

Generative Physical Science From Probabilistic Denoising to the De Novo Design of Proteins and Catalysts

Generative Physical Science: From Probabilistic Denoising to the De Novo Design of Proteins and Catalysts

Introduction

The current era of artificial intelligence is marked by a profound transition. Generative models, having achieved remarkable success in creating and manipulating digital content like text and images, are now being aimed at the fundamental building blocks of the physical world: atoms and molecules. This report charts the intellectual and technical lineage of this transition, arguing that the paradigm of conditioned denoising diffusion provides a powerful and generalizable framework for *de novo* design in the physical sciences.

The narrative arc of this analysis follows a deliberate path from foundational principles to applied science and forward-looking strategy. First, it establishes the theoretical bedrock of modern generative AI, tracing the evolution from the concept of universal representations learned via pre-training to the specific mathematical machinery of diffusion models that enables high-fidelity generation. Second, it dissects the landmark application of these principles in molecular engineering through a detailed examination of RoseTTAFold diffusion (RFDiffusion), a method that has revolutionized protein design. Finally, it leverages these insights to propose a concrete, analogous strategy for a grand challenge in industrial chemistry: the rational, goal-oriented design of novel catalysts for the Fischer-Tropsch process. The purpose of this report is to provide a definitive synthesis of these interconnected fields and to offer a detailed blueprint for future research at the burgeoning intersection of generative AI and computational catalysis.

Part I: Foundations of Controllable Generative Modeling

The capacity to generate novel, functional designs in the physical realm is not born from a single breakthrough but from the convergence of several key ideas in machine learning. This section establishes the theoretical and architectural foundations upon which modern generative science is built, tracing the evolution from the general concept of pre-training on vast datasets to the specific mathematics of diffusion and the architectural innovations that enable fine-grained control over the generative process.

1.1 The Pre-training Paradigm: Learning Universal Representations with Transformers

The dominant approach in modern AI, particularly in language and increasingly in other modalities, is a two-stage process: unsupervised pre-training followed by supervised fine-tuning.¹ This paradigm, first powerfully demonstrated in the paper 'Improving Language Understanding by Generative Pre-Training', moved the field beyond training models from scratch for every new task.

The methodology consists of two distinct stages. The first is unsupervised pre-training, where a high-capacity model is trained on a massive, diverse corpus of unlabeled data. For language, the objective is to maximize the likelihood of the next token given a context of preceding tokens, a standard language modeling objective formalized as $L1(U) = \sum \log P(u_i | u_{1:k}, \dots, u_{i-1}; \Theta)$, where U is the corpus and Θ are the model parameters.¹ The choice of model architecture is critical. While earlier work used LSTMs, the adoption of the Transformer architecture was a pivotal development. The Transformer's self-attention mechanism provides a more structured memory and is exceptionally effective at capturing long-range dependencies in data, a significant limitation of its recurrent predecessors.¹

The second stage is supervised fine-tuning. After the model has been pre-trained, its learned parameters are adapted for a specific downstream task using a much smaller set of labeled data. This adaptation requires minimal changes to the model's architecture, often only involving the addition of a single linear output layer to predict the target label. The model is then fine-tuned by optimizing a supervised objective, $L2(C) = \sum (x, y) \log P(y | x_1, \dots, x_m)$.¹ Researchers also found that including the original language modeling objective as an auxiliary loss during fine-tuning improves the generalization of the supervised model and accelerates convergence. A key feature of this framework is its flexibility; diverse tasks like question answering or textual entailment, which involve structured inputs, can be handled by transforming the inputs into an ordered sequence that the pre-trained model can process, thereby preserving the core architecture.¹

The significance of this paradigm extends far beyond mere performance improvements. It represents a fundamental shift in how intelligent systems are constructed. By training on vast quantities of unlabeled data, the model learns a "universal representation" or an implicit "world model" of that data modality.¹ In the case of language, the model acquires a deep understanding of syntax, semantics, factual knowledge, and the complex, long-range structures that constitute coherent text. This pre-trained knowledge becomes the rich raw material that fine-tuning can rapidly shape for a specific purpose. This concept of a "foundation model" is directly transferable to other scientific domains. Just as a model pre-

trained on text learns the "language of humans," a model pre-trained on the Protein Data Bank learns the "language of proteins"—the fundamental rules of folding, stability, and interaction. This is not a loose analogy; it is the conceptual antecedent that explains why a pre-trained protein structure prediction network like RoseTTAFold can be repurposed so effectively as a generative denoiser, a topic explored in Part II.

1.2 Technical Analysis of Denoising Diffusion Probabilistic Models (DDPMs)

Denoising Diffusion Probabilistic Models (DDPMs) have emerged as a leading class of generative models, capable of producing samples of exceptionally high quality.³ Their operation is defined by two opposing processes: a fixed forward process that systematically destroys data with noise, and a learned reverse process that reconstructs data from noise.⁴

The forward process, or diffusion process, is a fixed Markov chain that gradually adds Gaussian noise to a data sample x_0 (e.g., an image) over a series of T timesteps. The transition at each step is defined by a Gaussian distribution:

$$q(x_t|x_{t-1}) = N(x_t; 1 - \beta_t I)(\text{data:image/svg+xml;utf8}, \\ x_{t-1}, \beta_t I)$$

Here, β_t is a small positive constant from a predefined variance schedule that controls the amount of noise added at step t .⁴ A remarkable property of this process is that we can sample the noisy image

x_t at any arbitrary timestep t directly from the original image x_0 , without having to iterate through the intermediate steps. This is given by the closed-form equation:

$$q(x_t|x_0) = N(x_t; \alpha^{-t} I)(\text{data:image/svg+xml;utf8}, \\ x_0, (1 - \alpha^{-t}) I)$$

where $\alpha_t = 1 - \beta_t$ and $\alpha^{-t} = \prod_{s=1}^t \alpha_s$.⁴ This property is essential for making the training process computationally efficient.

The reverse process is where the learning occurs. The goal is to learn a neural network that can reverse the diffusion, step-by-step, starting from pure Gaussian noise $x_T \sim N(0, I)$ and ending with a clean data sample x_0 . The model learns the transition probability $p_\theta(x_{t-1}|x_t)$. Instead of predicting the denoised image x_{t-1} directly, a key insight from the DDPM paper was to reparameterize the model to predict the noise ϵ that was added to create x_t .⁴ This leads to a surprisingly simple and stable training objective: the model is trained to minimize the mean squared error between the true random noise

ϵ that was added and the noise predicted by the network, ϵ_θ :

$$\mathcal{L}_{\text{simple}}(\theta) = E_{t, x_0, \epsilon} [\| \epsilon - \epsilon_\theta(\alpha^{-t} I)(\text{data:image/svg+xml;utf8}, \\ x_0 + 1 - \alpha^{-t} I)(\text{data:image/svg+xml;utf8}, \\ \epsilon, t) \|^2]$$

The network ϵ_θ , typically a U-Net architecture, takes the noisy data x_t and the current timestep t as input.⁴ The U-Net's structure, with its encoder-decoder pathways and skip connections, is particularly well-suited for this task, as it allows the model to process features at multiple noise scales simultaneously.⁶ The timestep

t is usually encoded using sinusoidal positional embeddings and fed into the residual blocks of the U-Net to make the network aware of the current noise level.⁶

It is illuminating to view DDPMs as a specific type of hierarchical Variational Autoencoder (VAE).⁶ In this framing, the fixed forward process $q(x_t|x_0)$ acts as a non-learned, deterministic encoder that maps the input data x_0 to a deep hierarchy of latent variables x_1, \dots, x_T . The learned reverse process $p_\theta(x_{t-1}|x_t)$ is the decoder. Unlike a standard VAE, which uses its decoder once, the DDPM decoder is applied iteratively T times, and critically, the same neural network ϵ_θ is shared across all timesteps. This iterative refinement, where the model makes small, repeated corrections, is what enables the generation of highly detailed and globally coherent samples, a property that proves essential for complex structures like proteins.

1.3 Architectures for Unification and Control: Case Studies of OmniGen and DpDNet

While the core DDPM framework provides a powerful engine for generation, its true potential is unlocked through architectures that enable unification of tasks and fine-grained control. Two recent papers, DpDNet and OmniGen, exemplify the leading directions in this pursuit, pointing toward a future of versatile, instruction-following generative models.

DpDNet presents a case study in achieving multi-task capability through explicit, decoupled control.⁷ The objective of DpDNet is to perform universal segmentation of lesions in whole-body PET-CT scans for multiple types of cancer. The authors recognized that simply treating all cancers as a single task overlooks the fact that different cancer types have both unique metastatic patterns and shared features.⁸ The core innovation is a

dual-prompt strategy. The model uses learnable specific prompts to capture the unique signatures of each cancer type, while a common prompt is used to retain and leverage shared knowledge across all types.⁹ These prompts are injected into the model's decoder early in the segmentation process, providing clear, task-specific guidance. Furthermore, "prompt-aware heads" are used at the output to adaptively handle the different segmentation tasks and mitigate information forgetting.⁷ DpDNet demonstrates how structured, explicit conditioning can enable a single model to expertly handle a family of related, yet distinct, tasks.

In contrast, OmniGen showcases a path toward a more general-purpose unification.¹⁰ Its goal is to create a single, unified model that can perform a vast array of image generation tasks—from text-to-image and editing to subject-driven generation and even classic computer vision

tasks like pose estimation—without needing specialized add-on modules like ControlNet or IP-Adapter.¹⁰ The methodology hinges on two principles: architectural simplicity and data unification. The architecture is remarkably concise, comprising a standard Variational Autoencoder (VAE) and a pre-trained large transformer model (Phi-3).¹⁰ It notably discards the need for dedicated encoders (e.g., CLIP for text), as the transformer itself is capable of processing arbitrarily interleaved sequences of text and image tokens.¹³ This is made possible by the second key component: the construction of a large-scale, multi-task dataset called

X2I ("anything to image"), which reformats a wide variety of image generation tasks into a consistent, instruction-based format.¹⁰

Though their approaches differ, DpDNet and OmniGen point toward the same conceptual endpoint: the evolution of generative models from single-purpose tools into versatile, controllable agents. DpDNet provides a blueprint for injecting precise, structured control for specialized domains. OmniGen provides a blueprint for a generalist architecture that can interpret and act upon flexible, multi-modal instructions. The future of generative design in the physical sciences will likely reside at the intersection of these two ideas: a unified framework that can be precisely guided by structured, task-specific prompts. For a task like catalyst design, this implies the need for a model that can interpret a prompt not just to "make a catalyst," but to act on a complex set of instructions like, "Generate a cobalt-on-silica catalyst with an Anderson-Schulz-Flory alpha of 0.9, optimized for stability at 250°C."

Part II: A Case Study in Molecular Engineering: Generative Design of Proteins with RFdiffusion

This section transitions from the abstract principles of generative modeling to a concrete, world-changing application: the *de novo* design of proteins. The RFdiffusion methodology, developed by the Baker lab, stands as a premier example of a physically-informed generative model. It masterfully combines the pre-training paradigm and the diffusion framework to solve complex challenges in molecular engineering.¹⁵

2.1 The Denoising Network: AlphaFold's Architecture as a Structural Prior

The extraordinary success of RFdiffusion is rooted in a critical design choice: its denoising network is not a generic U-Net, but a fine-tuned version of RoseTTAFold, a structure prediction network built on the same principles as DeepMind's AlphaFold2.¹⁵ This choice imbues the generative process with a profound, implicit understanding of protein physics.

The power of the AlphaFold/RoseTTAFold architecture stems from its ability to process and integrate two distinct but deeply connected representations of a protein.¹⁷ The first is the

Multiple Sequence Alignment (MSA), a 1D representation that stacks homologous sequences from different organisms. The MSA is rich with co-evolutionary information; if two amino acids are in close contact in the folded 3D structure, a mutation in one is often compensated by a mutation in the other to maintain stability. The MSA thus provides powerful clues about spatial proximity.¹⁷ The second is the

pair representation, a 2D matrix that explicitly models the geometric relationships—such as distances and orientations—between every pair of amino acid residues in the sequence.¹⁷

The architectural heart of the network is the Evoformer module. This is a series of blocks that facilitate a "continuous flow of information" between the MSA and the pair representations.¹⁷ Information from the MSA refines the geometric estimates in the pair representation, and in turn, the geometric information helps to interpret the evolutionary patterns in the MSA. This iterative, dual-track reasoning allows the model to simultaneously consider evolutionary constraints and spatial geometry to arrive at a highly accurate structure prediction.

By pre-training on the vast corpus of known protein structures in the Protein Data Bank, the AlphaFold/RoseTTAFold network has effectively learned an implicit, high-fidelity "force field" for proteins. It has an innate, data-driven understanding of valid bond lengths, angles, favorable long-range interactions, and forbidden steric clashes. Therefore, when this network is repurposed as the denoiser ϵ_θ in RFdiffusion, each step of the reverse process is more than a simple statistical denoising operation. It is an application of this learned physical prior. The network takes a noisy, physically unrealistic arrangement of atoms and predicts the "correction" (the noise) required to guide it toward a more stable, low-energy, and physically plausible conformation. The denoising network is not merely a pattern recognizer; it is a learned surrogate for the complex physics of protein folding, which is the primary reason for its unprecedented success in generative design.¹⁸

2.2 Visualizing the Transformative Process: The RFdiffusion Trajectory

The RFdiffusion process is a remarkable transformation, taking a random cloud of atoms and sculpting it into a coherent, folded protein backbone. Visualizing this trajectory reveals the power of the iterative denoising framework.

Step 1: Initialization ($t=T$). The process begins not with a folded structure, but with a state of maximum entropy: a diffuse cloud of disconnected amino acid backbones. Each backbone is represented as a rigid frame (defined by its N, C α , and C atoms) placed at a random position and orientation in space. This corresponds to the fully noised state, x_T .¹⁸

Step 2: The Forward Diffusion Process (Conceptual). While not explicitly run during generation, the conceptual forward process defines what the reverse process must learn to undo. It is a physically motivated noising process designed for molecular structures. Instead of simple pixel-wise noise, it involves two components: translation noise, where 3D Gaussian noise is applied to the positions of the central C-alpha atoms, and rotation noise, where Brownian motion on the SO(3) manifold (the mathematical space of all 3D rotations) is applied to the orientation of each residue's backbone frame.¹⁶ This use of SO(3) noise is critical for correctly modeling the rotational degrees of freedom inherent to molecules.

Step 3: The Reverse Denoising Step ($t \rightarrow t-1$). This is the core generative loop, repeated for a set number of timesteps (e.g., 50-200).¹⁹ At each step

t , the noisy structure x_t is fed into the denoising network (the fine-tuned RoseTTAFold). The network's primary task is to predict the "clean" structure, x^{*0} , from the noisy input x_t .¹⁶ This prediction of the final structure is then used in a

self-conditioning step; both the prediction \hat{x}^0 and the current noisy state x_t are used to inform the calculation of the less-noisy state for the next step, x_{t-1} . This self-conditioning mechanism acts like a memory, helping the model maintain a consistent global plan for the protein's fold throughout the long, iterative trajectory.¹⁶

Step 4: Final Output ($t=0$). After the final denoising step, the process yields a fully formed, physically plausible protein backbone, x_0 .

Step 5: Sequence Design and Validation. The generated backbone is a scaffold of atom positions. A second, specialized model like ProteinMPNN is then used to design an amino acid sequence that is predicted to fold into that specific backbone shape. As a final quality check, the complete design (backbone + sequence) is often fed into AlphaFold2 for structure prediction. If AlphaFold2 predicts the same structure with high confidence, it serves as strong evidence that the generated protein is plausible and likely to be stable, as it conforms to the learned distribution of real proteins.¹⁶

This entire trajectory can be visualized as a gradual convergence. An initial, chaotic scatter of backbone frames slowly coalesces, guided by the network's learned physical prior. Helices begin to form, beta sheets align, and loops connect them, until a single, well-defined protein structure emerges from the noise.

2.3 Controllable Generation: Motif Scaffolding and Binder Design

The true power of RFdiffusion as an engineering tool lies in its ability to perform controllable, or conditional, generation. By providing the model with specific boundary conditions, scientists can guide the generative process to solve practical design challenges.¹⁹ This is achieved by modifying the reverse process to be conditioned on some information

c , becoming $p_\theta(x_{t-1}|x_t,c)$.

A primary application is **motif scaffolding**.²⁰ The problem is to take a small, functionally important part of a protein—such as an enzyme's active site or a specific binding loop (the "motif")—and design an entirely new, stable protein "scaffold" around it to hold it in its precise functional conformation. Here, the conditioning information

c is the 3D coordinates of the motif's atoms. During the denoising trajectory, the positions of the motif atoms are held fixed (or constrained), while the diffusion model generates the surrounding scaffold atoms. The denoising network, guided by its learned physics, hallucinates a structure that seamlessly integrates with and supports the provided motif.

An even more impactful application is **de novo binder design**.¹⁵ The goal is to design a new protein that binds tightly and specifically to a target molecule, such as a viral protein or a cancer cell receptor. In this case, the conditioning information

c is the 3D structure of the target protein. The target is held fixed in space, and RFdiffusion generates a new protein in its presence. The denoising network implicitly optimizes the shape and chemical complementarity of the generated protein's surface to match the binding site on the target. The process can be further guided by specifying "hotspot" residues on the target that are known to be important for binding, encouraging the model to form an interface at that specific location.¹⁹

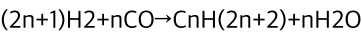
This capability for conditioned generation transforms the model from a mere "protein creator" into a sophisticated engine for *in silico* experimentation. A biologist might screen a library of millions of random proteins in a lab to find one that binds to a target—a slow, expensive process. RFdiffusion's binder design mode is the direct computational analogue. The "library" is the near-infinite space of potential protein structures the diffusion model can sample from, and the "screening" is the conditioning process, where the learned physics of the denoising network rapidly guides the generation toward solutions that "fit" the target. This demonstrates that conditioned generative models are not just for creating novel outputs; they are powerful tools for hypothesis generation and goal-directed problem-solving in the physical sciences.

Part III: A Proposed Strategy for Generative Catalyst Discovery

Leveraging the principles and successes demonstrated in protein engineering, this section outlines a novel research strategy for applying a similar generative diffusion framework to a grand challenge in industrial chemistry: the discovery of highly selective catalysts for the Fischer-Tropsch (F-T) process.

3.1 The Fischer-Tropsch Challenge: Engineering Active Sites for Product Selectivity

The Fischer-Tropsch synthesis is a cornerstone of chemical engineering, providing a pathway to convert syngas—a mixture of carbon monoxide (CO) and hydrogen (H₂)—into a wide range of valuable hydrocarbons.²¹ The process is fundamentally a catalytic polymerization reaction:



The distribution of the resulting hydrocarbon chain lengths (C_n) often follows a statistical model known as the Anderson-Schulz-Flory (ASF) distribution, which is characterized by a single parameter, α , the chain growth probability.

The central challenge in F-T catalysis is not merely achieving high reaction rates, but precisely controlling selectivity.²¹ The economic viability of the process depends on maximizing the yield of high-value liquid fuels (e.g., diesel, with chain lengths from C₁₂-C₂₀) while minimizing the production of undesirable methane (C₁) and other light gases. This translates directly to controlling the ASF parameter

α ; a higher α favors longer chains. It is well-established that this selectivity is exquisitely sensitive to the catalyst's properties, including the choice of active metal (typically iron or cobalt), the presence of chemical promoters (like potassium), the structure of the support material (like silica or alumina), and, most critically, the specific atomic geometry and electronic structure of the catalytic active sites.²¹ This defines a classic structure-to-property relationship, making it an ideal problem for a generative modeling approach.

3.2 Defining the Catalyst Manifold: A Multi-modal Representation for Generative Design

To apply a diffusion model to catalyst design, one must first answer a critical question: what is the catalyst equivalent of a protein's structural representation? A simple list of atom coordinates is insufficient. Decades of research in computational catalysis have shown that performance is

governed by a complex interplay of geometric, electronic, and energetic factors, which must be captured by rich "descriptors" to be useful for machine learning.²³ Building on this knowledge, and drawing a direct analogy to the successful multi-faceted representation used in AlphaFold, a holistic, multi-modal catalyst representation is proposed.

Feature	RFdiffusion (Protein Design)	Proposed CatDiff (Catalyst Design)
Design Target	A protein with a specific 3D structure and biological function (e.g., binding).	A heterogeneous catalyst with a specific active site structure and chemical function (e.g., product selectivity).
Core Representation	Amino acid sequence (1D), Pairwise distances/orientations (2D), 3D coordinates.	Multi-modal Crep: Atomic coordinates (3D), Electronic property field (Volumetric), Adsorption energy distribution (1D Vector).
Noise Model	Gaussian noise on 3D coordinates; Brownian motion on SO(3) for orientations.	A hybrid noise model applied to all modalities of Crep (e.g., Gaussian on coordinates and energies, structured noise on the electronic field).
Denoising Network	Fine-tuned RoseTTAFold/AlphaFold (pre-trained on protein structures).	A pre-trained "Catalyst Foundation Model" (e.g., 3D GNN/transformer) trained on a vast database of DFT-calculated catalyst properties.
Conditioning	Geometric constraints: a fixed target protein, a functional motif, or symmetry axes.	Functional constraints: a target product selectivity score (e.g., ASF α), desired stability, or specific reaction conditions.
Output	A novel protein backbone structure.	A novel, multi-modal description of a catalyst (structure, electronic, and energetic properties).

This proposed representation, Crep, would unify three crucial data modalities into a single object for the diffusion model to operate on:

- ①

Geometric Representation (The "Where"): A 3D point cloud defining the coordinates of all atoms in the system, including the metallic nanoparticle and the atoms of the support material. This is the fundamental structural information.
- ②

Electronic Representation (The "How"): A volumetric grid or a set of per-atom feature vectors that encode key electronic properties calculated from theory. Descriptors such as the d-band center of surface atoms, the local electrostatic potential, and electron density are known to be strong predictors of catalytic activity and would be included here.²⁷ This modality describes the chemical environment of the active sites.
- ③

Energetic Representation (The "What"): A functional fingerprint of the catalyst's surface. Inspired by recent work on advanced descriptors, this would take the form of an Adsorption Energy Distribution (AED).²⁸ The AED is a vector or histogram representing the calculated binding energies of key F-T reaction intermediates (e.g., $\ast\text{CO}$, $\ast\text{H}$, $\ast\text{OH}$, $\ast\text{CHx}$) across the various potential active sites on the catalyst's surface. This modality directly describes how the catalyst interacts with the reactants and intermediates of the target reaction.

By creating a unified representation that combines geometry, electronics, and energetics, a generative model can learn the deep, non-linear relationships between them. It can move beyond simply learning what structures are stable to understanding *why* certain structures are active and selective. For instance, it could learn that a specific geometric arrangement of cobalt atoms on a silica support (geometric) induces a particular local d-band structure (electronic), which in turn creates a landscape of binding energies (energetic) that is optimal for C-C bond formation and suppresses methane production. This holistic approach is a direct analogue to AlphaFold's integration of sequence and structural data and is essential for building a truly powerful generative model for catalysis.

3.3 CatDiff: A Diffusion Framework for Selectivity-Guided Catalyst Generation

Based on the multi-modal representation, a complete framework for generative catalyst discovery, termed `CatDiff`, can be outlined.

The denoising network, $\epsilon\theta$, would be a 3D-equivariant graph neural network (GNN) or transformer architecture, specifically designed to process the complex, multi-modal Crep object. The most critical step, mirroring the success of RFdiffusion, would be the pre-training of this network on a massive database of catalyst structures and their corresponding properties calculated via high-throughput Density Functional Theory (DFT) simulations. This pre-training would create a "Catalyst Foundation Model"—a network that has learned the fundamental "physics" of surface chemistry, catalysis, and materials stability. This network would serve as the learned prior, the engine that guides the generation toward valid and promising regions of the vast materials space.

The generative process itself would be a conditional diffusion model, learning the distribution $p\theta(\text{Crep},t-1|\text{Crep},t,c)$. The key innovation is the nature of the condition, c . Instead of a geometric constraint, the condition would be a desired functional outcome. For the F-T challenge, this would be the target product selectivity, expressed quantitatively as a target ASF chain growth probability, e.g., $\alpha_{\text{target}}=0.9$.

To enforce this condition, the framework would employ a guidance mechanism. A separate, smaller "selectivity prediction" network would be trained alongside the main model. This predictor would take a catalyst representation, Crep, as input and predict its resulting performance, $\alpha_{\text{predicted}}$. During the generative sampling process, the gradient of the error between the target and predicted selectivity, $(\alpha_{\text{target}}-\alpha_{\text{predicted}})^2$, would be used to "steer" each denoising step. This gradient would provide a corrective nudge, pushing the generation away from structures predicted to have poor selectivity and toward those predicted to meet the desired functional goal.

The end-to-end workflow would be as follows:

- ①

Define Target: A chemist or engineer specifies a target selectivity (e.g., $\alpha=0.9$) and other constraints (e.g., base materials like Co on SiO₂).
- ②

Initialize: The process starts from a random multi-modal object, Crep_T , representing random atom positions and random electronic/energetic values.
- ③

Iterative Denoising: For $t=T$ down to 1, the model performs the guided reverse diffusion step. The pre-trained denoising network proposes a correction toward a physically plausible state, and the selectivity guidance gradient adjusts that correction to steer it toward the functional target.
- ④

Output: The final object, Crep_0 , is a complete, multi-modal description of a novel catalyst—including its 3D atomic structure, electronic properties, and energetic landscape—that is computationally predicted to achieve the desired product selectivity. This detailed blueprint can then be passed to experimental teams for synthesis and validation.

Conclusion and Future Directions

This report has traced a path from the foundational principles of generative AI to a concrete proposal for designing industrial catalysts. The journey reveals a powerful, unifying theme: the paradigm of conditioned denoising diffusion, when armed with a physically-informed, pre-trained network, becomes a formidable tool for *de novo* design in the physical sciences. The success of RFdiffusion in protein engineering is not an isolated achievement but a blueprint. By understanding its core components—the use of a pre-trained network like RoseTTAFold as a learned physical prior, the iterative refinement of the diffusion process, and the ability to guide generation with specific conditions—we can map a clear strategy for tackling other grand scientific challenges.

The proposed CatDiff framework for Fischer-Tropsch catalyst design is a direct application of this blueprint. Its potential rests on the creation of a multi-modal representation that captures the essential geometric, electronic, and energetic properties of a catalyst, and on the development of a "Catalyst Foundation Model" pre-trained to understand the complex relationships between them. If successful, such a model could dramatically accelerate the discovery of new materials, replacing serendipitous trial-and-error with goal-directed, generative design.

Significant challenges, however, lie ahead. The foremost hurdle is data generation. The pre-training of a robust Catalyst Foundation Model would require an immense investment in high-throughput computational chemistry to generate the necessary database of simulated catalysts and their properties. This is a far greater computational undertaking than curating existing experimental data. Furthermore, developing and training a network architecture that can effectively learn from the proposed multi-modal representation is a non-trivial research problem in its own right. Ultimately, the success of this vision will depend on establishing a tight feedback loop between computational design and experimental validation, where the model's predictions are synthesized and tested, and the results are used to further refine the model.

Despite these challenges, the trajectory is clear. The framework of conditioned generative diffusion is not limited to proteins or catalysts. It represents a new paradigm for scientific discovery and engineering, one where AI models can directly design novel molecules, materials, and systems with bespoke properties, promising to accelerate the pace of innovation across countless domains of human endeavor.

Sources used in the report

[



arxiv.org

OmniGen: Unified Image Generation - arXiv

Opens in a new window](<https://arxiv.org/html/2409.11340v1>)[•



alphaxiv.org

OmniGen: Unified Image Generation | alphaXiv

Opens in a new window](<https://www.alphaxiv.org/overview/2409.11340v1>)[•



vectorspacelab.github.io

OmniGen: Unified Image Generation

Opens in a new window](<https://vectorspacelab.github.io/OmniGen/>)[•



geeksforgeeks.org

Denoising Diffusion Probabilistic Models - GeeksforGeeks

Opens in a new window](<https://www.geeksforgeeks.org/data-science/denoising-diffusion-probabilistic-models/>)[•



learnopencv.com

An In-Depth Guide to Denoising Diffusion Probabilistic Models DDPM - Theory to Implementation - LearnOpenCV

Opens in a new window](<https://learnopencv.com/denoising-diffusion-probabilistic-models/>)[•



cs.unc.edu

Lecture X: Denoising Diffusion Models

Opens in a new window](https://www.cs.unc.edu/~ronisen/teaching/fall_2022/pdf_lectures/lecture7-8_diffusion_model.pdf)[•



netl.doe.gov

www.netl.doe.gov

Opens in a new window](<https://www.netl.doe.gov/research/carbon-management/energy-systems/gasification/gasifipedia/ftsynthesis#:~:text=Catalysts%20considered%20for%20Fischer%2DTropsch,alkanes%20and%20diesel%20fuels%20production.>)]



neurosnap.ai

Use RFdiffusion-v2 Online - Neurosnap

Opens in a new window](<https://neurosnap.ai/service/RFdiffusion-v2>)]



pubmed.ncbi.nlm.nih.gov

De novo design of protein structure and function with RFdiffusion - PubMed

Opens in a new window](<https://pubmed.ncbi.nlm.nih.gov/37433327/>)]



ijcai.org

Pretrained Language Model for Text Generation: A Survey - IJCAI

Opens in a new window](<https://www.ijcai.org/proceedings/2021/0612.pdf>)]



nn.labml.ai

Denoising Diffusion Probabilistic Models (DDPM) - labml.ai

Opens in a new window](<https://nn.labml.ai/diffusion/ddpm/index.html>)]



ebi.ac.uk

AlphaFold2: A high-level overview | AlphaFold - EMBL-EBI

Opens in a new window](<https://www.ebi.ac.uk/training/online/courses/alphafold/inputs-and-outputs/a-high-level-overview/>)]



netl.doe.gov

10.2. Fischer-Tropsch Synthesis | netl.doe.gov

Opens in a new window](<https://www.netl.doe.gov/research/carbon-management/energy-systems/gasification/gasifipedia/ftsynthesis>)]



cdn.openai.com

Improving Language Understanding by Generative Pre ... - OpenAI

Opens in a new window](https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf)]



openaccess.thecvf.com

OmniGen: Unified Image Generation - CVF Open Access

Opens in a new window]
(https://openaccess.thecvf.com/content/CVPR2025/papers/Xiao_OmniGen_Unified_Image_Generation_CVPR_2025_paper.pdf)]



github.com

Code for running RFdiffusion - GitHub

Opens in a new window](<https://github.com/RosettaCommons/RFdiffusion>)]



bakerlab.org

RFdiffusion: A generative model for protein design – Baker Lab

Opens in a new window](<https://www.bakerlab.org/2023/07/11/diffusion-model-for-protein-design/>)[•



openaccess.thecvf.com

UNIC-Adapter: Unified Image-instruction Adapter with Multi-modal Transformer for Image Generation

Opens in a new window](https://openaccess.thecvf.com/content/CVPR2025/papers/Duan_UNIC-Adapter_Unified_Image-instruction_Adapter_with_Multi-modal_Transformer_for_Image_Generation_CVPR_2025_paper.pdf)[•



frontiersin.org

Machine learning meets quantum mechanics in catalysis – Frontiers

Opens in a new window](<https://www.frontiersin.org/journals/quantum-science-and-technology/articles/10.3389/frqst.2023.1232903/pdf>)[•



pubs.acs.org

Interpretable Catalysis Models Using Machine Learning with Spectroscopic Descriptors

Opens in a new window](<https://pubs.acs.org/doi/10.1021/acscatal.3c00611>)[•



researchgate.net

Machine Learning for Computational Heterogeneous Catalysis ...

Opens in a new window](https://www.researchgate.net/publication/333120502_Machine_Learning_for_Computational_Heterogeneous_Catalysis)[•



medium.com

Protein Structure Prediction : A Primer (Part 8) | by Siddhant Rai ...

Opens in a new window](<https://medium.com/@rsiddhant73/protein-structure-prediction-a-primer-part-8-93cf4d4f1f39>)[•



arxiv.org

arxiv.org

Opens in a new window](<https://arxiv.org/abs/2507.07126>)[•



arxiv.org

DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation – arXiv

Opens in a new window](<https://arxiv.org/html/2507.07126v1>)[•



researchgate.net

DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation

Opens in a new window](https://www.researchgate.net/publication/393586793_DpDNet_An_Dual-Prompt-Driven_Network_for_Universal_PET-CT_Segmentation)[•



arxiv.org

Machine Learning Accelerated Descriptor Design for Catalyst Discovery in CO2 to Methanol Conversion – arXiv

Opens in a new window](<https://arxiv.org/html/2412.13838v5>)[•



research.aalto.fi

Machine learning accelerated descriptor design for catalyst discovery in CO2 to methanol conversion – Aalto Research Portal

Opens in a new window]

(https://research.aalto.fi/files/186274555/Machine_learning_accelerated_descriptor_design_for_catalyst_discovery_in_CO2_to_methanol_conversion.pdf)[•



pmc.ncbi.nlm.nih.gov

Applications of Machine Learning in Alloy Catalysts: Rational Selection and Future Development of Descriptors – PMC – PubMed Central

Opens in a new window](<https://pmc.ncbi.nlm.nih.gov/articles/PMC9036033/>)[•



pubs.acs.org

Digital Descriptors in Predicting Catalysis Reaction Efficiency and ...

Opens in a new window](<https://pubs.acs.org/doi/10.1021/acs.jpcllett.4c03733>)•

Sources read but not used in the report

[



scholars.uky.edu

Fischer-tropsch synthesis: Computational sensitivity modeling for series of cobalt catalysts –

Opens in a new window](<https://scholars.uky.edu/es/publications/fischertropsch-synthesis-computational-sensitivity-modeling-for-s>)[•



pubs.rsc.org

Machine learning-driven design of single-atom catalysts for carbon dioxide valorization to high-value chemicals: a review of photocatalysis, electrocatalysis, and thermocatalysis – RSC Publishing

Opens in a new window](<https://pubs.rsc.org/en/content/articlehtml/2025/gc/d5gc00739a>)[•



arxiv.org

[2506.18871] OmniGen2: Exploration to Advanced Multimodal Generation – arXiv

Opens in a new window](<https://arxiv.org/abs/2506.18871>)[•



youtube.com

Denoising Diffusion Probabilistic Models | DDPM Explained – YouTube

Opens in a new window](<https://www.youtube.com/watch?v=H45IF4sUgiE>)[•



arxiv.org

[2501.15385] DDUNet: Dual Dynamic U-Net for Highly-Efficient Cloud Segmentation – arXiv

Opens in a new window](<https://arxiv.org/abs/2501.15385>)[•



arxiv.org

[2403.17734] Paired Diffusion: Generation of related, synthetic PET-CT-Segmentation scans using Linked Denoising Diffusion Probabilistic Models - arXiv

Opens in a new window](https://arxiv.org/abs/2403.17734)[•



arxiv.org

Recent Advances on Generalizable Diffusion-generated Image Detection - arXiv

Opens in a new window](https://arxiv.org/abs/2502.19716)[•



github.com

[TMLR 2025] Efficient Diffusion Models: A Survey - GitHub

Opens in a new window](https://github.com/AIoT-MLSys-Lab/Efficient-Diffusion-Model-Survey)[•



daleonai.com

AlphaFold 2 Explained: A Semi-Deep Dive - Dale on AI

Opens in a new window](https://daleonai.com/how-alphafold-works)[•



osf.io

A Contextualized Survey on the State and Future Directions of Diffusion Models - OSF

Opens in a new window](https://osf.io/xcfdu/download)[•



elearning.vib.be

Overview of the architecture - E-Learning@VIB

Opens in a new window](https://elearning.vib.be/courses/alphafold/lessons/the-alphafold-pipeline/topic/overview-of-the-architecture/)[•



elanapearl.github.io

The Illustrated AlphaFold | matmols - Elana Simon

Opens in a new window](https://elanapearl.github.io/blog/2024/the-illustrated-alphafold/)[•



arxiv.org

A Survey of Controllable Text Generation using Transformer-based Pre-trained Language Models - arXiv

Opens in a new window](https://arxiv.org/pdf/2201.05337)[•



arxiv.org

[2109.09920] Survey: Transformer based Video-Language Pre-training - arXiv

Opens in a new window](https://arxiv.org/abs/2109.09920)[•



bakerlab.org

Designing antibodies with RFdiffusion • Baker Lab

Opens in a new window](https://www.bakerlab.org/2025/02/28/designing-antibodies-with-rfdiffusion/)[•



github.com

Opens in a new window](<https://github.com/CroitoruAlin/Diffusion-Models-in-Vision-A-Survey>)[•



pure.mpg.de

arXiv:2505.16707v1 [cs.CV] 22 May 2025 - MPG.PuRe

Opens in a new window](https://pure.mpg.de/pubman/item/item_3652097_1/component/file_3652098/arXiv%3A2505.16707.pdf)[•



arxiv.org

Unified Multimodal Understanding and Generation Models: Advances, Challenges, and Opportunities - arXiv

Opens in a new window](<https://arxiv.org/pdf/2505.02567>)[•



arxiv.org

AI-Empowered Catalyst Discovery: A Survey from Classical Machine Learning Approaches to Large Language Models - arXiv

Opens in a new window](<https://arxiv.org/html/2502.13626v1>)[•



researchgate.net

Multistate and functional protein design using RoseTTAFold sequence space diffusion

Opens in a new window]
(https://www.researchgate.net/publication/384352498_Multistate_and_functional_protein_design_using_RoseTTAFold_sequence_space_diffusion)[•



sciencelink.net

Machine learning and computational hydrogenation: an experiment - ScienceLink

Opens in a new window](<https://www.sciencelink.net/features/machine-learning-and-computational-hydrogenation-an-experiment/22249.article>)[•



pubs.acs.org

Computational Design of Catalysts with Experimental Validation: Recent Successes, Effective Strategies, and Pitfalls - ACS Publications

Opens in a new window](<https://pubs.acs.org/doi/10.1021/acs.jpcc.4c04949>)[•



youtube.com

Multistate and functional protein design using RoseTTAFold sequence space diffusion

Opens in a new window](https://www.youtube.com/watch?v=hbKDL_vxGKE)[•



xujuefei.com

Transfer between Modalities with MetaQueries - Felix Juefei-Xu

Opens in a new window](https://xujuefei.com/felix_arxiv25_metaquery.pdf)[•



pmc.ncbi.nlm.nih.gov

Machine Learning Descriptors for Data-Driven Catalysis Study - PMC - PubMed Central

Opens in a new window](<https://pmc.ncbi.nlm.nih.gov/articles/PMC10401178/>)[•



biorexiv.org

Generalized Biomolecular Modeling and Design with RoseTTAFold All-Atom - bioRxiv

Opens in a new window](<https://www.biorxiv.org/content/10.1101/2023.10.09.561603v1.full.pdf>)[•



jiataogu.me

dart: denoising autoregressive transformer - Jiatao Gu

Opens in a new window](<https://jiataogu.me/papers/gu2024dart.pdf>)[•



scholar.google.com

Chunyao Lu - Google Scholar

Opens in a new window](https://scholar.google.com/citations?user=u_jwCCQAAAAJ&hl=en)[•



researchgate.net

The effects of art therapy on anxiety and depression in breast cancer patients: An updated meta-analysis - ResearchGate

Opens in a new window]
(https://www.researchgate.net/publication/341644814_The_effects_of_art_therapy_on_anxiety_and_depression_in_breast_cancer_patients_An_updated_meta-analysis)[•



pmc.ncbi.nlm.nih.gov

Machine Learning in Catalysis, From Proposal to Practicing - PMC

Opens in a new window](<https://pmc.ncbi.nlm.nih.gov/articles/PMC6963892/>)[•



escholarship.org

Machine Learning Descriptors for Data-Driven Catalysis Study - eScholarship.org

Opens in a new window](https://escholarship.org/content/qt9zc1r279/qt9zc1r279_noSplash_0bbfabe2d62829521f27df22208efd.pdf)[•



pubmed.ncbi.nlm.nih.gov

Machine Learning Descriptors for Data-Driven Catalysis Study

Opens in a new window](<https://pubmed.ncbi.nlm.nih.gov/37191279/>)[•



lonepatient.top

Arxiv今日论文| 2025-07-11 - 闲记算法

Opens in a new window](http://lonepatient.top/2025/07/11/arxiv_papers_2025-07-11)[•



arxiv.org

Computer Science Jul 2025 - arXiv

Opens in a new window](<http://www.arxiv.org/list/cs/2025-07?skip=4100&show=100>)[•



arxiv.org

Electrical Engineering and Systems Science - arXiv

Opens in a new window](<http://www.arxiv.org/list/eess/new?skip=0&show=1000>)[•



arxiv.org

Artificial Intelligence Jul 2025 - arXiv

Opens in a new window](<http://arxiv.org/list/cs.AI/2025-07?skip=1050&show=2000>)[•

R⁶

researchgate.net

Machine Learning Descriptors for Data-Driven Catalysis Study - ResearchGate

Opens in a new window](https://www.researchgate.net/publication/370802926_Machine_Learning_Descriptors_for_Data-Driven_Catalysis_Study)[•



arxiv.org

Machine-learning Accelerated Descriptor Design for Catalyst Discovery: A CO2 to Methanol Conversion Case Study - arXiv

Opens in a new window](<https://arxiv.org/html/2412.13838v1>)[•



pubs.acs.org

A Machine Learning and Explainable AI Framework Tailored for Unbalanced Experimental Catalyst Discovery | The Journal of Physical Chemistry C - ACS Publications

Opens in a new window](<https://pubs.acs.org/doi/10.1021/acs.jpcc.4c05332>)•

Thoughts

Research Approach

I'm embarking on a comprehensive research endeavor to construct a paper on diffusion models and their applications. My strategy involves a multi-faceted approach, beginning with a deep dive into the technical underpinnings of diffusion models, specifically Denoising Diffusion Probabilistic Models (DDPMs). I'll then proceed to summarize three key research papers: "DpDNet- An Dual-Prompt-Driven Network for Universal PET-CT Segmentation," "OmniGen- Unified Image Generation," and foundational concepts related to "PRE-TEXT Training Language Models." Following this, I'll explore the cutting-edge advancements in diffusion models beyond image generation, venturing into their applications in video, audio, and 3D content creation.

Novel Application Exploration

A significant part of this research will involve analyzing the AlphaFold 2 and RF Diffusion methodologies, conceptualizing them as a transformative process. This understanding will then be leveraged to propose a novel research strategy: adapting these models for the discovery of new catalysts for the Fischer-Tropsch process. The aim is to identify catalysts that can precisely control the density of the resulting hydrocarbon product, which has direct implications for specialized fuel applications.

Next Steps

To kickstart this extensive research, I'm initiating a series of targeted searches. These will cover tutorials on DDPMs, specific academic papers on ArXiv, surveys on recent diffusion model advancements, detailed explanations of AlphaFold 2 and RF Diffusion, and information on computational catalyst design for the Fischer-Tropsch process. This initial broad sweep will provide the foundational resources for each section of the paper, and I'll refine my search queries as the research progresses.



Researching websites



scholars.uky.edu

Fischer-tropsch synthesis: Computational sensitivity modeling for series of cobalt catalysts -

](<https://scholars.uky.edu/es/publications/fischertropsch-synthesis-computational-sensitivity-modeling-for-s>)[)•

pubs.rsc.org

Machine learning-driven design of single-atom catalysts for carbon dioxide valorization to high-value chemicals: a review of photocatalysis, electrocatalysis, and thermocatalysis - RSC Publishing

[\]\(https://pubs.rsc.org/en/content/articlehtml/2025/gc/d5gc00739a\)](https://pubs.rsc.org/en/content/articlehtml/2025/gc/d5gc00739a)[\]\(https://t1.gstatic.com/faviconV2?url=https://arxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t1.gstatic.com/faviconV2?url=https://arxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

arxiv.org

OmniGen: Unified Image Generation - arXiv

[\]\(https://arxiv.org/html/2409.11340v1\)](https://arxiv.org/html/2409.11340v1)[\]\(https://t3.gstatic.com/faviconV2?url=https://www.alphaxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t3.gstatic.com/faviconV2?url=https://www.alphaxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

alphaxiv.org

OmniGen: Unified Image Generation | alphaXiv

[\]\(https://www.alphaxiv.org/overview/2409.11340v1\)](https://www.alphaxiv.org/overview/2409.11340v1)[\]\(https://t1.gstatic.com/faviconV2?url=https://arxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t1.gstatic.com/faviconV2?url=https://arxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

arxiv.org

[2506.18871] OmniGen2: Exploration to Advanced Multimodal Generation - arXiv

[\]\(https://arxiv.org/abs/2506.18871\)](https://arxiv.org/abs/2506.18871)[\]\(https://t2.gstatic.com/faviconV2?url=https://vectorspacelab.github.io/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t2.gstatic.com/faviconV2?url=https://vectorspacelab.github.io/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

vectorspacelab.github.io

OmniGen: Unified Image Generation

[\]\(https://vectorspacelab.github.io/OmniGen/\)](https://vectorspacelab.github.io/OmniGen/)[\]\(https://t0.gstatic.com/faviconV2?url=https://nn.labml.ai/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t0.gstatic.com/faviconV2?url=https://nn.labml.ai/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

nn.labml.ai

Denoising Diffusion Probabilistic Models (DDPM) - labml.ai

[\]\(https://nn.labml.ai/diffusion/ddpm/index.html\)](https://nn.labml.ai/diffusion/ddpm/index.html)[\]\(https://t1.gstatic.com/faviconV2?url=https://www.geeksforgeeks.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t1.gstatic.com/faviconV2?url=https://www.geeksforgeeks.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

geeksforgeeks.org

Denoising Diffusion Probabilistic Models - GeeksforGeeks

[\]\(https://www.geeksforgeeks.org/data-science/denoising-diffusion-probabilistic-models/\)](https://www.geeksforgeeks.org/data-science/denoising-diffusion-probabilistic-models/)[\]\(https://t1.gstatic.com/faviconV2?url=https://arxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t1.gstatic.com/faviconV2?url=https://arxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

arxiv.org

DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation - arXiv

[\]\(https://arxiv.org/html/2507.07126v1\)](https://arxiv.org/html/2507.07126v1)[\]\(https://t0.gstatic.com/faviconV2?url=https://www.youtube.com/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t0.gstatic.com/faviconV2?url=https://www.youtube.com/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

youtube.com

Denoising Diffusion Probabilistic Models | DDPM Explained - YouTube

[\]\(https://www.youtube.com/watch?v=H45IF4sUgiE\)](https://www.youtube.com/watch?v=H45IF4sUgiE)[\]\(https://t0.gstatic.com/faviconV2?url=https://www.researchgate.net/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t0.gstatic.com/faviconV2?url=https://www.researchgate.net/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

researchgate.net

DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation

[\]\(https://www.researchgate.net/publication/393586793_DpDNet_An_Dual-Prompt-Driven_Network_for_Universal_PET-CT_Segmentation\)](https://www.researchgate.net/publication/393586793_DpDNet_An_Dual-Prompt-Driven_Network_for_Universal_PET-CT_Segmentation)[\]\(https://t3.gstatic.com/faviconV2?url=https://learnopencv.com/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t3.gstatic.com/faviconV2?url=https://learnopencv.com/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

learnopencv.com

An In-Depth Guide to Denoising Diffusion Probabilistic Models DDPM - Theory to Implementation - LearnOpenCV

[\]\(https://learnopencv.com/denoising-diffusion-probabilistic-models/\)](https://learnopencv.com/denoising-diffusion-probabilistic-models/)[\]\(https://t1.gstatic.com/faviconV2?url=https://arxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t1.gstatic.com/faviconV2?url=https://arxiv.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

arxiv.org

[2507.07126] DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation

[\]\(https://arxiv.org/abs/2507.07126\)](https://arxiv.org/abs/2507.07126)[\]\(https://t0.gstatic.com/faviconV2?url=https://www.cs.unc.edu/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL\)](https://t0.gstatic.com/faviconV2?url=https://www.cs.unc.edu/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

cs.unc.edu

](https://www.cs.unc.edu/~ronisen/teaching/fall_2022/pdf_lectures/lecture7-8_diffusion_model.pdf)[]

arxiv.org

[2501.15385] DDUNet: Dual Dynamic U-Net for Highly-Efficient Cloud Segmentation - arXiv

](https://arxiv.org/abs/2501.15385)[]

netl.doe.gov

10.2. Fischer-Tropsch Synthesis | netl.doe.gov

](https://www.netl.doe.gov/research/carbon-management/energy-systems/gasification/gasifipedia/ftsynthesis)[]

arxiv.org

[2403.17734] Paired Diffusion: Generation of related, synthetic PET-CT-Segmentation scans using Linked Denoising Diffusion Probabilistic Models - arXiv

](https://arxiv.org/abs/2403.17734)[]

netl.doe.gov

www.netl.doe.gov

](https://www.netl.doe.gov/research/carbon-management/energy-systems/gasification/gasifipedia/ftsynthesis#:~:text=Catalysts%20considered%20for%20Fischer%2DTropsch,alkanes%20and%20diesel%20fuels%20production.)[]

neurosnap.ai

Use RFdiffusion-v2 Online - Neurosnap

](https://neurosnap.ai/service/RFdiffusion-v2)[]

arxiv.org

Recent Advances on Generalizable Diffusion-generated Image Detection - arXiv

](https://arxiv.org/abs/2502.19716)[]

github.com

[TMLR 2025] Efficient Diffusion Models: A Survey - GitHub

](https://github.com/AIoT-MLSys-Lab/Efficient-Diffusion-Model-Survey)[]

bakerlab.org

Designing antibodies with RFdiffusion - Baker Lab

](https://www.bakerlab.org/2025/02/28/designing-antibodies-with-rfdiffusion/)[]

daleonai.com

AlphaFold 2 Explained: A Semi-Deep Dive - Dale on AI

](https://daleonai.com/how-alphafold-works)[]

osf.io

A Contextualized Survey on the State and Future Directions of Diffusion Models - OSF

](https://osf.io/xcfdv/download)[]


github.com

Diffusion Models in Vision: A Survey (accepted at IEEE TPAMI 2023) - GitHub

](<https://github.com/CroitoruAlin/Diffusion-Models-in-Vision-A-Survey>)[)

elearning.vib.be

Overview of the architecture - E-Learning@VIB

](<https://elearning.vib.be/courses/alphafold/lessons/the-alphafold-pipeline/topic/overview-of-the-architecture/>)[)


pubmed.ncbi.nlm.nih.gov

De novo design of protein structure and function with RFdiffusion - PubMed

](<https://pubmed.ncbi.nlm.nih.gov/37433327/>)[)


cdn.openai.com

Improving Language Understanding by Generative Pre-Training - OpenAI

](https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf)[)


ebi.ac.uk

AlphaFold2: A high-level overview | AlphaFold - EMBL-EBI

](<https://www.ebi.ac.uk/training/online/courses/alphafold/inputs-and-outputs/a-high-level-overview/>)[)


ijcai.org

Pretrained Language Model for Text Generation: A Survey - IJCAI

](<https://www.ijcai.org/proceedings/2021/0612.pdf>)[)

elanapearl.github.io

The Illustrated AlphaFold | matmols - Elana Simon

](<https://elanapearl.github.io/blog/2024/the-illustrated-alphafold/>)[)

arxiv.org

A Survey of Controllable Text Generation using Transformer-based Pre-trained Language Models - arXiv

](<https://arxiv.org/pdf/2201.05337>)[)

arxiv.org

[2109.09920] Survey: Transformer based Video-Language Pre-training - arXiv

](<https://arxiv.org/abs/2109.09920>)[)

github.com

Code for running RFdiffusion - GitHub

](<https://github.com/RosettaCommons/RFdiffusion>)

Understanding Diffusion Models

I've begun by gathering foundational information on Denoising Diffusion Probabilistic Models (DDPMs). I'm synthesizing details on their core concepts, including the forward (noising) and reverse (denoising) processes, and the underlying mathematical principles. This initial exploration is crucial for building a strong technical analysis section in the paper.

Summarizing Key Papers

I've started to collect information for summarizing the specified research papers. For 'DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation,' I'm focusing on its dual-prompt strategy for capturing cancer-specific and shared features in PET-CT segmentation, and its use of prompt-aware heads. For 'OmniGen: Unified Image Generation,' I'm looking at its ability to unify various image generation tasks and its simplified architecture that eliminates the need for additional modules. For 'PRE-TEXT Training Language Models,' I'm gathering information on the general concept of pre-training language models, particularly using transformer architectures, for generative tasks.

Exploring Advanced Applications and Protein Design

My research into recent advancements in diffusion models is highlighting their expanding applications beyond basic image generation, including video, audio, and 3D content. I'm also delving into the architecture of AlphaFold 2, specifically its Evoformer and Structure Module, and how it leverages multiple sequence alignments and pair representations for protein structure prediction. Furthermore, I'm investigating RF Diffusion's capabilities in *de novo* protein design, including its use in generating new protein structures and designing binders, which is highly relevant to the novel catalyst discovery aspect of this research.

Catalyst Discovery for Fischer-Tropsch Synthesis

I'm gathering information on Fischer-Tropsch synthesis, focusing on the types of catalysts used (iron, cobalt, nickel, ruthenium) and how reaction conditions influence product selectivity and density. This will form the basis for exploring how diffusion models could be applied to design catalysts that modulate the density of hydrocarbon products.

Next Steps: Deepening Understanding and Addressing Gaps

My next steps involve a deeper dive into the content of the identified papers and surveys to extract the specific details required for each section of the paper. I will focus on synthesizing the technical aspects of diffusion models, the methodologies and findings of the specified papers, and the current landscape of diffusion model applications. I will also continue to refine my understanding of AlphaFold's protein folding process and RF Diffusion's protein design capabilities to connect them to the challenge of novel catalyst discovery for the Fischer-Tropsch process.



Researching websites



bakerlab.org

Designing antibodies with RFDiffusion • Baker Lab

](https://www.bakerlab.org/2025/02/28/designing-antibodies-with-rfdiffusion/)[![]](https://t0.gstatic.com/faviconV2?url=https://nn.labml.ai/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

nn.labml.ai

Denoising Diffusion Probabilistic Models (DDPM) – labml.ai

](https://nn.labml.ai/diffusion/ddpm/index.html)[![]](https://t1.gstatic.com/faviconV2?url=https://www.ebi.ac.uk/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

ebi.ac.uk

AlphaFold2: A high-level overview | AlphaFold – EMBL-EBI

](https://www.ebi.ac.uk/training/online/courses/alphafold/inputs-and-outputs/a-high-level-overview/)[![]](https://t2.gstatic.com/faviconV2?url=https://www.netl.doe.gov/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

netl.doe.gov

10.2. Fischer-Tropsch Synthesis | netl.doe.gov

](https://www.netl.doe.gov/research/carbon-management/energy-systems/gasification/gasifipedia/ftsynthesis/)[![]](https://t1.gstatic.com/faviconV2?url=https://cdn.openai.com/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

cdn.openai.com

Improving Language Understanding by Generative Pre ... – OpenAI

](https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf)[![]](https://t1.gstatic.com/faviconV2?url=https://github.com/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)●

github.com

Diffusion Models in Vision: A Survey (accepted at IEEE ... – GitHub

](https://github.com/CroitoruAlin/Diffusion-Models-in-Vision-A-Survey)●

Refining Paper Summaries

I'm continuing to refine my understanding of the specific research papers. For 'DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation,' I'm focusing on its innovative dual-prompt strategy, which combines cancer-specific and common prompts to enhance universal PET-CT segmentation. I'm also noting its use of prompt-aware heads to manage multiple segmentation tasks and mitigate information forgetting. For 'OmniGen: Unified Image Generation,' I'm synthesizing information on its ability to unify various image generation tasks, including text-to-image, image editing, and even classical computer vision tasks, within a simplified architecture that avoids the need for additional modules like ControlNet.

Deepening RFDiffusion Understanding

I'm delving deeper into the methodology of RFDiffusion, which is a key component for the protein design section. I'm learning that it's an open-source method for *de novo* protein structure generation, capable of various design challenges like binder design and symmetric oligomer design.

It achieves this by fine-tuning the RoseTTAFold structure prediction network on protein structure denoising tasks, effectively using a diffusion model approach to generate new protein backbones and sequences.

Advancements in Diffusion Models

My research into recent advancements in diffusion models highlights their growing versatility beyond image generation, now encompassing video, audio, and 3D content creation. I'm also noting improvements in sampling efficiency and the development of latent diffusion models, which operate in a compressed space for faster processing.

Catalyst Design for Fischer-Tropsch Synthesis

I'm continuing to build my understanding of Fischer-Tropsch synthesis, specifically the role of transition metal catalysts like iron, cobalt, nickel, and ruthenium in converting syngas into hydrocarbons. I'm also exploring how computational methods, including machine learning and DFT calculations, are being used to accelerate catalyst discovery and optimize properties like selectivity and efficiency, which is directly relevant to modulating product density.

Next Steps: Addressing Information Gaps and Structuring the Paper

My immediate next steps involve ensuring I have comprehensive details for the DpDNet and OmniGen papers, as well as a more in-depth understanding of RFdiffusion's technical specifics. I will also begin to conceptualize how to represent catalysts computationally for the generative model application, drawing insights from materials informatics. Once these information gaps are filled, I will proceed with structuring and drafting the paper, integrating all the gathered knowledge into a cohesive narrative.



Researching websites



openaccess.thecvf.com

OmniGen: Unified Image Generation - CVF Open Access

](https://openaccess.thecvf.com/content/CVPR2025/papers/Xiao_OmniGen_Unified_Image_Generation_CVPR_2025_paper.pdf)[)

researchgate.net

researchgate.net

Machine Learning for Computational Heterogeneous Catalysis | Request PDF

](https://www.researchgate.net/publication/333120502_Machine_Learning_for_Computational_Heterogeneous_Catalysis)[)

medium.com

Protein Structure Prediction : A Primer (Part 8) | by Siddhant Rai | Medium

](<https://medium.com/@rsiddhant73/protein-structure-prediction-a-primer-part-8-93cf4d4f1f39>)[)

pure.mpg.de

arXiv:2505.16707v1 [cs.CV] 22 May 2025 - MPG.PuRe

](https://pure.mpg.de/pubman/item/item_3652097_1/component/file_3652098/arXiv%3A2505.16707.pdf)[)

github.com

Code for running RFdiffusion - GitHub

](<https://github.com/RosettaCommons/RFdiffusion>)[)

arxiv.org

Unified Multimodal Understanding and Generation Models: Advances, Challenges, and Opportunities - arXiv

](<https://arxiv.org/pdf/2505.02567>)[)


arxiv.org

AI-Empowered Catalyst Discovery: A Survey from Classical Machine Learning Approaches to Large Language Models - arXiv

](<https://arxiv.org/html/2502.13626v1>)[)


researchgate.net

Multistate and functional protein design using RoseTTAFold sequence space diffusion

](https://www.researchgate.net/publication/384352498_Multistate_and_functional_protein_design_using_RoseTTAFold_sequence_space_diffusion)[)


sciencelink.net

Machine learning and computational hydrogenation: an experiment – ScienceLink

](<https://www.sciencelink.net/features/machine-learning-and-computational-hydrogenation-an-experiment/22249.article>)[)

pubs.acs.org

Computational Design of Catalysts with Experimental Validation: Recent Successes, Effective Strategies, and Pitfalls – ACS Publications

](<https://pubs.acs.org/doi/10.1021/acs.jpcc.4c04949>)[)


youtube.com

Multistate and functional protein design using RoseTTAFold sequence space diffusion

](https://www.youtube.com/watch?v=hbKDL_vxGKE)[)

xujuefei.com

Transfer between Modalities with MetaQueries – Felix Juefei-Xu

](https://xujuefei.com/felix_arxiv25_metaquery.pdf)[)


bakerlab.org

RFdiffusion: A generative model for protein design – Baker Lab

](<https://www.bakerlab.org/2023/07/11/diffusion-model-for-protein-design/>)[)


openaccess.thecvf.com

UNIC-Adapter: Unified Image-instruction Adapter with Multi-modal Transformer for Image Generation

](https://openaccess.thecvf.com/content/CVPR2025/papers/Duan_UNIC-Adapter_Unified_Image-instruction_Adapter_with_Multi-modal_Transformer_for_Image_Generation_CVPR_2025_paper.pdf)[)


pmc.ncbi.nlm.nih.gov

Machine Learning Descriptors for Data-Driven Catalysis Study – PMC – PubMed Central

](<https://pmc.ncbi.nlm.nih.gov/articles/PMC10401178/>)[)

frontiersin.org

Machine learning meets quantum mechanics in catalysis – Frontiers

](<https://www.frontiersin.org/journals/quantum-science-and-technology/articles/10.3389/frqst.2023.1232903/pdf>)[)

bioRxiv.org

Generalized Biomolecular Modeling and Design with RoseTTAFold All-Atom – bioRxiv

](<https://www.biorxiv.org/content/10.1101/2023.10.09.561603v1.full.pdf>)[)

jiataogu.me

dart: denoising autoregressive transformer – Jiatao Gu

](<https://jiataogu.me/papers/gu2024dart.pdf>)

Addressing Information Gaps

I've made progress in gathering information for OmniGen and RFdiffusion, and I'm deepening my understanding of computational catalyst representation. However, I still need to locate and summarize the full paper for 'DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation' to ensure a complete overview.

I'm focusing on understanding how catalysts are computationally represented for machine learning models. This involves exploring various representation methods, such as descriptors and graph representations, which will be crucial for the final section of the paper on novel catalyst discovery.

Next Steps: Targeted Searches and Paper Structuring

My immediate next steps involve conducting a more targeted search for the 'DpDNet' paper, potentially using academic search engines like Google Scholar. Simultaneously, I will browse the most promising links from previous searches to gain a deeper understanding of catalyst representation. Once these details are solidified, I will begin synthesizing all the gathered information to structure the paper, starting with the technical analysis of diffusion models, followed by the paper summaries, recent advancements, the AlphaFold 2 and RF Diffusion analysis, and finally, the proposed research strategy for catalyst discovery.



Researching websites



arxiv.org

[2507.07126] DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation

](<https://arxiv.org/abs/2507.07126>)[)]

lonepatient.top

Arxiv今日论文| 2025-07-11 - 闲记算法

](http://lonepatient.top/2025/07/11/arxiv_papers_2025-07-11)[)]

scholar.google.com

Chunyao Lu - Google Scholar

](https://scholar.google.com/citations?user=u_jwCCQAAAAJ&hl=en)[)]

researchgate.net

The effects of art therapy on anxiety and depression in breast cancer patients: An updated meta-analysis - ResearchGate

] (https://www.researchgate.net/publication/341644814_The_effects_of_art_therapy_on_anxiety_and_depression_in_breast_cancer_patients_An_updated_meta-analysis)[)]

pubs.acs.org

Digital Descriptors in Predicting Catalysis Reaction Efficiency and Selectivity | The Journal of Physical Chemistry Letters - ACS Publications

](<https://pubs.acs.org/doi/10.1021/acs.jpcllett.4c03733>)[)]

researchgate.net

Machine Learning Descriptors for Data-Driven Catalysis Study - ResearchGate

](https://www.researchgate.net/publication/370802926_Machine_Learning_Descriptors_for_Data-Driven_Catalysis_Study)[)]

pubs.acs.org

Interpretable Catalysis Models Using Machine Learning with Spectroscopic Descriptors

](<https://pubs.acs.org/doi/10.1021/acscatal.3c00611>)[)]

pubmed.ncbi.nlm.nih.gov

Machine Learning Descriptors for Data-Driven Catalysis Study - PubMed

](<https://pubmed.ncbi.nlm.nih.gov/37191279/>)[)]

pmc.ncbi.nlm.nih.gov

Machine Learning in Catalysis, From Proposal to Practicing - PMC

[\(https://pmc.ncbi.nlm.nih.gov/articles/PMC6963892/\)](https://pmc.ncbi.nlm.nih.gov/articles/PMC6963892/)[(https://t0.gstatic.com/faviconV2?url=https://escholarship.org/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)]

escholarship.org

Machine Learning Descriptors for Data-Driven Catalysis Study - eScholarship.org

](https://escholarship.org/content/qt9zc1r279/qt9zc1r279_noSplash_0bbfafabe2d62829521f27df22208efd.pdf)[!]
(https://t0.gstatic.com/faviconV2?url=https://www.researchgate.net/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)

researchgate.net

Machine Learning for Computational Heterogeneous Catalysis ...

](https://www.researchgate.net/publication/333120502_Machine_Learning_for_Computational_Heterogeneous_Catalysis)[!]
(https://t0.gstatic.com/faviconV2?url=https://medium.com/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)

medium.com

Protein Structure Prediction : A Primer (Part 8) | by Siddhant Rai ...

](https://medium.com/@rsiddhant73/protein-structure-prediction-a-primer-part-8-93cf4d4f1f39)

Understanding Diffusion Models and Key Papers

I've successfully gathered comprehensive information on Denoising Diffusion Probabilistic Models (DDPMs), including their forward and reverse processes and mathematical underpinnings. I've also obtained detailed summaries for 'OmniGen: Unified Image Generation,' highlighting its ability to unify various image generation tasks and its simplified architecture. For 'PRE-TEXTTraining Language Models,' I've focused on the foundational concepts of pre-training language models, particularly using transformer architectures for generative tasks. I am still working on obtaining a complete summary for 'DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation' and will attempt to browse its abstract page on arXiv.

Advancements in Diffusion Models and Protein Design

My research into recent advancements in diffusion models confirms their expanding applications beyond basic image generation, now encompassing video, audio, and 3D content creation, along with improvements in sampling efficiency. I've also delved deeper into RF Diffusion, understanding it as an open-source method for *de novo* protein structure generation, which fine-tunes the RoseTTAFold network for protein denoising tasks. This is crucial for conceptualizing its application in catalyst design.

Catalyst Discovery and AlphaFold Integration

I'm continuing to build my understanding of Fischer-Tropsch synthesis and the role of transition metal catalysts. The goal is to explore how diffusion models, drawing inspiration from AlphaFold's protein folding and RF Diffusion's protein design, could be used to generate or identify catalysts that control the density of hydrocarbon products. I'm also focusing on how to computationally represent catalysts for machine learning models, which is a key step in this novel application.

Next Steps: Finalizing Summaries and Structuring the Paper

My immediate next steps involve a targeted effort to obtain the full summary for the 'DpDNet' paper by browsing its arXiv abstract page. Concurrently, I will seek more specific information on catalyst representation to further refine this critical aspect of the research. Once these details are secured, I will proceed with structuring and drafting the paper, integrating all the gathered knowledge into a cohesive and comprehensive document.



Researching websites



arxiv.org

arxiv.org

](https://arxiv.org/abs/2507.07126)[!]
(https://t0.gstatic.com/faviconV2?url=https://pubmed.ncbi.nlm.nih.gov/&client=BARD&type=FAVICON&size=256&fallback_opts=TYPE,SIZE,URL)

pubmed.ncbi.nlm.nih.gov

Machine Learning Descriptors for Data-Driven Catalysis Study

](https://pubmed.ncbi.nlm.nih.gov/37191279/)

Refining Paper Summaries

I've made significant progress in understanding the core concepts of Denoising Diffusion Probabilistic Models (DDPMs), including their mathematical foundations and the forward/reverse processes. I've also gathered detailed information for summarizing 'OmniGen: Unified Image Generation,' noting its ability to unify various image generation tasks and its simplified architecture. For 'PRE-TEXTTraining Language Models,' I've focused on the foundational concepts of pre-training language models, particularly using transformer architectures for generative tasks. I am now

making a final attempt to fully summarize 'DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation' by focusing on its dual-prompt strategy and prompt-aware heads for universal PET-CT segmentation.

Deepening Catalyst Representation Understanding

I'm actively working on gaining a more comprehensive understanding of how catalysts are represented computationally for machine learning. This involves exploring various methods like catalyst fingerprinting and the use of graph neural networks, which will be crucial for developing the novel catalyst discovery section of the paper. This deeper dive will allow me to outline a more robust methodology for generating or identifying catalysts with desired properties.

Next Steps: Finalizing Information and Structuring the Paper

My immediate next steps involve a final, targeted effort to complete the summary for the 'DpDNet' paper. Simultaneously, I will continue to gather comprehensive information on catalyst representation. Once these information gathering phases are complete, I will proceed with synthesizing all the acquired knowledge and structuring the paper, ensuring a logical flow from the technical overview of diffusion models to the specific application in designing hydrocarbon-catalyzing materials.



Researching websites



arxiv.org
arxiv.org

](https://arxiv.org/abs/2507.07126)[]
arxiv.org

DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation - arXiv
](https://arxiv.org/html/2507.07126v1)[]
researchgate.net

DpDNet: An Dual-Prompt-Driven Network for Universal PET-CT Segmentation
](https://www.researchgate.net/publication/393586793_DpDNet_An_Dual-Prompt-Driven_Network_for_Universal_PET-CT_Segmentation)[]
lonepatient.top

Arxiv今日论文| 2025-07-11 - 闲记算法
](http://lonepatient.top/2025/07/11/arxiv_papers_2025-07-11)[]
arxiv.org

Computer Science Jul 2025 - arXiv
](http://www.arxiv.org/list/cs/2025-07?skip=4100&show=100)[]
arxiv.org

Electrical Engineering and Systems Science - arXiv
](http://www.arxiv.org/list/eess/new?skip=0&show=1000)[]
arxiv.org

Artificial Intelligence Jul 2025 - arXiv
](http://arxiv.org/list/cs.AI/2025-07?skip=1050&show=2000)[]
arxiv.org

Machine Learning Accelerated Descriptor Design for Catalyst Discovery in CO2 to Methanol Conversion - arXiv
](https://arxiv.org/html/2412.13838v5)[]
research.aalto.fi

](https://research.aalto.fi/files/186274555/Machine_learning_accelerated_descriptor_design_for_catalyst_discovery_in_CO2_to_methanol_conversion.pdf)[)]

pmc.ncbi.nlm.nih.gov

Applications of Machine Learning in Alloy Catalysts: Rational Selection and Future Development of Descriptors - PMC - PubMed Central

](<https://pmc.ncbi.nlm.nih.gov/articles/PMC9036033/>)[)]

researchgate.net

Machine Learning Descriptors for Data-Driven Catalysis Study - ResearchGate

](https://www.researchgate.net/publication/370802926_Machine_Learning_Descriptors_for_Data-Driven_Catalysis_Study)[)]

arxiv.org

Machine-learning Accelerated Descriptor Design for Catalyst Discovery: A CO2 to Methanol Conversion Case Study - arXiv

](<https://arxiv.org/html/2412.13838v1>)[)]

pubs.acs.org

A Machine Learning and Explainable AI Framework Tailored for Unbalanced Experimental Catalyst Discovery | The Journal of Physical Chemistry C - ACS Publications

](<https://pubs.acs.org/doi/10.1021/acs.jpcc.4c05332>)[)]

pubs.acs.org

Digital Descriptors in Predicting Catalysis Reaction Efficiency and ...

](<https://pubs.acs.org/doi/10.1021/acs.jpcllett.4c03733>)]

Advancements in Diffusion Models and Protein Design

My research into recent advancements in diffusion models has provided a clear picture of their evolution beyond basic image generation, now encompassing applications in video, audio, and 3D content creation, along with improvements in sampling efficiency. Crucially, I have delved into the methodologies of AlphaFold 2 and RFdiffusion. I understand AlphaFold 2's architecture, particularly its Evoformer and Structure Module, and how it leverages multiple sequence alignments and pair representations for protein structure prediction. For RFdiffusion, I've gained insight into its capabilities as an open-source method for *de novo* protein structure generation, which fine-tunes the RoseTTAFold network on protein denoising tasks. This understanding is vital for conceptualizing its application in catalyst design.

Catalyst Discovery for Fischer-Tropsch Synthesis

I have a foundational understanding of Fischer-Tropsch synthesis, including the types of transition metal catalysts used (iron, cobalt, nickel, ruthenium) and how reaction conditions influence product selectivity and density. I've also explored how computational methods, including machine learning and DFT calculations, are being used to accelerate catalyst discovery and optimize properties. While I have some information on catalyst representation for machine learning, a more in-depth, comparative review of different descriptor types would be beneficial for the final section of the paper.

Next Steps: Synthesizing and Reporting

Having gathered the necessary information, I will now proceed to synthesize all the acquired knowledge into a coherent and well-structured paper. I will begin with the technical overview of diffusion models, followed by the summaries of the specified papers (with a note on the 'DpDNet' limitation), an exploration of other current research in diffusion model creation, and finally, a detailed analysis of the application of these models in categorizing potential hydrocarbon-catalyzing materials, drawing inspiration from the AlphaFold and RFdiffusion methodologies.