TeBaAb: Text-Based Antigen-Conditioned Antibody Redesign via Directed Evolution

Anonymous Author(s)
Affiliation
Address
email

Abstract

The design of antibodies with high affinity and specificity for target antigens is a cornerstone of therapeutic and diagnostic innovation. Traditional optimization strategies, such as phage or yeast display and directed evolution, remain resource-intensive and limited in their ability to integrate contextual information. Recent AI-driven approaches have accelerated protein engineering, but most rely exclusively on structured inputs, overlooking the potential of natural language as a flexible design interface. In this work, we introduce **TeBaAb**, a novel text-based antigen-conditioned framework for antibody redesign that combines generative modeling with iterative optimization inspired by directed evolution. TeBaAb integrates a Conditional Variational Autoencoder (CVAE) jointly conditioned on antigen sequences and textual descriptions of antibody properties, coupled with a two-stage binding affinity predictor and an iterative enrichment loop. To support this approach, we curated AbDes, a new dataset of 7,800 text–antibody–antigen pairs with accompanying structural and binding information. Experimental evaluations demonstrate that TeBaAb improves the predicted binding affinity by an average of 15.5% compared to the original antibodies, while preserving structural confidence (RMSPE < 1.0Å) and generating sequences that are diverse and novel. By enabling text-conditioned antigen-specific antibody design, TeBaAb provides a promising new paradigm for accelerating therapeutic antibody discovery and expanding the antibody design space beyond traditional methods.

1 Introduction

Antibodies are Y-shaped glycoproteins produced by the adaptive immune system to recognize and neutralize antigens. Their high specificity arises from complementarity-determining regions (CDRs)—hypervariable loops in the variable domains of the heavy and light chains—that bind to distinct epitopes on the antigen surface. This mechanism underlies their essential role in immune defense and has been widely exploited in biotechnology and medicine. Monoclonal antibodies, in particular, are a leading class of biopharmaceuticals used in targeted therapies for cancer, autoimmune diseases, and infectious pathogens [4]. Given their clinical impact, substantial effort has gone into designing antibodies with improved biophysical and functional properties [11].

Traditional antibody engineering methods, such as phage display and directed evolution [24], are effective but often slow, labor-intensive, and limited by the scale of experimental screening. These approaches struggle to explore the vast combinatorial space of antibody sequences and to optimize multiple objectives simultaneously. Recently, artificial intelligence (AI) has enabled data-driven models that predict sequence–function relationships, generate novel antibodies, and estimate biophysical properties in silico—paving the way for faster, more scalable, and more targeted antibody discovery pipelines [11].

Submitted to Multi-modal Foundation Models and Large Language Models for Life Sciences workshop (NeurIPS 2025).

Despite the success of AI-driven approaches in antibody design, most existing computational methods focus on generating or optimizing antibody sequences conditioned solely on structured inputs, such as antigen sequences or 3D structures. These models typically lack the ability to incorporate flexible, high-level design intents specified by users, such as functional requirements, therapeutic context, or developability constraints, especially when such information is only available in natural language form. This limitation restricts the usability and adaptability of current systems in real-world therapeutic discovery workflows, where expert knowledge is often communicated through free-text annotations or design briefs.

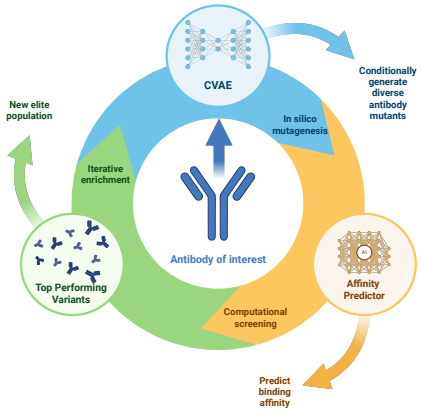

Figure 1: The TeBaAb pipeline integrates generative modeling, predictive evaluation, and iterative refinement for antibody redesign. A Conditional Variational Autoencoder (CVAE) generates candidate sequences conditioned on antigen sequences and textual descriptions. A two-stage affinity predictor evaluates binding strength, and top variants are iteratively fed back into the CVAE, forming an in-silico directed evolution loop.

Recently, there has been a growing interest in using natural language as a conditioning modality for protein generation [17]. Text-conditioned design offers a promising path toward more intuitive and controllable protein engineering, allowing users to specify desired properties or functional behaviors directly through human language. Inspired by these advances, we introduce **TeBaAb**, a text-based, antigen-conditioned framework for antibody redesign that allows guided optimization through natural language prompts and directed evolution strategies. A broader overview of related work in text-conditioned protein and antibody design is provided in Appendix A.

In summary, this work makes the following key contributions:

- We propose a novel framework for antibody design. Our framework integrates two key components: Conditional VAE for antibody generation and an optimization pipeline to achieve a higher binding affinity antibody that is guided by our predictor model.
- Although comprehensive data on antibody-antigen pairs are scarce, we have aggregated data from multiple sources, annotated antibodies, to create AbDes, a uniform dataset for antibody design. The dataset and the source code will be publicly released upon acceptance of the paper.
- Our experiments demonstrate the feasibility of designing antibodies based on descriptions while still targeting specific antigens, opening up a new research direction in the field of antibody discovery.

2 Methods

The TeBaAb framework operates in two coordinated phases, as illustrated in Figure 1. First, a Conditional Variational Autoencoder (CVAE) generates candidate antibody sequences conditioned on both the antigen sequence and a natural language description of the desired antibody properties. This conditioning ensures that the generated variants are biologically relevant and compatible with the original scaffold. In the second phase, a two-stage deep learning predictor estimates the binding affinity of each candidate, enabling rapid in silico prioritization without requiring immediate experimental validation.

To further refine the candidates, TeBaAb incorporates an iterative optimization loop inspired by directed evolution [25]. In each round, the top-performing sequences, those with the highest predicted binding affinities, are reintroduced as inputs to the CVAE, guiding the generation toward progressively improved variants. Over successive iterations, this loop mimics natural selection by converging on antibody sequences with enhanced predicted performance, while operating entirely within a computational framework.

Detailed descriptions of each component, including model architectures, training protocols, and input encoding strategies, are provided in Appendix B.

3 Experiments

3.1 Dataset

Text-conditioned antibody design requires a multimodal dataset combining antibody sequences, antigen context, and descriptive language. While existing resources such as SAbDab [7], AbSet [2], and ABSD [20] provide rich antibody–antigen data, they lack the natural language annotations needed for training models like TeBaAb.

To address this, we curated **AbDes**, a new dataset of 7,800 antibody–antigen pairs with aligned textual descriptions. Starting from AbSet, we filtered entries with complete heavy/light chains and antigen sequences, then retrieved associated metadata from the Protein Data Bank (PDB) [5] using each entry's PDB ID. Extracted fields include titles, source organisms, expression systems, and classification labels, forming lightweight but informative natural language annotations. The structure of each entry and representative examples are provided in Appendix C.2.

To train the binding affinity predictor used in TeBaAb's optimization loop, we extracted experimentally determined ΔG values from SAbDab. Due to data scarcity, only 400 such entries were available with complete sequence triplets and affinity labels. Despite the limited size, these high-quality records were essential for supervising the affinity oracle.

3.2 Results

In this section, we present a comprehensive evaluation of the TeBaAb framework, demonstrating its efficacy in the redesign of antibody sequences. As our method introduces a novel approach to text-conditioned antibody design, direct baseline comparisons with existing models are not feasible. Instead, we evaluate TeBaAb across several key metrics defined in Appendix C.1: **Maximum Binding Affinity (MBA), diversity, novelty, and structural confidence**. In addition, we analyze the influence of natural language descriptions on the design outcomes to assess how textual conditioning guides optimization. These results are provided in Appendix D.

3.2.1 Binding Affinity Optimization

Our primary objective is to enhance the binding affinity of antibodies to antigens. Table 1 summarizes the average predicted binding affinity for the original antibodies and their corresponding TeBaAb optimized variants. We report the predicted ΔG values, where lower values indicate stronger binding. The improvement in binding affinity is evident from the decrease in ΔG for optimized sequences.

Table 1: Average Predicted Binding Affinity (ΔG) for Original vs. TeBaAb-Optimized Antibodies.

Metric	**Original ΔG (Avg.)**	**Optimized ΔG (Avg.)**	**Average Improvement (%)**
Binding affinity	-10.04	-11.60	15.5

Note: Improvement (%) is calculated as $(|\text{Optimized } \Delta G| - |\text{Original } \Delta G|)/|\text{Original } \Delta G| \times 100$, reflecting the increase in absolute affinity.

As shown in Figure 2, TeBaAb consistently generates antibody sequences with improved predicted binding affinities, highlighting the framework's ability to effectively optimize this critical property.

3.2.2 Sequence Diversity and Novelty

To evaluate the generative capabilities of TeBaAb beyond affinity optimization, we evaluated the diversity and novelty of the antibody sequences generated. Diversity quantifies the variation within the redesigned set, while novelty measures their distinctiveness from the training data. Table 2 presents these metrics.

Table 2: Sequence Diversity and Novelty of TeBaAb-Generated Antibodies.

Metric	Value
Average Levenshtein Distance (Diversity)	87.73
Average Minimum Levenshtein Distance to Training Set (Novelty)	28.95

Note: Higher values for both metrics indicate greater diversity and novelty, respectively.

The calculated average Levenshtein distance [15] indicates that TeBaAb is capable of producing a diverse set of antibody sequences, which is crucial for exploring a wide range of potential solutions. Furthermore, the average minimum Levenshtein distance to the training set of 28.95 demonstrates the framework's ability to generate novel sequences, minimizing overlap with existing data and potentially leading to innovative designs.

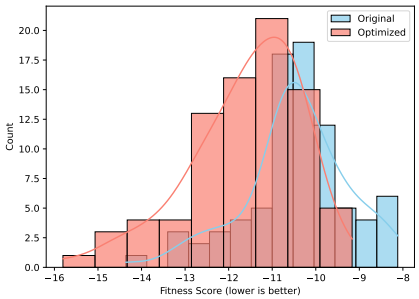

Figure 2: Histogram & KDE: Original vs Optimized Binding Affinity.

3.2.3 Structural Confidence

The maintenance of the structural integrity of the antibody scaffold is paramount for its stability and function. We use ABodyBuilder2 [1] to forecast the three-dimensional configurations of modified antibody sequences, utilizing its ensemble of four models to assess the reliability of the predictions. In Table 3, we report the root mean prediction error (RMSPE) for different regions of antibody. The prediction errors are **below 1.0 Å**, demonstrating the consistency and quality of the predicted structures, further suggesting that optimized sequences maintain structural characteristics that are well modeled by standard prediction tools.

Table 3: Average Structural Confidence (RMSPE) for TeBaAb-Optimized Antibodies (Å).

Region	Original Pred. Error	Optimized Pred. Error
Framework H-chain	0.32	0.72
CDR-H1	0.30	0.45
CDR-H2	0.19	0.27
CDR-H3	0.19	0.29
Framework L-chain	0.24	0.51
CDR-L1	0.25	0.74
CDR-L2	0.18	0.35
CDR-L3	0.24	0.36
Average	0.24	0.48

4 Conclusion

We introduced TeBaAb, a novel framework for text-based, antigen-conditioned antibody redesign that integrates a CVAE generation model with an Oracle-guided optimization pipeline. To support this, we curated AbDes, a dataset of 7,800 text-antibody-antigen triplets, enabling learning from natural language descriptions. Experimental results show that TeBaAb consistently improves predicted binding affinity while maintaining high structural confidence, demonstrating its potential for practical therapeutic design. By enabling controllable antibody optimization from descriptive prompts, TeBaAb opens a new direction in antibody engineering. Future work will extend textual conditioning to a wider range of properties, incorporate structural constraints, and pursue wet-lab validation of designed candidates.

References

[1] Brennan Abanades, Wing Ki Wong, Fergus Boyles, Guy Georges, Alexander Bujotzek, and Charlotte M Deane. ImmuneBuilder: Deep-Learning models for predicting the structures of immune proteins. *Communications Biology*, 6(1):575, May 2023.

[2] Diego S. Almeida, Matheus V. Almeida, Jean V. Sampaio, Eduardo M. Gaieta, Andrielly H. S. Costa, Francisco F. A. Rabelo, César L. Cavalcante, Geraldo R. Sartori, and João H. M. Silva. Abset: A standardized data set of antibody structures for machine learning applications. *Journal of Chemical Information and Modeling*, 65(10):4767–4774, 2025. PMID: 40349368.

[3] Iz Beltagy, Kyle Lo, and Arman Cohan. Scibert: A pretrained language model for scientific text, 2019.

[4] Mitchell Berger, Vidya Shankar, and Abbas Vafai. Therapeutic applications of monoclonal antibodies. *The American journal of the medical sciences*, 324(1):14–30, 2002.

[5] Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. The protein data bank. *Nucleic Acids Research*, 28(1):235–242, 01 2000.

[6] Fengyuan Dai, Yuliang Fan, Jin Su, Chentong Wang, Chenchen Han, Xibin Zhou, Jianming Liu, Hui Qian, Shunzhi Wang, Anping Zeng, Yajie Wang, and Fajie Yuan. Toward de novo protein design from natural language. *bioRxiv*, 2024.

[7] James Dunbar, Konrad Krawczyk, Jinwoo Leem, Terry Baker, Angelika Fuchs, Guy Georges, Jiye Shi, and Charlotte M. Deane. Sabdab: the structural antibody database. *Nucleic Acids Research*, 42(D1):D1140–D1146, 11 2013.

[8] Ahmed Elnaggar, Michael Heinzinger, Christian Dallago, Ghalia Rehawi, Yu Wang, Llion Jones, Tom Gibbs, Tamas Feher, Christoph Angerer, Martin Steinegger, Debsindhu Bhowmik, and Burkhard Rost. ProtTrans: Toward understanding the language of life through Self-Supervised learning. *IEEE Trans Pattern Anal Mach Intell*, 44(10):7112–7127, September 2022.

[9] Hao Fu, Chunyuan Li, Xiaodong Liu, Jianfeng Gao, Asli Celikyilmaz, and Lawrence Carin. Cyclical annealing schedule: A simple approach to mitigating KL vanishing. *CoRR*, abs/1903.10145, 2019.

[10] Thomas Hayes, Roshan Rao, Halil Akin, Nicholas J. Sofroniew, Deniz Oktay, Zeming Lin, Robert Verkuil, Vincent Q. Tran, Jonathan Deaton, Marius Wiggert, Rohil Badkundri, Irhum Shafkat, Jun Gong, Alexander Derry, Raul S. Molina, Neil Thomas, Yousuf A. Khan, Chetan Mishra, Carolyn Kim, Liam J. Bartie, Matthew Nemeth, Patrick D. Hsu, Tom Sercu, Salvatore Candido, and Alexander Rives. Simulating 500 million years of evolution with a language model. *Science*, 387(6736):850–858, 2025.

[11] Alissa M Hummer, Brennan Abanades, and Charlotte M Deane. Advances in computational structure-based antibody design. *Current opinion in structural biology*, 74:102379, 2022.

[12] Yuran Jia, Bing He, Tianxu Lv, YangXiao, Tianyi Zhao, and Jianhua Yao. De novo design of antigen-specific antibodies using structural constraint-based generative language model. In *ICLR 2025 Workshop on Generative and Experimental Perspectives for Biomolecular Design*, 2025.

[13] Ruofan Jin, Qing Ye, Jike Wang, Zheng Cao, Dejun Jiang, Tianyue Wang, Yu Kang, Wanting Xu, Chang-Yu Hsieh, and Tingjun Hou. Attabseq: an attention-based deep learning prediction method for antigen–antibody binding affinity changes based on protein sequences. *Briefings in Bioinformatics*, 25(4):bbae304, 07 2024.

[14] Henry Kenlay, Frédéric A. Dreyer, Aleksandr Kovaltsuk, Dom Miketa, Douglas Pires, and Charlotte M. Deane. Large scale paired antibody language models, 2024.

[15] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *Dokl. Akad. Nauk SSSR*, 163:845–848, 1965.

[16] Nuowei Liu, Jiahao Kuang, Yanting Liu, Changzhi Sun, Tao Ji, Yuanbin Wu, and Man Lan. Protein design with dynamic protein vocabulary, 2025.

[17] Shengchao Liu, Yanjing Li, Zhuoxinran Li, Anthony Gitter, Yutao Zhu, Jiarui Lu, Zhao Xu, Weili Nie, Arvind Ramanathan, Chaowei Xiao, et al. A text-guided protein design framework. *Nature Machine Intelligence*, pages 1–12, 2025.

[18] Shengchao Liu, Yanjing Li, Zhuoxinran Li, Anthony Gitter, Yutao Zhu, Jiarui Lu, Zhao Xu, Weili Nie, Arvind Ramanathan, Chaowei Xiao, Jian Tang, Hongyu Guo, and Anima Anandkumar. A text-guided protein design framework. *Nature Machine Intelligence*, 7(4):580–591, March 2025.

[19] Ali Madani, Bryan McCann, Nikhil Naik, Nitish Shirish Keskar, Namrata Anand, Raphael R. Eguchi, Po-Ssu Huang, and Richard Socher. Progen: Language modeling for protein generation, 2020.

[20] Simon Malesys, Rachel Torchet, Bertrand Saunier, and Nicolas Maillet. Antibody sequence database. *NAR Genomics and Bioinformatics*, 6(4):lqae171, 12 2024.

[21] Alexander Rives, Joshua Meier, Tom Sercu, Siddharth Goyal, Zeming Lin, Jason Liu, Demi Guo, Myle Ott, C. Lawrence Zitnick, Jerry Ma, and Rob Fergus. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, 118(15):e2016239118, 2021.

[22] Rohit Singh, Samuel Sledzieski, Bryan Bryson, Lenore Cowen, and Bonnie Berger. Contrastive learning in protein language space predicts interactions between drugs and protein targets. *Proceedings of the National Academy of Sciences*, 120(24):e2220778120, 2023.

[23] Cheng Tan, Yijie Zhang, Zhangyang Gao, Yufei Huang, Haitao Lin, Lirong Wu, Fandi Wu, Mathieu Blanchette, and Stan Z Li. DyAb: Flow matching for flexible antibody design with AlphaFold-driven pre-binding antigen. *Proc. Conf. AAAI Artif. Intell.*, 39(1):782–790, April 2025.

[24] Kathryn E Tiller and Peter M Tessier. Advances in antibody design. *Annual review of biomedical engineering*, 17(1):191–216, 2015.

[25] Thanh V. T. Tran and Truong Son Hy. Protein design by directed evolution guided by large language models. *IEEE Transactions on Evolutionary Computation*, 29(2):418–428, 2025.

[26] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2023.

[27] Chaohao Yuan, Songyou Li, Geyan Ye, Yikun Zhang, Long-Kai Huang, Wenbing Huang, Wei Liu, Jianhua Yao, and Yu Rong. Functional protein design with local domain alignment. *CoRR*, abs/2404.16866, 2024.

A Related work

Text-Guided Protein Design. Recent advances in protein language models (PLMs) have laid the foundation for text-guided protein design by modeling amino acid sequences analogously to natural language. Transformer-based architectures such as ESM [21], ProtTrans [8], and various BERT-based models have been pretrained on large-scale protein databases, capturing complex syntactic and functional patterns in protein sequences. Building on these representations, recent approaches have demonstrated the feasibility of generating or editing proteins from natural language prompts. Pinal [6], a 16B-parameter model trained on 1.7B protein–text pairs, uses a two-stage pipeline to map text to structural motifs and then generate sequences, validated experimentally. ProteinDT [18] introduces a three-stage framework: Contrastive alignment (ProteinCLAP), a facilitator network and a decoder trained in 441K text-protein pairs for zero-shot generation and editing. Other models, such as PAAG [27] and ProDVa [16], target functional domain design, and fragment-based generation, respectively. Large-scale models such as ProGen [19] and ESM-3 [10] enable prompt-guided generation and prediction of fitness. These works demonstrate the growing potential of natural language as a controllable interface for protein design.

Challenges in Antibody Design. Despite advances in general protein design, direct application to antibody engineering faces unique challenges. Antibodies require precise three-dimensional arrangements in their Complementarity Determining Regions (CDRs) while maintaining structural integrity and minimal immunogenicity, constraints fundamentally different from general proteins. Critical limitations include data scarcity: Although general protein-text datasets contain millions of entries, detailed antibody annotations linking natural language descriptions to binding properties and therapeutic functions remain limited [2, 23]. Additionally, structural constraints in antibody design are exceptionally stringent, as function depends exquisitely on precise CDR conformations. Current text-based models may struggle to capture these antibody-specific constraints, potentially generating non-functional or immunogenic sequences [12]. These challenges highlight the need for antigen-conditioned antibody design frameworks that bridge general text-based capabilities with specialized antibody requirements. Our TeBaAb framework addresses this gap by combining text-based property specification with explicit antigen conditioning and iterative optimization tailored for antibody engineering.

B Detailed Methodology

B.1 Conditional Variational Autoencoder

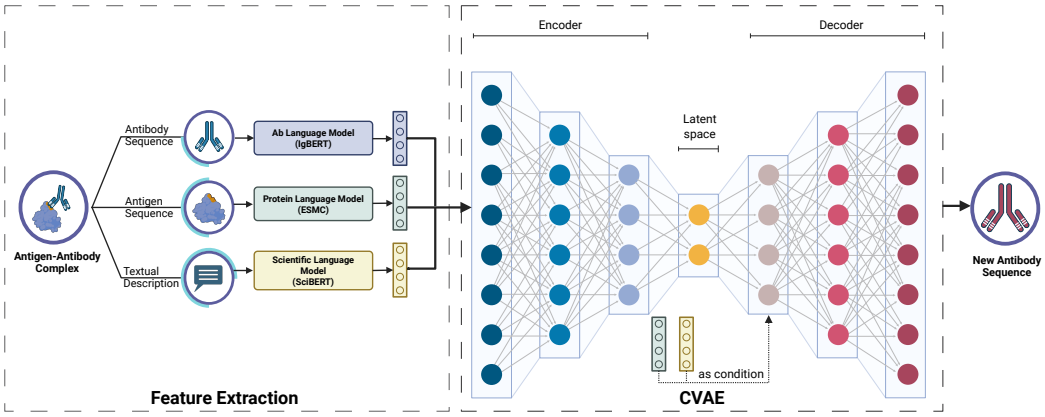

Figure 3: The CVAE integrates antibody embeddings (IgBERT), antigen embeddings (ESMC), and textual description embeddings (SciBERT) to generate antibody sequences. The encoder produces a latent distribution, and the transformer-based decoder reconstructs sequences conditioned on both latent variables and contextual inputs, enabling antigen-specific and text-guided antibody design.

Let the embedding of the antibody be denoted by $\mathbf{x} \in \mathbb{R}^{d_{\text{ab}}}$, representing the encoded representation of the antibody sequence obtained from IgBERT [14], a specialized antibody pre-trained language

model. The conditioning vector is defined as: $\mathbf{c} = \left[\mathbf{c}_{\text{antigen}};\, \mathbf{c}_{\text{description}}\right]$, where $\mathbf{c}_{\text{antigen}} \in \mathbb{R}^{d_{\text{ag}}}$ represents the antigen embedding extracted using ESMC [10], a protein language model specifically designed to understand protein sequences, and $\mathbf{c}_{\text{description}} \in \mathbb{R}^{d_{\text{des}}}$ denotes the textual description embedding generated by SciBERT [3], a domain-adapted language model for scientific text.

Encoder. The encoder, parameterized by ϕ, takes the concatenated vector $[\mathbf{x}; \mathbf{c}]$ and outputs parameters of the approximate posterior:

$$q_\phi(\mathbf{z} \mid \mathbf{x}, \mathbf{c}) = \mathcal{N}\big(\mathbf{z}; \mu_\phi(\mathbf{x}, \mathbf{c}),\, \mathrm{diag}(\sigma^2_\phi(\mathbf{x}, \mathbf{c}))\big),$$

where μ_ϕ and $\log \sigma^2_\phi$ are learned via neural networks. We then sample the latent representation using the reparameterization trick:

$$\mathbf{z} = \mu_\phi(\mathbf{x}, \mathbf{c}) + \sigma_\phi(\mathbf{x}, \mathbf{c}) \odot \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}).$$

Decoder. The decoder, implemented as a transformer network parameterized by θ, models the antibody sequence autoregressively conditioned on $\mathbf{z}$ and $\mathbf{c}$:

$$p_\theta(\mathbf{x}_{\text{seq}} \mid \mathbf{z}, \mathbf{c}) = \prod_{i=1}^{L} p_\theta(x_i \mid x_{<i},\, \mathbf{z}, \mathbf{c}),$$

where $\mathbf{x}_{\text{seq}} = (x_1, \ldots, x_L)$ denotes the amino acid sequence of length L, with each x_i represented as a one-hot vector.

Objective. The training objective minimizes a weighted sum of the reconstruction loss and the KL divergence loss. This is equivalent to maximizing the Evidence Lower Bound (ELBO):

$$\mathcal{L}_{\text{CVAE}}(\phi, \theta; \mathbf{x}, \mathbf{c}) = \underbrace{-\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x},\mathbf{c})}\left[\log p_\theta(\mathbf{x}_{\text{seq}} \mid \mathbf{z}, \mathbf{c})\right]}_{\text{Reconstruction Loss}} + \underbrace{\beta\, D_{\text{KL}}\big(q_\phi(\mathbf{z} \mid \mathbf{x}, \mathbf{c}) \parallel p(\mathbf{z})\big)}_{\text{KL Divergence Loss}},$$

where $p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$ is the prior distribution over the latent space, and β is a dynamically adjusted weight for the KL divergence term, controlled by a proportional-integral (PI) controller [9].

The two components of the loss are:

- **Reconstruction Loss** ($\mathcal{L}_{\text{recon}}$): Implemented as a token-level cross-entropy loss, it measures how well the model reconstructs the input sequence $\mathbf{x}_{\text{seq}}$ given the latent variable $\mathbf{z}$ and conditions $\mathbf{c}$.

  $$\mathcal{L}_{\text{recon}} = -\sum_{i=1}^{L} \log p_\theta(x_i \mid x_{<i}, \mathbf{z}, \mathbf{c}),$$

  where L is the sequence length, x_i is the i-th token, and $x_{<i}$ denotes the tokens preceding x_i. Padding tokens are masked (ignored) during loss computation.
- **KL Divergence Loss** (D_{KL}): This term regularizes the latent space by minimizing the Kullback-Leibler divergence between the approximate posterior $q_\phi(\mathbf{z} \mid \mathbf{x}, \mathbf{c})$ (learned by the encoder) and the standard normal prior $p(\mathbf{z})$. This ensures a meaningful and well-structured latent space, facilitating effective sampling for generation.

The PI controller adaptively modulates β during training, encouraging an optimal balance between latent compression (smaller KL divergence) and reconstruction fidelity (lower reconstruction error).

B.2 Binding Affinity Predictor

Inspired by recent advances in protein interaction prediction [22, 13], to predict antibody-antigen binding affinity, we employ a two-stage deep learning framework. First, a contrastive model aligns the embeddings of antibodies and antigens in a shared latent space. Then, a cross-attention model predicts the binding affinity by capturing their interactions. The contrastive model $\mathcal{M}_{\text{cont}}$, uses encoders ϕ_{ab} and ϕ_{ag} to project antibody ($\mathbf{ab}$) and antigen ($\mathbf{ag}$) embeddings into a common space. For a batch of N pairs $\{(\mathbf{ab}_i, \mathbf{ag}_i)\}_{i=1}^{N}$, projections are:

$$\mathbf{p}_{\text{ab},i} = \text{normalize}(\phi_{\text{ab}}(\mathbf{ab}_i)),$$
$$\mathbf{p}_{\text{ag},i} = \text{normalize}(\phi_{\text{ag}}(\mathbf{ag}_i)).$$

The contrastive loss is defined as:

$$\mathcal{L}_{\text{cont}} = -\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{\exp(\mathbf{p}_{\text{ab},i}\cdot\mathbf{p}_{\text{ag},i}/\tau)}{\sum_{j=1}^{N}\exp(\mathbf{p}_{\text{ab},i}\cdot\mathbf{p}_{\text{ag},j}/\tau)}\right).$$

The affinity predictor, $\mathcal{M}_{\text{aff}}$, uses these encoders and applies multi-head cross-attention as follows:

$$\begin{aligned}\mathbf{H}_{\text{ab}\to\text{ag}} &= \text{MultiHeadAttention}(\mathbf{proj}_{\text{ab}}, \mathbf{proj}_{\text{ag}}, \mathbf{proj}_{\text{ag}}),\\ \mathbf{H}_{\text{ag}\to\text{ab}} &= \text{MultiHeadAttention}(\mathbf{proj}_{\text{ag}}, \mathbf{proj}_{\text{ab}}, \mathbf{proj}_{\text{ab}}),\\ \mathbf{H} &= \mathbf{H}_{\text{ab}\to\text{ag}} + \mathbf{H}_{\text{ag}\to\text{ab}}, \quad \hat{y} = \theta(\mathbf{H}).\end{aligned}$$

Finally, the loss is the mean squared error defined as:

$$\mathcal{L}_{\text{aff}} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2.$$

B.3 Directed Evolution

Our antibody redesign strategy employs a directed evolution approach to iteratively improve antibody sequences towards higher binding affinity. This method mimics natural selection, progressively enriching sequences that exhibit superior fitness. The core idea is to explore the sequence space around an initial set of antibody sequences, evaluate the fitness of newly generated variants, and then select the most promising ones to serve as the basis for the next generation of exploration. This iterative refinement process aims to converge on sequences with optimal binding characteristics.

The directed evolution process commences by selecting K initial antibody sequences. These sequences are then subjected to an iterative optimization loop spanning G generations. In each generation, for every selected antibody sequence, a set of B neighboring sequences are generated. This generation step is powered by our CVAE. Given an antibody sequence embedding, along with the corresponding antigen embedding ($\mathbf{c}_{\text{antigen}}$) and a desired property embedding ($\mathbf{c}_{\text{description}}$), CVAE samples novel antibody sequences, yet contextually relevant. This allows for exploration of the sequence space while maintaining desired characteristics.

The fitness of these newly generated sequences is then rigorously evaluated. For this, we utilize our pre-trained Binding Affinity Predictor. This predictor acts as a computational oracle, providing an estimated binding affinity ($\hat{y}$) for each antibody-antigen pair. A lower predicted binding affinity (for example, a more negative value of ΔG) signifies a stronger binding interaction, thus indicating a higher fitness. The selection criterion for advancing sequences to the next generation is based on these predicted affinity values. Specifically, the top K antibody sequences, exhibiting the most favorable (lowest) predicted binding affinities, are chosen from the combined pool of parent and newly generated sequences to propagate to the subsequent generation. This ensures that the population progressively shifts towards higher-affinity regions of the sequence space.

The fitness score $F(S, A)$ for an antibody sequence S against a target antigen A is directly defined by its predicted binding affinity ($\hat{y}(S, A)$) calculated using our binding affinity predictor $F(S, A) = \hat{y}(S, A)$, where S represents the antibody sequence being evaluated and A represents the target antigen sequence. The objective of the directed evolution process is to minimize $F(S, A)$, thus identifying antibody sequences with the strongest predicted binding affinity to the specific target antigen.

B.4 Decoder model

The decoder is built on a Transformer architecture [26] and serves as the core module for generating antibody sequences (see Fig. 4). It reconstructs sequences from a rich, combined representation that merges latent features, descriptive metadata, and antigen-specific embeddings. These inputs are concatenated and projected into a sequence of hidden states, forming the "memory" for cross-attention.

During generation, the target tokens are first embedded and enriched with positional encodings, then passed through a stack of Transformer decoder layers. Within each layer, self-attention captures dependencies within the generated sequence so far, while cross-attention draws on the memory to integrate contextual information from the input.

Training employs teacher forcing, periodically replacing predicted tokens with their ground-truth counterparts at a fixed probability to stabilize learning. When teacher forcing is not applied, the decoder generates tokens sequentially using greedy decoding. Finally, hidden states are projected onto vocabulary logits, enabling step-by-step prediction of the next amino acid in the sequence.

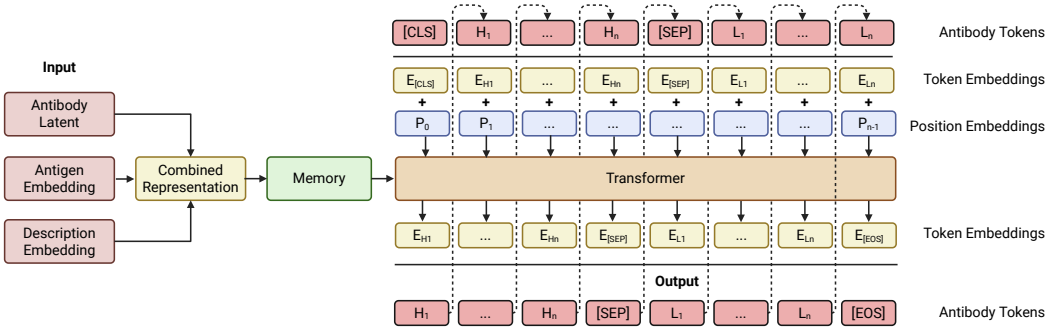

Figure 4: This figure illustrates the design of Transformer-based decoder for antibody sequence generation. The decoder integrates latent features, antigen embeddings, and textual description embeddings into a shared representation that serves as memory for cross-attention. Antibody sequences are reconstructed autoregressively: input tokens (heavy and light chains) are embedded with positional encodings, processed through Transformer layers, and predicted step by step. Self-attention captures dependencies within the sequence, while cross-attention injects contextual information, enabling controllable generation of antibody sequences aligned with antigen specificity and descriptive properties.

Table 4: Comparative Performance: TeBaAb (with Description) vs. TeBaAb (without Description).

Metric	**TeBaAb (w Description)**	**TeBaAb (w/o Description)**
Binding Affinity Improvement (Avg ΔG %)	15.5	8.9
Diversity	87.73	86.70
Novelty	28.95	26.11
Average Optimized Pred. Error (Å)	0.48	0.36

C Implementation Details

C.1 Evaluation metrics

To comprehensively assess the performance of our antibody redesign framework, we employ four primary evaluation metrics: maximum binding affinity (MBA), sequence diversity, novelty, and structural confidence.

Maximum Binding Affinity (MBA). The MBA metric quantifies the highest predicted binding affinity between the redesigned antibody and its specific antigen, reflecting the functional efficacy of the generated sequences. We use the free-energy-change scoring function to estimate MBA. Formally, MBA is defined as:

$$\text{MBA} = \max_{s \in S} \{-\Delta G(s, a)\}, \tag{1}$$

where S represents the set of redesigned antibody sequences, s denotes a single antibody sequence, and a denotes the target antigen. A lower predicted free energy (ΔG) indicates a higher binding affinity.

Diversity. To ensure the generation of varied antibody sequences rather than mere repetitions or minor modifications of known sequences, we measure diversity within the set of redesigned antibodies. Sequence diversity is computed using the pairwise Levenshtein distance [15] averaged over

all pairs of generated sequences, formally represented as:

$$\text{Diversity} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \text{Lev}(s_i, s_j), \tag{2}$$

where N is the number of antibody sequences and $\text{Lev}(s_i, s_j)$ denotes the Levenshtein distance between two antibody sequences s_i and s_j.

Novelty. Novelty assesses the degree to which the antibodies generated differ from known sequences in our training dataset. We calculate novelty as the average minimum sequence distance between each generated antibody sequence and all sequences in the training dataset ($\mathcal{D}_{train}$). The novelty score is defined as:

$$\text{Novelty} = \frac{1}{|S|} \sum_{s \in S} \min_{s' \in \mathcal{D}_{train}} \text{Lev}(s, s'). \tag{3}$$

Higher novelty scores indicate more distinctive and potentially innovative antibody sequences.

Structural Confidence. Assessing the structural quality of the redesigned antibodies is critical to ensure stability and functionality. We used ABodyBuilder2 [1] to predict the three-dimensional structures of redesigned antibody sequences, leveraging its ensemble of four models to estimate the reliability of the prediction. The root mean squared predicted error (RMSPE) (Å), derived from the diversity among these predictions, serves as a structure-related metric:

$$\text{RMSPE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \text{Var}(\mathbf{x}_i)}, \tag{4}$$

where N is the number of residues, and $\text{Var}(\mathbf{x}_i)$ represents the variance in the predicted coordinates of residue i in the ensemble. Lower RMSPE values indicate greater confidence in the predicted structure, ensuring structural reliability for redesigned antibodies without requiring a reference scaffold.

C.2 AbDes Dataset

In this appendix, we provide example entries from the AbDes dataset to illustrate its structure and the type of information it contains. The dataset is designed to facilitate text-conditioned antibody design by linking antibody sequences and their target antigens with rich textual descriptions of antibody properties.

The AbDes dataset comprises entries with the following key fields:

- **Antibody Sequence**: The complete heavy- and light-chain amino acid sequences of the antibody (e.g. '$heavy\_chain|light\_chain$').
- **Antigen Sequence**: The target antigen to which the antibody binds. Following [7], we use '$/$' to separate different antigen fragments.
- **Description**: A free-text description of the antibody's properties. This field represents our unique contribution, derived from PDB annotations and literature.
- ΔG **(Binding Affinity)**: The predicted binding free energy, where available. This value indicates the strength of the antibody-antigen interaction (lower values signify stronger binding). Note that ΔG values are available for a subset of the dataset.

D Additional Results: Impact of Textual Description

This section presents an additional evaluation designed to specifically highlight the contribution of the textual description input to the performance of the TeBaAb framework. Although our primary evaluation focuses on the full TeBaAb model (which incorporates textual descriptions), understanding the isolated impact of this novel conditioning is crucial. To achieve this, we compare the performance of our best performing TeBaAb configuration (with input of textual description) against

Table 5: Example Entries from the AbDes Dataset (General Structure).

Antibody Sequence	**Antigen**	**Description**	**Binding Affinity**
$QVQLV...\|QSALT...$	$VVKFMDVY...$	Vascular endothelial growth factor in complex with a neutralizing antibody, classified as an immune system, derived from mus musculus and expressed in escherichia coli, forms a Hetero 6-mer with Cyclic - C2 symmetry.	-11.55
$QVQLQ...\|QVQLQ...$	$KVFGRCEL...$	Hen egg white lysozyme, d18a mutant, in complex with mouse monoclonal antibody d1.3, classified as a complex (immunoglobulin/hydrolase), derived from mus musculus and expressed in escherichia coli, forms a Hetero 3-mer with Asymmetric - C1 symmetry, and has pseudo-symmetry of Asymmetric - C1 with Hetero 3-mer stoichiometry.	-10.45
$QIQLVQ...\|DIVMT...$	$IRDFNNLT...$	Refined crystal structure of the influenza virus n9 neuraminidase-nc41 fab complex, classified as a hydrolase(o-glycosyl), derived from influenza a virus (a/tern/australia/g70c/1975(h11n9)), forms a Hetero 12-mer with Cyclic - C4 symmetry.	-11.02

Note: Sequences are truncated for brevity.

a variant of TeBaAb where the textual description component ('des_rep') is excluded during both training and inference. This allows us to isolate the specific benefits conferred by conditioning the generation on textual properties.

All other hyperparameters and training procedures remained consistent with the TeBaAb model that performs the best described in Section 3. This ensures a direct comparison of the impact of the description input. Table 4 presents a comparative overview of key evaluation metrics: Binding Affinity, Sequence Diversity, Novelty, and Structural Confidence, between the full TeBaAb framework and its variant without textual description input.