

AI for Sustainable Food Futures: From Data to Dinner

Bianca Datta¹, Markus J. Buehler², Yvonne Chow³, Kristina Gligorić⁴,
Dan Jurafsky⁵, David L. Kaplan⁶, Rodrigo Ledesma-Amaro⁷, Giorgia Del Missier⁷,
Lisa Neidhardt⁷, Karim Pichara⁸, Benjamin Sanchez-Lengeling⁹, Miek Schlangen¹⁰,
Skyler R. St. Pierre¹¹, Ilias Tagkopoulos¹², Anna Thomas⁵, Nik Watson^{13,14},
Ellen Kuhl¹¹

¹ Good Food Institute, Washington, DC 20090, USA.

² Department of Civil & Environmental Engineering; Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

³ Singapore Institute of Food & Biotechnology Innovation SIFBI, Agency for Science, Technology and Research A*STAR, Nanos, 138669, Singapore.

⁴ Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218, USA.

⁵ Department of Computer Science, Stanford University, Stanford, CA 94305, USA.

⁶ Department of Biomedical Engineering, Tufts University, Medford, MA 02155, USA.

⁷ Department of Bioengineering; Bezos Centre for Sustainable Protein; Microbial Food Hub; Centre for Engineering Biology, Imperial College London, London, SW7 2AZ, UK.

⁸ NotCo, San Francisco, CA 94110, USA.

⁹ Department of Chemical Engineering and Applied Chemistry; Vector Institute for Artificial Intelligence, University of Toronto, Toronto, ON M5S 3E5, Canada.

¹⁰ Department of Green Technology, University of Southern Denmark, 5230 Odense, Denmark.

¹¹ Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, USA.

¹² USDA/NIFA AI Institute for Next-Generation Food Systems, University of California, Davis, Davis, CA 95616, USA.

¹³ School of Food Science and Nutrition, University of Leeds, Leeds LS2 9JT, UK.

¹⁴ National Alternative Protein Innovation Centre NAPIC, UK.

Contributing authors: biancad@gfi.org; mbuehler@mit.edu; yvonne.chow@a-star.edu.sg;
gligoric@jhu.edu; jurafsky@stanford.edu; david.kaplan@tufts.edu;
r.ledesma-amaro@imperial.ac.uk; giorgia.del-missier22@imperial.ac.uk; ln327@cam.ac.uk;
karim@thenotcompany.com; ben.sanchez@utoronto.ca; mies@igt.sdu.dk; sstpie@stanford.edu;
itagkopoulos@ucdavis.edu; thomasat@stanford.edu; n.j.watson@leeds.ac.uk;
ekuhl@stanford.edu;

Abstract

Global food systems must deliver nutritious and sustainable foods while sharply reducing environmental impact. Yet, food innovation remains slow, empirical, and fragmented. Artificial intelligence (AI) now offers a transformative path with the potential to link molecular composition to functional performance, bridge chemical structure to sensory outcomes, and accelerate cross-disciplinary innovation across the

entire production pipeline. Here we outline AI for Food as an emerging discipline that integrates ingredient design, formulation development, fermentation and production, texture analysis, sensory properties, manufacturing, and recipe generation. Early successes demonstrate how AI can predict protein performance, map molecules to flavor, and tailor consumer experiences. But significant challenges remain: lack of standardization, scarce multimodal data, cultural and nutritional diversity, and low consumer confidence. We propose three priorities to unlock the field: treating food as a programmable biomaterial, building self-driving laboratories for automated discovery, and developing deep reasoning models that integrate sustainability and human health. By embedding AI responsibly into the food innovation cycle, we can accelerate the transition to sustainable protein systems and chart a predictive, design-driven science of food for our own health and the health of our planet.

Keywords: artificial intelligence, machine learning, human health, planetary health, alternative protein

Motivation

Food is one of humanity's most consequential technologies for human health, planetary sustainability, and cultural identity [1]; yet, research and development for food remains slow, fragmented, and driven by trial-and-error rather than rational design. Radical, urgent, and creative food solutions are needed to combat the current climate crisis [2]. Strikingly, the food sector contributes 35% of the global greenhouse emissions across farmland, livestock, and production [3] and contributes heavily to biodiversity loss and public health challenges [4]. To meet climate goals and feed a growing population, we need alternative sources of protein that are delicious, affordable, scalable, and nutritionally compelling [5]. Today, however, the scientific bottlenecks in texture, flavor, scalability, and cost pose barriers to adoption [6]. Artificial intelligence (AI) has transformed the scientific landscape, entirely reshaping domains from protein design and drug discovery to materials engineering [7]. AI is now poised to revitalize how we grow and make food – from optimizing plant-based meat formulations to transforming menus with climate-conscious recipes [8]. Enabling protein diversification will have profound implications, and AI can substantially accelerate scientific progress if applied responsibly and towards the correct challenges [9]. But this acceleration is not guaranteed. If deployed in a scattershot way, AI risks generating low-quality outputs, increasing environmental and resource burdens, or undermining consumer trust, aspects that are especially sensitive in food [10]. The goal is not to deploy AI indiscriminately. Instead, we need to partner with AI where it measurably reduces time, cost, and uncertainty for problems that matter [11]. Responsible application-focused on the right challenges—is essential. In this perspective, we identify areas across the food innovation where we can constructively engage with AI tools, we provide background

on the current state of the art, outline new opportunities, and discuss the challenges facing AI to accelerate the future of food.

Historical Milestones in AI for Food. The intellectual and technological trajectory that has shaped the emergence of AI for Food as a discipline spans more than seven decades, weaving together advances in artificial intelligence, food science, and biotechnology. From the birth of symbolic AI in the 1950s [12] and the rise of machine learning and neural networks in the 1980s–90s [13], to the deep learning breakthroughs of the 2010s [14], and the advent of generative models [15] and foundation models [16] in the 2020s, AI has progressively expanded its ability to model, design, and predict complex systems. In parallel, developments in genomics [17], computational chemistry [18], fermentation technologies and personalized nutrition [19] created unprecedented data resources and biological insights directly relevant to food design. Their convergence has given rise to a new field in which AI is now central to ingredient discovery, recipe formulation, manufacturing optimization, and consumer personalization [11]. Figure 1 summarizes this evolution and highlights twelve pivotal milestones that trace the path from early AI to today's emerging discipline of AI for Food.

The Promise of AI for Food. Designing food materials poses unique challenges: its properties emerge from multiscale interactions among proteins, carbohydrates, lipids, and their processing conditions [20]. Yet, AI has moved food innovation from slow trial-and-error to data-driven discovery. Figure 2 illustrates four successful AI-powered food categories ranging from meat and dairy alternatives to snacks and condiments. Modern AI models now link chemical composition to sensory outcomes, enabling targeted formulation and rapid iteration. Machine-learning frameworks have predicted consumer appreciation directly from chemical-sensory

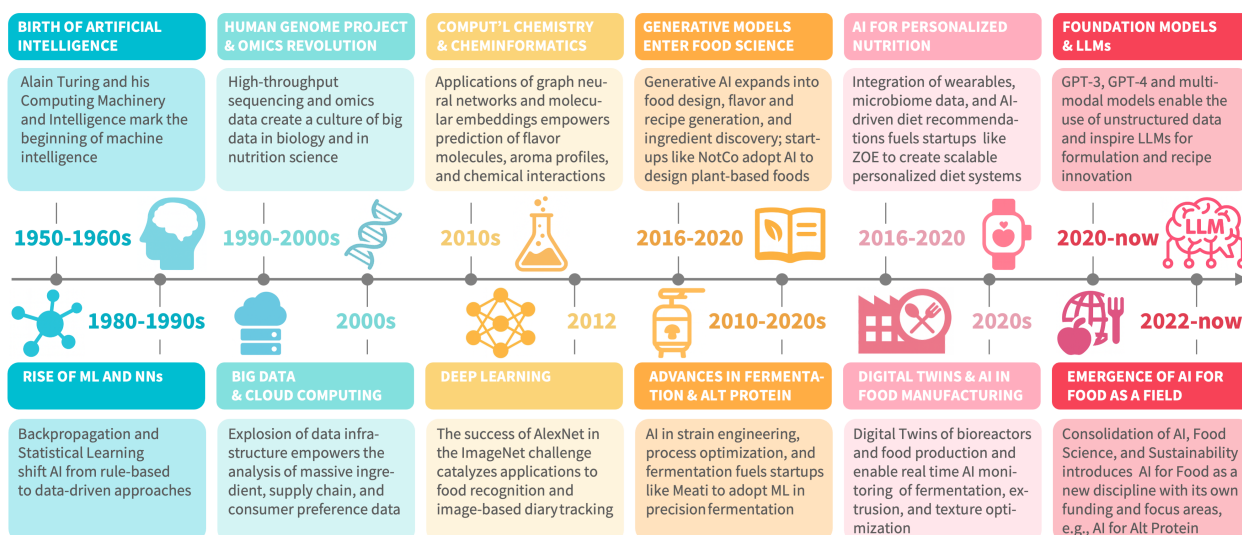


Fig. 1 Historical milestones in AI for Food. Enabling discoveries in artificial intelligence (AI) for food include the birth of artificial intelligence, the rise of machine learning (ML) and neural networks (NN), the era of the human genome project, and the time of big data and cloud computing laid the foundation for AI for food. Since the beginning of the 21st century, developments in computational chemistry, deep learning, generative models, and fermentation have inspired the use of these technologies for food innovation. The current decade is witnessing the first use of artificial intelligence for personalized nutrition, the application of digital twins in food manufacturing, and the use of foundation models and large language models (LLM) for food, which are collectively fueling the emergence of AI for Food as its own scientific discipline.

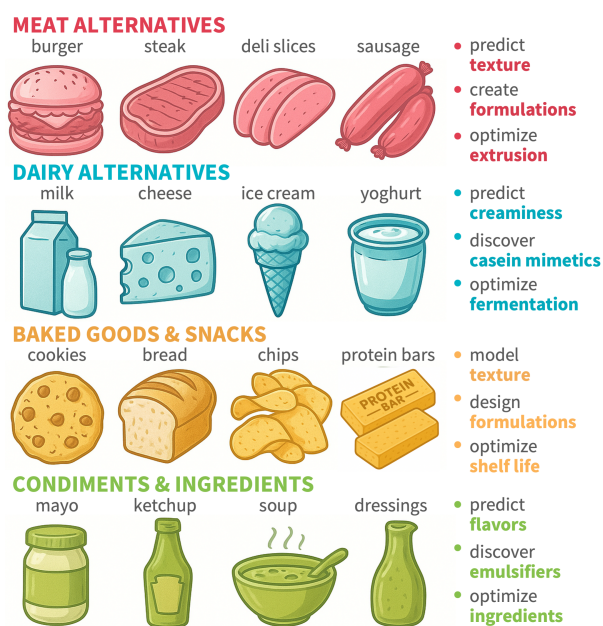


Fig. 2 The promise of AI for Food. AI is already transforming food innovation. Representative AI-powered foods include meat alternatives in pink: burgers, steaks, deli slices, and sausages; dairy alternatives in blue: milk, cheese, ice cream, and yoghurt; baked goods and snacks in orange: cookies, bread, chips, and protein bars; and condiments and ingredients in green: mayonnaise, ketchup, soups, and salad dressings. Together, these categories highlight AI's role in creating healthier, tastier, and more sustainable foods.

panels in beverages [21], while graph neural networks have mapped molecular structure to odor quality, offering generalizable approaches to flavor design [22]. Advances in protein modeling and molecular machine learning accelerate the discovery and engineering of functional ingredients [23, 24], while synthetic-biology-enabled precision fermentation provides scalable routes to novel proteins and textures [25]. Importantly, these innovations begin to address long-standing challenges: creating plant-based burgers and steaks with the fibrous texture of animal muscle, designing plant-based milk and cheese with the emulsification and meltability of dairy proteins, and producing condiments like mayonnaise without egg-derived stabilizers. Crucially, start-ups including NotCo, Perfect Day, Meati, and Climax Foods have translated these methods into marketable foods [26]. This shift from speculative promise to market reality exemplifies how AI now functions as a powerful engine of culinary innovation and sustainability.

The AI-Powered Production Cycle. Our proposed development model for plant-based, fermentation-derived, or cultivated products follows an AI-powered cycle that moves through ingredient design, formulation development, fermentation and production, texture analysis, sensory properties, manufacturing, and recipe generation. Figure 3 illustrates this iterative cycle, in which each step

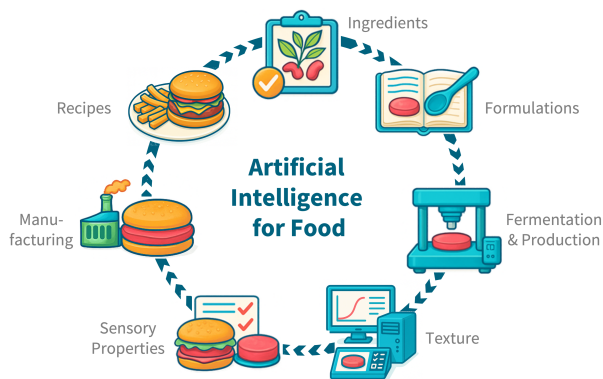


Fig. 3 The AI-powered production cycle. Seven key stages where artificial intelligence can accelerate the time to market: ingredient design, formulation development, fermentation and production, texture analysis, sensory properties, manufacturing, and recipe generation. The example of a plant-based burger illustrates the tight integration of food science, biotechnology, materials science, behavior science, and artificial intelligence to guide the development of plant-based burgers from raw ingredients to commercial products.

presents an opportunity for AI to link experimental data with human perception and create a faster and more tightly connected development process. Traditional trial-and-error methods cannot scale to the complexity of modern food design, but AI can unlock design principles that open entirely new pathways for food innovation. By closing the loop between biology, food science, engineering, and consumer perception, this cycle enables continuous refinement from raw ingredients to commercial launch. In the following sections, we highlight how state-of-the-art approaches, from generative models to digital twins, can drive innovation across every step of this process.

AI for Ingredients

Motivation. Novel food ingredients must balance sustainability, nutrition, sensory appeal, and functionality. Functionality refers to techno-functional traits such as water-holding, emulsification, gelation, viscosity, and stability, which shape processing and final structure. Many novel ingredients are designed to replicate animal-derived functionalities, for example, foaming capacity in egg replacers, yet achieving comparable performance remains a significant challenge. Structure-function relationships, well established in fields such as medical research, link molecular architecture to activity. We propose extending this concept to food innovation, whereby the macromolecular features of proteins, lipids, and fibers encode functionality of the ingredient as a whole. *Can AI bridge this scale gap by translating molecular features and interactions into predictions*

of techno-functional traits, enabling a functionality-first design–build–test–learn cycle?

Background and State of the Art. The success of animal-free ingredients depends heavily on their functionality and the extent to which formulation and consumer requirements can be met [27–29]. These traits arise from molecular features, e.g., sequence and conformation; from environmental factors, e.g., pH, ionic strength, and temperature; and from interactions with fats, fibers, and solutes [28]. Proteins, lipids, and fibres provide the core techno-functional properties, with proteins offering the greatest versatility. Wheat gluten and soy-based proteins currently dominate due to their animal-like functionality. Gluten’s disulfide-linked networks, for example, convert mesoscopic structures into fibrous textures under extrusion and shear [27]. To move beyond these staples, attention is turning to legumes, oilseeds, grains, algae, and leaves [30, 31]. In addition to input sources, processing strongly influences functionality: dry fractionation yields protein-enriched flours with largely native structures, while wet fractionation with isoelectric precipitation produces concentrates and isolates that shape solubility, gelation, and dispersibility [30, 32]. Lipids and fibers provide complementary functionality. Oleosome emulsions enhance interfacial stability [33], while fibres modulate viscosity, emulsions, and lubrication, offering routes to mimic animal fat [34, 35]. However, the inherent heterogeneity of plant-derived inputs complicates predictable performance. To address this, ingredients are often modified chemically, physically, or enzymatically. Increasingly, a shift is underway toward functionality-enriched fractions rather than purified isolates, aiming for more consistent and cost-effective performance [28, 30, 31]. In parallel, bottom-up strategies such as colloid design, tuning droplet size, interfacial layering, and emulsion/foam-gel networks, enable direct control of macroscale texture and stability [36].

Applications and Opportunities. AI and informatics now enable functionality-driven ingredient design by linking molecular and mesoscopic features to macroscopic outcomes under real processing constraints. Classification models support ingredient discovery and repurposing; for example, a quantitative structure–activity relationship model identified glycyrrhizin as a natural emulsifier, later validated experimentally [37]. Regression models enable rapid quality control by predicting pectin viscosity from simple parameters, lowering costs and accelerating formulation adjustments [38]. Functionality coupled to sensory design is increasingly supported by deep learning [39]. One magic

optimized starch-based gels for saltiness perception and gel strength by integrating e-tongue data with oral-processing measurements [40]. These gels are already used to reduce fat and sodium in animal-derived products and offer opportunities for animal-free alternatives. Explainable AI adds interpretability, for instance by showing how ionic conditions shape sesame protein rheology and thereby affect shear stress and viscosity [41]. Multi-objective optimization frameworks now predict solubility, gel strength, and emulsification capacity across diverse crops, providing scalable strategies for formulation [34]. Together, these advances enable closed-loop design–build–test–learn cycles, an iterative process of designing, constructing, testing, and refining ingredients. High-throughput screening, active learning, and digital twins operationalize AI models, while state-of-the-art methods combine latent-space learning such as autoencoders, conditional generation, and constraint-aware optimization to target multiple criteria and bridge lab-to-plate gaps.

Challenges and Open Questions. Ingredient design still struggles with consistent performance across diverse food matrices. Several major bottlenecks remain: first, model transferability poses a major challenge because AI models trained on specific ingredients or conditions often fail to generalize across new food matrices. Second, nonlinear system behavior complicates prediction, as multi-component food systems exhibit emergent interactions that require hybrid approaches combining physics-informed machine learning with empirical data. Third, texture and sensory preferences vary across cultures and regions, which means AI models must be trained on culturally diverse datasets to ensure consumer relevance. Fourth, interpretability and scalability remain essential, since models need to transparently link molecular features such as sequence, charge, or hydrophobicity to techno-functional performance. Resolving these challenges requires open, collaborative tools for functionality-first ingredient design, with the ultimate vision of an AI-powered discovery ecosystem that predicts techno-functional performance from molecular features.

AI for Formulations

Motivation. Ingredient functionality sets the foundation, but effective food design depends on how those ingredients interact in complete systems; formulation brings these components together. Meeting the dual mandate of sustainability and sensory fidelity in alternative proteins requires moving beyond slow, trial-and-error formulation development. AI offers a compelling approach to uncover

hidden patterns that connect ingredients, formulations, production, texture, and sensory properties, and propose candidate formulations that are predicted to meet nutritional, functional, or sensorial targets [42–45]. Today, AI practitioners can draw on a broad collection of public resources including Recipe1M+ [46], FlavorGraph [47], FlavorDB [48], USDA FoodData Central, and FAO/INFOODS, as well as domain-specific databases such as Phenol-Explorer, LipidBank, and the Volatile Compounds in Food database. Inputs for AI models include molecular descriptors, nutritional and functional properties, recipe structures, and process sequences; outputs are complete candidate formulations aligned with target criteria. Figure 4 illustrates state-of-the-art methods that combine latent-space learning through autoencoders with conditional generation, constraint-aware optimization, and expert-in-the-loop refinement. *Given the diversity of raw material information sources needed, can AI capture the complex, non-linear interactions between ingredients, processes, and perception—even under sparse and noisy data—and unlock a new generation of food formulations?*

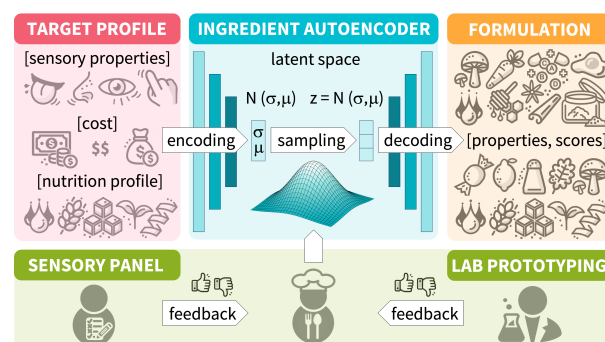


Fig. 4 AI for formulations. An iterative workflow to discover new formulations encodes target profiles such as sensory properties, cost, and nutritional profiles; generates formulation sets with deep learning; predicts properties and scores; and integrates laboratory prototyping and sensory panels in continuous feedback loops.

Background and State of the Art. Computational gastronomy has long shown that recipes and flavor chemistry form structured networks that can be mined for predictive patterns. The Flavor Network formalized ingredient compatibility via shared volatiles [49], while FlavorGraph integrated one million recipes with $\sim 1,500$ flavor molecules to generate graph embeddings for pairing prediction and ingredient clustering [47]. These studies demonstrated that high-dimensional, multimodal food representations can capture the complexity of culinary design. In alternative proteins, industrial

disclosures from NotCo patents [42–45] illustrate hybrid generative optimization pipelines. These systems embed ingredients and recipes in latent spaces using autoencoders and recurrent networks, condition sampling on feature vectors of target product profiles, refine candidates with constraint solvers such as sparse regression, and iteratively incorporate expert sensory feedback. This architecture mirrors multimodal generative trends in other fields and enables analogue discovery. It can identify non-obvious ingredient substitutions while maintaining process feasibility. Other companies, including Cli-max Foods and Eat Just, deploy similar proprietary approaches, while academic research continues to build on Recipe1M+ [46], FlavorGraph [47], and FlavorDB [48]. Together, these efforts highlight a state of the art that blends representation learning with constraint-aware generation to accelerate formulation discovery.

Applications and Opportunities. Generative AI opens new opportunities across protein diversification: It can help design plant-based cheeses that melt and stretch by identifying plant proteins with casein-like properties, since casein is the abundant milk protein that gives animal cheeses their characteristic structure and meltability. It can discover novel ingredient blends to replicate meat textures or pinpoint volatile combinations that evoke familiar flavors. It can brainstorm ideas for improving plant-based meat products in response to sensory panel feedback, and can even match an expert food scientist in predicting sensory panel rankings [98]. Large-scale virtual screenings allow researchers to focus on a handful of promising prototypes rather than having to rely on hundreds of trials, which shortens the production cycle. In the near future, these tools could tailor formulations to individual nutritional needs or taste preferences and unlock the full potential of personalized food design.

Challenges and Open Questions. Significant challenges remain before AI can transform formulation discovery at scale. High-quality sensory datasets are limited, constraining generalization across different food types. Even with detailed chemical and nutritional data, predicting human perception remains an open problem. Modeling ingredient performance under varying heat, shear, or storage conditions is incomplete. Cost, supply chain, allergen control, and clean-label requirements must be built into models as hard constraints. Trust among regulators, industry, and consumers will depend on transparent, auditable design processes. The near-term path is hybrid, where AI proposes and humans refine, toward transparent formulation platforms that integrate constraints and predict performance at scale.

AI for Fermentation & Production

Motivation. While molecular and formulation-level design establishes the building blocks of alternative proteins, realizing their potential requires equally sophisticated control at the cellular and microbial scale. Fermentation is central to protein diversification and sustainable food production, yet its success depends on tightly coupled factors such as strain stability, metabolic fluxes, nutrient composition, and bioreactor dynamics including pH, oxygen, and agitation. The integration of domain knowledge with historical and real-time data through AI-driven strategies holds the potential to overcome technical bottlenecks in optimizing yields with reduced resources. Still, fermentation today often depends on trial and error because datasets across microbial strain engineering, metabolic responses, and upscaling are complex, fragmented, and difficult to integrate. *Can AI fuse these diverse datasets with mechanistic insight to enable predictive, scalable, and cost-effective fermentation and production strategies for alternative proteins?*

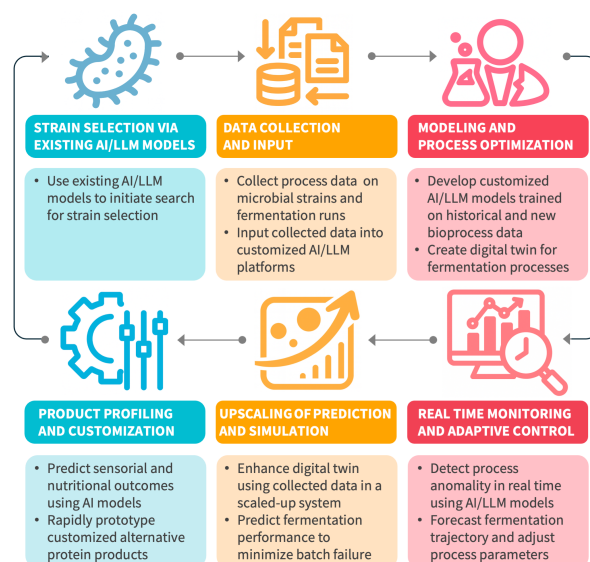


Fig. 5 AI for fermentation and production. Artificial intelligence and large language models can accelerate the fermentation and production cycle during microbial strain selection; data collection; process optimization; real time monitoring and adaptive control; upscaling of prediction and simulation; and product profiling and customization.

Background and State of the Art. Optimizing fermentation production requires a deep understanding of bioprocesses to maximize yield and efficiency while maintaining homogeneity, sterility, and scalability. Kinetic bioprocess models remain widely

used but are constrained by steady-state assumptions, and metabolic flux models for engineered strains often face computational bottlenecks. Experimental optimization through orthogonal design [50] and response surface methodology [51] remains the industry standard, while multiscale feature engineering is still limited [52]. Machine learning is valuable for modeling nonlinear fermentation dynamics [53] and monitoring anaerobic digestion [54], but shallow models lack predictive power. Inline monitoring tools such as biosensors or in situ microscopy remain expensive and difficult to scale [56], and advanced control strategies rarely transition into industrial practice [55]. AI applications to sustainable protein fermentation—including strain engineering, process optimization, and real-time monitoring—are emerging [57], and reinforcement learning has been explored for co-culture optimization [58]. However, large-scale adoption of AI, and especially LLMs, remains limited.

Applications and Opportunities. AI can shift fermentation and production from reactive optimization to predictive management, with customized strains and self-optimizing processes. Machine learning models can optimize strain design, nutrient feeds, and bioreactor control, and deep learning approaches can capture non-linear interactions better than traditional regression [59]. Automation and AI-enhanced imaging [60], soft sensors [61], and spectroscopy [57] can improve monitoring and adaptive control. Algorithms such as support vector machines [62] and fuzzy inference systems [63] support robust predictions, while evolutionary optimization [64] aids process tuning. Hybrid mechanistic-AI models [65] and explainable AI frameworks [66] can improve transparency and decision-making. Training LLMs on unstructured bioprocess data—including text, images, and sensor logs—could support knowledge retrieval, troubleshooting, and experimental design and bridging biology, chemistry, engineering and design through actionable insights, like predicting protein structures or cell morphology [67]. Figure 5 highlights the impact AI and large language models can have during the fermentation and production cycle. Collectively, these advances open opportunities for guided modeling, smart bioreactors, anomaly detection, forecasting, and multi-objective optimization across cost, nutrition, and sensory outcomes.

Challenges and Open Questions. To realize this potential, the field must overcome the key barriers of data scarcity, data integration, and model transferability. Generating large experimental datasets

remains labor intensive, though scaled-down bioreactors and small-data algorithms offer new possibilities. Robust biosecurity and ethical frameworks will be critical for LLM deployment. Integration with genome-scale metabolic models and digital twins of bioreactor performance could create predictive platforms for design and scale-up. *The ultimate goal is an automated, multimodal bioprocessing platform that integrates diverse models to cut time and cost while ensuring reliable, explainable, and human-aligned decision-making.*

AI for Texture

Motivation. Alongside optimized strains and processes, texture is a key quality attribute that shapes consumer acceptance. It arises from a complex interplay of physical, mechanical, and sensory properties and spans multiple scales—from molecular structure to oral processing—making it difficult to quantify and predict. Current methods rely heavily on costly sensory panels and labor-intensive mechanical tests such as cutting tests, tension, compression, and shear tests, food rheology, and texture profile analysis. The challenge is to bridge intuitive descriptors like crispiness, chewiness, or fibrousness with measurable and predictable parameters that ultimately drive product design. *Can AI integrate multimodal datasets—from force-deformation curves and rheological measurements—to predict and ultimately engineer food textures that accelerate the acceptance of sustainable proteins?*

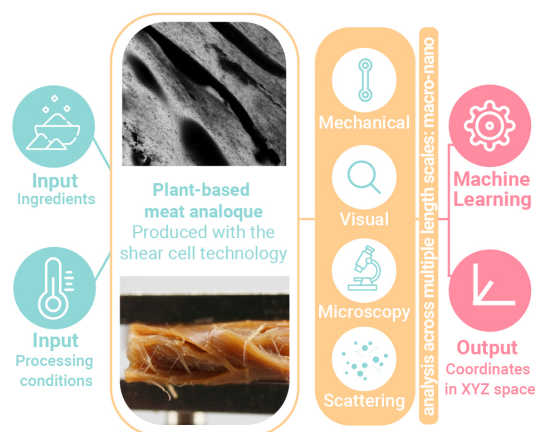


Fig. 6 AI for texture. Machine learning can integrate multimodal data, from ingredients to process conditions, to optimize plant-based meat analogues, from microstructural appearance to macrostructural texture. Future machine learning tools could help us reverse engineer ingredients and process conditions to achieve desired appearance and texture.

Background and State of the Art. Every ingredient and product possesses a characteristic texture, which we can engineer through processing techniques such as extrusion, shear cell structuring, or 3D printing to mimic animal-based foods. Recent advances show the promise of machine learning for decoding texture–structure relationships: constitutive neural networks can map mechanical signatures of plant-based meat alternatives onto interpretable and generalizable physics-based models [68, 69]. Autoencoders can predict sensory texture attributes from rheological data and advance data analysis beyond traditional correlation methods [39]. Classifiers such as support vector machines and neural networks can already identify snack freshness from mechanical and acoustic signals with up to 92% accuracy [70]. Researchers are using machine learning to predict viscosity, gelation, and foaming of plant proteins [34, 71]. Meanwhile, texture analysis from images is rapidly growing. Transfer learning with models like VGGNet, ResNet, and DenseNet allows us to predict the cooking times for fish with potential extension to plant-based foods [72], while ML-driven microscopic image analysis can help us identify droplets in colloidal systems to link microstructure to texture [73]. Beyond academic applications, commercial platforms already harness AI for texture optimization: For example, NotCo [42–45] and Climax Foods successfully deploy AI to mine ingredient databases, combine plant ingredients, and design formulations that replicate animal textures.

Applications and Opportunities. AI accelerates prototyping by predicting textural outcomes before physical trials and reduces time- and resource-intensive trial-and-error cycles. In addition, ingredient design and process optimization can balance target textures with environmental impact, for example, to reduce greenhouse gas emissions. Another important application is personalization, for example designing softer foods for elderly populations to improve swallowing safety. LLMs can also convert consumer feedback such as “too chewy” or “not crispy enough” into quantifiable parameters for product adjustment and bridge the gap between perception and engineering [74]. To address data sparsity, LLMs can generate synthetic training data by pairing realistic texture descriptors with mechanical properties. Ideally, AI can integrate data across multiple length scales, from molecular interactions to macroscopic structure, to capture the hierarchical origins of texture. Figure 6 illustrates how machine learning can integrate multimodal data, from ingredients to process conditions, to predict outcomes across the scales, from

microstructural appearance to macrostructural texture. Inversely, future machine learning tools could help us reverse engineer ingredients and process conditions to achieve a desired texture. More broadly, generative design tools, similar to those used for inorganic materials [7], hold the potential to propose ingredient–process combinations that meet desired textural constraints.

Challenges and Open Questions. Significant obstacles remain before AI-driven texture engineering will truly become routine. The largest bottleneck is the lack of large, open datasets that couple textural, sensory, and processing data. Industrial data sharing also remains limited. Data fusion remains technically challenging, as mechanical, rheological, visual, and compositional data are inherently multimodal and often unstructured. Foundation models trained on multiple modalities may help, but progress requires standardized protocols for measurement, evaluation, and reporting. To succeed, the field must establish robust multimodal datasets, shared standards and interoperable platforms. *The ultimate vision is an AI system for inverse design—with a specific target texture as input and optimal ingredients and processes as output—that transforms texture development into a predictive, model-guided science.*

AI for Sensory Properties

Motivation. Ultimately, food is not only a material design challenge, but a sensory experience. Aroma, flavor, and texture are the strongest determinants of consumer acceptance. The central challenge is to map measurable physical and chemical properties of ingredients to the multidimensional space of human perception. Machine learning offers a powerful framework to learn this mapping, enabling both the prediction of sensory profiles and the computational design of novel ingredients with desired characteristics. *How far are AI models from accurately predicting human sensory perception, and from enabling the design of novel flavor and aroma experiences?*

Background and State of the Art. Traditional sensory science has relied on human panels for perceptual evaluation and on analytical instruments such as gas chromatography–mass spectrometry for chemical composition. Panels remain the gold standard but are slow and expensive, while instruments offer precision without perceptual insight. Early machine learning approaches combined biomimetic sensors, such as E-noses, with classifiers [75], but required heavy feature engineering and were limited to narrow classification tasks [76]. The field

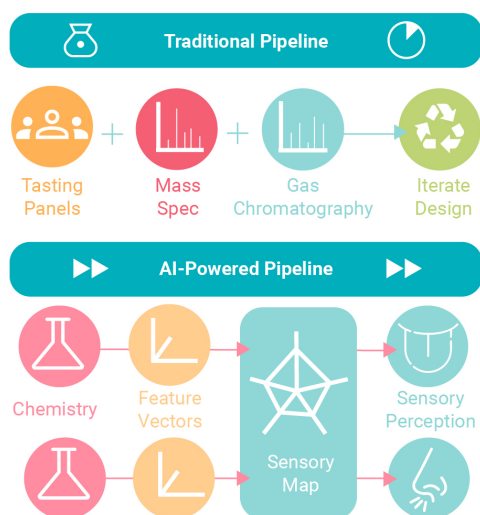


Fig. 7 AI for sensory properties. Optimizing the sensory perception of a food product is traditionally a costly, time-consuming process involving expert tasting panels and instrumental analysis including mass spectrometry and gas chromatography. In contrast, an AI-powered sensory pipeline has the potential to predict people's sensory perceptions from only the chemical or molecular make-up of the food, eliminating the need for costly human panels and vastly reducing time spent iterating product formulations.

has since shifted toward deep representation learning [77], which learns general-purpose sensory maps directly from raw data. The Principal Odor Map exemplifies this trend: a graph neural network organizes molecules into a vector space where distance encodes perceptual similarity, achieving human-level accuracy in odor labeling [22]. Researchers employ attention-based architectures to extend this mapping to molecular mixtures [78, 79]. Figure 7 highlights these advances: Algorithmic tools for predictive sensory modeling are maturing rapidly, but progress is constrained by the scarcity and cost of high-quality perceptual datasets.

Applications and Opportunities. Learned sensory representations open powerful opportunities for food innovation. High-throughput virtual screening of molecular libraries can identify novel flavor compounds for plant-based meats, accelerating product development [11]. Integrating predictive models with low-cost, miniaturized sensors enables real-time quality monitoring in manufacturing and packaging [80], for example with wireless freshness indicators that reduce food waste [81]. These tools also provide systematic frameworks for biomimicry, allowing formulators to predict how ingredient or process modifications alter the final sensory profile. Still, a feedback loop remains essential: computational predictions must be continuously validated and refined with real-world experiments to ensure consumer relevance.

Challenges and Open Questions. Today, the bottleneck is no longer algorithmic capacity but high-quality data acquisition [82, 83]. Deep learning models require large multimodal datasets, yet sensory panels cannot generate data at the necessary scale to capture the vast chemical space of food. The path forward lies in closing the loop between modeling and experiment. Generative models such as transformers and diffusion architectures are algorithmically ready, but their success depends on new data pipelines. Developing robust, low-cost hardware and embedding it into automated platforms will be critical to unlock the full potential of AI for sensory science. The ultimate vision is a Self-Driving Lab that unites synthesis, characterization, and machine learning into platforms to autonomously generate and analyze multimodal data and accelerate de novo food design.

AI for Manufacturing

Motivation. Even the most well-tuned designs will fail if they cannot be produced consistently and affordably. Manufacturing convincing meat analogues is inherently complex because technical goals such as texture, flavor, appearance, and shelf life interact non-linearly with environmental and economic constraints. Relevant datasets span ingredient composition and structure, rheology, thermal behavior, processing steps such as extrusion or fermentation, as well as telemetry, sensory, and consumer data. Through digital twins, computer vision, and other modes, AI is emerging as a key technology to integrate these multimodal datasets and stream-line process optimization. Typical model inputs include formulation and process profiles, while out-puts can range from texture, color, and flavor to cost and environmental impact. How can we ensure that AI-driven manufacturing pipelines deliver reliable, scalable, and cost-effective production of alternative proteins without compromising sensory quality or sustainability goals?

Background and State of the Art. AI and ML are deployed across the manufacturing pipeline—from formulation and process design to in-line quality control. For example, high-moisture extrusion can generate fibrous, anisotropic structures that mimic muscle tissue, but outcomes are highly sensitive to process parameters such as moisture, temperature, and screw configuration. Bayesian optimization and other data-driven strategies explore this parameter space more efficiently than traditional trial-and-error [85, 86]. Researchers can use informatics to link protein structure and functionality, for example, gelation, emulsification, water or fat

binding, to downstream performance. This ultimately allows designers to move beyond commodity proteins, but economic viability remains a major hurdle. The current price gap may be addressed with higher throughput, improved yields, cheaper inputs, and greater capacity utilization [87]. Meanwhile, digital twins show promise for soft sensing, scenario testing, and closed-loop control, though most food-sector deployments remain at pilot scale [88]. Data fragmentation persists as a key obstacle: measurements are dispersed across proprietary plant logs, supplementary materials, and discipline-specific repositories with inconsistent metadata. FAIR data principles tailored to agri-food [89], coupled with large language models for literature-scale information extraction [90], offer pathways to build structured knowledge graphs that connect formulation–process–structure–sensory chains for meat analogs.

Applications and Opportunities. AI is opening multiple opportunities for advancing alternative protein manufacturing. Engineers can use machine learning and Bayesian optimization to tune high-moisture extrusion parameters to target fibrousness, yield, and energy efficiency, while digital twins provide what-if simulations for scale-up and enable real-time sensing [84, 85, 87]. Food scientists can use informatics pipelines to rank plant proteins and design blends at scale using rheology–functionality mappings and literature-derived property tables, allowing ingredient discovery to expand well beyond commodity sources [86, 89]. Large language models further act as formulation copilots, transform unstructured texts into structured knowledge graphs, and suggest experiments to resolve uncertainty [89]. At the same time, computer vision and hyperspectral analytics can advance in-line quality assurance, linking surface color and structure to sensory proxies to enable continuous assessment during production [87]. Together, these approaches illustrate how AI can accelerate prototyping, reduce waste, and bridge the gap between laboratory development and industrial manufacturing.

Challenges and Open Questions. Scaling these tools requires high-quality, shareable datasets that link formulations, processes, structures, and sensory outcomes with standardized metadata. Proprietary and fragmented data currently hinder model transfer and benchmarking, underscoring the need for domain-specific FAIR implementations, shared testbeds, and community baselines [88]. Ingredient variability remains a crucial bottleneck—robust hybrid models are needed to achieve reliable process control [84, 85]. Industrial adoption also demands

capital investment, workforce training, and validation on real production lines—most food digital twins have not been tested at scale [87]. The ultimate vision is to establish data standards and LLM-enabled knowledge base that make AI-assisted formulation and control routine in manufacturing for plants to deliver consistent sensory performance at—or below—price parity with conventional meat, with lower cost, waste, and energy demand.

AI for Recipes

Motivation. Beyond production, food preparation determines consumer experience of alternative proteins. When considering this, both individuals and foodservice operations must balance enjoyment with health, cost, convenience, and allergens, while advancing broader goals like protein diversification and sustainability. Large Language Models are promising aids, as they are already trained on vast collections of recipes and can integrate data from sources such as Recipes1M, USDA FoodData Central, the HESTIA climate database, and Food.com reviews. How can we best integrate human knowledge into AI systems to ensure that recipe models generalize across diverse groups of users, dietary needs, and cultural traditions?

Background and State of the Art. Recent research has evaluated and refined LLMs for recipe-related tasks, including generation, revision, and preference prediction. Improvements rely on several strategies: fine-tuning with domain-specific datasets, retrieval-augmented generation that combines LLMs with external knowledge at inference time, and integration with optimization algorithms or external tools. A study benchmarking models on the Recipe1M dataset found off-the-shelf LLMs unreliable due to hallucinations [91], echoing anecdotal reports [92, 93]. However, fine-tuning on Recipe1M produced state-of-the-art results [91, 94], and supervised fine-tuning with direct preference optimization has enabled recipe revision tasks [95]. Other work demonstrates the benefits of retrieval-augmented generation for recipes [96, 97], while one study directly targeted sustainable menu design [98]. Encouragingly, some LLMs now match expert food scientists in preference modeling for alternative protein applications [98]. Combining these models with combinatorial optimization algorithms has the potential to generate recipes and menus that satisfy nutritional and sustainability goals while complementing expert chefs.

Applications and Opportunities. The ability to generate recipes under explicit constraints creates opportunities across individual and institutional

contexts. For individuals, fine-tuned LLMs could power AI health coaches that teach users how to modify meals to be more sustainable, nutritious, or culturally appropriate, while respecting restrictions such as allergies or religious requirements. Chefs and students could practice recipe generation in educational applications, receiving AI-driven feedback to refine their skills [99]. In institutional settings, LLMs could support foodservice operations in offices, universities, or hospitals, tailoring descriptions, ingredients, and preparation methods under strict constraints. This is particularly valuable in programs such as “Food as Medicine,” which deliver cost- and time-sensitive meals for patients with chronic conditions after hospital discharge [100]. In all these scenarios, AI could accelerate constrained recipe generation, broaden access to healthy and sustainable foods, and scale public health interventions that rely on tailored meal planning.

Challenges and Open Questions. Despite this promise, important challenges remain. Off-the-shelf LLMs perform unevenly across cultural contexts, often fail to capture diverse user preferences, and can generate outputs that are numerically inconsistent, unsafe, or infeasible. They also lack molecular-level understanding of food and have limited knowledge of food safety. Training LLMs from scratch remains prohibitively resource-intensive [101]. These shortcomings raise risks of culturally inappropriate, impractical, or even dangerous recipes, while incurring environmental costs. Mitigation strategies include equipping LLMs with external datasets, relying on open-source pre-trained models, and carefully integrating optimization tools. Beyond technical concerns, broader societal impacts must be addressed: LLMs should promote user agency and creativity rather than replace them, and they should be designed to complement, not displace, human expertise. The vision is a human-centered approach in which LLMs assist with recipe drafts and revisions while empowering users to adapt meals to their own needs, preferences, and cultures, ultimately supporting sustainable and healthy food choices.

Summary and Outlook

AI is beginning to transform food science across the entire innovation cycle in Figure 3—from ingredient design, formulation development, and fermentation and production to texture analysis, sensory properties, manufacturing, and recipe generation. Yet, the field’s true potential lies not only in optimizing today’s processes, but also in charting a bold trajectory for the future of food as a programmable, predictive science. Building on recent advances, we

envision three major areas that define the frontier of AI for sustainable food production:

Future Materials for Food. Treating food as a programmable biomaterial opens new opportunities to design ingredients from first principles. Figure 8 highlights how properties such as texture, functionality, stability, and sensory perception emerge from multiscale interactions among proteins, carbohydrates, lipids, and their processing conditions [20]. We may be able to optimize edible films and gels

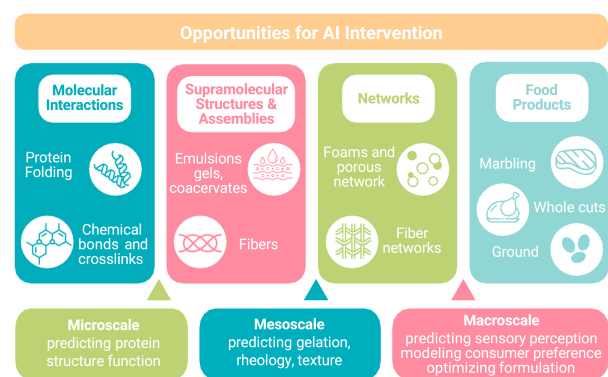


Fig. 8 Future Materials for Food. Treating food as a programmable multiscale material—from molecular interactions, supramolecular structures and assemblies, and networks to the final food product—opens new opportunities to design ingredients from first principles. In this biomaterials approach, properties such as texture, functionality, stability, and sensory perception emerge from multiscale interactions among proteins, carbohydrates, lipids, and their processing conditions.

for controlled nutrient release; or engineer plant-based meat analogues with tailored mechanical and sensory performance to endow such new components with both exceptional sensory experiences and nutritional value. Advances in protein design, including AlphaFold and diffusion-based generative models [102], suggest the possibility of de novo edible proteins with tunable performance. Early examples from cellular agriculture—such as engineered porcine fat cells with modified fatty acid composition [103, 104] or bovine muscle cells enriched in antioxidants or vitamins [105]—demonstrate the promise, while microbial and fungi-based fermentation highlight a complementary path. Mycelium fermentation, for instance, can upcycle agricultural side-streams such as brewers’ spent grain or fruit pomace, converting low-value byproducts into high-quality protein and fiber-rich ingredients with unique textures and link innovation with circularity and sustainability [106]. An opportunity lies ahead in integrating molecular design with tunable processing, including fermentation and upcycling, to create

predictive platforms that generate new food materials with tailored properties while reducing waste.

Automation and Self-Driving Labs. Empirical trial-and-error approaches that use rheological measurements, sensory panels, and compositional analysis have traditionally dominated food science [107]. These methods, while valuable, are slow, costly, and difficult to scale. Generative AI, robotics, and high-throughput experimentation point toward self-driving labs that autonomously design, test, and refine food materials. This embeds the design-build-test-learn principle directly into experimentation, enabling systematic exploration across scales and accelerating innovation. Recent advances in automated model discovery with custom-designed constitutive neural networks, show how AI can autonomously identify the best models, parameters, and even experimental protocols directly from data [108]—a method already successfully applied to compare plant-based and animal meats [69], deli products [109], and fungi-based steak [106]. Agentic AI systems such as Sparks [110], BioDiscoveryAgent [111], and the Virtual Lab [112] already demonstrate the capacity to generate hypotheses, perform simulations, and report results without human intervention. In food, such systems could integrate genome-scale metabolic models, digital twins of bioreactors, and in-line quality sensors [67, 113]. *A fully automated discovery pipeline that links synthesis, characterization, and modeling has the potential to accelerate innovation in sustainable, personalized, and scalable food production.*

Deep Reasoning Models. AI models that can reason will become a key part of self-driving lab infrastructure. Figure 9 highlights our vision for next-generation AI systems for food. We anticipate that the emerging discipline of AI for Food will benefit from ongoing rapid progress in general-purpose reasoning models as well as agentic systems that can autonomously design and perform tasks with the aid of additional tools [114]. While current generative models can propose new molecules or recipes, they remain limited by static datasets and lack the ability to reason about feasibility, manufacturability, or emergent properties. Bridging this gap requires an AI system that can generate new protein sequences, simulate their folding and mechanical behavior, and reason about emergent sensory properties, e.g., flavor, taste - all within a self-improving, interpretable loop. Deep reasoning models represent the next frontier: AI systems that move beyond pattern recognition to actively hypothesize, simulate, and generate new data. These models would integrate protein sequences, structural mechanics, metabolic fluxes, and sensory outcomes, reasoning

across scales to predict how molecular composition translates into taste, texture, and stability [20, 115]. Crucially, they could bridge the simulation-to-reality gap by iteratively validating predictions against experimental data and refining models in real time. These approaches could leverage auto-mated model discovery to identify the best experiments and models for alternative protein products [68], seamlessly integrate the results into finite element simulations [116] to simulate real-life chewing processes and aromas from cooking, and predict sensory perception directly from physical principles [69]. By embedding sustainability constraints, cultural diversity, and personalized health data, such systems could design foods that are not only functional but also socially and nutritionally relevant. The ultimate vision is a closed-loop, reasoning-based AI that autonomously proposes, evaluates, and refines edible materials—transforming food science into a predictive, design-driven discipline for future foods.

From Vision to Implementation. To translate the emerging opportunities of AI-enabled food systems into lasting impact, the community must unite around shared goals, standards, and responsibilities. Specifically, we will need to: i) establish common standards and open data repositories for the field with aligned incentives, security, and risk mitigation; ii) overcome the language barrier between wet-lab and computational scientists; iii) develop clear evaluation frameworks to match AI tools to food challenges and measure outputs; and iv) build consumer trust and prepare for evolving regulatory landscapes. But this field can only progress alongside parallel advancements and such a paradigm shift will require support across all sectors: Researchers must push real-time imaging, spectroscopy, and miniaturized hardware forward to generate high-quality, high-throughput data streams. Engineers and scientists need to expand automation and robotics so that self-driving laboratories and closed-loop cycles become practical. Developers of AI must improve compute sustainability to ensure models run energy-efficiently and align with environmental goals. Computational social scientists must work to understand the impact of current sustainable proteins and guide AI-enabled sustainable protein developers toward sensory and nutritional profiles and educational messages likely to increase displacement effects. Policymakers, ethicists, and scientists together must build governance frameworks in ethics, regulation, and cultural inclusivity so that AI-enabled food systems earn consumer trust and deliver equitable benefits across geographies. Funders and institutions must invest in

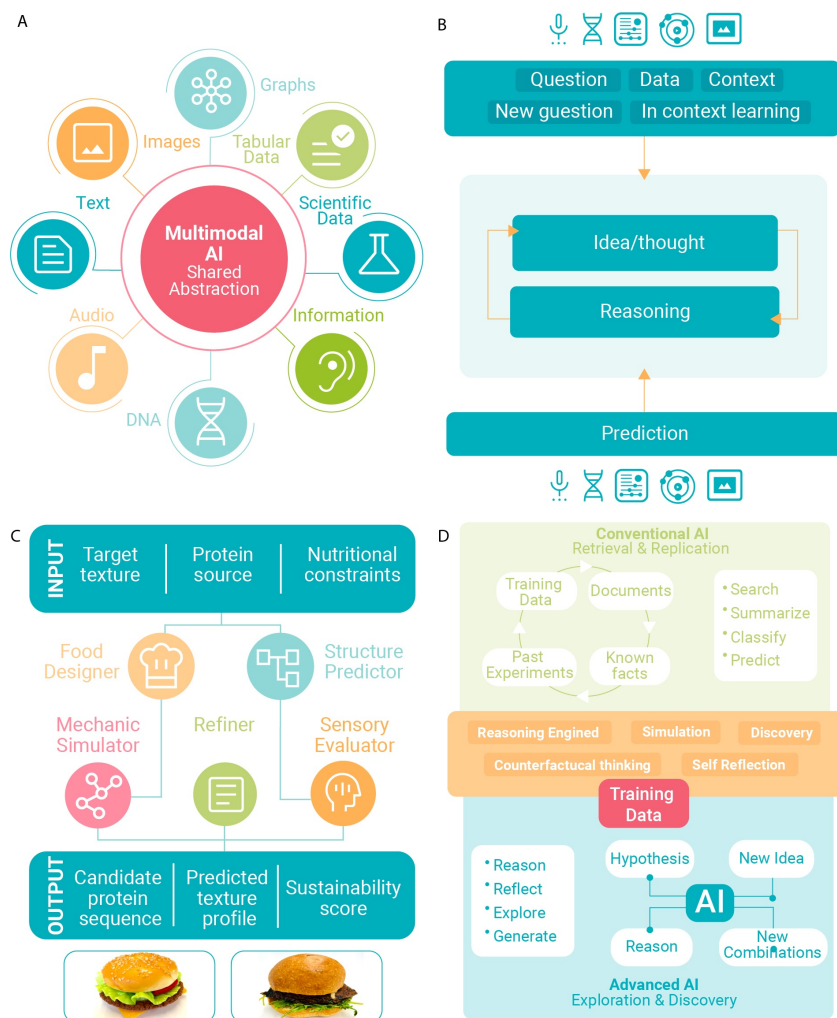


Fig. 9 Next-generation AI systems for scientific discovery in food. (A) Multimodal AI integrates diverse data types such as text, images, graphs, and molecular structures; (B) first-principles AI reasoning models perform iterative thinking and hypothesis generation; (C) a possible workflow applies these capabilities to edible protein design and predicts structure, mechanics, and sensory outcomes by explicitly modeling disparate concepts and their relationships; and (D) the transition from conventional retrieval-based AI to reasoning-enabled, self-reflective AI enables exploration, discovery, and ultimately, the generation of novel, functional foods.

the needed infrastructure—shared repositories, standards, and sensing and automation hardware—to support the interdisciplinary centers and fellowships needed to carry out this work. Together, we must ground this field in responsible innovation by designing evaluation frameworks, building regulatory clarity, and embedding transparency, interpretability, and inclusivity. These infrastructural investments echo the lessons from drug discovery and protein design, where progress accelerated only once community benchmarks, standards, and testbeds were in place. Moving forward, we need to think critically about engaging AI, recognizing the environmental footprint, and ensuring that automation responsibly complements existing scientific endeavors.

Outlook. Taken together, these three areas sketch a future where AI enables food innovation at unprecedented speed and precision. From molecular design of new ingredients, to automated laboratories that close the loop between modeling and experiment, to reasoning-based systems that integrate cultural, nutritional, and sensory dimensions, AI has the potential to reshape how we design, produce, and consume food. Realizing this vision will require interdisciplinary collaboration, open data sharing, and the creation of inclusive platforms that balance automation with human creativity and oversight. If successful, AI will not simply optimize food—it will redefine the future of food as a programmable, predictive, and deeply human-centered science for a sustainable global future.

Acknowledgments We thank Dr. Heideh Fattaey for stimulating discussions on the future of food, Ryan Yow for helping to design Figure 5, and Goran Atanasovski for illustrating Figures 6 to 9. This project was supported by the USDA Grant FA9550-23-1-0606 to DLK, by the BBSRC Grant BB/Y008510/1, the ERC Grant DEUSBIO-949080, and the Bezos Earth Fund through the Bezos Centre for Sustainable Protein BCSP/IC/001 to RLA, by Novo Nordisk Foundation grant NNF23OC0085919 to MS, by the NSF Graduate Research Fellowship to SRSP, by the UK National Alternative Protein Innovation Centre NAPIC, an Innovation and Knowledge Centre funded by the Biotechnology and Biological Sciences Research Council BBSRC and Innovate UK Grant BB/Z516119/1 to NW, and by the Food@Stanford Snack Grant, the NSF CMMI Award 2320933, and the ERC Advanced Grant 101141626 to EK.

Conflicts of interest All authors declare no financial competing interests.

Availability of data and materials All data and materials will be made available upon request.

Authors' contributions BD and EK designed the layout, BD, MJB, YC, KG, DJ, DLK, RLA, GDM, LN, KP, BSL, MS, SRSP, IT, AT, NW, and EK wrote and edited the paper.

References

- [1] Lappé, F.M. Diet for a Small Planet. Ballantine Books, New York (1971).
- [2] Clark, M. A., et al. Global food system emissions could preclude achieving the 1.5 and 2C climate change targets. *Science*. **370**, 705-708 (2020).
- [3] Xu, X. et al. Global greenhouse gas emissions from animal-based foods are twice those of plant-based foods. *Nature Food*. **2**, 724-732 (2021).
- [4] Keasing, F. Diet for a small footprint. *Proc. Nat. Acad. Sci.* **119**, e2204241119 (2022).
- [5] Friedrich, B. Transforming a 12,000-year-old technology. *Nature Food* **3** 807-808 (2022).
- [6] Barabasi, A.L. et al. The unmapped chemical complexity of our diet. *Nature Food* **1** 33-37 (2020).
- [7] Zeni, C. et al. A generative model for inorganic materials design. *Nature*. **639** 624-632 (2025).
- [8] Al-Sarayreh, M. et al. Inverse design and AI/Deep generative networks in food design: A comprehensive review. *Trends Food Sci. Tech.* **138** 215-228 (2023).
- [9] Datta, A. et al. Computer-aided food engineering. *Nature Food*. **3** 894-904 (2022).
- [10] King, A. Four ways to power-up AI for drug discovery. *Nature*. doi: 10.1038/d41586-025-00602-5 (2025).
- [11] Kuhl, E. AI for food: accelerating and democratizing discovery and innovation. *npj Sci. Food*. **9**, 82 (2025).
- [12] Turing, A. M. Computing machinery and intelligence. *Mind*. **LIX**, 433-460 (1950).
- [13] Rumelhart, D. E. et al. Learning representations by back-propagating errors. *Nature*. **323**, 533-536 (1986).
- [14] Krizhevsky, A. et al. ImageNet classification with deep convolutional neural networks. *Comm. ACM* **60**, 84-90 (2017).
- [15] Goodfellow, I. et al. Generative adversarial networks. in: *Advances in Neural Information Processing Systems (NeurIPS)*. **27**, 2672-2680 (2014).
- [16] Vaswani, A. et al. Attention is all you need. in: *Advances in Neural Information Processing Systems (NeurIPS)*. **30**, 5998-6008 (2017).
- [17] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome.. *Nature* **409**, 860-921 (2001).
- [18] Gilmer, J. et al. Neural message passing for quantum chemistry. in: *Proc. 34th Int. Conf. Machine Learning (ICML)*. **34**, 1263-1272 (2017).
- [19] Zeevi, D. et al. Personalized nutrition by prediction of glycemic responses. *Cell* **163**, 1079-1094 (2015).
- [20] Gordon, E. B. et al. Biomaterials in cellular agriculture and plant-based foods for the future. *Nat. Rev. Materials* **10**, 500-518 (2025).
- [21] Schreurs, M. et al. Predicting and improving complex beer flavor through machine learning. *Nat. Comm.* **15**, 2368 (2024).
- [22] Lee, B.-K. et al. A principal odor map unifies diverse tasks in olfactory perception. *Science*. **381**, 999-1006 (2023).
- [23] Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. **596**, 583-589 (2021).
- [24] Butler, K. T. et al. Machine learning for molecular and materials science. *Nature*. **559**, 547-555 (2018).

- [25] Graham, A. E. & Ledesma-Amaro, R. The microbial food revolution. *Nat. Comm.* **14**, 2231 (2023).
- [26] Lurie-Luke, E. Alternative protein sources: science-powered startups to fuel food innovation. *Nat. Comm.* **15**, 4425 (2024).
- [27] Kyriakopoulou, K. et al. Plant-based meat analogues. in: *Sustainable Meat Production and Processing*, Academic Press, 103–126 (2019).
- [28] Loveday, S. M. Food proteins: Technological, nutritional, and sustainability attributes of traditional and emerging proteins. *Ann. Rev. Food Sci. Tech.* **10**, 311–339 (2019).
- [29] Sha, L. & Xiong, Y. L. Plant protein-based alternatives of reconstructed meat: Science, technology, and challenges. *Trends Food Sci. & Tech.* **102**, 51–61 (2020).
- [30] Day, L. Proteins from land plants—Potential resources for human nutrition and food security. *Trends Food Sci. & Tech.* **32**, 25–42 (2013).
- [31] Fasolin, L. H. et al. Emergent food proteins—Towards sustainability, health and innovation. *Foods*. **10**, 600 (2021).
- [32] Haque, M. A. et al. Food proteins, structure, and function. in: *Reference Module in Food Science.*, Elsevier (2016).
- [33] McClements, D. J. Development of next-generation nutritionally fortified plant-based milk substitutes: Structural design principles. *Foods*. **9**, 421 (2020).
- [34] Lie-Piang, A. et al. Quantifying techno-functional properties of ingredients from multiple crops using machine learning. *Curr. Res. Food Sci.* **7**, 100601 (2023).
- [35] Siddiqui, H. et al. A review of the health benefits, functional properties, and ultrasound-assisted dietary fiber extraction. *Bioact. Carbohyd. Diet. Fibre.* **30**, 100356 (2023).
- [36] Dickinson, E. Colloids in food: Ingredients, structure, and stability. *Ann. Rev. Food Sci. Tech.* **6**, 211–233 (2015).
- [37] Liu, M.-Q. et al. Digging natural emulsifiers based on machine learning and exploration their performance for stabilizing dairy products. *SSRN*. <https://ssrn.com/abstract=4994781> (2025).
- [38] Siejak, P. et al. The prediction of pectin viscosity using machine learning based on physical characteristics—case study: Aglupectin HS-MR. *Sustainability*. **16**, 5877 (2024).
- [39] Kraessig, P. M. et al. Sensory-biased autoencoder enables prediction of texture perception from food rheology. *Food Res. Int.* **205**, 116007 (2025).
- [40] Meng, Y. et al. Advancing salt reduction technologies: AI-assisted structural design of starch-based emulsion gel systems for next-generation low sodium food formulations. *Trends Food Sci. & Tech.* doi:10.1016/j.tifs.2025.105234 (2025).
- [41] Yilmaz, M. T. et al. Explainable AI-driven evaluation of plant protein rheology using tree-based and Gaussian process machine learning models. *Ain Shams Eng. J.* **16**, 103565 (2025).
- [42] NotCo. Latent space method of generating food formulas. *US 10,915,818*. Feb 9, 2021.
- [43] NotCo. Neural network method of generating food formulas. *US 10,957,424*. Mar 23, 2021.
- [44] NotCo. Systems and methods to mimic target food items using artificial intelligence. *US 11,164,478*. Nov 2, 2021.
- [45] NotCo. Controllable formula generation. *US 11,164,069*. Nov 2, 2021.
- [46] Martin, J. et al. Recipe1M+: A dataset for learning cross-modal embeddings for cooking recipes and food images. *IEEE Trans. Pattern Anal. Mach. Intell.* **43** (2021).
- [47] Park, D. et al. FlavorGraph: a large-scale food-chemical graph for generating food representations and recommending food pairings. *Sci. Rep.* **11**, 931 (2021).
- [48] Garg, N. et al. FlavorDB: a database of flavor molecules. *Nucl. Acids Res.* **46**, D1210–D1216 (2018).
- [49] Ahn, Y.-Y. et al. Flavor network and the principles of food pairing. *Sci. Rep.* **1**, 196 (2011).
- [50] Feng, X. et al. Enhanced lipid production by *Chlorella pyrenoidosa* through magnetic field pretreatment of wastewater and treatment of microalgae-wastewater culture solution: Magnetic field treatment modes and conditions. *Bioresour. Technol.* **206**, 123102 (2020).
- [51] Sharma, D. et al. Response surface methodology and artificial neural network modelling for enhancing maturity parameters during vermicomposting of floral waste. *Bioresour. Technol.* **324**, 124672 (2021).
- [52] Liao X. et al. Artificial Intelligence: A solution to involution of Design-Build-Test-Learn cycle *Curr. Op. Biotechnol.* **75**, 102712 (2022).

- [53] Liu, L. et al. Modeling and optimization of microbial hyaluronic acid production by *Streptococcus zooepidemicus* using radial basis function neural network coupling quantum-behaved particle swarm optimization algorithm. *Biotechnol. Prog.* **25**, 1819–1825 (2009).
- [54] Jia, R. et al. Exploration of deep learning models for real-time monitoring of state and performance of anaerobic digestion with online sensors. *Bioresour. Technol.* **363**, 127908 (2022).
- [55] Rathore, A. S. et al. Bioprocess control: Current progress and future perspectives *Life.* **11**, 557 (2025).
- [56] Reardon, K. F. Practical monitoring technologies for cells and substrates in biomanufacturing. *Curr. Opin. Biotechnol.* **71**, 225–230 (2021).
- [57] Cheng, Y. et al. Artificial intelligence technologies in bioprocess: Opportunities and challenges. *Biores. Tech.* **369**, 128451 (2023).
- [58] Treloar, N.J. et al. Deep reinforcement learning for the control of microbial co-cultures in bioreactors. *PLoS Comput. Biol.* **16**, e1007783 (2020).
- [59] Singhal, A. et al. Pretreatment of *Leucaena leucocephala* wood by acidified glycerol: optimization, severity index and correlation analysis. *Bioresour. Technol.* **265**, 214–223 (2018).
- [60] Austerjost, J. et al. A machine vision approach for bioreactor foam sensing. *SLAS. Technol.* **26**, 408–414 (2021).
- [61] Butean, A. et al. A review of artificial intelligence applications for biorefineries and bioprocessing: From data-driven processes to optimization strategies and real-time control. *Processes.* **13**, 2544 (2025).
- [62] Alejo, L. et al. Effluent composition prediction of a two-stage anaerobic digestion process: machine learning and stoichiometry techniques. *Environ. Sci. Pollut. Res. Int.* **25**, 21149–21163 (2018).
- [63] Akinade, O. O. & Oyedele, L. O. Integrating construction supply chains within a circular economy: An ANFIS-based waste analytics system (A-WAS). *J. Clean. Prod.* **229**, 863–873 (2019).
- [64] Peng, J. et al. Time-dependent fermentation control strategies for enhancing synthesis of marine bacteriocin 1701 using artificial neural network and genetic algorithm. *Bioresour. Technol.* **138**, 345–352 (2013).
- [65] Sharma, D. & Singh, K. AI-enhanced bioprocess technologies: machine learning implementations from upstream to downstream operations. *World J. Microbio. Biotech.* **41**, 278 (2025).
- [66] Holzinger A. et al. AI for Life: Trends in Artificial Intelligence for biotechnology. *New Biotech.* 16-24 (2023).
- [67] Ghafarollahi, A. & Buehler, M. J. ProtAgents: protein discovery via large language model multi-agent collaborations combining physics and machine learning. *Digit. Discovery.* **3**, 1389-1409 (2024).
- [68] St. Pierre, S. R. et al. Discovering the mechanics of artificial and real meat. *Comp. Meth. Appl. Mech. Eng.* **415**, 116236 (2023).
- [69] St. Pierre, S. R. et al. The mechanical and sensory signature of plant-based and animal meat. *npj Sci. Food.* **8**, 94 (2024).
- [70] Sanahuja, M. et al. Classification of puffed snacks freshness based on crispiness-related mechanical and acoustical properties. *J. Food Eng.* **226**, 53–64 (2018).
- [71] Dahl, J. F. et al. Predicting rheological parameters of food biopolymer mixtures using machine learning. *Food Hydrocoll.* **160**, 110786 (2025).
- [72] Zhang, Y. et al. Construction of an intelligent recognition system for the cooking doneness of deep-fried golden pompano (*Trachinotus ovatus*) based on deep learning. *Food Biosci.* **68**, 106644 (2025).
- [73] Saalbrink, J. et al. Quantifying microscopic droplets in colloidal systems through machine learning-based image analysis. *Food Hydrocoll.* **166**, 111301 (2025).
- [74] Dunne, R. A. et al. Texture profile analysis and rheology of plant-based and animal meat. *Food Res. Int.* **205**, 115876 (2025).
- [75] Munekata, P. E. S. et al. Applications of electronic nose, electronic eye and electronic tongue in quality, safety and shelf life of meat and meat products: A review. *Sensors.* **23**, 672 (2023).
- [76] Karakaya, D. et al. Electronic nose and its applications: A survey. *Int. J. Auto. Comp.* **17**, 179–209 (2019).
- [77] Bengio, Y. et al. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Machine Intell.* **35**, 1798-1828 (2012).
- [78] Feng, D. et al. SMELLNET: A large-scale dataset for real-world smell recognition. *arXiv* doi:10.48550/arXiv.2506.00239 (2025).

- [79] Tom, G. et al. From molecules to mixtures: Learning representations of olfactory mixture similarity using inductive biases. *arXiv* doi:10.48550/arXiv.2501.16271 (2025).
- [80] Istif, E. et al. Miniaturized wireless sensor enables real-time monitoring of food spoilage. *Nature Food*. **4**, 427–436 (2023).
- [81] Watson, N. J. et al. Intelligent sensors for sustainable food and drink manufacturing. *Sustain. Food Syst.* **5**, 642786 (2021).
- [82] Rombach, R. et al. High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 10684-10695 (2022).
- [83] Bechler-Speicher, M. et al. Position: Graph learning will lose relevance due to poor benchmarks. *arXiv* doi:10.48550/arXiv.2502.14546 (2025).
- [84] Watson, N. & Rafiq, S. NAPIC Partner Engagement Workshop Report. <https://doi.org/10.48785/100/320>. accessed 15 Aug. 2025.
- [85] Dinali, M. et al. Fibrous structure in plant-based meat: High-moisture extrusion. *J. Food Process Preserv.* **40**, 2940-2968 (2024).
- [86] Jiang, Y. et al. Plant-based protein extrusion optimization: Comparison between machine learning and conventional experimental design. *Curr. Res. Food Sci.* **11**, 101157 (2025).
- [87] Reducing the Price of Alternative Proteins. <https://gfi.org/reducing-the-price-of-alternative-proteins>. accessed 15 Aug. 2025.
- [88] Abdurrahman, E. E. M. et al. Digital Twin applications in the food industry: a review. *Front. Sustain. Food Syst.* **9**, 1538375 (2025).
- [89] Top, J. et al. Cultivating FAIR principles for agri-food data. *Comput. & Electron. Agric.* **196**, 106909 (2022).
- [90] Bölücü, N. et al. An evaluation of Large Language Models for supplementing a food extrusion dataset. *Foods*. **14**, 1355 (2025).
- [91] Mohbat, F. & Mohammed, J. Z. Llava-chef: A multi-modal generative model for food recipes. *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. (2024).
- [92] Wells, P. How A.I. is changing how chefs cook. *The New York Times*. www.nytimes.com/2025/06/02/dining/ai-chefs-restaurants.html, accessed 15 Aug. 2025.
- [93] Ray, J. AI is a lousy chef. *Wired*. <https://www.wired.com/story/dishgen-ai-recipes-tested>, accessed 15 Aug. 2025.
- [94] Li, P. et al. Cheffusion: Multimodal foundation model integrating recipe and food image generation. *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management* (2024).
- [95] Senath, T. et al. Large Language Models for ingredient substitution in food recipes using supervised fine-tuning and direct preference optimization. *arXiv* doi:10.48550/arXiv.2412.04922 (2024).
- [96] Liu, G. et al. Retrieval augmented recipe generation. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (2025).
- [97] Mizrahi, M. et al. Cooking up creativity: A cognitively-inspired approach for enhancing LLM creativity through structured representations. *Trans. Assoc. Comp. Ling.* (2025).
- [98] Thomas, A. T. et al. What can large language models do for sustainable food? *International Conference on Machine Learning*. (2025).
- [99] Yang, D. et al. Social skill training with large language models. *arXiv*. doi:10.48550/arXiv.2404.04204 (2024).
- [100] Rosas, L. G. et al. The effectiveness of Recipe4Health: a quasi-experimental evaluation. *Am. J. Prevent. Med.* **68**, 377-390 (2025).
- [101] Rillig, M. C. et al. Risks and benefits of large language models for the environment. *Environ. Sci. Tech.* **57**, 3464-3466 (2023).
- [102] Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3 *Nature*. **630**, 493-500 (2024).
- [103] Yuen, J. S. K. et al. Aggregating in vitro-grown adipocytes to produce macroscale cell-cultured fat tissue with tunable lipid compositions for food applications. *eLife*. **12**, e82120 (2023).
- [104] Sugama, N. et al. Modulation of nutritional composition and aroma volatiles in cultivated pork fat by culture media supplementation. *bioRxiv*. doi: 10.1101/2025.07.29.667495.
- [105] Strout, A. J. et al. Engineering carotenoid production in mammalian cells for nutritionally

- enhanced cell-cultured foods. *Metab. Eng.* **62**, 126–137 (2020).
- [106] Vervenne, T. et al. Probing mycelium mechanics and taste: The moist and fibrous signature of fungi steak. *Acta Biomat.* **202**, 341-351 (2025).
- [107] Potter, N. N. & Hotchkiss, J. H. *Food Science*. 5th Edition, Springer (1995).
- [108] Linka, K. & Kuhl, E. A new family of Constitutive Artificial Neural Networks towards automated model discovery. *Comp. Meth. Appl. Mech. Eng.* **403**, 115731 (2023).
- [109] St. Pierre, S. R. et al. Biaxial testing and sensory texture evaluation of plant-based and animal deli meat. *Curr. Res. Food Sci.* **10**, 101080 (2025).
- [110] Ghafarollahi, A. & Buehler, M. J. Sparks: Multi-agent artificial intelligence model discovers protein design principles. *arXiv*. doi:10.48550/arXiv.2504.19017 (2025).
- [111] Roohani, Y., et al. BioDiscoveryAgent: An AI agent for designing genetic perturbation experiments. International Conference on Learning Representations. *arXiv*. doi:10.48550/arXiv.2405.17631 (2024).
- [112] Swanson, K., et al. The Virtual Lab of AI agents designs new SARS-CoV-2 nanobodies. *Nature*. doi:10.1038/s41586-025-09442-9 (2025).
- [113] Buehler, M. J. Accelerating scientific discovery with generative knowledge extraction, graph-based representation, and multimodal intelligent graph reasoning *Mach. Learn.: Sci. Technol.* **5**, 035083 (2024).
- [114] Boiko, D. A., et al. Autonomous chemical research with large language models. *Nature* **624**, 570-578 (2023).
- [115] Bagler, G. & Goel, M. Computational gastronomy: capturing culinary creativity by making food computable *NPJ Syst Biol Appl.* **10**, 72 (2024).
- [116] Peirlinck M, et al. On automated model discovery and a universal material subroutine for hyperelastic materials. *Comp. Meth. Appl. Mech. Eng.* **418**, 116534 (2024).