



A review of the theoretical atomistic modelling of amorphous carbon in the artificial intelligence era

Zhang Shengbin^{1,2}, Wang Yilin^{1,2}, Wang Jiang^{1,2}, Li Xiaoming¹,
Chen Jingpeng¹, Yi Zonglin^{1,*}, Su Fangyuan^{1,*}

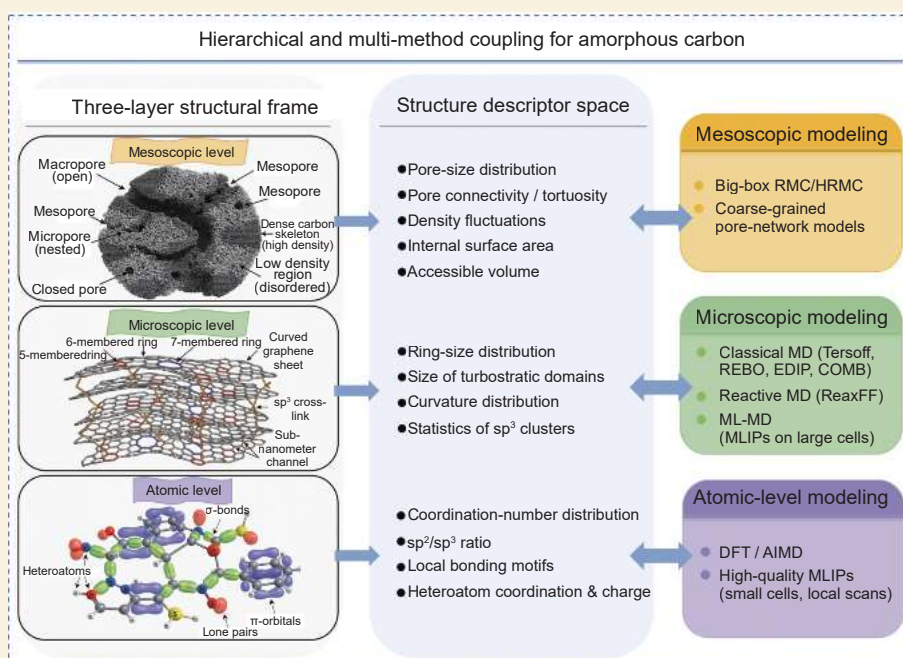
(1. Shanxi Key Laboratory of Carbon Materials, Institute of Coal Chemistry, Chinese Academy of Sciences, Taiyuan 030001, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Amorphous carbons are metastable materials with diverse local bonding and pore structures that strongly influence their properties. Because diffraction and spectroscopy provide averaged, low-dimensional structural information, constructing realistic atomistic models remains challenging. Methods for modelling amorphous carbon are summarised, from geometric and topological models to density functional theory, classical and reactive molecular dynamics, machine-learning interatomic potentials, reverse Monte Carlo and hybrid re-

verse Monte Carlo methods. These approaches are compared in terms of accuracy, system size, simulation time, and consistency with experimental data. We also describe amorphous-carbon structures spanning atomic bonding, microscopic organization and mesoscopic pore structure, and outline approaches that combine simulation with experimental constraints. Finally, we discuss how curated structure–process–property databases and AI-assisted methods can support reusable models and the rational design of amorphous carbon materials.

Key words: Amorphous carbon; Atomistic modelling; Molecular dynamics; Machine-learning interatomic potentials; Reverse Monte Carlo



1 Introduction

Amorphous carbon (a-C) denotes a broad class of metastable carbon materials that lack long-range translational symmetry but retain short-range bonding order dominated by sp^2 and sp^3 hybridizations^[1–4]. Depending on the sp^2/sp^3 ratio, hydrogen content and porosity, a-C includes diamond-like carbon (DLC), hydrogenated amorphous carbon (a-C:H), tetrahedral amorphous carbon (ta-C), hard carbon and porous carbon. High- sp^3 DLC and ta-C exhibit high hardness,

low friction and chemical stability, and are used in protective coatings, magnetic storage and biomedical surfaces^[5–7]. In contrast, sp^2 -rich hard carbons and porous carbons contain turbostratic nanodomains and pore structures that make them important electrode

Received: April 30, 2026

Revised: July 23, 2026

Accepted: July 24, 2026



materials for sodium-ion batteries and supercapacitors^[8–11]. However, the very structural diversity that underpins such extensive utility also imposes immense challenges on the precise resolution of their three-dimensional atomic configurations^[12]. Although this review mainly discusses amorphous-carbon modelling in the context of energy storage, especially hard-carbon and porous-carbon electrodes, the underlying atomistic modelling strategies are transferable to broader a-C systems, including DLC coatings, biomedical carbon surfaces, and tribological films.

The structural diversity of amorphous carbon originates from the competition between planar sp^2 bonding and tetrahedral sp^3 bonding^[2,13–14]. These bonding preferences cannot be fully satisfied within a continuous disordered network, leading to topological frustration. As a result, amorphous carbon preserves local bonding rules but lacks long-range periodicity. Its medium-range order is instead expressed through distorted graphitic clusters, non-hexagonal rings, coordination defects and curved carbon networks^[1,15].

Robertson's ternary phase diagram provides a useful framework for classifying amorphous carbons according to sp^2 fraction, sp^3 fraction and hydrogen content (Fig. 1)^[1,16]. In the hydrogen-free sp^2 -rich region, the structure approaches graphite-like amorphous carbon and glassy carbon, where distorted nanographitic domains are connected by defects such as five- and seven-membered rings^[17]. With increasing hydrogen content, H terminates and disrupts the C–C network, producing a-C:H or DLC with sp^2 clusters embedded in a hydrogen-containing sp^3 -rich matrix^[16]. At very high hydrogen content, the network becomes polymer-like. In the hydrogen-free sp^3 -rich region, dense tetrahedral networks form ta-C, whose density and hardness can approach those of diamond^[18]. Hard carbon lies within the sp^2 -rich domain but is distinguished by its “house-of-cards” texture and pore-containing mesostructure, which are critical for ion storage^[19–22]. Thus, amorphous carbon should be viewed as a family of metastable structures distributed across a broad bonding and compositional space rather than as a single phase.

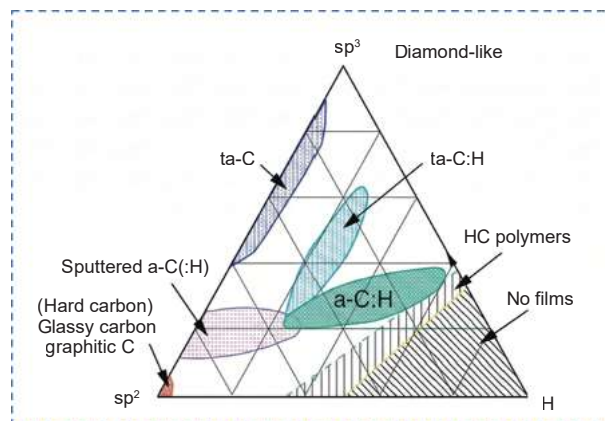


Fig. 1 Amorphous carbon materials' triangular phase diagram^[1].

(Copyright 2002, Elsevier)

Despite this classification, determining the three-dimensional atomic structure of amorphous carbon remains difficult. Raman spectroscopy, X-ray photoelectron spectroscopy, and X-ray/neutron total scattering provide averaged information on hybridization, cluster size, pair correlations and bonding environments, but they do not directly determine the full atomic connectivity of the carbon network^[12,15,23–26]. Therefore, reconstructing atomistic structures from experimental data is an ill-posed inverse problem: different atomic configurations may reproduce similar spectra or scattering profiles. Multimodal characterization can narrow the possible structural space, but it usually cannot yield a unique three-dimensional model.

Atomistic modelling addresses this limitation by connecting experimentally accessible signals with atomic configurations that cannot be directly resolved^[24,27–28]. When constrained and validated by experimental data, modelling can help identify local motifs, pore environments and defect structures responsible for measured properties. This is particularly important for energy-storage carbons, where target properties such as sodium-storage capacity, rate capability and ion-accessible pore size depend on both atomic bonding and pore-network architecture^[29–31]. Atomistic models therefore provide a basis for establishing structure–property relationships and for guiding the design of amorphous carbon materials beyond empirical trial-and-error.

This review summarizes the development of

amorphous-carbon modelling methods from geometric and topological models to density functional theory (DFT), classical and reactive molecular dynamics (MD), machine-learning interatomic potentials (MLIPs), reverse Monte Carlo (RMC) and hybrid reverse Monte Carlo (HRMC). These methods are compared in terms of accuracy, accessible length and time scales, and agreement with experimental constraints. We further discuss hierarchical structural descriptions and data-driven workflows for linking processing, structure and properties in amorphous carbon. Through this systematic review and critical analysis, we propose future directions for the advancement of the field, aspiring for this work to serve as a powerful tool for researchers to both deeply understand and rationally design amorphous carbon materials.

2 Early explorations: geometric constraints and topological construction

Before high-performance computing enabled dynamic simulations, atomistic models of amorphous carbon were mainly constructed from geometric intuition and topological constraints^[32–33]. These early approaches addressed a basic structural question: how can atoms be arranged into a disordered but chemically plausible network using only local bonding rules. Although models such as continuous random networks (CRNs) and the Polk model lacked explicit thermodynamic and kinetic descriptions, they introduced structural descriptors—coordination defects, bond-angle distributions and ring statistics—that remain widely used in amorphous-carbon modelling^[34–35].

2.1 From random packing to continuous random networks

Early on, researchers adopted an intuitive idea: to approximate amorphous solids as a Random Close Packing (RCP) of atoms^[36–37]. In this approach, carbon atoms are placed randomly in space, and bonds are assigned when interatomic distances fall below a prescribed cutoff, producing a three-dimensional (3D) network without long-range order^[38]. However, purely random packing neglects carbon bonding chemistry.

The resulting structures often contain unrealistic coordination environments and bond-angle distributions that deviate from those expected for sp^2 or sp^3 hybridization^[39]. They also show poor agreement with experimental density, radial distribution functions $g(r)$, and structure factors $S(Q)$ ^[40].

In this context, Zachariasen's continuous random network (CRN) theory for oxide glasses became a crucial conceptual breakthrough^[32]. In this theory, long-range periodicity is absent, but local structural units follow crystal-chemical rules in a statistical sense^[41–43]. This idea was later extended to amorphous semiconductors and amorphous carbon. A typical CRN model imposes a specified coordination number, restricts nearest-neighbor distances, and then relaxes the geometry so that calculated structural descriptors, especially $g(r)$, approach experimental data^[34,44–45].

Compared with RCP models, CRNs provide a more realistic description of coordination-number distributions, density and pair correlations^[46,47]. They changed the picture of amorphous carbon from a random atomic aggregate to a continuous network constrained by local valence rules. Medium-range disorder in such models arises from the topological diversity that remains after local bonding constraints are enforced. Subsequent, more refined geometric models can essentially be seen as attempts to locally modify and refine this baseline network.

2.2 Geometric constraint models: the polk model

Within the CRN framework, Polk's work on amorphous Ge/Si represents a peak of physically motivated structural construction. Building on Zachariasen's random network concept, Polk used a physical ball-and-stick model to manually construct a tetrahedral network of about 440 atoms. During construction, strict constraints were imposed: bond-length deviations were kept below 1%, bond-angle deviations within $\pm 10^\circ$, and dangling bonds were eliminated as far as possible^[34,48]. Although this approach relied heavily on intuition, the resulting model exhibited striking physical plausibility.

Based on this model, Polk and co-workers performed systematic topological analyses and found that the network was primarily composed of 5- and 6-

membered rings, with an approximate ratio of 1 : 4. This characteristic later became known as the “Polk ring statistics”^[44–45]. Subsequent studies (e.g., by Walrafen, Weaire and others) introduced simplified elastic energy terms and performed force-field relaxations on the Polk model, further smoothing local stress distributions and significantly improving the agreement of $g(r)$ with experimental data^[44–45]. These studies used coordination-number distributions, bond-length and bond-angle statistics, and ring-size distributions to characterize short- and medium-range structures^[39].

Polk-type models are important because they established a practical set of structural descriptors for disordered covalent networks. These descriptors, especially coordination and ring statistics, are still used to evaluate amorphous-carbon models generated by DFT, MD, MLIPs and RMC-based methods.

2.3 The Wooten–Winer–Weaire algorithm: geometric cost functions and bond switching

The Wooten–Winer–Weaire (WWW) algorithm extended CRN-based modelling from manual construction to computer-automated network optimization. Wooten and co-workers proposed a Monte Carlo method based on local bond-switching operations and generated near-perfect amorphous semiconductor networks^[35].

The key feature of the WWW algorithm is a geometric penalty function. This function quantifies bond-length and bond-angle distortions as deviations from ideal values. Although it is not a physical potential, it acts as a pseudo-potential for ranking candidate structures. During the simulation, local bonds are switched, for example by replacing A–B and C–D bonds with A–C and B–D bonds. The new topology is accepted or rejected using a Metropolis criterion based on the change in the penalty function^[35,49]. Repeated bond switching reduces geometric strain while largely preserving the target coordination.

The WWW algorithm was significant because it introduced a quantitative optimization criterion into topological network construction. Its strategy—random topology updates guided by a cost function—an-

ticipated later structure-generation methods based on physical potential-energy surfaces, including MD and machine-learning-assisted searches.

2.4 Geometric skeleton models for hard carbons and porous carbons

CRN, Polk and WWW models are most suitable for dense amorphous networks such as high-density a-C and ta-C. Hard carbons, glassy carbons and porous carbons require additional structural descriptions because their key features include turbostratic stacking, curved graphitic fragments and mesoscopic pores. Experimental studies indicate that these materials consist of distorted nanographitic domains whose random stacking and three-dimensional mismatch generate curved skeletons and non-hexagonal defects^[17,50]. In-situ TEM also shows that during the conversion of polymer precursors to glassy carbon, the formation, expansion, and rearrangement of nano-scale graphene sheets are central to microstructural evolution^[51–52].

Several geometric skeleton models have been developed to describe these features. One approach represents hard or glassy carbon as a “house-of-cards” assembly of finite graphene fragments. These fragments are treated as rigid or semi-rigid units that are randomly oriented, stacked and cross-linked in space. Such models naturally generate irregular interlayer voids and connected pore channels, and can reproduce anisotropy and pore-size distributions observed by neutron diffraction and SAXS^[50,53].

Another approach describes polymer-to-carbon conversion using simplified geometric evolution rules. Starting from an organic precursor network, operations such as volatilization, chain shrinkage and cross-linking densification are applied to create pores and carbon-rich frameworks^[54]. Although these models do not include explicit reaction kinetics, they can generate mesoporous topologies and closed/open pore ratios that qualitatively agree with experimental observations in hard carbons^[55]. Models based on fullerene-like shells or Schwarzschild minimal surfaces have also been used to explain non-hexagonal rings and negative-curvature connections in porous carbon networks^[17,56].

Overall, this generation of geometric skeleton models served as “structural hypothesis generators” in amorphous carbon research. At a time when accurate descriptions of interatomic interactions were unavailable, they captured key mesoscopic features of the material. While they cannot answer detailed questions about electronic structure or chemical reaction mechanisms, they furnished crucial initial frameworks and topological constraints for later, more refined modelling based on realistic potential energy surfaces.

3 Density functional theory: the accuracy benchmark for modelling

In the history of modelling amorphous carbon, the introduction of density functional theory (DFT) marks a fundamental paradigm shift from geometric rules to a quantum-mechanical description. Unlike the CRN or Polk models that rely on empirical geometric parameters, DFT solves the Kohn–Sham equations and, within a unified quantum-mechanical framework, yields the total energy, atomic forces, and electronic structure of the system. This allows not only accurate determination of atomic positions, but also quantitative analysis of bond strengths, local hybridization, and electronic density of states^[57–58]. Despite its scale limitations, DFT thus provides the accuracy benchmark that all subsequent empirical and machine-learning potentials must respect.

3.1 DFT-based structure generation and local chemical environments

For intrinsic amorphous carbon, *ab initio* melt–quench (AIMQ) is a common DFT-based protocol for generating energetically consistent structures. In this method, a finite carbon supercell is heated to a liquid-like state, equilibrated, quenched and then relaxed on the DFT potential-energy surface. The resulting structure represents a metastable configuration trapped from the high-temperature liquid state.

Using this approach, Bauschlicher and co-workers studied bulk amorphous carbon and its surfaces over a range of densities, generating models from low-density graphite-like carbon to high-density diamond-like carbon^[59]. These studies showed that dens-

ity strongly controls the sp^2/sp^3 ratio. Similar AIMQ strategies have been used to examine high-pressure amorphous phases. For example, simulations of quenchable amorphous diamond showed that quenching high-pressure liquid carbon followed by relaxation at ambient pressure can produce a disordered network with nearly complete fourfold coordination and a diamond-like density^[60]. These results support the existence of dense sp^3 -rich amorphous carbon without long-range crystallinity.

In intermediate-density amorphous carbon, DFT models are useful for quantifying the relation between hybridization and electronic properties. By constructing models with different sp^3 fractions and calculating band structures, density of states and optical constants, DFT studies have clarified how mixed sp^2/sp^3 bonding affects electronic and optical behavior^[61] (Fig. 2). These calculations provide atomistic links among density, bonding configuration and electronic structure.

DFT-based molecular dynamics has also been used to study structural evolution during carbon-network formation. Such simulations show that 5- and 7-membered rings can form naturally during liquid cooling and help release local strain^[62]. In doped amorphous carbons, DFT further identifies the preferred incorporation sites and coordination environments of heteroatoms such as N and S, and explains how they modify electronic states near the Fermi level^[63–64].

Overall, DFT provides atomically resolved information on density-dependent hybridization, coordination, ring statistics, heteroatom environments and electronic structure. Compared with purely geometric models, DFT-generated structures are not only topologically plausible but also energetically consistent at the quantum-mechanical level.

3.2 Linking DFT models to experimental characterization

Because experimental spectra and scattering data do not uniquely determine 3D atomic structures, DFT is often used for forward modelling: a proposed atomic model is used to calculate measurable signals, which are then compared with experiment. This

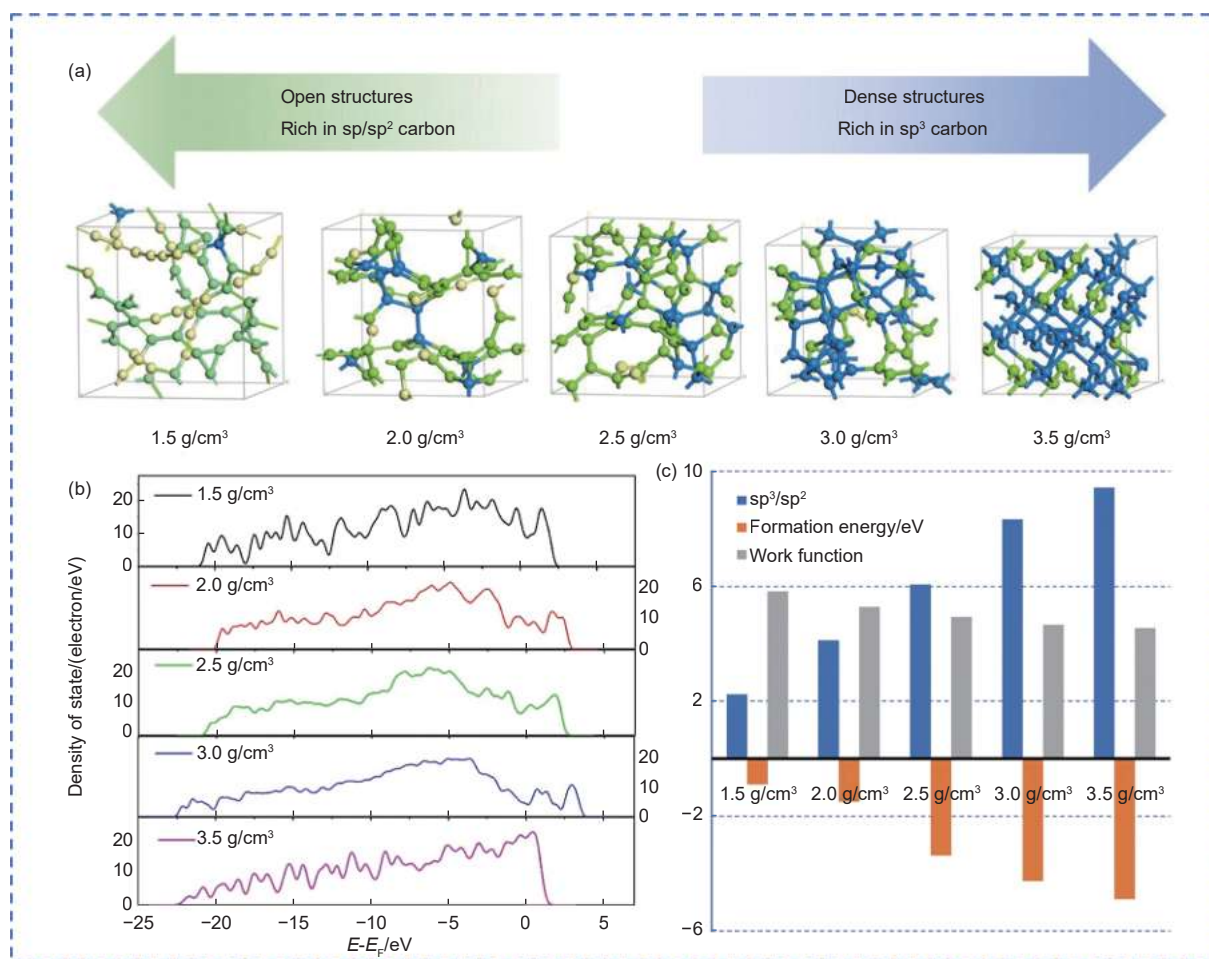


Fig. 2 DFT links between density, bonding hybridization, and electronic properties in amorphous carbon^[61]. (a) Ab initio melt-quench A-C models ($\rho = 1.5\text{--}3.5\text{ g cm}^{-3}$) relaxed by DFT; atoms colored by local hybridization (sp, sp^2 , sp^3). (b) DFT total density of states (DOS) for the corresponding models. (c) DFT-computed sp^3/sp^2 ratio, formation energy and work function as a function of density. (Copyright 2019, Springer Nature)

strategy helps connect local atomic environments with spectroscopic and electrochemical observables.

In spectroscopic analysis, DFT has been used to relate optical gaps, sp^2 cluster size and sp^3 content^[61,65]. Walter and co-workers calculated C 1s binding energies for amorphous-carbon environments with different sp^2/sp^3 ratios, bond-angle distortions and hydrogen terminations^[57]. They found that local-environment variations in disordered carbon can produce C 1s shifts exceeding 1 eV. This helps explain peak broadening and shifts in experimental XPS spectra and indicates that simple multi-peak fitting may not fully capture structural disorder in amorphous carbon.

DFT has also been combined with machine learning to interpret X-ray absorption spectra. Aarva et al. generated local structural motifs and correspond-

ing carbon K-edge spectral fingerprints from DFT-level amorphous-carbon models^[66] (Fig. 3a and b). Machine-learning models were then trained to map structural descriptors to spectral shapes. This example shows that DFT can act both as a spectral simulator and as a data source for learning structure-spectrum relationships.

In electrochemical systems, DFT is widely used to analyze sodium-storage mechanisms in hard carbons^[67] (Fig. 3c-e). Because real hard carbons contain complex turbostratic domains and hierarchical pores, DFT studies usually use simplified but physically meaningful models, such as bilayer or multilayer graphitic fragments, oxygen-containing surfaces, interlayer regions and defect sites. Au and co-workers calculated Na adsorption energies at defects, interlayer sites and pore walls, providing insight into possible

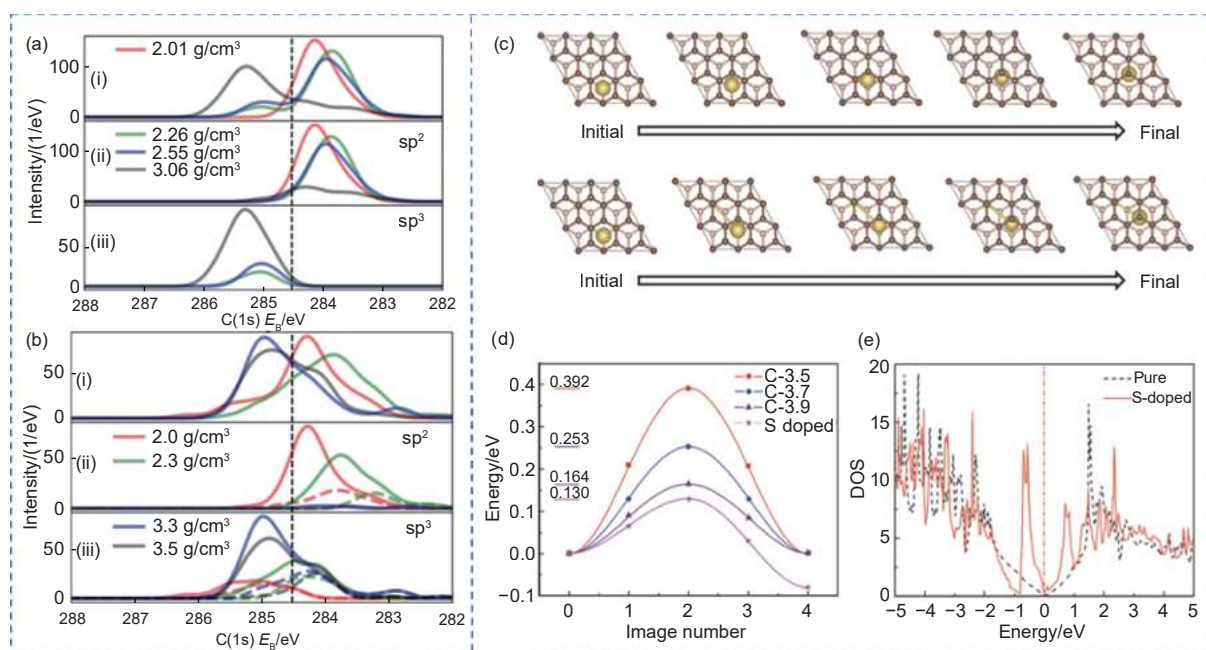


Fig. 3 DFT-based interpretation of C(1s) binding-energy shifts and Na-storage mechanism in S-doped hard carbon. (a,b) Simulated C(1s) spectra for hydrogen-free a-C and hydrogenated a-C:H, obtained from site-resolved C(1s) binding energies with Gaussian broadening^[57]. (Copyright 2025 Phys. Rev. Materials) (c) Minimum-energy Na⁺ migration pathways along interlayer voids in pristine and S-doped hard carbon^[64]. (d) Na⁺ diffusion barriers versus interlayer spacing and S-doping state^[64]. (e) DFT density of states (DOS) of pristine and S-doped hard carbon^[64]. (Copyright 2018, WILEY-VCH)

Na-storage locations^[29]. Jia and co-workers used DFT and ab initio MD to examine Na diffusion in curved and defective graphene models, showing that curvature and vacancies increase adsorption energies and can convert free 2D diffusion into restricted or trapping behavior^[68–69]. In S-doped hard carbons, DFT further reveals that appropriate sulfur doping and interlayer expansion can both lower Na diffusion barriers and enhance electronic conductivity via increased density of states near the Fermi level, consistent with experimentally observed improvements in rate capability and cycle life^[64].

Taken together, these examples illustrate DFT's bridging role among structure, spectroscopy, and performance. Given a structural model, it can simultaneously provide energies, electronic states, and spectral signatures, enabling amorphous carbon modelling to move beyond “drawing a plausible atomic picture” to an iterative process in which models are repeatedly calibrated and constrained by comparison with experimental spectra, thereby gradually approaching the behavior of real materials.

3.3 The accuracy–scale contradiction

The main limitation of DFT is computational

cost. Typical DFT models of amorphous carbon contain about 100 atoms and rarely exceed a few hundred atoms, while ab initio molecular dynamics is usually limited to tens of picoseconds. These scales are sufficient for local bonding rearrangements, short-range order and dopant effects, but they are too small to represent mesoscopic pore networks, long-range defect connectivity or stress–porosity coupling over tens of nanometers.

This limitation is especially important for hard carbons and porous carbons. In these materials, sodium storage, ion confinement and transport are controlled not only by local adsorption sites but also by the statistical distribution and connectivity of pores, interlayers and graphitic domains. DFT therefore often reduces the real structure to idealized motifs, such as slit-like interlayers, pore mouths, surfaces or isolated defects^[64]. These models are useful for identifying local mechanisms, but they cannot directly determine the abundance, spatial correlation or connectivity of these motifs in real disordered carbons. Similarly, DFT cannot readily access slow processes such as extended cooling, large-scale phase separation, graphitic-domain growth or gas-release-driven

pore formation, which occur over much longer time and length scales^[23,49,59].

For this reason, the role of DFT in multiscale modelling has shifted from direct large-scale structure generation to reference-data generation and validation. DFT calculations are widely used to calibrate empirical potentials. For example, Pastewka and co-workers refined Brenner-type bond-order potentials by matching selected energies and forces to DFT results, improving the description of sp^2/sp^3 balance, ring statistics and fracture behavior^[70]. More recently, machine-learning interatomic potentials for amorphous carbon, including GAP and neural-network potentials, have been trained on DFT data sets covering liquid carbon, low-density graphite-like structures, high-density diamond-like structures, defects and heteroatom-containing configurations^[66,71–73].

In sum, DFT occupies an irreplaceable foundational position in amorphous carbon modelling. It should not be regarded as merely a high-accuracy but impractical tool constrained by scale, but rather as the most basic and crucial ground-truth provider in the entire modelling chain. Only by clarifying, at the DFT level, the fundamental physics of hybridization competition and electronic-state evolution can subsequent large-scale simulations rest on a firm physical basis.

4 Molecular dynamics: the evolution of interatomic potentials

Density functional theory provides accurate local bonding and electronic information, but its accessible system size and time scale are limited. Molecular dynamics (MD) extends amorphous-carbon simula-

tions to larger systems and longer trajectories by integrating classical equations of motion with an interatomic potential^[74–75]. For amorphous carbon, the reliability of MD is therefore mainly determined by the chosen potential rather than by the integration algorithm itself^[58,76].

Interatomic potentials for carbon can be broadly divided into three categories. The first includes classical empirical potentials, such as Tersoff, REBO/AIREBO, EDIP, LCBOP, ABOP and COMB, which are efficient and widely used for structural and mechanical simulations. The second is represented by ReaxFF, which introduces bond-order-dependent reactivity and charge equilibration, enabling simulations of bond breaking, bond formation, gas evolution and functional-group transformation. The third category consists of machine-learning interatomic potentials (MLIPs), such as GAP and neural-network-based potentials, which are trained on DFT data to approach first-principles accuracy at much lower computational cost.

4.1 Classical MD: empirical potentials and melt–quench modeling

Classical MD uses analytic empirical potentials to approximate interatomic interactions and simulate structural evolution at relatively low computational cost (Table 1). To capture the strongly covalent, multi-hybrid, and environment-sensitive behavior of carbon, potential development has primarily followed two ideas: bond-order potentials and environment-dependent potentials.

Tersoff's bond-order potential is one of the starting points for classical potentials in carbon^[77]. By in-

Table 1 Classical interatomic potentials employed in simulations of amorphous carbon

Potential	Functional type	Chemical space	Cost	Ref.
Tersoff family	Short-range bond-order potential with 2- and 3-body terms	Mainly C, Si, C–Si	Low	[76]
Brenner REBO	Reactive empirical bond-order (REBO)	C/H hydrocarbons	Low	[82]
AIREBO	REBO + non-bonded (LJ or Morse) + torsional terms	C/H (some C/O/H variants)	Moderate	[83,87]
EDIP-C	Environment-dependent bond-order potential	Pure C	Low–moderate	[88]
LCBOP	Long-range reactive bond-order potential	Pure C (diamond, graphite, liquid carbon)	Moderate–high	[89]
ABOP	Analytic bond-order potential framework	C and selected carbides (TiC, SiC, etc.)	Low–moderate	[90,91]
ReaxFF	Reactive many-body force field with bond-order and dynamic charges	Broad C/H/O/N and multicomponent systems	High	[92,93]
COMB3	Charge-optimized many-body potential	Many elements including C/O/H and metals	High	[94]

roducing a coordination-dependent bond-order term, it describes energy changes associated with different local environments, including crystalline, defective and disordered structures^[78–79]. Early simulations using the Tersoff potential showed the influence of ion energy on amorphous-carbon film formation, although its description of metastable amorphous structures remained limited^[80].

For hydrocarbon systems, Brenner developed the reactive empirical bond-order (REBO) potential based on the Tersoff formalism and later proposed a second-generation REBO potential with improved parameterization of C–C and C–H bond lengths, angles and dissociation energies^[81–82]. REBO can reproduce common carbon structures such as graphite and diamond and describe selected nonequilibrium processes. AIR-REBO further extends REBO by adding long-range Lennard–Jones-type interactions and torsional terms, allowing simulations of interlayer interactions in graphitic systems and torsional behavior in hydrocarbon chains^[83–84].

To better describe high-density amorphous carbon and complex local environments, EDIP was extended from silicon to carbon^[85]. Marks and co-workers modified EDIP by including π -bonding effects, dihedral-angle penalties and π -repulsion terms, improving the description of sp^2/sp^3 hybridization, ring structures and liquid–amorphous transitions^[85–86]. Simulated liquid and amorphous carbons under EDIP show elastic constants and structure factors in reasonable agreement with experimental data.

De Tomas and co-workers systematically compared six mainstream potentials, including Tersoff, REBO-II, EDIP and LCBOP-I^[71,95] (Fig. 4a–c). They found marked differences in how these potentials describe the sp^3 fraction, radial distribution functions, and graphitization behavior of amorphous carbon. Among them, EDIP showed the best performance in modelling graphitization from disordered to ordered graphitic structures, more accurately capturing bond-breaking and bond-formation dynamics. At the same time, COMB-type potentials embed variable-charge models into a many-body framework, allowing car-

bon charges to be determined self-consistently via charge equilibration. This combination of covalent, ionic, and metallic features makes COMB promising for multicomponent carbon-based systems containing elements such as O, N and Si^[94].

In terms of modelling protocols, these classical potentials are most often used in conjunction with the melt–quench method. Starting from a random or ordered initial structure, the system is heated to a high-temperature “liquid” regime under the chosen potential, allowing atoms to mix and relax thoroughly, then cooled at a prescribed rate and relaxed at the target temperature to obtain an amorphous network. By adjusting target density, temperature schedule, cooling rate, and external pressure, one can generate a series of a-C/ta-C models and analyze their radial distribution functions, coordination-number distributions, and sp^2/sp^3 ratios^[98–100]. As simulation scale and accuracy have increased, our understanding of amorphous carbon classification has undergone a qualitative shift. Zhu and co-workers recently proposed a “discontinuous phase diagram” for amorphous carbons, based on extensive MD simulations^[97] (Fig. 4d). Using topological characteristics, they divided amorphous carbons into 6 representative phases: disordered graphitic network (DGN), high-density amorphous carbon (HDAC), amorphous dual-phase carbon (a-DG), amorphous diamond (a-D), paracrystalline diamond (p-D), and nanopolycrystalline diamond (NPD). Their results indicate that the transformation from graphite-like to diamond-like amorphous carbon is not a smooth continuum but involves discontinuous topological transitions, challenging the traditional view of amorphous carbon structures as continuously varying.

The advantages of classical MD are clear. It can handle systems with up to hundreds of thousands of atoms over tens of picoseconds or longer, enabling statistical studies of relationships among density, hybridization, and mechanical properties. The analytic form of the potentials and the relatively low computational cost make it straightforward to perform parameter sweeps across different systems and conditions and to conduct sensitivity analyses.

From a review perspective, however, several lim-

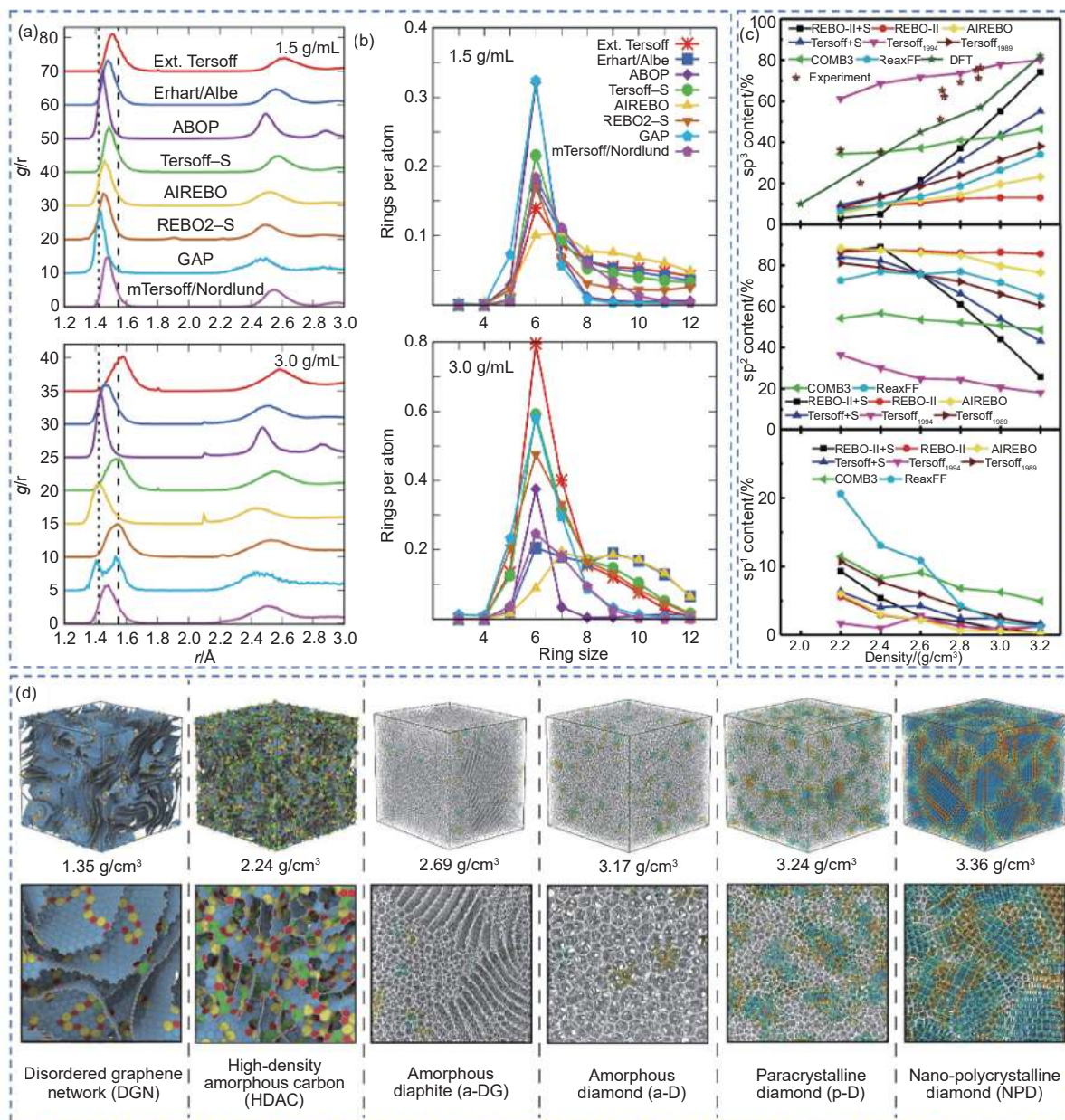


Fig. 4 Classical interatomic potentials for the atomistic description of amorphous carbon. (a) $g(r)$ for liquid-quenched a-C at $\rho=1.5$ and 3.0 g cm^{-3} predicted by representative potentials (extended Tersoff, Erhart–Albe, ABOP, Tersoff S, AIREBO, REBO2 S, etc)^[71]. (b) Ring-size distributions for the same structures^[71]. (Copyright 2019, Elsevier) (c) Density dependence of $sp^1/sp^2/sp^3$ fractions predicted by different potentials, compared with experimental/DFT trends^[96]. (Copyright 2020, Elsevier) (d) EDIP configurations at different densities corresponding to distinct amorphous-carbon phases (DGN, HDAC, a-DG, a-D, p-D, NPD)^[97]. (Copyright 2024, Oxford University Press)

itations must be emphasized. Most classical potentials do not explicitly treat charge transfer or complex reaction pathways; chemical reactions, gas evolution, and functional-group transformations are either absent or can only be captured in very limited scenarios. Potential transferability is also limited. For example, Liu and co-workers used eight potentials to model amorphous carbon via melt–quench and found that even under identical thermal protocols, the resulting

densities, sp^2/sp^3 ratios, ring statistics, and graphitization tendencies differed significantly, revealing “preferences” of each potential for certain regions of configuration space^[96](Fig. 4c). For this reason, classical MD is best viewed as an efficient tool for structural exploration and trend analysis. The resulting models, especially when used to interpret specific chemical mechanisms or high-precision spectroscopic data, typically require further calibration against DFT or ex-

periment.

4.2 Reactive MD: ReaxFF and explicit carbonization pathways

Reactive force fields extend beyond classical potentials by allowing the topology itself to evolve self-consistently during simulation, rather than merely describing the energy of a fixed bonding pattern. ReaxFF combines the bond-order formalism with a charge-equilibration scheme so that bond breaking and formation, charge transfer, and nonbonded interactions are treated on a unified potential-energy surface; this makes it possible to explicitly simulate chemical reaction pathways.

Li and co-workers systematically compared Tersoff, REBO-II, and ReaxFF for amorphous carbon^[101]. They found that ReaxFF showed superior structural fidelity: the sp^3 fraction in ReaxFF-based MD matched DFT data very well, and the radial distribution functions—both peak positions and shapes—were closer to experimental results than those obtained with traditional potentials. ReaxFF parameters are typically derived from extensive quantum-chemistry and DFT data, including reaction energies for small molecules, bond-dissociation curves, adsorption energies, and energies of selected solid phases, which confers a degree of transferability across broad chemical spaces^[93,102].

A large body of ReaxFF-based work on amorphous carbons falls into two routes. One follows the classical melt–quench scheme but replaces the potential with ReaxFF: starting from pure carbon or simple carbon networks, one melts the system and quenches it to produce amorphous carbon and then exploits ReaxFF's reactivity to study high-temperature annealing, graphitization, sp^3 to sp^2 conversion, and crack initiation. Parameter studies have shown that ReaxFF can effectively simulate graphitization, capturing the growth and alignment of graphene-like fragments during annealing^[103] (Fig. 5a-e). Ranganathan and co-workers analyzed how quench rate and cooling protocol affect the resulting structure, demonstrating that ReaxFF-based melt–quench can generate a wide variety of amorphous carbons with realistic chemical and

structural features^[104].

The second route starts from chemically realistic precursors and simulates thermolytic carbonization. Initial models may represent phenolic resins, asphaltenes, lignin, cellulose, coal-tar pitch or other organic precursors. Under staged heating, ReaxFF-MD can follow main-chain scission, side-group elimination, cross-linking, dehydrogenation, deheteroatomization and carbon enrichment^[105] (Fig. 5f-h). This allows the transformation from organic precursor to amorphous carbon to be analyzed at the atomistic level. AlAreeqi and co-workers reviewed ReaxFF applications in O— and H—containing carbon systems, including biomass and coal pyrolysis, combustion, coke formation and catalyst coking^[106]. Zhao and co-workers emphasized that realistic precursor models are necessary for predicting pore development and graphitization in porous carbons^[107]. Their work illustrates how precursor dehydrogenation, bond cleavage and fragment recombination can be connected to the formation of porous carbon frameworks.

ReaxFF extends amorphous-carbon modelling from static structures to reaction pathways, linking precursor chemistry with carbon-network formation. Its main limitations are parameter dependence, limited transferability across chemical environments and higher computational cost than classical potentials. Different ReaxFF parameter sets may give different quantitative results for graphitization, oxidation or pyrolysis. Therefore, ReaxFF simulations should be validated against DFT calculations, experimental product distributions, structural descriptors or spectroscopic data whenever possible.

4.3 Machine-learning interatomic potentials: high-accuracy energy surfaces for amorphous carbon

Machine-learning interatomic potentials (MLIPs) provide another route for improving the accuracy–scale balance in amorphous-carbon simulations. Classical potentials are efficient but limited by fixed analytic forms and manual parameterization. ReaxFF introduces reactivity, but its complex functional form and large parameter set make transferabil-

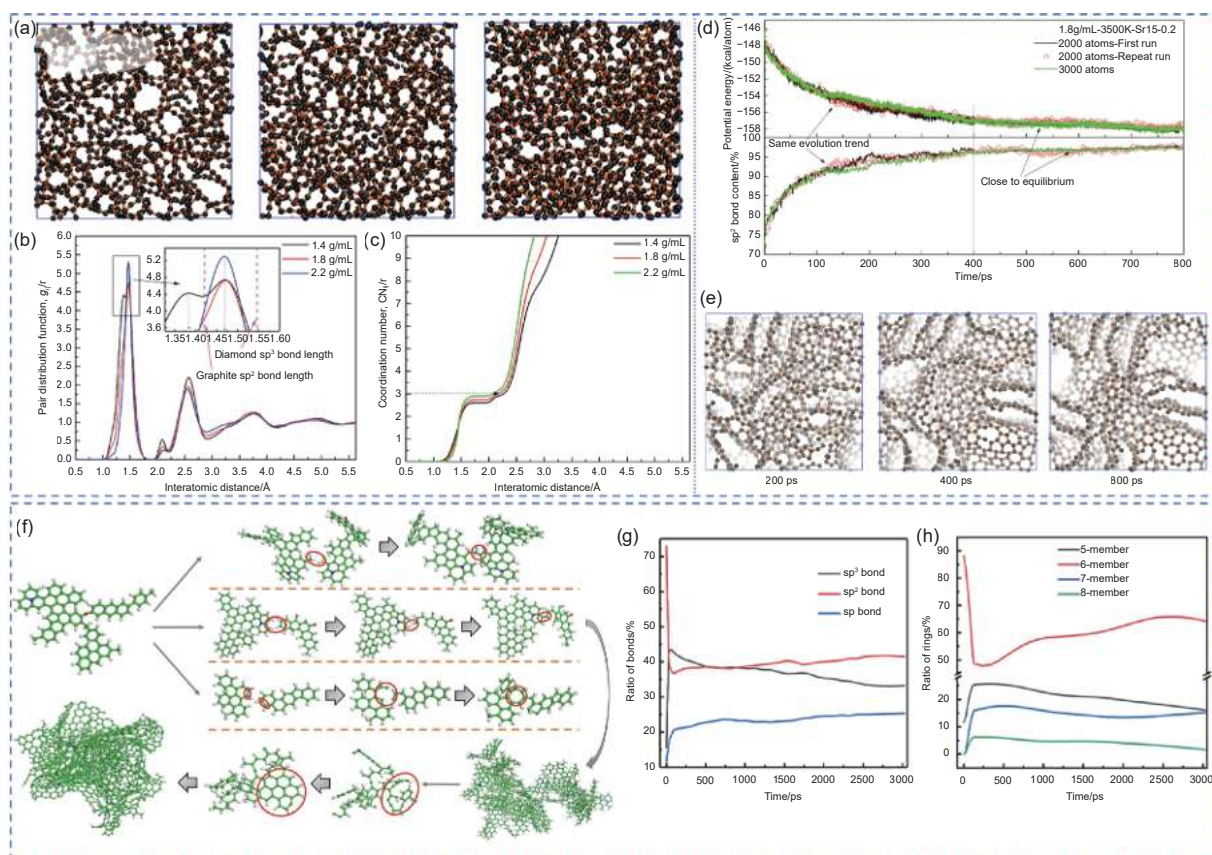


Fig. 5 ReaxFF workflows for melt-quench a-C and annealing-driven graphitization and precursor pyrolysis/carbonization. (a-e) Melt-quench a-C models, radial distribution functions, coordination curves, sp^2 -bond evolution and graphene-like fragment formation during annealing^[103]. (Copyright 2018, American Chemical Society) (f-h) ReaxFF-MD pyrolysis of modified coal-tar pitch, showing the precursor model, $sp/sp^2/sp^3$ bond evolution and ring-size evolution during carbonization^[105]. (Copyright 2023, Elsevier)

ity difficult. MLIPs instead learn the potential-energy surface from quantum-mechanical data and can therefore approximate DFT-level energies and forces over selected regions of configuration space^[108–109].

The construction of MLIPs generally involves three steps: data generation, structural representation and regression. First, DFT or ab initio MD calculations are used to generate representative configurations with corresponding energies and forces. For amorphous carbon, the training set should include liquid carbon, amorphous structures over a range of densities, crystalline phases such as graphite, diamond and graphene, and relevant defect or surface structures^[71,110]. This coverage is necessary to reduce extrapolation errors. Second, atomic environments are encoded using descriptors that are invariant to translation, rotation and permutation of identical atoms. Common descriptors include pair distances, angular functions, smooth overlap of atomic positions

(SOAP), symmetry functions and graph-based embeddings^[110]. For amorphous carbon, descriptors must distinguish different coordination states, ring environments, local curvature and medium-range order. Third, a regression model, such as Gaussian process regression, a neural network or a tensor-based model, maps these descriptors to atomic energy contributions and forces^[72,111].

Among MLIPs, Deringer and Csányi's Gaussian Approximation Potential (GAP) for condensed carbon is particularly noteworthy. In their work, they constructed a GAP trained on DFT data for liquid carbon and amorphous carbon at multiple densities, using a hierarchy of two-body, three-body, and SOAP descriptors^[110,112] (Fig. 6a-e). Their analysis showed that for carbon, a short-range potential constructed with appropriate descriptors can approximate the DFT energy surface within a well-defined accuracy bound. Tests demonstrated that the resulting GAP faithfully

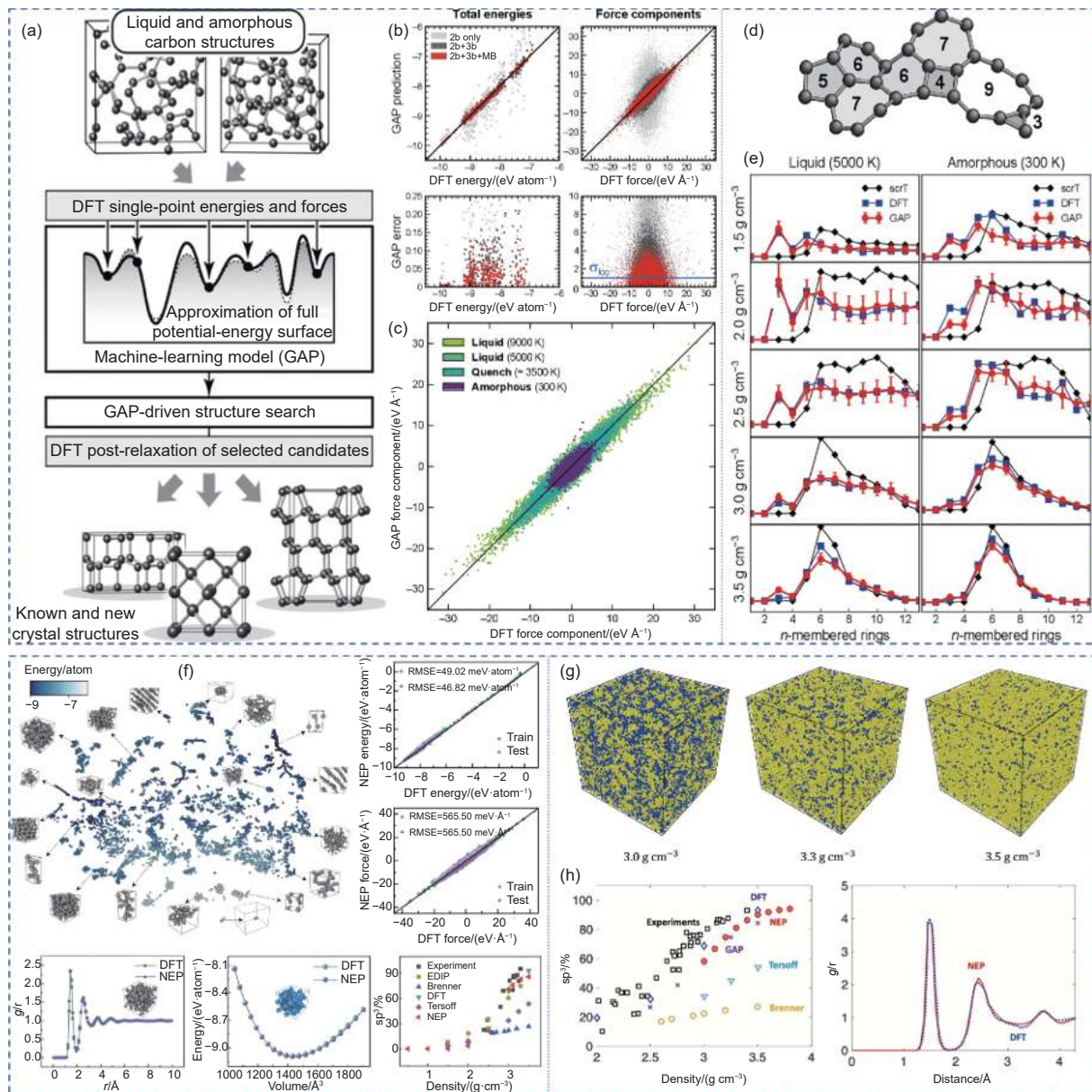


Fig. 6 Machine-learning interatomic potentials (MLIPs) for amorphous-carbon simulations: GAP and NEP examples. (a) Workflow for constructing a carbon GAP from DFT energies and forces^[112]. (Copyright 2017, Wiley-VCH) (b,c) GAP-DFT energy and force comparisons for liquid and amorphous carbon^[110]. (d,e) Ring-definition scheme and ring-size distributions predicted by empirical potentials, DFT and GAP^[110]. (Copyright 2017, Physical Review B) (f) NEP training-set coverage and NEP-DFT energy/force comparisons^[113]. (Copyright 2025, Springer Nature) (g-i) Large-scale NEP-generated amorphous-carbon structures, density-dependent sp² fractions and g(r) comparison with DFT^[114]. (Copyright 2025, Springer Nature)

reproduced DFT behavior over a wide range of amorphous configurations, rather than merely interpolating around a few sampled points.

Building upon this, Rowe and co-workers merged amorphous-carbon data with extensive DFT results for crystalline and defect structures to develop the more general “GAP-20” potential for carbon^[115]. GAP-20 accurately describes lattice constants, bond lengths, formation energies and phonon dispersions for diverse carbon allotropes. For amorphous-carbon

modelling, its importance lies in placing a-C, ta-C, graphite, diamond and defective carbon within a unified potential-energy framework. This enables simulations of liquid quenching, deposition, annealing, graphitization and amorphous–crystalline interfaces using a single potential.

More recently, neural-network potentials and the neuroevolution potential (NEP) framework have improved computational efficiency while maintaining near-DFT accuracy^[111,114] (Fig. 6f-i). While sharing the

same overall philosophy as GAP, these models differ in descriptor design, network architecture, and implementation details^[116]. For example, NEP uses compact descriptors based on Chebyshev and Legendre polynomials and employs evolutionary algorithms to optimize network structure, achieving near-DFT accuracy with MD-level computational cost on GPUs^[117]. Using NEP, Liu and co-workers conducted simulations on the order of 10^5 atoms and constructed temperature–density phase diagrams that reveal hidden medium-range order in amorphous carbon^[113]. They mapped out multiple amorphous sub-phases in the sp^3/sp^2 –density plane and traced how local topology and medium-range order evolve with temperature and density, in effect extending the “phase-diagram problem” for amorphous carbon to very large system sizes at near-DFT accuracy.

However, MLIPs are interpolative models. Their accuracy depends on the quality and coverage of the training data, and simulations outside the training domain can produce uncontrolled errors or nonphysical forces. For amorphous carbon, this risk is particularly relevant for extreme densities, unusual defects, reactive environments, heteroatom chemistry and electrochemical interfaces. Active learning provides one solution: MD trajectories are monitored for configurations that are far from the training set, and additional DFT calculations are performed to update the model^[118].

MLIPs should therefore be viewed as high-accuracy intermediate-scale tools rather than universal potentials. They are well suited for structure generation, relaxation and property prediction in chemically similar systems covered by the training data. For processes involving large compositional changes or complex reaction networks, such as pyrolysis, oxidation and electrochemical cycling, MLIPs should be combined with ReaxFF, DFT and experimental constraints rather than used alone.

4.4 Practical access to representative interatomic potentials

In practical MD simulations, selecting a potential also means selecting a specific parameter or model file. Different parameterizations within the same potential family can produce different amorphous-carbon structures under the same simulation protocol. The potential file should therefore be reported as part of the model definition.

Common sources include the NIST Interatomic Potentials Repository, OpenKIM, libAtoms/GAP databases, original publications, supplementary information and author-maintained repositories (Table 2).

For reproducibility, authors should report the exact potential file name, source, version or commit identifier when available, software package and version, unit system, element mapping, charge-equilibration settings for reactive force fields, and at least one

Table 2 Practical access to representative interatomic potentials for amorphous-carbon simulations

Potential class	Source / Format	Implementation	Key considerations
Tersoff-type bond-order potentials	.tersoff files; NIST IPR, OpenKIM, original publications	LAMMPS <i>pair_style tersoff</i> ; element mapping via <i>pair_coeff</i>	Parameters fitted for crystalline carbon (graphite, diamond, graphene, SiC) may not transfer to porous a-C, ta-C, or highly strained films
REBO / AIREBO / LCBO / EDIP-type potentials	Built-in potential styles; original parameter files; supplementary materials	LAMMPS built-in pair styles or package-specific implementations	Characteristic bias in sp^2/sp^3 balance, ring statistics, graphitization tendency, and nonbonded interactions
ReaxFF	<i>ffield.reax</i> files; published ReaxFF parameter repositories; supplementary information	LAMMPS <i>pair_style reaxff</i> with charge equilibration	Chemistry-specific parameter sets; a hydrocarbon, combustion, oxide, or aqueous-interface parameter set should not be assumed transferable across all carbonization or electrochemical environments
GAP and other MLIPs	XML/model files; libAtoms/GAP databases; author/institutional repositories; supplementary datasets	QUIP/LAMMPS <i>pair_style quip</i> ; ASE/quippy; model-specific MD engines	Reliability controlled by training-domain coverage; extrapolation should be monitored via uncertainty indicators, DFT spot checks, or active learning
NEP / NNP / MACE-type models	Model files and training datasets; author repositories; public model repositories; supplementary materials	GPUMD; ASE/LAMMPS interfaces; model-specific engines	Code version, descriptor settings, cutoff, training data, and validation metrics must be reported for reproducibility

validation benchmark relevant to the target amorphous-carbon system. Recommended benchmarks include radial distribution functions, coordination-number distributions, sp^2/sp^3 fractions, density dependence, elastic response, graphitization trajectory, or comparison with selected DFT calculations.

5 Data-driven reconstruction: Reverse Monte Carlo methods

The methods discussed above mainly follow a forward-modelling strategy: a potential-energy surface is defined, and the atomic structure evolves on that surface. In many experiments, the problem is reversed. Structural factors, radial distribution functions and related statistical descriptors are measured first, and the task is to construct 3D atomic models consistent with these data. For amorphous carbon, total scattering is one of the main experimental routes for probing bulk structure, but the measured information is much lower in dimensionality than the full atomic configuration space. Structural reconstruction from scattering data is therefore an ill-posed inverse problem^[119].

Reverse Monte Carlo (RMC) and hybrid reverse Monte Carlo (HRMC) address this problem by using experimental data as direct structural constraints^[23,120]. In retrospect, this “given data and constraint functions, let the algorithm fit the structure” paradigm represents an early form of data-driven modelling in materials science—primitive compared with modern deep learning, yet already embodying elements of data-driven optimization and multi-objective trade-offs.

5.1 Reverse Monte Carlo and Hybrid Reverse Monte Carlo

Classical RMC is a stochastic optimization method for constructing atomic configurations from experimental data. Given the number of atoms, chemical composition, simulation cell and target data, RMC minimizes the difference between calculated and experimental structural functions. McGreevy and Pusztai first formalized this approach for modelling liquids from neutron total scattering, and it has since

been widely used for liquids, glasses and amorphous solids^[23].

In a typical RMC simulation, one first generates an initial configuration, either by random placement or by using a heuristic structure derived from MD/DFT. One then computes the model radial distribution function $g_{\text{calc}}(r)$ (or structure factor $S_{\text{calc}}(Q)$) and compares it with the experimental curve $g_{\text{exp}}(r)$ (or $S_{\text{exp}}(Q)$). An error function such as

$$\chi^2 = \sum_{i=1}^{N_{\text{exp}}} \omega_i [g_{\text{calc}}(r_i) - g_{\text{exp}}(r_i)]^2$$

is defined, where ω_i are weight factors. When multiple data sets are fit simultaneously (e.g., X-ray and neutron scattering, PDF plus EXAFS), the corresponding χ^2 contributions are typically combined into a single overall cost function.

During the simulation, an atom is randomly displaced and the change in χ^2 is evaluated. If $\Delta\chi^2 < 0$, the move is accepted unconditionally; if $\Delta\chi^2 > 0$, the move is still accepted with probability $\exp(-\Delta\chi^2/2)$, following a Metropolis-like rule that allows the algorithm to escape local minima. After many such trial-and-accept/reject steps, χ^2 decreases and stabilizes, and the system converges to a statistical ensemble of three-dimensional configurations that reproduce the experimental scattering within its uncertainty.

As discussed earlier, pure RMC uses only the data-fidelity term in the cost function, which makes it highly flexible in fitting experimental curves but also accentuates the ill-posedness of the inverse problem: within a finite Q -range and resolution, many structurally different configurations—some of which may be unphysical at the local level—can yield equally good fits. To alleviate this, HRMC introduces physical potentials as energy regularizers, unifying data fitting and structural plausibility in a single optimization framework^[120–121] (Fig. 7). A typical HRMC cost function can be written schematically as

$$\Phi = \chi^2 + \lambda_E E_{\text{pot}} + \sum_k \lambda_k C_k$$

where χ^2 quantifies the misfit between calculated and experimental scattering, E_{pot} is the total potential energy evaluated with an empirical potential (e.g., EDIP,

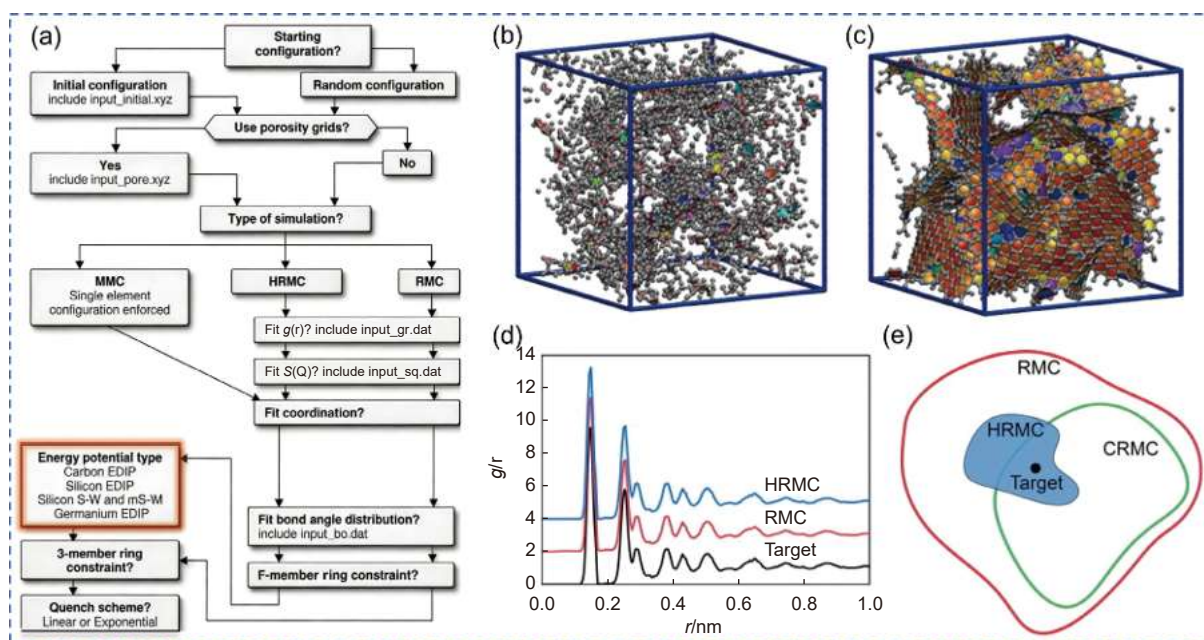


Fig. 7 Workflows and key distinction between RMC and HRMC for constructing atomistic models of disordered carbons. (a) Workflow showing selection of experimental constraints (e.g., $S(Q)$, $g(r)$, coordination, bond angles, pore volume) and inclusion of an energy term in HRMC^[122]. (Copyright 2013, Elsevier). (b,c) Disordered nanoporous carbon (DNC) structures generated by RMC and HRMC^[123]. (d) $g(r)$ comparison for an experimental target, a pure RMC fit, and HRMC^[123]. (e) Conceptual configuration-space picture for RMC, constraint-based RMC, and HRMC relative to the target solution^[123]. (Copyright 2012, Elsevier)

Tersoff, ReaxFF), C_k represents geometric or macroscopic constraints (minimum C—C distances, coordination-number windows, density, porosity, internal surface area), and λ_E, λ_k are weights that balance these terms.

In HRMC, the acceptance criterion is accordingly generalized to penalize increases in the total cost function Φ . A proposed move that deteriorates the data fit significantly or raises the potential energy or violates geometric constraints is unlikely to be accepted; only moves that keep all contributions within acceptable bounds have a high acceptance probability. This effectively narrows the structural search space from “all configurations compatible with scattering” to “configurations compatible with both scattering and local physical plausibility.” For amorphous carbons, one typically chooses a potential that has been validated for carbon (e.g., EDIP or ReaxFF) as E_{pot} , and combines it with constraints on minimum C—C distances, reasonable coordination-number ranges, porosity, and density. The result is an ensemble of structures that maintain high-quality fits to total scattering while suppressing nonphysical features such as ghost atoms, improper coordination, and pathological ring

patterns^[122,124–125].

5.2 Applications of RMC/HRMC in amorphous carbon modeling

For amorphous carbon, RMC/HRMC are mainly used to construct three-dimensional structures that are consistent with scattering data and basic chemical constraints. A typical workflow includes four steps. First, the target experimental data are selected, such as X-ray or neutron total scattering, PDF, density, H/C ratio, porosity or adsorption data. Second, a simulation cell is built with the correct composition and density. Third, RMC or HRMC is used to optimize atomic positions under data-fitting, geometric and energetic constraints. Finally, structural descriptors such as coordination statistics, ring distributions, sp^2 -domain size, curvature, pore-size distribution and pore connectivity are extracted from the converged ensemble^[23,121].

This approach is useful because structural features in the final model can be traced back to experimental constraints. For porous carbons, HRMC has been used to reconstruct atomistic frameworks that reproduce both scattering data and porosity-related

measurements. For example, Farmahini and co-workers developed a two-stage HRMC strategy for SiC-derived nanoporous carbon^[126] (Fig. 8a-f).

Importantly, RMC/HRMC structures do not represent the endpoint of modelling. The most practical and widely used downstream application is to treat these reconstructions as initial configurations or reference frameworks for subsequent DFT, ReaxFF, or ML-MD simulations. In this workflow, RMC/HRMC first builds a “statistically realistic” backbone under scattering constraints, and this backbone is then relaxed or evolved on a chosen potential-energy surface

under conditions that introduce additional physics, such as high-temperature graphitization, gas or electrolyte infiltration, and ion insertion^[128]. The resulting simulations remove residual unphysical stress concentrations in raw RMC models and, more crucially, allow dynamical and transport processes to be studied in networks fully consistent with experimental scattering, rather than in idealized slit-pore or hand-constructed graphitic models. Studies of porous-carbon adsorption and Na-storage mechanisms in hard carbons have repeatedly demonstrated the value of such experimentally anchored models^[127] (Fig. 8g-i).

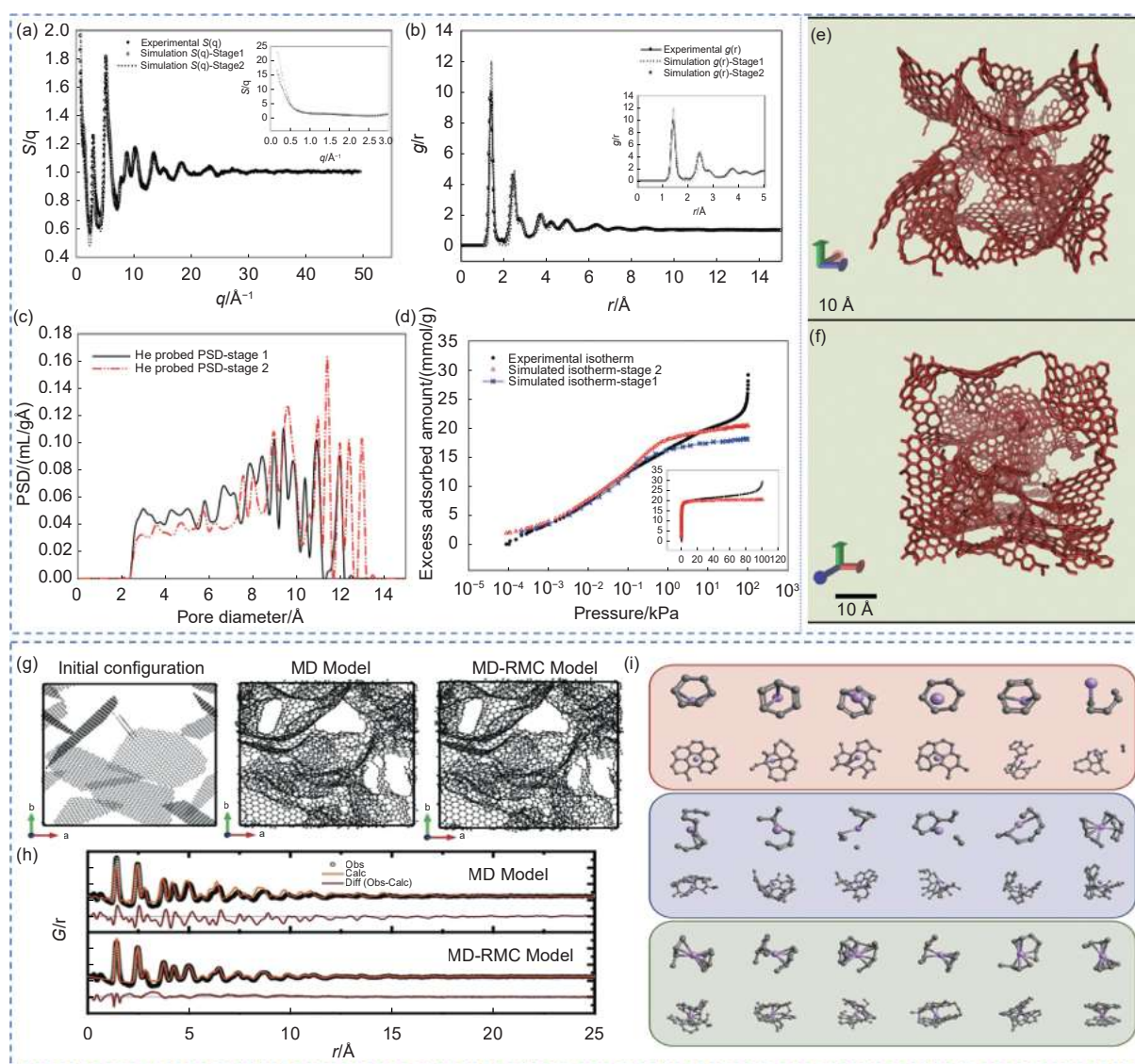


Fig. 8 Examples of how RMC/HRMC are used to construct atomistic models of amorphous carbon. (a–f) Two-stage HRMC modelling of SiC-derived nanoporous carbon, including fits to, pore-size distributions and adsorption isotherms, and the corresponding reconstructed frameworks^[126]. (Copyright 2013, American Chemical Society). (g–i) MD–RMC refinement of hard-carbon models against neutron total-scattering PDF data and the resulting Na-storage motifs after Na insertion^[127]. (Copyright 2022, Wiley)

Beyond this, using RMC/HRMC structures to directly calibrate potentials is less developed in amorphous carbons but already established in other glassy and liquid systems. For instance, Gereben and Pusztai's RMCMD and RMC_POT frameworks combine experimental scattering with MD-derived partial RDFs to assess the extent to which certain potentials produce local structures compatible with experiment^[129]. Analogous strategies may be extended to amorphous carbons, using HRMC as a structural “filter” to identify which regions of configuration space preferred by a given potential are actually supported by scattering data.

Given that RMC/HRMC models are labeled by specific samples, processing conditions, and scattering experiments, and have passed energy and geometry constraints, they are particularly attractive as high-confidence samples for constructing structure–property databases or training AI models. Although current data sets remain limited in size and standardization, it is natural to view RMC/HRMC reconstructions as valuable ingredients for future large-scale databases of amorphous carbon structures and properties.

5.3 Limitations of RMC/HRMC and the need for experimental advances

Despite their unique contributions, RMC/HRMC suffer from intrinsic limitations largely dictated by the dimensionality and quality of the input data rather than by algorithmic sophistication. The essence of the inverse problem is “many unknowns, few constraints”: total scattering provides one-dimensional data, whereas the number of atomic degrees of freedom grows linearly with atom count and typically far exceeds the number of independent experimental observations^[130]. Under these conditions, even using mature programs such as RMCProfile cannot guarantee a unique solution^[131]. Pure RMC, which lacks energetic regularization, tends to converge toward maximum-entropy configurations—networks that are as disordered as possible while still matching the data. Such models often underrepresent hidden medium-range order and may deviate from real materials in key statist-

ics such as ring distributions and coordination environments^[132]. For amorphous carbons, where scattering is only weakly sensitive to sp^2/sp^3 distributions and pore morphologies, these issues are even more acute.

HRMC moderates this problem by reintroducing energy and geometric constraints into the optimization, reducing local artifacts and bringing the reconstructed structures closer to those seen in forward MD or DFT models^[121,133]. However, the improvement is tied to the quality of the chosen potential and constraints. If the potential has systematic biases for certain local environments, HRMC may compromise between reproducing experimental data and adhering to the potential, pushing the model toward structures favored by the latter. As a result, RMC/HRMC structures are best regarded as ensembles of plausible candidates consistent with both data and model assumptions, rather than unique “true structures.”

These limitations highlight a deeper issue: the ultimate bottleneck for RMC/HRMC lies in the information content of available experimental data. As long as the input is restricted to one-dimensional $S(Q)$ or $g(r)$ with limited Q -range and signal-to-noise ratio, no algorithm can fully resolve the ill-posedness. The key to narrowing the solution space is to increase the dimensionality and richness of experimental constraints.

Recent progress has moved in this direction by merging multiple modalities. Wider- Q -range X-ray and neutron total scattering can better constrain long-range density fluctuations; combined fitting of PDFs, $S(Q)$, and Bragg peaks in tools like RMCProfile allows simultaneous control of local and medium-range structure in multi-phase materials; integrating EXAFS and XANES brings local coordination and bonding statistics into the same optimization framework; and combining total scattering with solid-state NMR, Raman, and even three-dimensional imaging techniques opens the door to multi-modal data fusion^[132]. In amorphous carbons, the development of high-resolution total scattering instrumentation and dedicated workflows for three-dimensional reconstruction of carbonaceous materials further illustrates that al-

gorithmic advances must go hand in hand with improvements in experimental data^[134].

Within the broader modelling picture proposed in this review, RMC/HRMC occupy a clearly defined role: they are neither traditional potential-based methods nor purely data-fitting black boxes, but experimental-anchored structural reconstruction tools sitting between the two. They are most suitable for generating large-scale candidate structures that are tightly constrained by scattering and basic physical plausibility and then serving as starting points and constraints for DFT, classical MD, ReaxFF and ML-MD. To enhance the credibility of amorphous carbon models, future efforts must therefore focus not only on refining RMC/HRMC algorithms but also on advancing experimental techniques to provide richer, higher-quality constraints.

6 Summary and outlook

6.1 Revisiting the accuracy–scale–fidelity triangle

Throughout this review, the main methods have been compared in terms of accuracy, accessible scale and experimental fidelity. Geometric and topological models can generate large continuous random networks at low computational cost and allow direct control of structural motifs. However, their reliability depends strongly on the assumed construction rules and they do not include explicit energetic criteria.

DFT provides the highest level of accuracy for local bonding, charge redistribution and electronic structure, but it is limited to relatively small cells and short simulation times. It is most suitable for studying representative local motifs, dopants, defects, adsorption sites, migration barriers and spectroscopic signatures. Classical MD with potentials such as Tersoff, REBO/AIREBO, EDIP and COMB extends simulations to larger systems and longer times, supporting melt–quench, deposition and annealing simulations. Its accuracy, however, depends on the transferability of the potential to the target density, hybridization state and chemical environment.

ReaxFF adds explicit reactivity and is useful for

carbonization, oxidation, dehydrogenation, gas evolution and electrochemical interfacial processes. Its main limitations are parameter dependence and higher computational cost. MLIPs occupy an intermediate region between DFT and classical MD. When the training data adequately cover the relevant configuration space, MLIPs can reproduce DFT-level energies and forces for much larger amorphous-carbon systems. Their reliability, however, decreases when simulations enter poorly sampled regions.

RMC and HRMC primarily improve experimental fidelity. RMC searches for structures that reproduce experimental $S(Q)$, $g(r)$ or PDF data, while HRMC adds potentials and geometric constraints to reduce unphysical configurations. Because experimental data are incomplete, RMC/HRMC structures should be treated as ensembles of plausible candidates rather than unique atomic structures.

The practical conclusion is that method selection should be problem-oriented. DFT is preferred for local bonding, charge redistribution and spectral interpretation. Classical MD and MLIPs are suitable for density-dependent structural trends, mechanical response and large-cell amorphization. ReaxFF is required when bond-breaking chemistry or precursor carbonization is central. RMC/HRMC should be used when consistency with scattering or other experimental data is the main constraint. In most cases, a reliable workflow combines several methods: DFT defines local reference physics, MD or MLIP generates statistically meaningful configurations, and RMC/HRMC or multimodal experiments constrain the structural solution space.

6.2 Hierarchical structural view and multi-method coupling

A useful way to organize multi-method modelling is to describe amorphous carbon as a hierarchical structure. Instead of enforcing one universal workflow for all problems, methods can be assigned according to structural level. Here, we propose a three-level framework consisting of atomic, microscopic and mesoscopic descriptions (Fig. 9).

At the atomic level (analogous to the material's

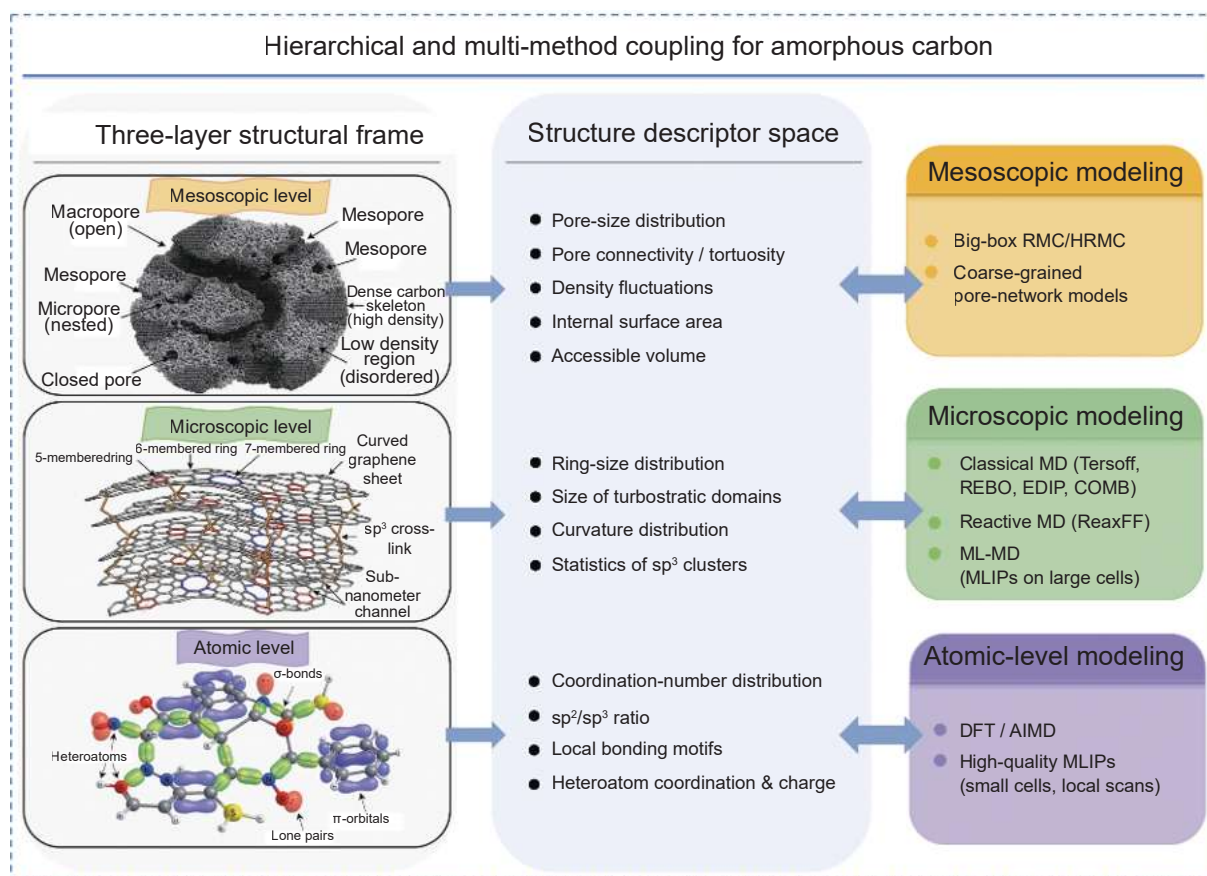


Fig. 9 Hierarchical structural picture and multi-method coupling framework for amorphous carbon

genetic code), the focus is on individual atoms and their nearest-neighbor environments: coordination-number distributions, sp^2/sp^3 hybridization, the coordination of heteroatoms such as N, O and S, and associated local energies and electronic structures. DFT and high-quality MLIPs are indispensable here as ground-truth providers for bonding and electronic states.

At the microscopic level, atomic units assemble into continuous networks that exhibit medium-range order: ring statistics (including 5/7-membered defects), sizes and curvatures of turbostratic graphitic domains, distributions of sp^3 cross-linked regions, and small pore cavities and subnanometer channels. Classical MD, ReaxFF, and ML-MD dominate this level, helping to elucidate how atomic motifs link and rearrange over nanometer scales.

At the mesoscopic level, attention shifts to overall texture and morphology: hierarchical pore structures, pore-size distributions and connectivity, inter-

penetrating density domains, intra-particle multiphase skeletons, and film-current-collector interfaces. HRMC and coarse-grained pore-network models play key roles here, integrating total scattering and adsorption data to reconstruct statistically realistic porous skeletons.

Within this framework, multi-method coupling is essentially the cross-level transmission of structural descriptors. Mesoscopic models provide density fluctuations, pore statistics, and other boundary conditions for network-level simulations; microscopic-level models feed back ring statistics, fragment-size distributions, and curvature spectra that refine mesoscopic constraints; atomic-level calculations define energies and electronic signatures for representative local environments sampled from microscopic-level configurations.

Ideally, atomic, microscopic, and mesoscopic levels are all described within a compatible descriptor system so that results generated at different scales and

by different methods can be mapped into a common structural parameter space. For energy-storage-oriented amorphous carbons, the practical role of atomistic modelling is to translate hidden structural features into descriptors that can be connected with synthesis and electrochemical performance. In hard-carbon and porous-carbon electrodes, these descriptors include interlayer spacing, graphitic-domain curvature, defect density, heteroatom coordination, closed- and accessible-pore statistics, ion-adsorption sites, and diffusion barriers. They do not directly prescribe a synthesis recipe, but they identify which structural variables should be tuned during precursor selection, stabilization, carbonization, activation, doping, or post-treatment.

Although the examples discussed in this review are mainly centered on energy-storage carbons, the same hierarchical descriptor framework has broader implications for amorphous-carbon materials. In DLC and ta-C coatings, the descriptor focus shifts toward sp^3 fraction, density, residual stress, hydrogen termination, graphitic clustering, interfacial bonding, and friction-induced structural evolution. In biomedical a-C/DLC surfaces, the relevant descriptors further shift to surface termination, roughness, hydrophilicity, interfacial water structure, ion adsorption, and biomolecular adsorption motifs. Therefore, the general value of atomistic modelling lies in establishing a transferable descriptor language that connects processing, disordered structure, and application-specific performance across different amorphous-carbon systems.

6.3 Toward data-centric modelling in the AI era

The role of artificial intelligence in amorphous-carbon modelling should be viewed with both expectation and caution. At present, the most concrete progress has been made in machine-learning interatomic potentials (MLIPs), where DFT-trained models such as GAP, neural-network potentials, and NEP-type potentials have already extended the accessible scale of near-DFT amorphous-carbon simulations. In contrast, broader AI technologies—including active learning, high-throughput screening, large-language-model

(LLM)-assisted literature mining, and multimodal data fusion—remain less mature for amorphous carbon than for crystalline materials. For a-C systems, these technologies are better regarded as emerging opportunities and possible infrastructure for future modelling, rather than as fully established workflows.

A first promising direction is active learning for carbon-specific MLIP development. Amorphous carbon spans liquid carbon, sp^2 -rich turbostratic networks, sp^3 -rich tetrahedral networks, porous frameworks, surfaces, defects, doped structures, and non-equilibrium carbonization pathways. A static training set is unlikely to cover all these environments. Therefore, future MLIP development could benefit from iterative active-learning protocols: preliminary MLIP-driven simulations would explore candidate structures; configurations with high uncertainty, unusual coordination, or extrapolative local environments would be selected for DFT labelling; and the model would then be retrained until the target configuration space is adequately sampled. For amorphous carbon, such protocols should deliberately include high-temperature liquids, low-density porous networks, high-density ta-C-like structures, curved graphitic fragments, under-coordinated surfaces, heteroatom-containing motifs, and strained interfacial environments.

A second direction is high-throughput generation and screening of amorphous-carbon structural models. Unlike crystalline materials, amorphous carbons cannot be enumerated by space group, composition, and lattice parameters. Their model space is controlled by processing-related variables such as precursor chemistry, density, hydrogen content, carbonization temperature, quench rate, pressure, dopant type, pore volume, and graphitic-domain size. A future high-throughput framework would therefore need to generate ensembles of plausible structures rather than single deterministic models. For each model, descriptors such as coordination distribution, ring statistics, sp^2/sp^3 ratio, graphitic-domain curvature, pore-size distribution, accessible surface area, Na adsorption-energy distribution, diffusion barriers, and elastic response could be extracted. These descriptors may

then be used to construct process–structure–property maps for energy-storage-oriented amorphous carbons. However, such a framework requires standardized model-generation protocols and consistent descriptor definitions before reliable data-driven screening can be achieved.

A third opportunity lies in large-language-model-assisted knowledge extraction^[135]. The amorphous-carbon literature contains extensive information on precursors, stabilization treatments, carbonization conditions, Raman parameters, XPS-derived bonding states, NMR signatures, pore descriptors, adsorption data, electrochemical performance, and simulation protocols. Much of this information is scattered across heterogeneous publications and is not directly machine-readable. In principle, large language models and materials-specific natural-language-processing tools could help extract and organize these data into structured databases. However, for scientific use, such extraction must be coupled with human curation, source verification, uncertainty annotation, and consistent terminology. LLMs should therefore be regarded as assistants for database construction rather than as independent sources of physical knowledge.

A fourth direction is AI-assisted multimodal data fusion^[136]. The non-uniqueness of amorphous-carbon structure reconstruction cannot be fully resolved by fitting a single scattering curve. Total scattering constrains pair correlations; Raman and XPS provide in-

formation on bonding and hybridization; NMR probes local chemical environments; SAXS/SANS and gas adsorption constrain pore structures; microscopy and tomography provide real-space morphological information. Future data-fusion strategies could combine these observables with atomistic descriptors from DFT, MD, MLIP-MD and RMC/HRMC models, thereby narrowing the feasible structural ensemble more effectively than any single technique. In this sense, multimodal learning may become a useful extension of experimentally constrained modelling. Nevertheless, this requires careful treatment of experimental resolution, incompatible length scales, sample-to-sample variability, and the non-unique mapping between spectra and structures.

Overall, the AI era should not be interpreted as a replacement of physics-based modelling by black-box prediction. For amorphous carbon, a more realistic path is a physically constrained, data-centric framework in which DFT provides local reference physics, MD and MLIPs generate statistically meaningful configurations, RMC/HRMC and experiments constrain structural fidelity, and AI tools assist in sampling, data organization, descriptor extraction, and uncertainty-aware pattern discovery. The immediate challenge is therefore not only to introduce more complex algorithms, but to build curated datasets, standardized descriptors, reproducible workflows, and validation protocols suitable for the intrinsically disordered and metastable nature of amorphous carbon.

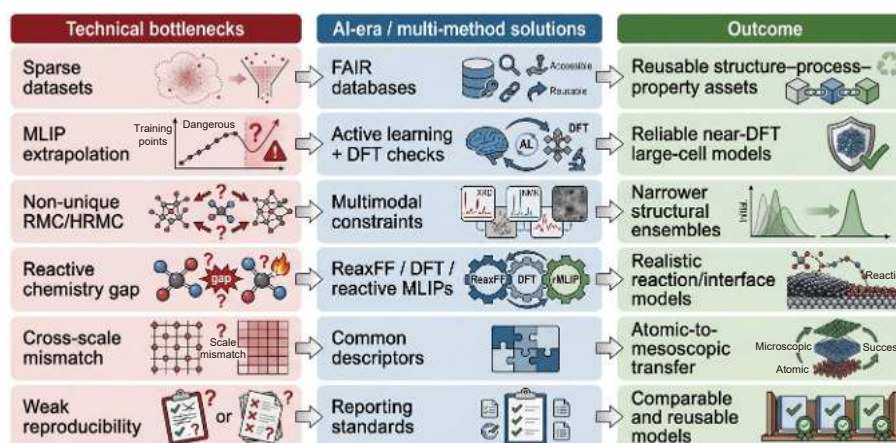


Fig. 10 Technical bottlenecks and solution roadmap for AI-era amorphous-carbon modelling. The main bottlenecks include data sparsity, MLIP extrapolation, non-unique RMC/HRMC reconstruction, incomplete reactive chemistry, cross-scale descriptor mismatch and weak reproducibility. The corresponding solutions include curated databases, active learning, multimodal experimental constraints, reactive modelling, standardized descriptors and reproducible reporting

6.4 Technical bottlenecks and solution roadmap

Despite rapid methodological progress, AI-era amorphous-carbon modelling still faces several technical bottlenecks (Fig. 10). The first is data sparsity and inconsistency: processing conditions, structural descriptors, potential files, experimental constraints, and performance metrics are often reported in incompatible formats. This limits the construction of transferable structure–process–property models. The second is MLIP extrapolation. Amorphous carbon spans liquid, graphitic, tetrahedral, porous, hydrogenated, doped, surface, and interfacial environments, and a model trained on only part of this space may generate nonphysical structures outside its training domain. The third is the non-uniqueness of inverse reconstruction. Total scattering alone cannot uniquely determine the 3D topology of a-C, so RMC/HRMC models must be further constrained by Raman, XPS, NMR, SAXS/SANS, gas adsorption, microscopy, and energetic regularization. The fourth is the incomplete description of reactive chemistry, especially in carbonization, oxidation, tribochemistry, and electrochemical interfaces. This requires coordinated use of ReaxFF, DFT reaction-pathway calculations, reactive MLIPs, and experimentally constrained reconstruction. Finally, cross-scale descriptor incompatibility and weak reporting standards still prevent direct comparison among models generated by different groups.

A practical solution is to establish a closed, data-centric modelling workflow: DFT provides local reference physics; classical MD, ReaxFF and MLIPs generate statistically meaningful configurations and process trajectories; RMC/HRMC and multimodal experiments constrain structural fidelity; active learning identifies missing regions of configuration space; and standardized descriptors connect atomic bonding, network topology, pore architecture, and application performance. Such a roadmap would move amorphous-carbon modelling from isolated case studies toward reproducible, transferable, and application-guided digital materials design.

Acknowledgements

This work was financially supported by the au-

tonomous research project of SKLCM (SCJC-XCL-202501, SCJC-XCL-202503) and the Joint Funds of the National Natural Science Foundation of China (U25B20110).

References

- [1] Robertson J. Diamond-like amorphous carbon[J]. *Materials Science and Engineering: R: Reports*, 2002, 37(4-6): 129-281.
- [2] Falcao E H, Wudl F. Carbon allotropes: beyond graphite and diamond[J]. *Journal of Chemical Technology & Biotechnology*, 2007, 82(6): 524-531.
- [3] McKenzie D R. Tetrahedral bonding in amorphous carbon[J]. *Reports on Progress in Physics*, 1996, 59(12): 1611-1664.
- [4] Galli G, Martin R M, Car R, et al. Structural and electronic properties of amorphous carbon[J]. *Physical Review Letters*, 1989, 62(5): 555-558.
- [5] Grill A. Tribology of diamondlike carbon and related materials: an updated review[J]. *Surface and Coatings Technology*, 1997, 94-95: 507-513.
- [6] Roy R K, Lee K R. Biomedical applications of diamond-like carbon coatings: a review[J]. *Journal of Biomedical Materials Research Part B: Applied Biomaterials*, 2007, 83B(1): 72-84.
- [7] Casiraghi C, Robertson J, Ferrari A C. Diamond-like carbon for data and beer storage[J]. *Materials Today*, 2007, 10(1-2): 44-53.
- [8] Stevens D A, Dahn J R. High capacity anode materials for rechargeable sodium-ion batteries[J]. *Journal of The Electrochemical Society*, 2000, 147(4): 1271.
- [9] Wang Y, Yi Z, Xie L, et al. Releasing free radicals in precursor triggers the formation of closed pores in hard carbon for sodium-ion batteries[J]. *Advanced Materials*, 2024, 36(26): 2401249.
- [10] Simon P, Gogotsi Y. Materials for electrochemical capacitors[J]. *Nature Materials*, 2008, 7(11): 845-854.
- [11] Mao Y, Yi Z, Xie L, et al. Elimination of hydrogen bonds in cellulose enables high-performance disordered carbon anode in sodium-ion batteries[J]. *Energy Storage Materials*, 2024, 73: 103845.
- [12] Billinge S J L, Levin I. The problem with determining atomic structure at the nanoscale[J]. *Science*, 2007, 316(5824): 561-565.
- [13] Robertson J. Amorphous carbon[J]. *Advances in Physics*, 1986, 35(4): 317-374.
- [14] LiBassi A, Ferrari A C, Stolojan V, et al. Density, sp³ content and internal layering of DLC films by X-ray reflectivity and electron energy loss spectroscopy[J]. *Diamond and Related Materials*, 2000, 9(3-6): 771-776.
- [15] Ferrari A C, Robertson J. Interpretation of Raman spectra of disordered and amorphous carbon[J]. *Physical Review B*, 2000, 61(20): 14095-14107.
- [16] Ohtake N, Hiratsuka M, Kanda K, et al. Properties and classification of diamond-like carbon films[J]. *Materials*, 2021, 14(2): 315.

- [17] Martin J W, De Tomas C, Suarez-Martinez I, et al. Topology of disordered 3D graphene networks[J]. *Physical Review Letters*, 2019, 123(11): 116105.
- [18] Fallon P J, Veerasamy V S, Davis C A, et al. Properties of filtered-ion-beam-deposited diamondlike carbon as a function of ion energy[J]. *Physical Review B*, 1993, 48(7): 4777-4782.
- [19] Franklin R E. Crystallite growth in graphitizing and non-graphitizing carbons[J]. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1951, 209(1097): 196-218.
- [20] Li C C, Wang Y L, Yi Z L, et al. Enhanced Na storage performance through in situ depolymerization-repolymerization in bamboo-derived hard carbon anodes[J]. *Journal of Materials Chemistry A*, 2025, 13(22): 16533-16546.
- [21] Song M, Xie L, Su F, et al. New insights into the effect of hard carbons microstructure on the diffusion of sodium ions into closed pores[J]. *Chinese Chemical Letters*, 2024, 35(6): 109266.
- [22] Liu Y, Xue J S, Zheng T, et al. Mechanism of lithium insertion in hard carbons prepared by pyrolysis of epoxy resins[J]. *Carbon*, 1996, 34(2): 193-200.
- [23] McGreevy R L. Reverse monte carlo modelling[J]. *Journal of Physics: Condensed Matter*, 2001, 13(46): R877-R913.
- [24] Drabold D A. Topics in the theory of amorphous materials[J]. *The European Physical Journal B*, 2009, 68(1): 1-21.
- [25] Yuan R, Guo Y, Gurgan I, et al. Raman spectroscopy analysis of disordered and amorphous carbon materials: a review of empirical correlations[J]. *Carbon*, 2025, 238: 120214.
- [26] Chu P K, Li L. Characterization of amorphous and nanocrystalline carbon films[J]. *Materials Chemistry and Physics*, 2006, 96(2-3): 253-277.
- [27] Zunger A. Inverse design in search of materials with target functionalities[J]. *Nature Reviews Chemistry*, 2018, 2(4): 0121.
- [28] Curtarolo S, Hart G L W, Nardelli M B, et al. The high-throughput highway to computational materials design[J]. *Nature Materials*, 2013, 12(3): 191-201.
- [29] Au H, Alptekin H, Jensen A C S, et al. A revised mechanistic model for sodium insertion in hard carbons[J]. *Energy & Environmental Science*, 2020, 13(10): 3469-3479.
- [30] Kaukonen M, Nieminen R M. Atomic-scale modeling of the ion-beam-induced growth of amorphous carbon[J]. *Physical Review B*, 2000, 61(4): 2806-2811.
- [31] McCulloch D G, McKenzie D R, Goringe C M. Ab initio simulations of the structure of amorphous carbon[J]. *Physical Review B*, 2000, 61(3): 2349-2355.
- [32] Zachariasen W H. The atomic arrangement in glass[J]. *Journal of the American Chemical Society*, 1932, 54(10): 3841-3851.
- [33] Kakinoki J, Komura Y, Ino T. Electron diffraction study of evaporated carbon films[J]. *Acta Crystallographica*, 1960, 13(3): 171-179.
- [34] Polk D E. Structural model for amorphous silicon and germanium[J]. *Journal of Non-Crystalline Solids*, 1971, 5(5): 365-376.
- [35] Wooten F, Winer K, Weaire D. Computer generation of structural models of amorphous Si and Ge[J]. *Physical Review Letters*, 1985, 54(13): 1392-1395.
- [36] Bernal J D. Geometry of the structure of monatomic liquids[J]. *Nature*, 1960, 185(4706): 68-70.
- [37] Finney J L. Random packings and the structure of simple liquids. I. The geometry of random close packings[J]. *Proceedings of the Royal Society of London A: Mathematical and Physical Science*, 1970, 319(1539): 479-493.
- [38] Bennett C H. Serially deposited amorphous aggregates of hard spheres[J]. *Journal of Applied Physics*, 1972, 43(6): 2727-2734.
- [39] Connell G A N, Temkin R J. Modeling the structure of amorphous tetrahedrally coordinated semiconductors. I[J]. *Physical Review B*, 1974, 9(12): 5323-5326.
- [40] Etherington G, Wright A C, Wenzel J T, et al. A neutron diffraction study of the structure of evaporated amorphous germanium[J]. *Journal of Non-Crystalline Solids*, 1982, 48(2-3): 265-289.
- [41] Jiang Z H, Zhang Q Y. The structure of glass: a phase equilibrium diagram approach[J]. *Progress in Materials Science*, 2014, 61: 144-215.
- [42] Thorpe M F, Chubynsky M V, Jacobs D J, et al. Non-randomness in network glasses and rigidity[J]. *Glass Physics and Chemistry*, 2001, 27(2): 160-166.
- [43] Wright A C. The great crystallite versus random network controversy: a personal perspective[J]. *International Journal of Applied Glass Science*, 2014, 5(1): 31-56.
- [44] Polk D E, Boudreaux D S. Tetrahedrally coordinated random-network structure[J]. *Physical Review Letters*, 1973, 31(2): 92-95.
- [45] Steinhardt P, Alben R, Duffy M G, et al. Relaxed continuous-random-network models[J]. *Physical Review B*, 1973, 8(12): 6021-6023.
- [46] Wright A C, Thorpe M F. Eighty years of random networks[J]. *physica status solidi (b)*, 2013, 250(5): 931-936.
- [47] Barkema G T, Mousseau N. High-quality continuous random networks[J]. *Physical Review B*, 2000, 62(8): 4985-4990.
- [48] Lewis L J. Fifty years of amorphous silicon models: the end of the story?[J]. *Journal of Non-Crystalline Solids*, 2022, 580: 121383.
- [49] Zheng Z, Zheng F, Wu Y N, et al. A generalized bond switching Monte Carlo method for amorphous structure generation[J]. *Computational Materials Today*, 2025, 6: 100031.
- [50] Shiell T B, McCulloch D G, Bradby J E, et al. Neutron diffraction discriminates between models for the nanoarchitecture of graphene sheets in glassy carbon[J]. *Journal of Non-Crystalline Solids*, 2021, 554: 120610.
- [51] Sharma S, Shyam Kumar C N, Korvink J G, et al. Evolution of glassy carbon microstructure: in situ transmission electron microscopy of the pyrolysis process[J]. *Scientific Reports*, 2018, 8(1): 16282.

- [52] Chai Y, Guo J, Hong C, et al. Reducing steric hindrance to enhance the oxidation reactivity of coal precursors for high-performance hard carbons in sodium-ion batteries[J]. *Energy Storage Materials*, 2025, 83: 104678.
- [53] Shibazaki Y, Kono Y, Shen G. Compressed glassy carbon maintaining graphite-like structure with linkage formation between graphene layers[J]. *Scientific Reports*, 2019, 9(1): 7531.
- [54] Tian Y R, Yi Z L, Su F Y, et al. Regulating the pore structure of activated carbon by pitch for high-performance sodium-ion storage[J]. *ACS Applied Materials & Interfaces*, 2024, 16(14): 17553-17562.
- [55] Hu M, He J, Zhao Z, et al. Compressed glassy carbon: an ultrastrong and elastic interpenetrating graphene network[J]. *Science Advances*, 2017, 3(6): e1603213.
- [56] Shiell T B, McCulloch D G, McKenzie D R, et al. Graphitization of glassy carbon after compression at room temperature[J]. *Physical Review Letters*, 2018, 120(21): 215701.
- [57] Walter M, Mangolini F, McClimon J B, et al. Origin of C(1s) binding energy shifts in amorphous carbon materials[J]. *Physical Review Materials*, 2025, 9(3): 035601.
- [58] Allen M P. Computer simulation of liquid crystals[M]//Ferrario M, Ciccotti G, Binder K, eds. Computer simulations in condensed matter systems: From materials to chemical biology volume 2. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006: 191-210.
- [59] Bauschlicher C W, Lawson J W. Amorphous carbon and its surfaces[J]. *Chemical Physics*, 2010, 374(1-3): 77-82.
- [60] Zeng Z, Yang L, Zeng Q, et al. Synthesis of quenchable amorphous diamond[J]. *Nature Communications*, 2017, 8(1): 322.
- [61] Liu L, Lu F, Tian J, et al. Electronic and optical properties of amorphous carbon with different sp^3/sp^2 hybridization ratio[J]. *Applied Physics A*, 2019, 125(5): 366.
- [62] Steele B A, Bastea S, Kuo I F W. Ab initio structural dynamics of pure and nitrogen-containing amorphous carbon[J]. *Scientific Reports*, 2023, 13(1): 19657.
- [63] Zheng B, Zheng W T, Zhang K, et al. First-principle study of nitrogen incorporation in amorphous carbon[J]. *Carbon*, 2006, 44(5): 962-968.
- [64] Hong Z, Zhen Y, Ruan Y, et al. Rational design and general synthesis of S-doped hard carbon with tunable doping sites toward excellent Na-ion storage performance[J]. *Advanced Materials*, 2018, 30(29): 1802035.
- [65] Wang Y, Yi Z, Zhang S, et al. Lowering energy barriers of free radicals facilitates defect-suppressed carbon layers of hard carbon[J]. *Small*, 2025, 21(36): e06923.
- [66] Aarva A, Deringer V L, Sainio S, et al. Understanding X-ray spectroscopy of carbonaceous materials by combining experiments, density functional theory, and machine learning. part I: Fingerprint spectra[J]. *Chemistry of Materials*, 2019, 31(22): 9243-9255.
- [67] Yu H, Yi Z, Jia W, et al. Confinement-induced stability evolution of Na clusters in closed pores of hard carbon: a DFT and AIMD study[J]. *The Journal of Physical Chemistry A*, 2025, 129(14): 3231-3241.
- [68] Jia W, Yi Z, Yu H, et al. New insights into the effects of atomic-scale disordered structures in hard carbon anodes on Na diffusion behaviors: a DFT study[J]. *Surfaces and Interfaces*, 2025, 62: 106175.
- [69] Fan Y F, Yi Z L, Zhou Y, et al. The relationship between the high-frequency performance of supercapacitors and the type of doped nitrogen in the carbon electrode[J]. *New Carbon Materials*, 2024, 39(5): 1015-1026.
- [70] Pastewka L, Pou P, Pérez R, et al. Describing bond-breaking processes by reactive potentials: importance of an environment-dependent interaction range[J]. *Physical Review B*, 2008, 78(16): 161402.
- [71] De Tomas C, Aghajamali A, Jones J L, et al. Transferability in interatomic potentials for carbon[J]. *Carbon*, 2019, 155: 624-634.
- [72] Eyert V, Wormald J, Curtin W A, et al. Machine-learned interatomic potentials: Recent developments and prospective applications[J]. *Journal of Materials Research*, 2023, 38(24): 5079-5094.
- [73] Yi Z, Zhou Y, Liu H, et al. Predicting practical reduction potential of electrolyte solvents via computational hydrogen electrode and interpretable machine-learning models[J]. *npj Computational Materials*, 2025, 11(1): 135.
- [74] Frenkel D, Smit B, Ratner M A. Understanding molecular simulation: from algorithms to applications[J]. *Physics Today*, 1997, 50(7): 66-66.
- [75] van Gunsteren W F, Daura X, Hansen N, et al. Validation of molecular simulation: an overview of issues[J]. *Angewandte Chemie International Edition*, 2018, 57(4): 884-902.
- [76] Tersoff J. Empirical interatomic potential for carbon, with applications to amorphous carbon[J]. *Physical Review Letters*, 1988, 61(25): 2879-2882.
- [77] Tersoff J. Modeling solid-state chemistry: Interatomic potentials for multicomponent systems[J]. *Physical Review B*, 1989, 39(8): 5566-5568.
- [78] Sha Z D, Brancio P S, Pei Q X, et al. A modified Tersoff potential for pure and hydrogenated diamond-like carbon[J]. *Computational Materials Science*, 2013, 67: 146-150.
- [79] Stephan U, Haase M. A molecular dynamics study and the electronic properties of amorphous carbon using the Tersoff potential[J]. *Journal of Physics: Condensed Matter*, 1993, 5(49): 9157-9168.
- [80] Lee S H, Lee C S, Lee S C, et al. Structural properties of amorphous carbon films by molecular dynamics simulation[J].

- Surface and Coatings Technology, 2004, 177-178: 812-817.
- [81] Brenner D W, Shenderova O A, Harrison J A, et al. A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons[J]. *Journal of Physics: Condensed Matter*, 2002, 14(4): 783-802.
- [82] Brenner D W. Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films[J]. *Physical Review B*, 1990, 42(15): 9458-9471.
- [83] Stuart S J, Tutein A B, Harrison J A. A reactive potential for hydrocarbons with intermolecular interactions[J]. *The Journal of Chemical Physics*, 2000, 112(14): 6472-6486.
- [84] O'Connor T C, Andzelm J, Robbins M O. AIREBO-M: A reactive model for hydrocarbons at extreme pressures[J]. *The Journal of Chemical Physics*, 2015, 142(2): 024903.
- [85] Bazant M Z, Kaxiras E, Justo J F. Environment-dependent interatomic potential for bulk silicon[J]. *Physical Review B*, 1997, 56(14): 8542-8552.
- [86] Marks N A, Cooper N C, McKenzie D R, et al. Comparison of density-functional, tight-binding, and empirical methods for the simulation of amorphous carbon[J]. *Physical Review B*, 2002, 65(7): 075411.
- [87] Kum O, Ree F H, Stuart S J, et al. Molecular dynamics investigation on liquid-liquid phase change in carbon with empirical bond-order potentials[J]. *The Journal of Chemical Physics*, 2003, 119(12): 6053-6056.
- [88] Marks N A. Generalizing the environment-dependent interaction potential for carbon[J]. *Physical Review B*, 2000, 63(3): 035401.
- [89] Los J H, Fasolino A. Intrinsic long-range bond-order potential for carbon: performance in Monte Carlo simulations of graphitization[J]. *Physical Review B*, 2003, 68(2): 024107.
- [90] Zhou X W, Ward D K, Foster M E. An analytical bond-order potential for carbon[J]. *Journal of Computational Chemistry*, 2015, 36(23): 1719-1735.
- [91] Pettifor D G, Oleinik I I. Analytic bond-order potentials beyond Tersoff-Brenner. I. theory[J]. *Physical Review B*, 1999, 59(13): 8487-8499.
- [92] Srinivasan S G, Van Duin A C T, Ganesh P. Development of a ReaxFF potential for carbon condensed phases and its application to the thermal fragmentation of a large fullerene[J]. *The Journal of Physical Chemistry A*, 2015, 119(4): 571-580.
- [93] Van Duin A C T, Dasgupta S, Lorant F, et al. ReaxFF: A reactive force field for hydrocarbons[J]. *The Journal of Physical Chemistry A*, 2001, 105(41): 9396-9409.
- [94] Liang T, Shan T R, Cheng Y T, et al. Classical atomistic simulations of surfaces and heterogeneous interfaces with the charge-optimized many body (COMB) potentials[J]. *Materials Science and Engineering: R: Reports*, 2013, 74(9): 255-279.
- [95] De Tomas C, Suarez-Martinez I, Marks N A. Graphitization of amorphous carbons: A comparative study of interatomic potentials[J]. *Carbon*, 2016, 109: 681-693.
- [96] Liu Q, Li L, Jeng Y R, et al. Effect of interatomic potentials on modeling the nanostructure of amorphous carbon by liquid quenching method[J]. *Computational Materials Science*, 2020, 184: 109939.
- [97] Zhu Y, Fang Z, Zhang Z, et al. Discontinuous phase diagram of amorphous carbons[J]. *National Science Review*, 2024, 11(4): nwae051.
- [98] Palmer J C, Llobet A, Yeon S H, et al. Modeling the structural evolution of carbide-derived carbons using quenched molecular dynamics[J]. *Carbon*, 2010, 48(4): 1116-1123.
- [99] Anantharaman T R, Suryanarayana C. Review: A decade of quenching from the melt[J]. *Journal of Materials Science*, 1971, 6(8): 1111-1135.
- [100] Zhai M, Yoshioka T, Yang J, et al. Preparation and characterization of amorphous carbon (a-C) membranes by molecular dynamics simulation[J]. *Desalination and Water Treatment*, 2013, 51(25-27): 5231-5236.
- [101] Li L, Xu M, Song W, et al. The effect of empirical potential functions on modeling of amorphous carbon using molecular dynamics method[J]. *Applied Surface Science*, 2013, 286: 287-297.
- [102] Chenoweth K, Van Duin A C T, Goddard W A. ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation[J]. *The Journal of Physical Chemistry A*, 2008, 112(5): 1040-1053.
- [103] Li K, Zhang H, Li G, et al. ReaxFF molecular dynamics simulation for the graphitization of amorphous carbon: A parametric study[J]. *Journal of Chemical Theory and Computation*, 2018, 14(5): 2322-2331.
- [104] Ranganathan R, Rokkam S, Desai T, et al. Generation of amorphous carbon models using liquid quench method: A reactive molecular dynamics study[J]. *Carbon*, 2017, 113: 87-99.
- [105] You Z, Xiao J, Wang G, et al. Molecular representation and atomic-level coking evolution investigation of modified coal tar pitch via ¹³C NMR, MALDI-TOF-MS, SAXS, and ReaxFF MD[J]. *Fuel*, 2023, 348: 128561.
- [106] AlAreeqi S, Bahamon D, Polychronopoulou K, et al. Insights into the thermal stability and conversion of carbon-based materials by using ReaxFF reactive force field: recent advances and future directions[J]. *Carbon*, 2022, 196: 840-866.
- [107] Zhao L, Yuan B, Gong L, et al. Insight into the molecular structure construction and pyrolysis mechanism simulation of typical carbonaceous feedstocks for the rapid design and preparation of porous carbon-based materials[J]. *Fuel*, 2025, 379: 132893.
- [108] Mishin Y. Machine-learning interatomic potentials for materials science[J]. *Acta Materialia*, 2021, 214: 116980.
- [109] Wang G, Wang C, Zhang X, et al. Machine learning interatomic potential: Bridge the gap between small-scale models and realistic device-scale simulations[J]. *iScience*, 2024, 27(5): 109673.
- [110] Deringer V L, Csányi G. Machine learning based interatomic potential for amorphous carbon[J]. *Physical Review B*, 2017, 95(9): 094203.
- [111] Ko T W, Ong S P. Recent advances and outstanding challenges for machine learning interatomic potentials[J]. *Nature Computational*

- Science*, 2023, 3(12): 998-1000.
- [112] Deringer V L, Csányi G, Proserpio D M. Extracting crystal chemistry from amorphous carbon structures[J]. *ChemPhysChem*, 2017, 18(8): 873-877.
- [113] Liu Y, Gao T, Xiao Q, et al. Temperature-density dependent hidden order of amorphous carbon[J]. *Communications Physics*, 2025, 8(1): 420.
- [114] Moon J, Tian Z. Crystal-like thermal transport in amorphous carbon[J]. *npj Computational Materials*, 2025, 11(1): 137.
- [115] Rowe P, Deringer V L, Gasparotto P, et al. An accurate and transferable machine learning potential for carbon[J]. *The Journal of Chemical Physics*, 2020, 153(3): 034702.
- [116] Wang J, Shen H, Yang R, et al. A deep learning interatomic potential developed for atomistic simulation of carbon materials[J]. *Carbon*, 2022, 186: 1-8.
- [117] Fan Z, Zeng Z, Zhang C, et al. Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport[J]. *Physical Review B*, 2021, 104(10): 104309.
- [118] Deringer V L, Caro M A, Csányi G. Machine learning interatomic potentials as emerging tools for materials science[J]. *Advanced Materials*, 2019, 31(46): 1902765.
- [119] Playford H, Keen D, Tucker M. Local structure of crystalline and amorphous materials using reverse Monte Carlo methods[J]. *Neutron News*, 2016, 27(3): 17-21.
- [120] Pikunic J, Clinard C, Cohaut N, et al. Structural modeling of porous carbons: constrained reverse Monte Carlo method[J]. *Langmuir*, 2003, 19(20): 8565-8582.
- [121] Opletal G, Petersen T C, O'Malley B, et al. HRMC: Hybrid reverse Monte Carlo method with silicon and carbon potentials[J]. *Computer Physics Communications*, 2008, 178(10): 777-787.
- [122] Opletal G, Petersen T C, Snook I K, et al. HRMC_2. 0: Hybrid reverse Monte Carlo method with silicon, carbon and germanium potentials[J]. *Computer Physics Communications*, 2013, 184(8): 1946-1957.
- [123] Palmer J C, Gubbins K E. Atomistic models for disordered nanoporous carbons using reactive force fields[J]. *Microporous and Mesoporous Materials*, 2012, 154: 24-37.
- [124] Jain S K, Pellenq R J M, Pikunic J P, et al. Molecular modeling of porous carbons using the hybrid reverse Monte Carlo method[J]. *Langmuir*, 2006, 22(24): 9942-9948.
- [125] Nguyen T X, Cohaut N, Bae J S, et al. New method for atomistic modeling of the microstructure of activated carbons using hybrid reverse Monte Carlo simulation[J]. *Langmuir*, 2008, 24(15): 7912-7922.
- [126] Farmahini A H, Opletal G, Bhatia S K. Structural modelling of silicon carbide-derived nanoporous carbon by hybrid reverse Monte Carlo simulation[J]. *The Journal of Physical Chemistry C*, 2013, 117(27): 14081-14094.
- [127] Surta T W, Koh E, Li Z, et al. Combining experimental and theoretical techniques to gain an atomic level understanding of the defect binding mechanism in hard carbon anodes for sodium ion batteries[J]. *Advanced Energy Materials*, 2022, 12(25): 2200647.
- [128] Bhatia S K. Characterizing structural complexity in disordered carbons: From the slit pore to atomistic models[J]. *Langmuir*, 2017, 33(4): 831-847.
- [129] Gereben O, Pusztai L. Understanding the structure of molecular liquids via combinations of molecular dynamics simulations and reverse Monte Carlo modeling: handling information deficiency[J]. *Journal of Non-Crystalline Solids*, 2015, 407: 213-219.
- [130] McGreevy R L. RMC: progress, problems and prospects[J]. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1995, 354(1): 1-16.
- [131] Tucker M G, Keen D A, Dove M T, et al. RMCProfile: reverse Monte Carlo for polycrystalline materials[J]. *Journal of Physics: Condensed Matter*, 2007, 19(33): 335218.
- [132] Liu C, Zhang Z, Ding J, et al. On the reliability of using reverse Monte Carlo simulations to construct the atomic structure model of metallic glasses[J]. *Scripta Materialia*, 2023, 225: 115159.
- [133] Opletal G, Petersen T, O'Malley B, et al. Hybrid approach for generating realistic amorphous carbon structure using Metropolis and reverse Monte Carlo[J]. *Molecular Simulation*, 2002, 28(10-11): 927-938.
- [134] Yoshimoto M, Ito K, Omote K. Local structure of amorphous carbon investigated by X-ray total scattering and RMC modeling[J]. *Scientific Reports*, 2024, 14(1): 25298.
- [135] Jiang X, Wang W, Tian S, et al. Applications of natural language processing and large language models in materials discovery[J]. *npj Computational Materials*, 2025, 11(1): 79.
- [136] Moro V, Loh C, Dangovski R, et al. Multimodal foundation models for material property prediction and discovery[J]. *Newton*, 2025, 1(1): 100016.