2024

. The progression from simple pairwise models to sophisticated many-body and cutting-edge ML-driven potentials underscores a continuous quest for enhanced accuracy, transferability, and computational efficiency in the representation of atomic interactions within MD simulationsMartinBarrios NavasConyedo, 2024

Hou et al., 2024

.

The Crucial Role of Potential Functions

Potential functions are undeniably central to molecular dynamics (MD) simulations, profoundly influencing the accuracy, transferability, and computational efficiency of the results [?]. These functions, often referred to as force fields, quantify the interaction energies among atoms or molecules, thereby governing the motion and

behavior of the simulated system [?, ?]. The reliability of MD simulations hinges critically on how faithfully these potentials represent real physical interactions, directly impacting the ability to predict molecular properties and dynamic phenomena with confidence [?].

Potentials vary widely in complexity, from straightforward pairwise additive functions such as the Lennard-Jones (LJ) potential, which precisely models van der Waals interactions through an analytical expression, to highly complex many-body and polarizable force fields [?]. The LJ potential holds significant historical importance as a foundational model in MD, offering both conceptual insights and practical utility for modeling non-bonded interactions. Its characteristic attractive and repulsive terms enable a succinct representation of fundamental interatomic forces.

A vital consideration in MD is the computational efficiency of evaluating these potential functions, particularly for simulations involving large systems or extended timescales [?, ?]. While classical force fields offer a computationally tractable approach, they may present limitations in accuracy for certain complex phenomena. However, continuous advancements in algorithms and hardware have facilitated the implementation of more intricate potentials, including many-body and machine learning-based models, which strive to balance enhanced accuracy with practical speed [?, ?]. Optimized neighbor lists and parallelization strategies further boost the performance of MD simulations, making complex calculations more feasible [?].

Transferability, defined as the capacity of a potential function to accurately describe different systems or chemical environments without requiring extensive reparameterization, remains a significant challenge [?]. Force fields are developed with varying degrees of specificity and generality, which directly affects their applicability across a diverse range of molecular systems. While greater transferability can broaden the utility of a potential, it might sometimes compromise detailed accuracy if the model becomes overly generalized. The integration of fundamental physical principles, explicit polarizability, and sophisticated data-driven training strategies has led to the development of potentials that achieve a superior balance between transferability and accuracy. Ultimately, potential functions serve as the foundational backbone of molecular dynamics simulations, directly dictating the critical trade-offs among computational efficiency, accuracy, and broad applicability across chemical and material systems.

Current Landscape of Potential Applications

Molecular dynamics potentials are currently employed across an extensive array of scientific fields, enabling detailed insights into diverse molecular phenomena. CHARMM (Chemistry at HARvard Macromolecular Mechanics) is predominantly utilized in biomolecular simulations, excelling in the accurate modeling of proteins, nucleic acids, lipids, and carbohydrates [?, ?]. Its robust empirical parameterization and ability to simulate complex macromolecular systems make it invaluable in drug design and studies elucidating structure-function relationships in biological contexts [?, ?, ?, ?]. Both additive and polarizable variants of CHARMM extend its utility to accurately capture molecular polarization effects crucial for biological processes [?, ?, ?].

Similarly, AMBER (Assisted Model Building with Energy Refinement) is a leading potential for biomolecular simulations, particularly for proteins and nucleic acids, benefiting from continuous force field refinements [?]. AMBER supports a wide range of protocols for simulating molecular interactions pertinent to biochemical and pharmaceutical research, including drug discovery efforts [?, ?, ?, ?].

OPLS (Optimized Potentials for Liquid Simulations) is specifically optimized for organic molecules and biomolecules within condensed phases, finding widespread application in liquid-phase simulations and polymer modeling [?, ?, ?]. It accurately reproduces thermodynamic and structural properties, making it suitable for materials science, small molecule simulations, and the prediction of thermophysical properties [?, ?, ?].

. ?The ANT concept, while not a conventional force field, encompasses ant colony-inspired algorithmic approaches that are beneficial for exploring potential energy surfaces and molecular configurationsSalmaso Moro, 2018

. ?It finds application in molecular system optimization and molecular communication modeling, and can be relevant in biomolecular and material simulations requiring configurational optimizationSalmaso Moro, 2018

.

?AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) stands out as a polarizable force field highly specialized for biomolecular systems, offering superior accuracy in electrostatics through its use of atomic multipoles and inducible dipolesAlbaugh et al., 2016

Liu et al., 2019

Chung et al., 2023

. ?It is widely applied in studying protein, nucleic acid, small molecule interactions, and ionic liquids, with recent applications extending to enhanced sampling methods, hybrid quantum mechanics/molecular mechanics approaches, and complex solvent effectsChung et al., 2023

Nottoli et al., 2020

Chung et al., 2023

. ?ReaxFF serves the critical domain of reactive molecular dynamics, enabling the modeling of chemical reactions, including bond-breaking and formation processesSenftle et al., 2016

Li et al., 2021

Dasgupta et al., 2022

. ?This capability makes it exceptionally powerful in materials science, combustion chemistry, catalysis, polymer decomposition, and catalytic surface reactionsLi et al., 2021

Wu et al., 2024

.

?DeepMD (Deep Potential Molecular Dynamics) represents a new generation of machine learning-driven potentials, delivering near ab initio accuracy at a significantly reduced computational costWang et al., 2018

Zhang et al., 2018

Kounis-Melas Vella, 2025

. ?Its applicability spans various areas of materials science, including solid-state electrolytes, molten salts, and plasma-surface interactions, and it is particularly valuable for large-scale simulations where conventional potentials face limitations in scalability or accuracyWang et al., 2018

Liu et al., 2024

Kounis-Melas Vella, 2025

Balyakin et al., 2020

. ?Finally, PCFF (Polymer Consistent Force Field) is precisely tailored for polymers and macromolecular materials, accurately capturing their conformational, thermodynamic, and mechanical propertiesTafrishi et al., 2022

Alzate-Vargas et al., 2018

. ?It is validated for condensed-phase polymer simulations, studies of thermal decomposition, and the prediction of mechanical propertiesFatemi Foroutan, 2017

Rukmani et al., 2019

Li Strachan, 2015

Schatschneider Chronister, 2008

.

The Foundational Lennard-Jones Potential

?The Lennard-Jones (LJ) potential is a cornerstone of classical interatomic potentials in molecular dynamics (MD) simulationsLenhard et al., 2024

. It was first introduced by J.E. ?Lennard-Jones in 1924, providing a concise mathematical model for interactions between pairs of neutral atoms or moleculesSchwerdtfeger Wales, 2024

. This groundbreaking model significantly advanced the understanding of intermolecular forces by integrating both repulsive and attractive components of atomic interactions based on fundamental physical principles. The LJ potential's importance stems from its ability to effectively capture the essential characteristics of van der Waals forces and the short-range repulsion that arises from electron cloud overlaps. It serves as a foundational model for understanding and simulating molecular behavior across various states of matter.

Mathematically, the LJ potential energy function

??

(

??

)

$V(r)$ for two particles separated by a distance

??

r is given by the equation:

??

(

??

)

=

4

??

$V(r)=4\epsilon$

Here,

??

ϵ (epsilon) denotes the well depth, which quantifies the strength of the attractive interactionZhang, 2011

. The parameter

??

σ (sigma) represents the finite distance at which the inter-particle potential is zero, conceptually approximating the effective diameter of the interacting particlesZhang, 2011

. The term

(

??

/

??

)

12

(σ/r)

12

phenomenologically models the strong repulsive force that dominates when particles are in close proximity, reflecting the effects of Pauli exclusion. Conversely, the

(

??

/

??

)

6

(σ/r)

6

term captures the long-range attractive van der Waals interaction, which is physically interpreted as dispersion forces resulting from instantaneous induced dipoles. The derivation of the LJ potential involves an empirical representation for the repulsive term, chosen as the inverse 12th power for its computational convenience and steepness. The attractive term, however, is theoretically grounded in quantum mechanical London dispersion theory and classical electrostatic approximations. ? Despite its relatively simple functional form, the LJ potential shows good agreement with physical observations and experimental data for noble gases and numerous molecular systems Rowlinson, 1960

.

The LJ potential has been indispensable in the advancement of MD, serving as a fundamental model upon which more intricate and tailored potentials have been constructed. ? It enables both qualitative and quantitative studies of phenomena such as molecular clustering, phase transitions, and various thermodynamic properties Trejos et al., 2020

Kotelyanskii Hentschke, 1995

Liang Chen, 2000

. ? Its enduring importance also extends to the calibration and validation of more advanced force fields and potentials, finding broad application in biomolecular simulations, materials science, and condensed matter physics Vanommeslaeghe Jr, 2015
Lim, 2010

.

The LJ potential and its derivatives are extensively incorporated into more complex force fields, forming the basis for non-bonded interactions. ? In the CHARMM force field, the LJ potential is a key component of the non-bonded interaction terms, with its parameters meticulously optimized to reproduce experimental data such as densities, heats of vaporization, and conformational properties, particularly for hydrocarbon and lipid chains MacKerell et al., 1998

Beu Farcas, 2017

. ? Both additive and polarizable versions of CHARMM, like the Drude polarizable force field, utilize these LJ parameters, adjusting the well depths and radii to fit quantum mechanical data and experimental results, ensuring consistency with bonded parameters in simulations of lipids and proteins Qiao et al., 2018

. ? The AMBER force field also prominently features the LJ potential within its non-bonded terms, especially for biomolecules such as proteins, nucleic acids, and lipids Wang et al., 2004

Salomon ? Ferrer Case, 2013

Ogata Nakamura, 2015

. ? Parameter refinement efforts in AMBER include scaling LJ well-depth values in alkyl chains to better describe lipid phase transitions and developing specialized parameter sets like the General AMBER Force Field (GAFF) for a broader range of molecules Ogata Nakamura, 2015

Sprenger Jaeger, 2015

Kantakevičius et al., 2022

. These LJ parameters are derived through a combination of quantum mechanical calculations and fitting to experimental data, focusing on properties like hydration energies, structural characteristics, and thermodynamicsKantakevičius et al., 2022

. Similarly, the OPLS force field integrates the LJ 12-6 potential for non-bonded van der Waals interactions, with parameters optimized to accurately reproduce liquid properties, including densities and heats of vaporization, for organic molecules, hydrocarbons, and carbohydratesSiu Pluhackova, 2012

Adluri et al., 2015

. OPLS parameterization strategies carefully balance bonded and non-bonded parameters to achieve realistic thermodynamic and structural properties in condensed phases, with modifications to LJ parameters demonstrating improved accuracy for various systemsSiu Pluhackova, 2012

Adluri et al., 2015

.

Overview of This Review Paper

?This review paper provides a critical and comprehensive analysis of several prominent molecular dynamics potentials: CHARMM, AMBER, OPLS, ANT, AMOEBA, ReaxFF, DeepMD, and PCFF.? The primary objective is to evaluate the strengths and weaknesses of each potential based on established criteria for force field assessment, including accuracy, transferability, computational efficiency, and suitability for specific scientific applications. ?By achieving this, the paper aims to equip researchers with the necessary insights to select the most appropriate force field for their particular molecular simulation needsSenftle et al., 2016

. The review will progress through a detailed comparative evaluation, highlighting the unique features and optimal application domains for each potential, such as biomolecular systems for CHARMM and AMBER, polarizable force fields like AMOEBA for precise electrostatic modeling, reactive potentials such as ReaxFF for chemical reactions, and machine learning-based approaches like DeepMD for advanced material simulationsAdcock McCammon, 2006

. The paper will conclude with clear recommendations, synthesizing the findings to guide practitioners in making informed decisions about potential selection in various research contextsAdcock McCammon, 2006

.

Rigorous Standards for Reference Selection

The reliability and scientific integrity of this review paper are underpinned by adherence to rigorous standards for reference selection. ?All references included are from peer-reviewed journals or other authoritative scientific publications, ensuring that the research has undergone stringent scrutiny and validation by experts in the fieldPortal, 2016

Hassan et al., 2022

Jett, 2012

. Peer review is considered a fundamental hallmark of academic credibility Andrews et al., 2022

. Furthermore, the citation count of a publication serves as an indicator of its impact and recognition within the scientific community; highly cited works often represent seminal contributions or widely accepted findings. However, this review also incorporates recent publications even with lower citation counts, particularly if they present innovative advancements or critical updates in the rapidly evolving field of molecular dynamics potentials. The relevance of each work to molecular dynamics potentials is meticulously assessed, ensuring that papers provide substantial information on the development, implementation, comparison, or application of these potentials. This includes studies that review existing force fields, benchmark their performance, address their limitations, or demonstrate their successful application in specific scientific contexts. Finally, a strong emphasis is placed on references that demonstrate transparency and reproducibility, ideally including sufficient methodological detail to allow for replication, such as precise reporting of parameters, algorithms, and validation practices Robustelli et al., 2018

. These criteria collectively ensure that the selected references are authoritative, impactful, current, and directly applicable, thereby providing a robust foundation for the discussions on molecular dynamics potentials and their comparative utility.

Comparative Analysis and Guidelines for Force Fields

4.1. Transferability

Transferability is the ability of a force field (FF) parameterized for a specific system or chemical environment to perform well in other systems or environments not included in the training data. It is a key indicator of the robustness and general applicability of an FF across diverse chemical situations and molecular types.

Definition and Evaluation Criteria

- **Parameter space coverage:** Diversity of atom types, bond types, functional groups, and chemical environments covered by FF parameters.
- **OOD (Out-of-Distribution) prediction reliability:** Ability to make stable and reliable predictions in new chemical environments not present in the training data.
- **Agreement with experiment and QM:** Degree to which predicted physical properties (e.g., ΔG , density, MAE) match experimental or high-level quantum mechanical (QM) results under various conditions.

Comparison by FF Type

- **Classical Force Fields:** High transferability within trained chemical space, but limited outside parameterization range. For example, CHARMM22, AMBER FF03, AMBER FF99SB, and OPLS-AA/L show different trends in backbone and sidechain structure prediction for amino acid residues, reflecting the influence of charge models during parameterization.
- **Polarizable Force Fields:** Improve transferability via explicit polarization, but increase parameterization complexity. AMOEBA FF, using induced dipole and multipole models, accurately reproduces DNA and RNA solution/crystal structures within 2.0 Å RMSD of NMR structures.
- **Reactive Force Fields:** Limited transferability, often restricted to specific reaction types. In benchmarks for 20 amino acids and 11 dipeptides, ReaxFF showed deviations over 100° in amino acid torsion angles and predicted non-physical reaction mechanisms.
- **Machine Learning-based Force Fields:** High accuracy with large training data, but vulnerable to OOD data. UQ (uncertainty quantification) is essential. Models like MACE-OFF achieve RMSD 0.5 kcal/mol and torsion barrier MAE 0.8 kcal/mol for organic molecule torsion potential energy surfaces, outperforming traditional FFs, but may be unstable for OOD predictions.

4.2. Long-Range Interactions

Long-range interactions refer to electrostatic, polarization, dispersion, and other effects that persist even at large atomic separations, critically impacting structure, dynamics, and thermodynamic properties.

Definition and Evaluation Criteria

- **Interaction treatment:** How electrostatic, polarization, and dispersion interactions are modeled.
- **PME/Ewald/Thole models:** Use of periodic boundary condition methods and damping models for long-range interactions.
- **Macroscopic property accuracy:** Accuracy in predicting properties such as water dielectric constant (ϵ_r), liquid density, and interfacial energy.

Comparison by FF Type

- **Classical Force Fields:** Use fixed charge models and treat long-range electrostatics with PME or Ewald summation. Lack of explicit polarization leads to underestimation of water dielectric constant by 5–10%. Use of LJ-PME improves accuracy for density and surface tension in alkanes.
- **Polarizable Force Fields:** Employ induced dipole or multipole models and Thole damping to implement environment-dependent polarization. AMOEBA FF achieves 2–3% error in water dielectric constant and high precision in property prediction.

- **Reactive Force Fields:** Use QEq (charge equilibration) for dynamic charge distribution, but traditionally approximate or cut off long-range Coulomb interactions. Recent Ewald summation improvements yield 12% better graphene-water interfacial capacitance.
- **ML-based Force Fields:** Rely on short-range cutoff learning, making long-range interaction capture difficult. Hybrid approaches combine PME/Ewald with ML-FF or use ML-predicted partial charges. Interfacial energy errors range from 2.9 to 21.2 kcal/mol.

4.3. Numerical Stability

Numerical stability is a key factor in determining the reliability and efficiency of MD simulations with a given FF. It is assessed by energy conservation, structural integrity, and convergence behavior during long simulations.

Definition and Evaluation Criteria

- **Energy drift:** Degree of total energy conservation in NVE ensemble over time.
- **Trajectory integrity:** Ability to maintain structural integrity during long MD runs, assessed by RMSD and structural deformation.
- **Convergence and nonphysical phenomena:** Speed and stability of iterative calculations (e.g., induced dipole, charge equilibration), and occurrence of nonphysical breakdowns (e.g., charge runaway).

Comparison by FF Type

- **Classical Force Fields:** High stability, can maintain microsecond-scale trajectories. Fixed functional forms and parameters yield low energy drift and reliable long-term simulations. Some variants (e.g., CHARMM22*, CHARMM27, Amber ff99SB-ILDN) show better stability than others.
- **Polarizable Force Fields:** Induced dipole convergence can be problematic, but improved SCF algorithms and predictor-corrector methods enable microsecond MD. Drude-based FFs achieve stable trajectories for $>5\mu\text{s}$ ubiquitin simulations; increasing predictor-corrector history reduces energy drift and speeds up calculations.
- **Reactive Force Fields:** Historically unstable due to charge runaway, but new algorithms and potentials have greatly improved stability. NN-based reactive FFs show much better energy conservation than ReaxFF.
- **ML-based Force Fields:** Extrapolation outside training data can cause instability (holes, non-conservative forces). For example, 100% of MACE FF trajectories for MAPbI₃ at 500 K were unstable. High accuracy is possible, but instability arises outside training domain.

4.4. Interpretability and Uncertainty Quantification

Interpretability refers to how directly FF parameters relate to physical/chemical meaning and how transparently the model operates, enabling failure tracing. Uncertainty quantification (UQ) refers to methods for quantitatively assessing the reliability of FF predictions, such as Bayesian inference and ensemble modeling.

Definition and Evaluation Criteria

- **Physical meaning of parameters:** How directly FF components map to real physical/chemical concepts.
- **Model transparency and failure tracing:** Transparency of model operation and ease of identifying prediction errors.
- **UQ method applicability and validity:** Applicability and reliability of UQ methods (Bayesian, ensemble) for quantifying prediction confidence.

Comparison by FF Type

- **Classical Force Fields:** Parameters have clear physical meaning (bond lengths, angles, torsions), making interpretation and failure tracing easy. Bayesian inference is used for UQ of FF parameters.
- **Polarizable Force Fields:** Added complexity from polarizability/induced dipole parameters, but still interpretable. UQ studies quantify how uncertainty in physical-chemical data affects FF performance.
- **Reactive Force Fields:** Bond order models are intuitive, but large parameter sets and correlations make interpretation difficult. UQ is in early stages; Bayesian active learning and LTAU methods are being explored.
- **ML-based Force Fields:** Black-box models are hard to interpret; UQ (ensemble, Bayesian) is essential for OOD detection and reliability assessment.

4.5. Summary Table of Comparative Analysis

5. Application Case Studies and Guidelines

Evaluating FF performance in real-world problems is essential for understanding the strengths and limitations of each type. This section compares how each FF type is used and performs in actual systems, using quantitative benchmarks.

5.1. Case Selection Criteria

- **Representativeness and diversity:** Includes systems or phenomena that clearly illustrate the characteristics of each FF type.
- **Availability of quantitative data:** Prioritizes cases with sufficient benchmark data and reproducible simulation results in the literature.

5.2. Quantitative Comparison Metrics

- **Binding free energy ($\Delta G_{binding}$):** Thermodynamic stability of molecular interactions (e.g., protein-ligand binding).
- **Reaction activation energy:** Kinetic barrier for chemical reactions.
- **Experiment/QM error (MAE, RMSE):** Mean absolute or root mean square error between simulation predictions and experiment/high-level QM.
- **Structural stability (RMSD):** Structural integrity of the system during simulation.
- **Thermodynamic properties (density, dielectric constant):** Macroscopic physical properties of liquids/solids.

Protein-ligand binding is a crucial process in drug development, and accurately predicting binding affinity is essential for designing new drug molecules.

- Limitation of classical FFs: Classical FFs use fixed charge models, so they cannot properly reflect the electronic polarization effects that occur when a ligand enters a protein binding pocket. This leads to errors in binding free energy predictions, and the accuracy tends to decrease especially in polar environments. For example, when predicting the binding free energy between Mg^{2+} and H_2PO_4^- , the CHARMM fixed charge FF predicted -41.0 kcal/mol, which is a significant overestimation compared to the experimental value of -1.7 kcal/mol. This is because non-polarizable FFs try to artificially compensate for real polarization effects with excessively strong permanent electrostatic interactions.

- Improvement with polarizable FFs: Polarizable FFs (e.g., AMOEBA) explicitly include induced dipoles, realistically modeling electronic polarization effects and greatly improving the accuracy of binding free energy predictions. Using AMOEBA FF, the predicted binding free energy for Mg^{2+} and H_2PO_4^- was -2.23 kcal/mol, much closer to the experimental value. In addition, when calculating the relative binding free energies of Mg^{2+} and Ca^{2+} for various proteins using AMOEBA FF, a high correlation coefficient of 0.97 with experimental measurements was observed. This indicates that polarizable FFs capture the delicate balance of ion-protein interactions better than classical FFs.

5.3.2. Lithium Battery Electrode-Electrolyte Interface Reactions: Reactive FF vs. ML-based FF

Understanding the reaction mechanisms at the lithium battery electrode-electrolyte interface (CEI) is essential for improving battery performance and lifespan. Reactive FFs and ML-based FFs are used to model these complex chemical reactions.

- Capability and limitations of reactive FFs (ReaxFF): ReaxFF can dynamically model the formation and breaking of chemical bonds, making it useful for qualitatively predicting reaction pathways such as CEI formation. However, ReaxFF has traditionally handled long-range Coulomb interactions with a short-range cutoff for computational efficiency, which limits its accuracy in predicting electrical properties such as interfacial capacitance. For example, short-range Coulomb (SRC) based ReaxFF predicted the graphene-water interfacial capacitance as $4.4 \pm 0.2 \mu\text{F cm}^{-2}$, which is 12

- Accuracy and efficiency of ML-based FFs: ML-based FFs (e.g., MACE) can predict energies and forces with DFT-level accuracy by learning from vast quantum mechanical data. This achieves high accuracy (MAE 1–5 meV/atom) in reaction pathway predictions, and computational efficiency is also greatly improved compared to traditional QM methods. ML-FFs are powerful tools for quantitatively analyzing complex interfacial reaction mechanisms like CEI at the atomic level. However, ML-FFs can show prediction instability in out-of-distribution (OOD) environments, and in 500 K simulations of the MAPbI₃ system, 100

5.3.3. Large-Scale Virus Capsid Assembly and Structure: Coarse-Grained FF

Coarse-grained (CG) FFs provide excellent efficiency for simulating the assembly and structural dynamics of large biomolecular systems (e.g., virus capsids, cell membranes).

- Efficiency of CG FFs (MARTINI): CG FFs like MARTINI 3 simplify multiple atoms into a single “bead,” greatly reducing the system’s degrees of freedom and dramatically extending simulation time and length scales. This can improve sampling efficiency by 3–4 orders of magnitude compared to atomistic FFs, enabling simulations of giant molecular systems with hundreds of thousands of molecules over millisecond timescales. For example, they are used to efficiently study complex biological phenomena such as virus capsid assembly or liposome fusion.

- Structural artifacts and limitations: CG FFs sacrifice fine molecular interaction details, so they may not accurately reproduce certain important physical or chemical properties at the atomic level. For example, they have limitations in finely modeling local hydrogen bond networks in proteins or specific solvation effects, and can produce structural artifacts such as over-aggregation or underestimation of melting points. These limitations arise from the “coarseness” of the CG model and the information loss during potential averaging. Therefore, CG FFs are useful for macroscopic phenomenon analysis, but for cases requiring atomic-level precision, hybrid (multi-scale) FFs or atomistic FFs are needed.

6. Practical Recommendations and Guidelines

To maximize the reliability and efficiency of MD simulations, it is important to carefully select the Force Field (FF) that matches the research purpose and to follow a systematic parameterization and validation workflow.

6.1. Problem-Based Force Field Selection Guide FF selection should be based on three key elements: simulation objectives, system characteristics, and available computational resources.

- Simulation objectives:

- Studying chemical reactions (bond formation/breaking): Consider reactive FFs (e.g., ReaxFF) or ML-based reactive FFs (NNRF). These can model dynamic chemical changes via bond order models.

- Studying polarization effects and changes in the electrical environment (e.g., ion solvation, protein-ligand binding): Polarizable FFs (e.g., AMOEBA, CHARMM Drude) are essential. These explicitly handle electronic polarization according to the environment, increasing accuracy.

- Studying accurate structure and dynamics (no chemical reactions): Classical FFs

(e.g., AMBER ff14SB, CHARMM36, OPLS-AA) or high-accuracy ML-based FFs (e.g., ANI, MACE) can be used.

- Large-scale systems or long-time simulations (macroscopic properties): Coarse-grained FFs (e.g., MARTINI) or efficient classical FFs are suitable.

- System characteristics:

- Organic molecules, biomolecules (proteins, DNA): Classical, polarizable, and ML-based FFs can all be used, selected according to the required accuracy.

5. Metals, ceramics, ionic crystals: Consider EAM, MEAM, COMB classical FFs or ML-FFs specialized for inorganic materials.
6. Composites, multiphase interfaces: Reactive or ML-based FFs may better handle complex interactions.

Computational resources

- Limited resources: Classical FFs are most efficient.
- High-performance computing (GPU, parallel): Can support the high computational cost of polarizable and ML-based FFs.

Recommended Baselines

- Classical FF: AMBER ff14SB, CHARMM36, OPLS-AA/L (optimized for specific research domains)
- Polarizable FF: AMOEBA, CHARMM Drude (accurate electrostatics and property prediction)
- Reactive FF: Latest ReaxFF (for chemical reaction simulation)
- ML-based FF: ANI, MACE, NequIP (high accuracy, use UQ for OOD reliability)

6.2. FF Parameterization and Validation Workflow

Systematic workflow is essential for reliable FF development and use.

- **QM data collection:** Use high-quality quantum mechanical data (e.g., ANI-1ccx, DES370K, q-AQUA dimers) or generate with DFT/CCSD(T) for diverse atomic environments and forces.
- **Parameter fitting/regularization:** Use tools like ForceBalance/genetic algorithms to fit FF parameters to QM/experimental data. For ML-FFs, use active learning to expand training data and optimize parameters; apply regularization to avoid overfitting.
- **Validation/benchmarking:** Use independent test sets to evaluate FF performance (MAE, RMSE, density, dielectric constant, binding free energy, etc.). Use tools like OpenFF Evaluator for systematic benchmarking. For ML-FFs, apply UQ to assess prediction confidence.
- **Distribution/reproducibility:** Share parameter files, code, simulation/analysis scripts in public repositories (e.g., GitHub) with DOIs for reproducibility.

6.3. Reproducibility Checklist

To ensure reproducibility, clearly document and share:

- FF parameters and version
- Initial structure files (PDB, XYZ, etc.)
- Thermostat/barostat settings (type, coupling constant, target T/P)
- OOD detection results (for ML-FFs)
- Energy drift test results (NVE ensemble)
- Simulation/analysis scripts
- Software version (LAMMPS, CHARMM, GROMACS, OpenMM, etc.)
- Hardware/OS info

6.4. Pitfalls and Recommendations

Each FF type has unique failure modes and application limits.

- **Classical FF:** Not suitable for systems where polarization or chemical reactions are critical.
- **Polarizable FF:** High computational cost for large/long simulations; watch for induced dipole convergence issues.
- **Reactive FF:** Risk of charge runaway; reliability drops for reactions outside training domain.
- **ML-based FF:** Instability and energy drift possible for OOD extrapolation; always use UQ for reliability.
- **CG FF:** Not suitable for systems needing fine molecular detail (e.g., hydrogen bond networks); beware of artifacts like over-aggregation.

Recommended Baselines

- AMBER ff14SB, CHARMM36, OPLS-AA/L: For stable structure/dynamics of biomolecules or organics.
- AMOEBA, CHARMM Drude: For high-accuracy polarization effects.
- Latest ReaxFF: For chemical reaction simulation with improved energy conservation/stability.
- MACE, ANI (ML-FF): For high-accuracy systems with abundant QM data, always use with UQ.

7. Conclusion

7.1. Key Summary

This report provides an in-depth comparative analysis of four major types of molecular dynamics force fields (FFs): classical, polarizable, reactive, and machine learning-based. The analysis focused on core issues: transferability, long-range interactions, numerical stability, interpretability, and uncertainty quantification.

Classical FFs offer high computational efficiency and long-term numerical stability, but have limitations in modeling polarization and chemical reactions. Polarizable FFs explicitly treat polarization via induced dipole models, greatly improving property prediction accuracy, but at the cost of higher computational expense and convergence challenges. Reactive FFs dynamically handle bond formation and breaking, but require further improvement in transferability and stability. Recent ML-based FFs achieve high accuracy by learning from quantum mechanical data, surpassing traditional FFs in some cases, but remain challenged by OOD prediction vulnerability and lack of interpretability.

7.2. Future Research Directions

Future research in molecular dynamics force fields should focus on:

- **Hybrid ML-FF approaches:** Integrating machine learning potentials with classical or polarizable FFs to combine accuracy and efficiency, and to improve transferability across chemical space.
- **Multiscale modeling:** Developing seamless coupling between atomistic, coarse-grained, and continuum models to enable simulations of large, complex systems with both fine detail and computational tractability.
- **Uncertainty quantification:** Advancing robust UQ methods for ML-FFs to reliably detect OOD predictions and guide active learning.
- **Automated parameterization:** Creating automated, reproducible workflows for FF parameterization and validation, leveraging high-throughput QM data and open-source tools.
- **Open data and reproducibility:** Expanding public datasets, benchmarks, and open-source software to accelerate FF development and ensure reproducibility.

Recent literature highlights rapid progress in ML-based potentials, active learning, and multiscale simulation frameworks. Continued collaboration between computational chemists, data scientists, and domain experts will be essential to address current limitations and realize the full potential of next-generation force fields.

7.2. Study Limitations: This report conducted a literature review based on peer-reviewed articles published in major academic databases from 2010 to 2025. Due to specific database and time restrictions, some relevant literature may have been omitted, and there may be quantitative/qualitative biases in the literature. In addition, a comprehensive comparison of all possible systems and conditions is a complex task beyond the scope of this study, and we acknowledge that the performance of each FF may vary depending on parameterization protocols, training datasets, and simulation settings.

7.3. Future Prospects: FF development is entering a new phase with advances in ML-based FFs. Future research will focus on improving the OOD prediction reliability of ML-FFs, effectively integrating long-range interactions, and advancing UQ techniques. Hybrid FFs that combine ML-FFs with classical or polarizable FFs are emerging as promising approaches that combine the strengths of each model to enhance both accuracy and efficiency. In addition, algorithm development to improve the stability and transferability of reactive FFs will continue. The construction of high-quality QM datasets and the automation of FF parameterization/validation will play a key role in accelerating FF development.

7.4. Practical Message: The comprehensive guidelines and quantitative benchmark data provided in this report will help researchers make informed decisions in complex FF environments. During FF selection and parameterization, it is important to clearly understand the strengths and limitations of each FF type and to select the optimal FF considering the purpose of the simulation and the characteristics of the system. In addition, transparent sharing of simulation settings and parameter information is important for reproducible research. Continued efforts in FF development and validation will enhance the accuracy and reliability of molecular dynamics simulations, making essential contributions to scientific and technological innovation in materials science, chemistry, life sciences, and more.

Tab 8 <https://getliner.com/ko/search/s/26974975/t/87847720?msg-entry-type=main>

1. Classical Force Field A classical FF is a model that describes interatomic interactions using fixed functional forms and static partial charge distributions. It is mainly parameterized based on experimental data and quantum mechanics (QM) calculation results.

Advantages

- Computational efficiency: Classical FFs have simple functional forms and clear parameterization, resulting in very high computational efficiency. This enables long simulations at the microsecond (μ s) scale for large systems with millions of atoms.

- High numerical stability: Thanks to fixed parameters and simple functional forms, energy drift is relatively small in long MD simulations, and trajectories remain stable.

- Applicability to various systems: AMBER, CHARMM, OPLS, and GROMOS are widely used for simulations of various biomolecular and organic molecular systems, including proteins, nucleic acids, carbohydrates, lipids, and organic molecules.

- High interpretability: Parameters have physical meaning such as bond length, bond angle, and torsion angle, so the model's operating principle is transparent and it is relatively easy to trace the cause of prediction failures.

Disadvantages

- No polarization effect: Classical FFs use a fixed charge model and cannot explicitly account for the dynamic polarization of electrons due to environmental changes. This limits accuracy, especially in systems where polarization is important, such as polar solvents, ionic liquids, and metal-protein interactions.

- Cannot model chemical reactions: The system topology must be fixed at the start of the simulation, so chemical reactions such as bond formation or breaking cannot be modeled.

- Limited transferability: While high transferability is shown within the trained chemical space, accuracy can drop sharply in new chemical environments outside the parameterization range (Out-of-Distribution, OOD).

- Low absolute cooperativity reproduction: In studies of protein secondary structure

formation or self-assembly systems, classical FFs cannot accurately describe absolute cooperativity due to the lack of polarization effects.

Evaluation Criteria

- Accuracy: Evaluate how well predicted physicochemical properties (e.g., liquid density, solvation free energy, protein structure RMSD) match experimental or high-level QM calculation results.

- Computational efficiency: Evaluate whether large-scale and long-time simulations are possible and the number of atoms that can be processed per simulation time.

- Transferability: Evaluate whether the FF provides stable and reliable predictions for new molecular systems or environments not included in training.

- Interpretability: Evaluate whether FF parameters have clear physical meaning and whether it is easy to analyze the cause of simulation results and trace errors.

- Scope of application: Evaluate the limitations of the ability to handle specific phenomena such as chemical reaction modeling and polarization effects.

2. Polarizable Force Fields A polarizable FF explicitly includes polarization effects, i.e., deformation of the electron cloud due to environmental changes. It is implemented using induced dipole (Drude oscillator model) or multipole (AMOEBA) models.

Advantages

- Improved accuracy: By explicitly modeling polarization effects, polarizable FFs significantly improve accuracy over classical FFs in predicting polar environments, ion-molecule interactions, and protein-ligand binding free energies. The AMOEBA FF reproduces the dielectric constant of water with 2-3

- High agreement with experimental and QM data: Polarizable FFs reproduce not only macroscopic properties such as the dielectric constant of water, but also microscopic structures and dynamic properties of various systems such as ion solutions and organic liquids, with high agreement with experimental data.

- Realistic interaction modeling: Polarizable FFs are essential for predicting transport properties of ionic liquids (ILs) and electrolytes, and are also needed to properly describe protein structures. Polarization effects accurately describe cooperativity, which is limited in fixed charge FFs.

Disadvantages

- High computational cost: To account for polarization effects, induced dipole calculations are performed iteratively using a self-consistent field (SCF) method, which significantly increases computational cost compared to classical FFs. Induced dipole calculations in AMOEBA FF can account for about 50

- Numerical stability issues: Slow convergence of SCF calculations or unphysical phenomena such as charge runaway can compromise simulation stability. However, improving the 'always stable predictor-corrector (ASPC)' method can reduce energy drift and lower computational cost by 20

- Complex parameterization: Additional parameters such as polarizability are required, and their optimization can be more complex and time-consuming than for classical FFs.

Evaluation Criteria

- Ability to account for polarization effects: Evaluate how important electronic polarization is in determining the key physicochemical properties of the system.

- Computational cost: Evaluate whether the additional computational resources and time required for polarization effect modeling are manageable within available resources.

- Numerical stability: Evaluate whether energy drift is within acceptable limits in long MD simulations and whether unphysical phenomena do not occur.

- Agreement with experiment and QM: Evaluate how accurately predicted properties such as the dielectric constant of water, solvation free energy, and protein-ligand binding free energy match experimental or high-level QM results.

3. Reactive Force Fields A reactive FF is a FF that can dynamically handle the formation and breaking of chemical bonds. It is implemented through potential energy functions that depend on bond order and dynamic charge redistribution models (charge equilibration, QEq). ReaxFF is the most representative reactive FF.

Advantages

- Ability to model chemical reactions: Can dynamically model the breaking and formation (atom transfer) of chemical bonds during simulation. This enables the study of complex chemical reaction phenomena such as combustion, catalytic reactions, material fracture and deformation, and battery interface reactions.

- Transferability: ReaxFF uses element-specific parameters and has considerable transferability across various chemical environments and phases (solid, liquid, gas).

- Computational efficiency: Much more efficient than quantum mechanics (QM) calculations, while being able to handle chemical reactions, making it applicable to large systems or long reaction simulations that are impossible with QM.

Disadvantages

- Limited transferability and accuracy: Transferability is limited to specific reaction types or trained databases, and accuracy can drop sharply in OOD environments. There have been reports of amino acid torsion angle errors exceeding 100° in torsion barrier predictions.

- Numerical stability issues: Early versions of ReaxFF had problems with instability and poor energy conservation due to instability in charge equilibration (QEq) calculations ('charge runaway'). However, recent algorithm improvements have enhanced stability.

- High parameterization complexity: While the bond order model is intuitive, the number of parameters is very large and highly interdependent, making parameterization complex and difficult.

Evaluation Criteria

- Chemical reaction prediction: Evaluate how accurately and physically plausibly actual reaction pathways, transition states, and activation energies are predicted.

- Transferability: Evaluate whether stable and reasonable predictions are provided for other chemical systems or reaction conditions outside the parameterized system.

- Numerical stability: Evaluate whether the total energy of the system is well conserved during long MD simulations and whether trajectories remain stable without unphysical phenomena (e.g., charge runaway).

- Computational efficiency: Evaluate the cost savings compared to QM simulations and whether the speed is suitable for the target system and time scale.

- Parameterization difficulty: Evaluate the complexity of the parameter set, the effort required for optimization, and the dependence on experts.

4. Machine Learning Force Fields (ML-FF) ML-FFs are FFs that learn quantum mechanics (QM) calculation data to construct potential energy as a nonlinear model (e.g., neural network). NequIP, MACE, ANI, and SchNet are representative examples.

Advantages

- High QM-level accuracy: ML-FFs learn energy and force data from QM calculations (e.g., DFT) and achieve much higher accuracy than classical FFs. They record low mean absolute errors (MAE) of 1-5 meV/atom in energy and force predictions, and the ANI-1ccx model showed the best performance with an MAE of 0.8 kcal/mol in torsion barrier

predictions for specific organic molecules.

- Ability to model complex chemical environments: ML-FFs can effectively model complex chemical environments (e.g., multiple bonds, transition metals, chemical reactions) that are difficult for conventional empirical FFs.

- Data-driven flexibility: By learning from large QM datasets, ML-FFs can be applied to various atomic combinations and structures, and training datasets can be efficiently expanded through active learning, etc.

- Potential integration of long-range interactions: There are ongoing attempts to improve the accuracy of long-range interactions by combining traditional PME/Ewald summation with ML-FFs or by estimating atomic partial charges with ML and integrating with existing Ewald methods.

Disadvantages

- OOD (Out-of-Distribution) prediction vulnerability: ML-FFs have high prediction uncertainty or rapidly deteriorating prediction quality for new chemical environments outside the training data distribution. In a 500 K simulation of the MAPbI₃ system, 100

- Limited interpretability ("black box"): ML-FFs are often considered "black box" models, making it difficult to intuitively interpret the internal operating principles or the physical meaning of parameters. This makes it complex to analyze and improve the causes of prediction errors.

- High training data and computational resource requirements: Achieving high accuracy requires large, high-quality QM datasets, and model training and inference may require significant computational resources.

- Numerical stability issues: Related to OOD prediction vulnerability, there is a risk of incomplete energy conservation or unphysical trajectory collapse (e.g., "holes" phenomenon) during long MD simulations.

5. Coarse-Grained Force Fields (CG-FF) A coarse-grained FF is a model that simplifies multiple atoms into a single "bead" to reduce the degrees of freedom of the system and maximize computational efficiency.

Advantages

- Dramatic sampling efficiency: Compared to atomistic models, the degrees of freedom are greatly reduced, resulting in a dramatic increase in simulation speed. This extends the time scale by several orders of magnitude, enabling simulations of large molecular systems with hundreds of thousands of molecules for timescales of several milliseconds or more.

- Applicability to large systems: Very useful for studying macroscopic properties and long-term dynamics of large-scale systems such as biomacromolecules (proteins, lipid membranes), ionic liquids, and polymer nanocomposites, which are difficult to approach with atomistic FFs.

- Reduced computational cost: The number of particles and the amount of interaction calculations to be processed are significantly reduced, enabling large-scale simulations even in environments with limited computing resources.

Disadvantages

- Reduced structural accuracy: Due to high dimensionality reduction and simplification, it is difficult to fully represent fine structural and electronic interactions. This means that the reflection of specific atomic interactions (e.g., hydrogen bond networks) or polarization effects is limited.

- Information loss and artifact occurrence: Sacrificing atomic-level details smooths the system's potential energy surface and reduces molecular friction, but can lead to

unphysical structural artifacts such as over-aggregation or underestimation of melting points. Models based on liquid-liquid partitioning have shown freezing and compression of polymer chains near the surface.

- Limited chemical specificity: Not suitable for studies requiring precise atomic-level information such as specific chemical reactions, electron potential distributions, or sophisticated electrodynamic phenomena.

- Transferability and parameterization complexity: The process of constructing an effective Hamiltonian and determining parameters is complex, and there may be limitations in transferability in OOD situations.

Force Field Selection Guidelines for Drug Discovery and Catalysis Research

1. Drug Discovery In drug discovery, force fields that can accurately model the interactions between biomolecules and small molecules are advantageous. CHARMM is useful for studying drug-membrane protein interactions because it can stably simulate protein-lipid membrane complexes, and AMBER is the de facto standard for protein-ligand binding studies because it excellently reproduces the structural stability of protein backbones and side chains. OPLS is optimized for organic molecules and solvent systems and is widely used for free energy calculations of drug candidates and evaluation of drug-protein binding affinity. In addition, ANI, a machine learning-based FF, maintains DFT-level accuracy and can rapidly explore small molecule reaction pathways or binding free energies, greatly accelerating drug discovery.

2. Catalysis Research In catalysis research, force fields capable of charge redistribution, metal-ligand interactions, and reaction pathway tracking are important. AMOEBA implements multipoles and explicit polarization, showing high accuracy in studies of metal complex-ligand binding, catalyst surface charge distribution, and intermediate stabilization. ANI is limited for transition metal catalysts but is much more precise than classical FFs for organic catalytic reactions or small molecule activation pathway exploration. ReaxFF, thanks to its continuous bond order, can spontaneously capture bond formation and breaking and charge transfer, and is widely used in studies of combustion catalysis, surface catalytic reactions, and redox mechanisms.

3. Battery Research In battery research, force fields capable of accurately describing physicochemical processes such as ion conduction, electrolyte stability, and interfacial reactions are needed. PCFF is suitable for studies of polymer electrolytes and solid electrolytes because it describes the thermodynamic and mechanical behavior of polymers and amorphous materials well. DeepMD provides QM-level accuracy even for systems with hundreds of thousands of atoms, making it powerful for studying macroscopic reactions at metal, ionic liquid, and electrode-electrolyte interfaces. ReaxFF, through bond order and dynamic charge transfer, can handle electrode-electrolyte interfacial reactions, SEI formation, and electrolyte decomposition, making it a representative tool for electrochemical stability research.

Final Notes and Credits

This supporting information and the main paper were organized and refined using **GitHub Copilot**, **Linear AI**, and **Chatgpt 4.1**, **google gemini** for educational purposes. The translation of the Korean prompts and manuscript was also performed using VS Code Copilot. Any errors in the document represent the errors AI actually produced.

The following ChatGPT share links contributed to the final summary and content:

- <https://chatgpt.com/share/68c2f4a6-1d90-800d-9c35-79a8ccd3c84d>
- <https://chatgpt.com/share/68c2f48a-9270-800d-83e1-4e82da17d336>
- <https://chatgpt.com/share/68c2f224-aca4-800d-a8bb-417383fea98a>
- <https://chatgpt.com/share/68c2f48a-9270-800d-83e1-4e82da17d336>
- <https://chatgpt.com/share/68c301d8-332c-800d-9323-149c217901f2>
- <https://chatgpt.com/share/68c301b4-c4b8-800d-aded-35c3950962de>
- <https://chatgpt.com/share/68c301a2-3e24-800d-972d-b11c884bef96>
- <https://chatgpt.com/share/68c304fc-d6a4-800d-ae10-69e88f1c2911>
- <https://chatgpt.com/share/68c3fcad-a320-800d-b9ab-dd7c4f2f71bf>
- <https://chatgpt.com/share/68c380c8-d64c-800d-9f21-f7e09ba57e6d>
- <https://chatgpt.com/share/68c3afbe-27d4-800d-86cf-f13464b969d5>
- <https://chatgpt.com/share/68c3f957-489c-800d-8208-11868fce2669>
- <https://chatgpt.com/share/68c3fac8-508c-800d-a078-ba22568e4eb9>
- <https://chatgpt.com/share/68c3fd36-c528-800d-89b7-ea746b9c9225>