Diffusion Generative Models for Designing Efficient Singlet Fission Dimers

Lasse Kreimendahl,[†] Mikhail Karnaukh,[†] and Merle I. S. Röhr[*,†,‡]

†Institute of Physical and Theoretical Chemistry, Julius-Maximilians-Universität Würzburg,
Emil-Fischer-Str. 42, 97074 Würzburg, Germany

‡Center for Nanosystems Chemistry, Julius-Maximilians-Universität Würzburg, Theodor-Boveri Weg,
97074 Würzburg, Germany

E-mail: merle.roehr@uni-wuerzburg.de

Abstract

Diffusion generative models, a class of machine learning techniques, have shown remarkable promise in materials science and chemistry by enabling the precise generation of complex molecular structures. In this paper, we propose a novel application of diffusion generative models for stabilizing reactive molecular structures identified through quantum mechanical screening. Specifically, we focus on the design challenge presented by Singlet Fission (SF), a phenomenon crucial for advancing solar cell efficiency beyond theoretical limits. While theoretical chemistry has been successful in predicting intermolecular arrangements with enhanced SF coupling, the practical implementation of these configurations faces challenges due to discrepancies between favorable and stabilized structures. To address this gap, we introduce a three-step strategy combining quantum mechanical screening for identifying optimal molecular arrangements and diffusion generative models for predicting stabilizing linkers. Through a case study on cibalackrot dimers, a promising SF material, we demonstrate the efficacy of our approach in enhancing SF efficiency by stabilizing the desired molecular arrangements.

Diffusion models (DMs) are a class of generative machine learning (ML) techniques that model complex data distributions by learning to iteratively transform random noise into structured outputs through a guided stochastic process, enabling the precise generation of highly diverse and high-quality data.[1,2] In recent years, DMs have received considerable attention and have become state of the art for image generation tasks.[3,4] This progress has led to the development of popular tools such as StabilityAI's Stable Diffusion and Midjourney,

1

as well as the introduction of OpenAI's DALL·E.[5-7] In the realm of materials science and chemistry, these models have begun to demonstrate significant potential in the design and prediction of novel compounds and assemblies.[8,9] A major application of DMs in chemistry is the generation of novel molecular structures, which are essential for tasks such as molecular design and property prediction. These models exploit the ability of diffusion processes to capture long-range dependencies in complex molecular structures, allowing the generation of diverse and chemically plausible molecules.[10,11] In drug discovery and development, DMs can accelerate innovation by learning latent representations of molecular structures, facilitating the generation of novel compounds with desired traits such as high efficacy and low toxicity.[12] In this context, DMs have also been adapted to include the protein target binding pocket as a condition for the generation process, allowing for modeling of protein-ligand interactions.[13-15] Here, DMs for molecule discovery employ various neural network architectures, such as graph-neural networks (GNNs),[16-18] convolutional neural networks (CNNs),[19,20] transformers[11,21,22] or combinations of them,[23] and operate on different molecular representations based on SMILES or molecular graphs.[17,24] They can also integrate with reinforcement learning strategies, directing synthesis towards molecules with specific properties, thereby enhancing the efficiency of drug discovery efforts.[25-27] In addition to drug design, DMs have also been used to generate new polycyclic aromatic hydrocarbons with defined electronic properties for optoelectronic applications.[28] Further, DMs are employed to predict chemical reactions by modeling how reactant molecules diffuse through a reaction network. This ability allows them to estimate the likelihood of various reaction pathways and uncover novel reaction mechanisms.[29,30] Other applications of DMs for molecules include conformer generation,[18] molecular dynamics simulations[31,32] or linker design.[33]

Diffusion Generative Models

The basic idea of DMs — learning to map a simple distribution to a complex one using a parameterized transition kernel — is inspired by nonequilibrium thermodynamics.[1,2] Starting from a tractable distribution, such as a standard normal distribution, the goal is to transform it to the target distribution, represented by the training data set. To learn this mapping, a forward diffusion process is applied where a data point x_0 from the target distribution is progressively *diffused* from time $t = 0$ to $t = T$ by adding random Gaussian noise

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t|\sqrt{1 - \beta_t}x_{t-1}, \beta_t\mathbf{I}) \tag{1}$$

until all of its structure is lost. Here, $\beta_t \in (0, 1)$ controls how much of the data is retained and how much noise is added. In a reverse denoising process, the goal is to reverse the noise addition process by making a

parameterized distribution kernel

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}|\boldsymbol{\mu}_\theta(x_t, t), \sigma_t^2 \mathbf{I}) \tag{2}$$

as close as possible to the true but unknown reverse denoising transition probabilities $q(x_{t-1}|x_t)$. In practice, this is achieved by training a neural network $\boldsymbol{\epsilon}_\theta$ to predict the noise $\boldsymbol{\epsilon}_t$ added during the forward process, and thus the parameters θ of the neural network are typically learned by minimizing the mean squared error loss between the predicted and true noise

$$\mathcal{L} = \mathbb{E}_{t,x_0,\boldsymbol{\epsilon}} \left[\|\boldsymbol{\epsilon}_t - \boldsymbol{\epsilon}_\theta(x_t, t)\|^2 \right]. \tag{3}$$

This setting constitutes a powerful generative model, since, starting from a random sample x_T from the initial normal distribution, one can generate new and unseen data by iteratively removing noise from time $t = T$ to $t = 0$ using the trained neural network. The result is a completely new sample x_0 that follows the underlying distribution of the training data. While DMs have produced impressive results in text-to-image generation, advances in the development of efficient *equivariant* neural networks have significantly improved the ability of these models to also handle data with complex geometric structures, such as molecular data. In so-called geometric DMs, molecules are typically represented as point clouds of atoms, with the diffusion process acting on both atom positions and atom types. Geometric DMs have been successfully used in the past to generate molecular structures of small drug-like molecules, protein scaffolds and molecular linkers. [17,33] We provide a more detailed description of the mathematical background of (geometric) DMs in the SI.

Singlet Fission

Herein, we propose the application of diffusion generative models to predict (supra)molecular linkers capable of stabilizing reactive molecular structures, which have been identified by quantum chemical methods. The exploration of molecular architectures that promote or optimize specific processes is an essential topic in computational chemistry. Singlet Fission (SF), a light-induced phenomenon in electronically interacting molecules in which a single exciton is converted into two triplet excitons, [34] serves as an exemplary case in this research highlighting the complexities involved: besides the monomer-specific energy conditions, the efficiency of SF is determined by the highly structure-dependent electronic couplings and therefore sensitive to the intermolecular packing arrangement. [35] SF has attracted considerable interest due to its ability to potentially surpass the theoretical efficiency limits of single-junction solar cells, however, to date, only

a limited number of materials have demonstrated the capability to efficiently support SF, motivating active research in the theoretical development of SF materials. At the molecular level, theoretical design strategies have proven effective in aligning energy levels, also employing ML techniques.[36–42] However, at the intermolecular level, despite efficient quantum mechanical descriptions of the electronic couplings that allow the identification of optimal molecular (dimer) arrangements to enhance the SF rate,[43–49] there exists a significant discrepancy between these predicted configurations and their stable structures. Effective stabilization strategies are crucial to bridge this gap, ensuring that the theoretically ideal arrangements are translated into reliably structured materials with preserved electronic properties. In response to this gap, our research introduces a strategy that combines the conformational screening of desired electronic functionalities to identify optimal molecular arrangements with the use of diffusion generative models to predict (supra)molecular frameworks capable of stabilizing these structural motifs. Notably, the design of a stabilizing matrix that promotes reactive molecular configurations is a strategy also employed by nature, as exemplified by the Fenna–Matthews–Olson complex of green sulfur bacteria,[50] in which the protein matrix arranges the chlorophylls to modulate electronic couplings, enabling efficient excitation transport.[51] Recent synthetic work has successfully utilized a linker-stabilized strategy to produce defined dimer and trimer stacks of perylene bisimide (PBI) molecules with packing motifs favorable for SF resulting in enhanced SF.[52,53] Here, we rationalize this approach employing quantum mechanical screening along with diffusion generative models and exemplify it by predicting stabilized 2,2'-dihydro-cibalackrot dimer arrangements with enhanced SF efficiency. Cibalackrot has shown promise as an SF material due to its favorable T_1/S_1-ratio of approximately 0.5, alongside high extinction coefficients and robust molecular stability. However, intermolecular interactions in both crystalline and amorphous phases favor the stabilization of the S_1-state over the T_1-state, leading to predominantly endothermic SF and the formation of excimers and charge-separating trap sites. Therefore, cibalackrot is a promising candidate for the ligand-stabilized SF material design, in which the intermolecular interactions are favorably adjusted by the packing motif.

Linker Design Strategy

In our approach, we employ a fast and generally applicable three-step workflow for the prediction of supramolecular matrices capable of stabilizing reactive molecular structures, which we have adapted to the prediction of the linker-stabilized dimer packing motifs with enhanced effective SF coupling, as illustrated in Fig. 1 and described in the following.
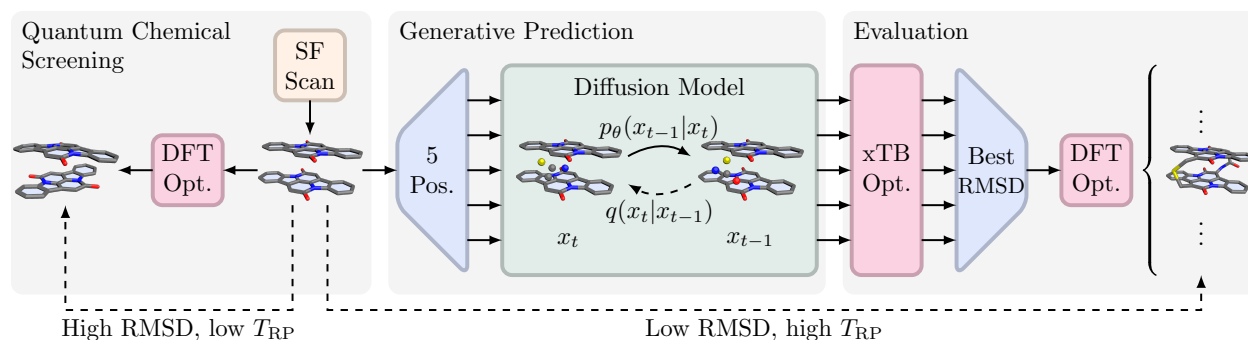
Figure 1: Schematic representation of the applied workflow for the stabilization of moelecular dimers with enhanced SF coupling. The workflow integrates screening of the SF coupling (yellow), geometry optimization (red), a geometric diffusion model (green) and evaluation steps (blue).

Quantum Chemical Screening

In the first, quantum mechanical targeting step, we identify the target structure by screening the suitable electronic descriptor representing the reactivity of interest as a function of molecular coordinates. In this work, we focus on the SF activity. According to Fermi's Golden Rule, the SF rate is proportional to the square of the SF electronic matrix element T_{RP}. Using Michl's molecular frontier orbital (FO) approach, we estimate the effective SF coupling based on diabatic states formed by the frontier orbitals of two monomers A and B. Each monomer has three electronic states: the ground state S_0, the singlet excited state S_1, and the triplet state T_1. The combined system includes seven singlet states: the ground state, two locally excited (LE) states, two charge transfer states (CA and AC), and two doubly excited states (TT and S_1S_1). The SF rate is assumed to be proportional to the squared matrix element of the diabatic Hamiltonian, which links the exciton to the biexciton triplet states through indirect coupling via charge transfer states:

$$T_{\mathrm{RP}} = \left| \frac{\langle \mathrm{LE}|\hat{H}|\mathrm{CA}\rangle\langle \mathrm{CA}|\hat{H}|\mathrm{TT}\rangle}{\Delta E(\mathrm{CA})} + \frac{\langle \mathrm{LE}|\hat{H}|\mathrm{AC}\rangle\langle \mathrm{AC}|\hat{H}|\mathrm{TT}\rangle}{\Delta E(\mathrm{AC})} \right| \tag{4}$$

with $\Delta E(\mathrm{AC}) = E(\mathrm{AC}) - E(S_0S_1)$. Considering the FO model and employing the zero differential overlap approximation, equation (4) reduces to

$$^{\mathrm{Fock}}T_{\mathrm{RP}} = \sqrt{\frac{3}{2}} \left| \frac{(l_A|\hat{F}|l_B)(l_A|\hat{F}|h_B)}{\Delta E(\mathrm{CA})} - \frac{(h_A|\hat{F}|h_B)(h_A|\hat{F}|l_B)}{\Delta E(\mathrm{AC})} \right|, \tag{5}$$

5

where $\hat{F}$ is the Fock matrix. $T_{\rm RP}$ was approximated by switching from Fock matrix elements to overlap matrix elements, represented as $\langle\phi|\hat{F}|\psi\rangle \approx \epsilon\langle\phi|\hat{S}|\psi\rangle$, where $\hat{S}$ denotes the overlap matrix. This approximation is consistent with the further simplification of Michl's model:[44]

$$^{\rm Overlap}T_{\rm RP} = \sqrt{\frac{3}{2}}\epsilon^2 \left| \frac{\langle l_A|\hat{S}|l_B\rangle\langle l_A|\hat{S}|h_B\rangle}{\Delta E(\text{CA})} - \frac{\langle h_A|\hat{S}|h_B\rangle\langle h_A|\hat{S}|l_B\rangle}{\Delta E(\text{AC})} \right| \tag{6}$$

In this work we have kept ϵ constant. $\Delta E(\text{CA})$ and $\Delta E(\text{AC})$ are both approximated to one. By neglecting the constant prefactor and ignoring the unit, we are effectively screening the SF coupling according to:

$$T_{\rm RP} = \left| \langle l_A|\hat{S}|l_B\rangle\langle l_A|\hat{S}|h_B\rangle - \langle h_A|\hat{S}|h_B\rangle\langle h_A|\hat{S}|l_B\rangle \right| \tag{7}$$

The decomposition of singlet fission rates into expressions based on Hartree–Fock (HF) orbital expansion coefficients was first introduced and validated by Josef Michl.[49] Building on this framework, we employed our recently developed semi-empirical implementation, which utilizes the AM1 method within the PYSEQM code, to compute the matrix elements efficiently.[54] This approach enables rapid screening of numerous molecules.[41,55] In previous work, we benchmarked these singlet fission coupling expressions against higher-level quantum chemistry methods, with additional details provided in the SI.[41] The associated Python code is openly accessible online.[56] Employing this quantum mechanical model, we analyzed the effective SF coupling as a function of intermolecular coordinates, focusing on both longitudinal and transversal slipping (as illustrated in Fig. 2 b)), while maintaining a rigid monomer structure. Notice that screening strategies that scan the entire inter- and even intramolecular coordinate space have been proposed and could be employed for this purpose.[45,49,57,58] The screening identified the configuration that maximizes SF coupling, termed the **SF structure**, presented in Fig. 2 b). The dimer motif exhibits a slipping of -0.4 Å in the x-direction and 0.6 Å in the y-direction relative to the perfectly parallel configuration. The vertical distance was set to 3.5 Å. The effective SF coupling amounts $T_{\rm RP}^2 = 1.47 \cdot 10^{-9}$. To determine the need for stabilizing the detected packing motif, we identified the minimum-energy conformation of the cibalackrot dimer using global optimization with the CREST tool at GFN2-xTB level[59–69] and subsequent Density Functional Theory (DFT) optimization at r^{2}SCAN-3c/def2-mTZVPP level.[63,70–77] This conformational search revealed the stable ground-state configuration, which we call the **opt structure** and is shown in Fig. 3 b). For the cibalackrot dimer, this structure differs drastically from the **SF structure** due to a pronounced relative rotation of the monomers of about 68° as well as a clear bowl-shaped curvature of the monomers, while the
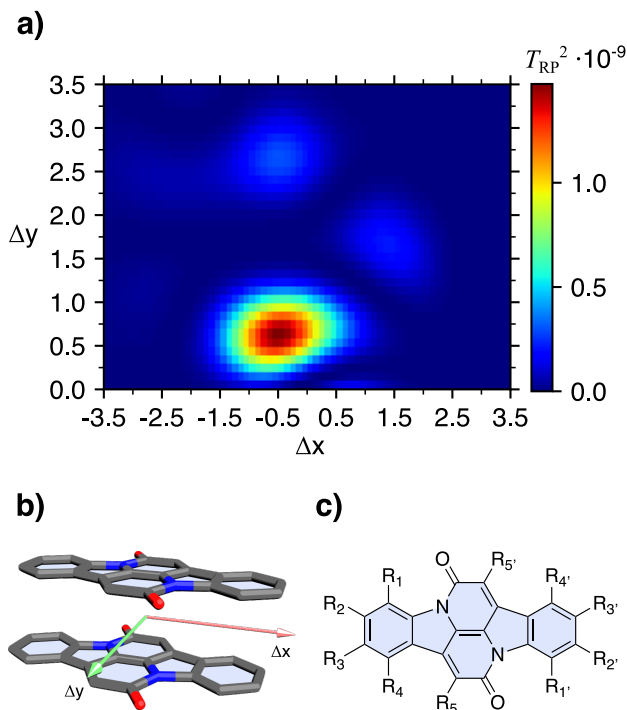
6

Figure 2: a) Two-dimensional scan of the effective SF coupling as a function of longitudinal and transversal slipping in the cibalackrot dimer. Due to the rotational symmetry, only one half of the total scan is shown. b) Illustration of the **SF structure**, indicating the scanning modes. c) Schematic representation of 2,2'-dihydro-cibalackrot, highlighting the potential anchor sites R_1 to R_5 for linkers connecting the two monomers.

vertical distance is reduced to 3.12 Å. We further quantified this difference by determining the Root-Mean-Square Deviation (RMSD), defined as

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \|r_i - r_i'\|^2}\,, \tag{8}$$

between the **opt structure** r and the **SF structure** r' as 2.393 Å. Here, we employed the Kabsch algorithm to ensure maximum overlap between the compared structures.[78,79] Importantly, the SF coupling of the optimized dimer yields $T_{\text{RP}}^2 = 2.79 \cdot 10^{-19}$ and is thus significantly lower compared to the **SF structure**, underscoring the need for effective stabilization.

Generative Prediction

In the second, generative step, the identified target structure is passed to a geometric diffusion model to predict supramolecular matrices that could stabilize the structural motif. In this example, we employed the pre-trained DiffLinker diffusion model to predict linking ligands between the monomers that serve to
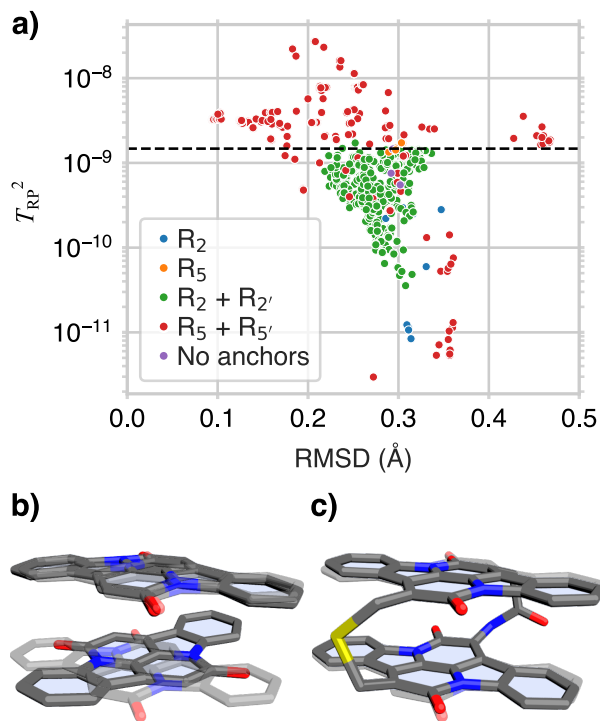
7

Figure 3: a) Distribution of the 1000 DFT-optimized linked dimer structures characterized by the squared effective SF coupling and the RMSD with respect to the **SF structure**, colored by the anchoring position of the generated linker. The dashed line indicates the squared effective SF coupling of the **SF structure** with optimal SF activity ($T_{\mathrm{RP}}^2 = 1.47 \cdot 10^{-9}$) determined from the two-dimensional scan. b) The minimum-energy **opt structure** of the unlinked cibalackrot dimer shows large conformational changes (RMSD = 2.393 Å) and low SF coupling ($T_{\mathrm{RP}}^2 = 2.79 \cdot 10^{-19}$) compared to the **SF structure** (transparent). c) With the generated linker, the dimer structure is effectively stabilized (RMSD = 0.094 Å) and a high SF coupling of $T_{\mathrm{RP}}^2 = 3.21 \cdot 10^{-9}$ is obtained.

stabilize the dimer packing motif.[33] DiffLinker is a geometric diffusion model that is capable of generating molecular linkers between two or more fragments by additionally conditioning the learned denoising process of the geometric DM on the 3D information of the surrounding fragments. In addition, the model is able to generate linkers between designated anchor atoms, as well as to predict the optimal position of the linkers themselves. In this process, the optimal number of linker atoms is determined by the model prior to the start of the generation process, allowing greater flexibility compared to other deep generative linker design models, such as DeLinker and 3DLinker.[80,81] For a detailed discussion of the DiffLinker model, as well as a comparison with the aforementioned models, please refer to the SI. We followed three strategies for linker generation: First, we prompted the DiffLinker model to generate linkers between anchor atoms R_2 and R_5, respectively, as these are described in the literature as the most suitable positions for linkers or other substituents.[82,83] The linker positions are numbered as shown in Fig. 2 c). For this purpose, we used model parameters from the DiffLinker model pre-trained on the GEOM dataset,[84] which contains 3D

8

structures of molecular conformations of small to medium-sized organic molecules. We emphasize at this point that neither the target structure we want to stabilize nor related singlet fission materials are present in this training dataset, but mainly drug-like organic molecules. In addition to the one-linker strategy, we also took into account the opposite positions $R_{2'}$ and $R_{5'}$, resulting in two stabilizing linkers per dimer. Here, the two linkers were generated sequentially by the DiffLinker model. Finally, we let the model determine the most suitable anchor atoms itself, which resulted in arbitrarily positioned linkers (no anchors). For each of the five cases, we sampled 1000 linker structures from the trained DiffLinker model. All generated structures and results obtained are available online.[85] We found that the average number of heavy atoms per linker determined by the model was 3.69, with 8 being the highest and 3 being the lowest. Note that the DiffLinker model produces only heavy atoms and we add hydrogen atoms afterwards to fill the valences. With 64.9%, carbon was amongst the most frequent atom types, followed by nitrogen (17.8%), sulfur (9.9%) and oxygen (7.4%). We also note that among these 5000 linkers, many chemically equivalent motifs were generated multiple times, which is why we present a detailed analysis of the composition and frequency of the generated structures in the SI.
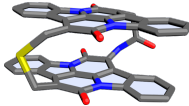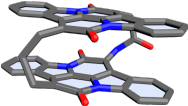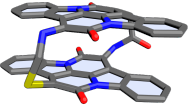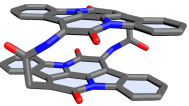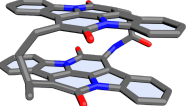
Evaluation

In the third, evaluation step, the individual quality of the generated structures in stabilizing the initially identified reactive structural motif is assessed through quantum chemical optimization and subsequent quantification of structural deviation. In this study, the 5000 generated structures underwent preliminary optimization using a low-level quantum chemistry method GFN2-xTB.[59,61] We evaluated these structures by determining their RMSD values relative to the **SF structure** (linker atoms excluded). For the 1000 structures with the smallest RMSD values, we performed a more refined optimization using higher-level DFT optimization at the r^{2}SCAN-3c/def2-mTZVPP level to ensure greater accuracy. We recalculated the RMSD and effective SF coupling for these structures and identified the best linker-stabilized configurations that promote improved SF efficiency.

Results and Discussion

In Fig. 3 a) we present the distribution of these structures characterized by their squared effective SF coupling and the RMSD with respect to the **SF structure**. This analysis shows that the "two-linker" strategy yields more favorable results, with anchoring positions $R_5 + R_{5'}$ in particular lead to the best performing structures. With positions $R_2 + R_{2'}$, the structures have a slightly lower coupling than the **SF structure**, but still significantlly larger than the unlinked **opt structure**. Furthermore, a clear trend between the RMSD and the

9

Table 1: Presentation of the top 5 structures obtained by the generative workflow, evaluated on the basis of their respective RMSD values and the strength of the effective SF coupling. All DFT-optimized structures exhibit a higher or comparable SF coupling strength relative to the **SF structure**, in each case several orders of magnitude larger than the unlinked **opt structure**.

Structure					
Position	$R_5 + R_{5'}$	$R_5 + R_{5'}$	$R_5 + R_{5'}$	$R_5 + R_{5'}$	$R_5 + R_{5'}$
$T_{RP}^2 \cdot 10^{-9}$	3.21	3.17	3.09	2.61	4.04
RMSD (Å)	0.094	0.124	0.133	0.139	0.166

effective SF coupling can be observed, i.e. low RMSD values generally correlate with high couplings, which supports our hypothesis that high SF activity is achieved by effective stabilization. The five best structures ranked by their RMSD values relative to the **SF structure** are presented in Table 1, which range from 0.094 to 0.166 Å, significantly smaller than those for the unlinked, optimized dimer structure. Furthermore, the calculated SF couplings are notably higher than those for the optimized "free dimer" and even exceed the **SF structure** obtained from the scan. We attribute this to the additional relaxed degrees of freedom, such as rotation or curvature, during optimization, which remain fixed in the scan. From these findings, we conclude that the chosen strategy of generating and evaluating linkers from deep generative models such as DMs to stabilize desired structural motifs has high potential. The fast and effective computational methods used in this work to evaluate the SF coupling and the energy complement the rapid sampling of the generative model and allow a large number of structures ($\approx$5000) to be screened quickly and effectively. It was shown that the generated linkers are able to stabilize the desired conformation, resulting in comparable or even higher SF activity. In addition, we confirmed the hypothesis that the ability of the linkers to stabilize the desired structure (measured in RMSD) is a good indicator of SF activity, as both quantities correlate (see Fig. 3 a)). The workflow also holds high potential for incorporating guided DMs, e.g. for assessments of synthetic accessibility and electronic neutrality into the linker generation process. Synthetic accessibility refers to the feasibility of synthesizing the proposed structures using established chemical methods, while electronic neutrality ensures that the electronic properties of the molecules remain unaffected by the generated linkers.

Conclusion

We therefore believe that deep generative models represent a promising avenue for addressing the challenges of designing high-performing materials for specific applications such as singlet fission, where the precise 3D structure of molecular aggregates plays a critical role. Traditional screening methods, which evaluate the properties of thousands of random samples *post hoc* from existing molecular libraries, are computationally intensive and often inefficient. This inefficiency is further compounded by the limited availability of structured data for specific tasks, such as linker design, where only a small number, or sometimes no candidates, meet the stringent design criteria. Moreover, adapting known linker structures to entirely new fragments remains a highly complex and ambiguous task, rendering such approaches either unsuccessful or impractically slow. In contrast, while generative models require extensive data for training, they possess a key advantage: the ability to learn the underlying structure and relationships within the data rather than merely evaluating pre-existing samples. A trained model, such as DiffLinker, can therefore generate novel structures by effectively exploring a larger, previously unknown portion of the chemical space. This capability allows the discovery of candidates that lie beyond the reach of traditional screening approaches. Furthermore, the generative process provides a unique advantage in linker design by accounting for the surrounding molecular context and facilitating exploration of non-intuitive solutions. This approach has been shown to yield structures with improved SF coupling, emphasizing the potential of generative models in enabling practical and scalable solutions for molecular design challenges. In the future, the integration of generative ML methods for the prediction of molecular matrices will further enhance the practical applicability of these models. However, the main challenge remains (as is likely the case for the entire field) to obtain a sufficiently large amount of structured and reliable data with which to train generative models such as DiffLinker and use them for specific purposes. This means that substantial experimental and computational data, including successful linker examples for dimer or oligomer stabilization, is required to bridge the gap to experimental implementation. In conclusion, by embedding synthetic feasibility and electronic neutrality assessments within the generative framework, alongside robust data integration, we can drive forward the practical utilization of these models in stabilizing complex molecular structures.

Acknowledgement

Supporting Information Available

The Supporting Information includes details on diffusion models, the DiffLinker model, the computational and statistical analysis of the generated linkers, and comparisons of the SF expression with higher-level methods.

References

(1) Sohl-Dickstein, J.; Weiss, E.; Maheswaranathan, N.; Ganguli, S. Deep Unsupervised Learning using Nonequilibrium Thermodynamics. Proceedings of the 32nd International Conference on Machine Learning. 2015; pp 2256–2265.

(2) Ho, J.; Jain, A.; Abbeel, P. Denoising Diffusion Probabilistic Models. Advances in Neural Information Processing Systems. 2020; pp 6840–6851.

(3) Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-Resolution Image Synthesis with Latent Diffusion Models. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022; pp 10674–10685.

(4) Oppenlaender, J. The Creativity of Text-to-Image Generation. Proceedings of the 25th International Academic Mindtrek Conference. New York, NY, USA, 2022; pp 192–202.

(5) Stability AI Stable Diffusion. `https://stablediffusionweb.com`, Accessed: 2024-05-22.

(6) Midjourney. `https://www.midjourney.com`, Accessed: 2024-05-22.

(7) OpenAI DALL-E 3. `https://openai.com/index/dall-e-3/`, Accessed: 2024-07-02.

(8) Yang, S.; Cho, K.; Merchant, A.; Abbeel, P.; Schuurmans, D.; Mordatch, I.; Cubuk, E. D. Scalable Diffusion for Materials Generation. 2024; `http://arxiv.org/abs/2311.09235`, arXiv:2311.09235 [cs.LG], Submitted: 2024-06-03, Accessed: 2024-12-09.

(9) Fu, X.; Xie, T.; Rosen, A. S.; Jaakkola, T.; Smith, J. MOFDiff: Coarse-grained Diffusion for Metal-Organic Framework Design. 2023; `http://arxiv.org/abs/2310.10732`, arXiv:2310.10732 [physics.chem-ph], Submitted: 2023-10-16, Accessed: 2024-12-09.

(10) Liu, C.; Fan, W.; Liu, Y.; Li, J.; Li, H.; Liu, H.; Tang, J.; Li, Q. Generative Diffusion Models on Graphs: Methods and Applications. Electronic proceedings of IJCAI 2023. 2023; pp 6702–6711.

(11) Huang, H.; Sun, L.; Du, B.; Lv, W. Learning Joint 2D & 3D Diffusion Models for Complete Molecule Generation. 2023; `http://arxiv.org/abs/2305.12347`, arXiv:2305.12347 [q-bio.BM], Submitted: 2023-06-04, Accessed: 2024-12-09.

(12) Alakhdar, A.; Poczos, B.; Washburn, N. Diffusion Models in De Novo Drug Design. *Journal of Chemical Information and Modeling* **2024**, *64*, 7238–7256.

(13) Schneuing, A.; Harris, C.; Du, Y.; Didi, K.; Jamasb, A.; Igashov, I.; Du, W.; Gomes, C.; Blundell, T. L.; Lio, P.; Welling, M.; Bronstein, M.; Correia, B. Structure-based drug design with equivariant diffusion models. *Nature Computational Science* **2024**, In press.

(14) Lin, H.; Huang, Y.; Zhang, O.; Ma, S.; Liu, M.; Li, X.; Wu, L.; Ji, S.; Hou, T.; Li, S. Z. Q. DiffBP: Generative Diffusion of 3D Molecules for Target Protein Binding. *Chemical Science* **2024**, In press.

(15) Guan, J.; Qian, W. W.; Peng, X.; Su, Y.; Peng, J.; Ma, J. 3D Equivariant Diffusion for Target-Aware Molecule Generation and Affinity Prediction. 2023; `http://arxiv.org/abs/2303.03543`, arXiv:2303.03543 [q-bio.BM], Submitted: 2023-03-06, Accessed: 2024-12-09.

(16) Satorras, V. G.; Hoogeboom, E.; Welling, M. E(n) Equivariant Graph Neural Networks. Proceedings of the 38th International Conference on Machine Learning. 2021; pp 9323–9332.

(17) Hoogeboom, E.; Satorras, V. G.; Vignac, C.; Welling, M. Equivariant Diffusion for Molecule Generation in 3D. Proceedings of the 39th International Conference on Machine Learning. 2022; pp 8867–8887.

(18) Xu, M.; Yu, L.; Song, Y.; Shi, C.; Ermon, S.; Tang, J. GeoDiff: a Geometric Diffusion Model for Molecular Conformation Generation. 2022; `http://arxiv.org/abs/2203.02923`, arXiv:2203.02923 [cs.LG], Submitted: 2022-03-06, Accessed: 2024-12-09.

(19) Schütt, K.; Kindermans, P.-J.; Sauceda Felix, H. E.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. Advances in Neural Information Processing Systems. 2017; pp 992–1002.

(20) Huang, L.; Zhang, H.; Xu, T.; Wong, K.-C. MDM: Molecular Diffusion Model for 3D Molecule Generation. *Proceedings of the AAAI Conference on Artificial Intelligence* **2023**, *37*, 5105–5112, Number: 4.

(21) Vignac, C.; Osman, N.; Toni, L.; Frossard, P. MiDi: Mixed Graph and 3D Denoising Diffusion for Molecule Generation. Machine Learning and Knowledge Discovery in Databases: Research Track:

European Conference, ECML PKDD 2023, Turin, Italy, September 18–22, 2023, Proceedings, Part II. 2023; pp 560–576.

(22) Vignac, C.; Krawczuk, I.; Siraudin, A.; Wang, B.; Cevher, V.; Frossard, P. DiGress: Discrete Denoising diffusion for graph generation. 2023; `http://arxiv.org/abs/2209.14734`, arXiv:2209.14734 [cs.LG], Submitted: 2023-05-23, Accessed: 2024-12-09.

(23) Huang, H.; Sun, L.; Du, B.; Lv, W. Conditional Diffusion Based on Discrete Graph Structures for Molecular Graph Generation. *Proceedings of the AAAI Conference on Artificial Intelligence* **2023**, *37*, 4302–4311, Number: 4.

(24) Peng, X.; Zhu, F. Hitting stride by degrees: Fine grained molecular generation via diffusion model. *Expert Systems with Applications* **2024**, *244*, 122949.

(25) Ragoza, M.; Masuda, T.; Koes, D. R. Generating 3D molecules conditional on receptor binding sites with deep generative models. *Chemical Science* **2022**, *13*, 2701–2713.

(26) Liu, M.; Luo, Y.; Uchino, K.; Maruhashi, K.; Ji, S. Generating 3D Molecules for Target Protein Binding. Proceedings of the 39th International Conference on Machine Learning. 2022; pp 13912–13924.

(27) Bai, Q.; Ma, J.; Xu, T. In *Applications of Generative AI*; Lyu, Z., Ed.; Springer International Publishing, 2024; pp 461–475.

(28) Weiss, T.; Mayo Yanes, E.; Chakraborty, S.; Cosmo, L.; Bronstein, A. M.; Gershoni-Poranne, R. Guided diffusion for inverse molecular design. *Nature Computational Science* **2023**, *3*, 873–882.

(29) Kim, S.; Woo, J.; Kim, W. Y. Diffusion-based generative AI for exploring transition states from 2D molecular graphs. *Nature Communications* **2024**, *15*, 341.

(30) Duan, C.; Du, Y.; Jia, H.; Kulik, H. J. Accurate transition state generation with an object-aware equivariant elementary reaction diffusion model. *Nature Computational Science* **2023**, *3*, 1045–1055.

(31) Wu, F.; Li, S. Z. DiffMD: A Geometric Diffusion Model for Molecular Dynamics Simulations. *Proceedings of the AAAI Conference on Artificial Intelligence* **2023**, *37*, 5321–5329, Number: 4.

(32) Han, J.; Xu, M.; Lou, A.; Ye, H.; Ermon, S. Geometric Trajectory Diffusion Models. 2024; `http://arxiv.org/abs/2410.13027`, arXiv:2410.13027 [cs.CV], Submitted: 2024-10-16, Accessed: 2024-12-09.

(33) Igashov, I.; Stärk, H.; Vignac, C.; Schneuing, A.; Satorras, V. G.; Frossard, P.; Welling, M.; Bronstein, M.; Correia, B. Equivariant 3D-conditional diffusion model for molecular linker design. *Nature Machine Intelligence* **2024**, *6*, 417–427.

(34) Smith, M. B.; Michl, J. Singlet Fission. *Chemical Reviews* **2010**, *110*, 6891–6936.

(35) Casanova, D. Theoretical Modeling of Singlet Fission. *Chemical Reviews* **2018**, *118*, 7164–7207.

(36) Padula, D.; Omar, O. H.; Nematiaram, T.; Troisi, A. Singlet fission molecules among known compounds: finding a few needles in a haystack. *Energy & Environmental Science* **2019**, *12*, 2412–2416.

(37) Zeng, T.; Ananth, N.; Hoffmann, R. Seeking Small Molecules for Singlet Fission: A Heteroatom Substitution Strategy. *Journal of the American Chemical Society* **2014**, *136*, 12638–12647.

(38) Zeng, W.; Bakouri, O. E.; Szczepanik, D. W.; Bronstein, H.; Ottosson, H. Excited state character of Cibalackrot-type compounds interpreted in terms of Hückel-aromaticity: a rationale for singlet fission chromophore design. *Chemical Science* **2021**, *12*, 6159–6171.

(39) Liu, X.; Tom, R.; Gao, S.; Marom, N. Assessing Zethrene Derivatives as Singlet Fission Candidates Based on Multiple Descriptors. *The Journal of Physical Chemistry C* **2020**, *124*, 26134–26143.

(40) Weber, F.; Mori, H. Machine-learning assisted design principle search for singlet fission: an example study of cibalackrot. *npj Computational Materials* **2022**, *8*, 1–14.

(41) Singh, A.; Humeniuk, A.; Röhr, M. I. S. Energetics and optimal molecular packing for singlet fission in BN-doped perylenes: electronic adiabatic state basis screening. *Physical Chemistry Chemical Physics* **2021**, *23*, 16525–16536.

(42) Schaufelberger, L.; Blaskovits, J. T.; Laplaza, R.; Jorner, K.; Corminboeuf, C. Inverse Design of Singlet-Fission Materials with Uncertainty-Controlled Genetic Optimization. *Angewandte Chemie International Edition* e202415056, In press.

(43) Jadhav, S. D.; Sasikumar, D.; Hariharan, M. Modulating singlet fission through interchromophoric rotation. *Physical Chemistry Chemical Physics* **2022**, *24*, 16193–16199.

(44) Buchanan, E. A.; Havlas, Z.; Michl, J. In *Advances in Quantum Chemistry*; Sabin, J. R., Brändas, E. J., Eds.; Advances in Quantum Chemistry: Ratner Volume; Academic Press, 2017; Vol. 75; pp 175–227.

(45) Tom, R.; Gao, S.; Yang, Y.; Zhao, K.; Bier, I.; Buchanan, E. A.; Zaykov, A.; Havlas, Z.; Michl, J.; Marom, N. Inverse Design of Tetracene Polymorphs with Enhanced Singlet Fission Performance by Property-Based Genetic Algorithm Optimization. *Chemistry of Materials* **2023**, *35*, 1373–1386.

(46) Mirjani, F.; Renaud, N.; Gorczak, N.; Grozema, F. C. Theoretical Investigation of Singlet Fission in Molecular Dimers: The Role of Charge Transfer States and Quantum Interference. *The Journal of Physical Chemistry C* **2014**, *118*, 14192–14199.

(47) Ryerson, J. L. et al. Structure and photophysics of indigoids for singlet fission: Cibalackrot. *The Journal of Chemical Physics* **2019**, *151*, 184903.

(48) Renaud, N.; Sherratt, P. A.; Ratner, M. A. Mapping the Relation between Stacking Geometries and Singlet Fission Yield in a Class of Organic Crystals. *The Journal of Physical Chemistry Letters* **2013**, *4*, 1065–1069.

(49) Zaykov, A.; Felkel, P.; Buchanan, E. A.; Jovanovic, M.; Havenith, R. W. A.; Kathir, R. K.; Broer, R.; Havlas, Z.; Michl, J. Singlet Fission Rate: Optimized Packing of a Molecular Pair. Ethylene as a Model. *Journal of the American Chemical Society* **2019**, *141*, 17729–17743.

(50) Fenna, R. E.; Matthews, B. W. Chlorophyll arrangement in a bacteriochlorophyll protein from Chlorobium limicola. *Nature* **1975**, *258*, 573–577.

(51) Renger, T.; Marcus, R. A. On the relation of protein dynamics and exciton relaxation in pigment–protein complexes: An estimation of the spectral density and a theory for the calculation of optical spectra. *The Journal of Chemical Physics* **2002**, *116*, 9997–10019.

(52) Hong, Y.; Kim, J.; Kim, W.; Kaufmann, C.; Kim, H.; Würthner, F.; Kim, D. Efficient Multiexciton State Generation in Charge-Transfer-Coupled Perylene Bisimide Dimers via Structural Control. *Journal of the American Chemical Society* **2020**, *142*, 7845–7857.

(53) Hong, Y.; Rudolf, M.; Kim, M.; Kim, J.; Schembri, T.; Krause, A.-M.; Shoyama, K.; Bialas, D.; Röhr, M. I. S.; Joo, T.; Kim, H.; Kim, D.; Würthner, F. Steering the multiexciton generation in slip-stacked perylene dye array via exciton coupling. *Nature Communications* **2022**, *13*, 4488.

(54) Zhou, G.; Nebgen, B.; Lubbers, N.; Malone, W.; Niklasson, A. M. N.; Tretiak, S. Graphics Processing Unit-Accelerated Semiempirical Born Oppenheimer Molecular Dynamics Using PyTorch. *Journal of Chemical Theory and Computation* **2020**, *16*, 4951–4962.

(55) Greiner, J. E.; Singh, A.; Röhr, M. I. S. Functionality optimization for effective singlet fission coupling screening in the full-dimensional molecular and intermolecular coordinate space. *Physical Chemistry Chemical Physics* **2024**, *26*, 19257–19265.

(56) SFast: GitHub Repository. `https://github.com/roehr-lab/SFast-Singlet-Fission-adiabatic-basis-screening` Accessed: 2024-12-09.

(57) Röhr, M. I. S. New theoretical methods for the exploration of functional landscapes. *International Journal of Quantum Chemistry* **2021**, *121*, e26747.

(58) Lindner, J. O.; Röhr, M. I. S. Metadynamics for automatic sampling of quantum property manifolds: exploration of molecular biradicality landscapes. *Physical Chemistry Chemical Physics* **2019**, *21*, 24716–24722.

(59) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation* **2019**, *15*, 1652–1671.

(60) Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Computational Molecular Science* **2018**, *8*, e1327.

(61) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Computational Molecular Science* **2021**, *11*, e1493.

(62) Neese, F. The ORCA program system. *WIREs Computational Molecular Science* **2012**, *2*, 73–78.

(63) Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Computational Molecular Science* **2022**, *12*, e1606.

(64) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *The Journal of Chemical Physics* **2020**, *152*, 224108.

(65) Neese, F. Approximate second-order SCF convergence for spin unrestricted wavefunctions. *Chemical Physics Letters* **2000**, *325*, 93–98.

(66) Grimme, S. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *Journal of Chemical Theory and Computation* **2019**, *15*, 2847–2862.

(67) Pracht, P.; Bohle, F.; Grimme, S. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Physical Chemistry Chemical Physics* **2020**, *22*, 7169–7192.

(68) Pracht, P.; Grimme, S. Calculation of absolute molecular entropies and heat capacities made simple. *Chemical Science* **2021**, *12*, 6551–6568.

(69) Pracht, P.; Grimme, S.; Bannwarth, C.; Bohle, F.; Ehlert, S.; Feldmann, G.; Gorges, J.; Müller, M.; Neudecker, T.; Plett, C.; Spicher, S.; Steinbach, P.; Wesołowski, P. A.; Zeller, F. CREST—A program for the exploration of low-energy molecular chemical space. *The Journal of Chemical Physics* **2024**, *160*, 114110.

(70) Grimme, S.; Hansen, A.; Ehlert, S.; Mewes, J.-M. r2SCAN-3c: A "Swiss army knife" composite electronic-structure method. *The Journal of Chemical Physics* **2021**, *154*, 064103.

(71) Neese, F. An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix. *Journal of Computational Chemistry* **2003**, *24*, 1740–1747.

(72) Neese, F. The SHARK integral generation and digestion system. *Journal of Computational Chemistry* **2023**, *44*, 381–396.

(73) Lehtola, S.; Steigemann, C.; Oliveira, M. J. T.; Marques, M. A. L. Recent developments in libxc — A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1–5.

(74) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *The Journal of Chemical Physics* **2017**, *147*, 034112.

(75) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *The Journal of Chemical Physics* **2019**, *150*, 154122.

(76) Caldeweyher, E.; Mewes, J.-M.; Ehlert, S.; Grimme, S. Extension and evaluation of the D4 London-dispersion model for periodic systems. *Physical Chemistry Chemical Physics* **2020**, *22*, 8499–8512.

(77) Grimme, S.; Bohle, F.; Hansen, A.; Pracht, P.; Spicher, S.; Stahn, M. Efficient Quantum Chemical Calculation of Structure Ensembles and Free Energies for Nonrigid Molecules. *The Journal of Physical Chemistry A* **2021**, *125*, 4039–4054.

(78) Kromann, J. Calculate Root-mean-square devia- tion (RMSD) of Two Molecules Using Rotation. `https://github.com/charnley/rmsd`, Accessed: 2024-09-02.

(79) Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A* **1976**, *32*, 922–923.

(80) Imrie, F.; Bradley, A. R.; van der Schaar, M.; Deane, C. M. Deep Generative Models for 3D Linker Design. *Journal of Chemical Information and Modeling* **2020**, *60*, 1983–1995.

(81) Huang, Y.; Peng, X.; Ma, J.; Zhang, M. 3DLinker: An E(3) Equivariant Variational Autoencoder for Molecular Linker Design. Proceedings of the 39th International Conference on Machine Learning. 2022; pp 9280–9294.

(82) Zeng, W.; Szczepanik, D. W.; Bronstein, H. Cibalackrot-type compounds: Stable singlet fission materials with aromatic ground state and excited state. *Journal of Physical Organic Chemistry* **2023**, *36*, e4441.

(83) Kaleta, J.; Dudič, M.; Ludvíková, L.; Liška, A.; Zaykov, A.; Rončević, I.; Mašát, M.; Bednárová, L.; Dron, P. I.; Teat, S. J.; Michl, J. Phenyl-Substituted Cibalackrot Derivatives: Synthesis, Structure, and Solution Photophysics. *The Journal of Organic Chemistry* **2023**, *88*, 6573–6587.

(84) Axelrod, S.; Gómez-Bombarelli, R. GEOM, energy-annotated molecular conformations for property prediction and molecular generation. *Scientific Data* **2022**, *9*, 185.

(85) SFLinker: GitHub Repository. `https://github.com/roehr-lab/SFLinker`, Accessed: 2024-12-09.