A deep learning approach for rational affinity maturation of anti-VEGF nanobodies

Gaëlle Verdon[1]   Laurent David[1]   Alexandre de Brevern[2,3]

Yasser Mohseni Behbahani[1,2,3*]

[1]Inria, [2]Université Paris Cité, [3]Inserm, Paris, France
*Corresponding author: yasser.mohseni-behbahani@u-paris.fr

Abstract

Nanobodies offer several advantages over conventional antibodies due to their lower immunogenicity, enhanced stability, and superior tissue penetration, making them promising candidates for cancer therapy. In this study, we employ deep learning algorithms to design anti-VEGF nanobodies via affinity maturation. Our approach integrates structure-guided mutational modeling and systematic measurement of binding affinity and stability for rational optimization of Complementarity Determining Regions. In addition, we developed a sequence-based melting temperature predictor for nanobodies, ensuring stability of the designed mutants. Our method achieves energy reductions up to -4.92 kcal/mol. Our melting temperature predictor demonstrated a Pearson correlation coefficient of 0.772. These findings emphasize the potential of computational approaches for nanobody affinity maturation and stability prediction, paving the way for more effective therapeutic designs.

1   Introduction

Nanobody or VHH (Fig. **1**) is the single variable domain of heavy-chain-only antibodies (HCAbs) which retains full binding capacity despite its compact size (15kDa, 10 times smaller than antibodies) [2]. Due to their small size, VHHs have superior stability, enhanced tissue permeability and rapid blood clearance compared to conventional antibodies [29, 17], making them particularly effective for targeting dense tumors and crossing the blood-brain barrier [29]. They rely solely on three complementarity-determining regions (CDRs), including an extended CDR3 loop (16–24 amino acids vs. $\sim$ 10 in antibodies) [9], to recognize antigens with high specificity and affinity. This elongated CDR3 allows VHHs to access sterically restricted or cryptic epitopes that are often inaccessible to bulkier antibodies, such as enzyme active sites or viral canyons [11, 29, 9, 26]. These favorable biophysical and pharmacokinetic properties make VHHs ideal candidates for anti-angiogenic therapeutic solutions [5, 13].

Vascular endothelial growth factor (VEGF) regulates angiogenesis and is often overexpressed in cancer, promoting tumor growth [25]. Blocking the interaction between VEGF and its receptor (VEGFR) is a key therapeutic approach in oncology [5]. A promising strategy involves designing protein-based binders, such as VHHs, to disrupt this interaction. Recent advancements in computational biology have transformed protein engineering, with deep learning approaches emerging as a powerful alternative to traditional experimental methods. These approaches accelerate affinity maturation between

Figure 1: VHH secondary structure.

the binder and the target and improve efficiency by exploring the mutational landscape to identify beneficial variations [14, 10]. They contribute to multiple stages of affinity maturation, including binder design, structural modeling, and assessment of mutation-induced changes in binding affinity ($\Delta\Delta G_{Bind} = \Delta G_{Bind}^{Mutant} - \Delta G_{Bind}^{Wild-type}$) between the wild type and the mutant complexes.
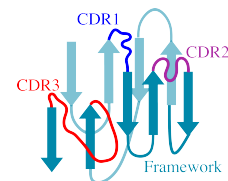
A class of deep learning-based methods is protein language models (pLM), which extract biophysical and evolutionary information that determine protein structure and function [7, 21, 15]. Sequence-based deep learning approaches like MuLAN rely on pLM and offer a fast and accurate prediction of $\Delta\Delta G_{Bind}$ in protein-protein interactions [16]. However, CDR sequences are highly diverse and are not concerned by the notion of conservation during evolution. This limits the ability of pLM to extract sufficient predictive information from interfaces where these regions are involved. Another class operates on protein design, addressing the inverse folding problem: identifying amino acid sequences that fold into a given 3D backbone structure. ProteinMPNN, a structure-based graph neural network, is an example of these methods applied to the problem of *de novo* binder design [3]. AbMPNN is a fine-tuned version of ProteinMPNN trained specifically for antibodies; it focuses on determining and optimizing the CDR3 [4]. For protein structure prediction, AlphaFold-Multimer [8] and AlphaFold3 [1] accurately reconstruct protein complexes and provide insights into binding mechanisms. ColabFold is a faster implementation of AlphaFold-Multimer, which optimizes inference and accessibility while maintaining similar performance [18]. Structure-based tools like DLA-Mutation [19] and DDgPredictor [24] accurately estimate $\Delta\Delta G_{Bind}$, with the latter being especially effective for antibody-antigen interactions.

Thermostability is essential for VHH functionality and design feasibility. It is shown that the VHH's CDR loops play a crucial role in its stability [6]. TemBERTure, a deep learning model, predicts protein melting temperature (T_m) by leveraging pLM protBERT-BFD [7], fine-tuned through an adapter-based approach [22]. It shows great promise in the T_m prediction of a wide range of proteins.

This study aims to develop a deep learning-based affinity maturation framework for the rational design of anti-VEGF nanobodies with improved binding affinity and thermal stability. ColabFold structural predictions enable us to evaluate the effect of sequence variants on the VHH-VEGF complex, and DDgPredictor provides precise estimates of $\Delta\Delta G_{Bind}$. To ensure the biophysical robustness of VHH variants, we trained a VHH-specific model derived from TemBERTure to predict the thermostability of new VHH designs. Our integrative approach enables us to explore the mutational landscape in the CDR3 loop and beyond at high throughput, balancing enhanced binding with stability preservation.

2 Materials and Methods

2.1 Affinity maturation pipelines

After selecting antibodies and VHH that bind VEGF, their CDRs are assigned using human expertise. This step is followed by introducing mutations on the CDR3 using ProteinMPNN and exhaustive mutational scanning. ProteinMPNN generates CDR3 sequences conditioned to the VHH-VEGF complex structure. The Structure of wild-type and mutant are predicted using ColabFold, which produces multiple models per sequence; due to computational limits, only the best structure is kept when the sequence does not change. The $\Delta\Delta G_{\text{Bind}}$ is then predicted using DDgPredictor to identify favorable mutants ($\Delta\Delta G_{Bind} < 0$). We developed three pipelines for affinity maturation.

Wild-type based maturation. Starting from the VHH-VEGF complex structure with identified CDR regions, we generated 10,000 variants by introducing multiple point mutations within the CDR3 using ProteinMPNN. A redundancy reduction step of the resulting mutant sequences is performed, keeping them unique. We also conducted exhaustive single-point mutations at every position of the CDR3 without considering ProteinMPNN (**Fig. S1**).

Structure-adapted maturation. We adapt the wild-type based maturation pipeline to account for mutation-induced conformational variability. At each position of CDR3, we generate a single-point mutation following the ProteinMPNN probability output. We predict the structure and $\Delta\Delta G_{Bind}$ following each positional mutation (**Fig. 2**). This allows us to adjust the structure along the design. However, since mutations are performed sequentially, one after the other, scanning the CDR3, the generation of mutations on the final amino acid is influenced by the preceding ones. As a result, not all combinations of mutations may be explored. To address this, we add an option to continue the process iteratively. Each iteration begins with the structure and sequence produced at the final step of the previous iteration (*i.e.* the last position of the CDR3 sequence). This iterative process explores more thoroughly the mutational landscape and eventually discovers more favorable sequences. We add another option to guide the mutations towards higher complex stability (guided maturation process). We retain mutations that lower $\Delta\Delta G_{Bind}$ before moving to the next position (**Fig. 2**).

2.2 Thermostability prediction

We finetuned the TemBERTure best model (model replica 2) for VHH's T_m prediction using the NbThermo dataset. The learning rate was optimized while the other parameters were kept at their default values (dropout in the adapter head is set to 0.2, no weight decay, and warmup ratio of 0). We also finetuned the prot-BERT model following the adapter architecture proposed by the TemBER-Ture. The learning rate, the number of layers, the dropout, and the activation function of the adapter head were optimized, while a warmup ratio of 0 and no weight decay were kept as default.

2.3 Data

Affinity maturation: The two nanobodies and the antibody used in this study come from SabDab-nano and SabDab, respectively, two publicly available databases containing experimentally determined antibody and nanobody structures [23]. The VHH-VEGF complexes 3P9W and 5FV2 were extracted using "VHH"

Figure 2: Structure-adapted maturation pipeline. A mutation is applied to VHH at position i of CDR3 following the amino acid substitution generated by ProteinMPNN. After each mutation, the pipeline determines the complex structure and predicts $\Delta\Delta G_{Bind}$. The subsequent amino acid substitution at position i+1 is determined by the mutant structure generated for position i. One can repeat the process several iterations. In each iteration, process k+1 starts with the structure and sequence of the mutant complex designed for position N from process k. In the guided maturation pipeline, the structure and the sequence of the mutant are kept only if the process finds a more favorable sequence ($\Delta = \Delta\Delta G_{Bind}^i - \Delta\Delta G_{Bind}^{Best} < 0$).

and "vascular endothelial growth factor" as *Antibody type* and *keyword query*, respectively. The Antibody-VEGF complex 6BFT was extracted using "Fv" and "vascular endothelial growth factor A" as *Antibody type* and *keyword query*, respectively. The 5FV2 complex consists of a VHH and a dimeric VEGF, while the 3P9W complex contains a VHH in interaction with a monomeric VEGF.

Thermostability prediction: For training our model, we used the NbThermo dataset, containing T_m, amino acid sequences, and annotated CDRs of 564 VHHs [27]. We kept 513 non-redundant VHH sequences with available T_m. To prevent data leakage, we ensured that there was no redundancy of CDRs across our training, validation, and test sets. This resulted in the test, validation, and training sets of 52, 51, and 410 samples, respectively.

3 Experiments and Results

We applied the wild-type based approach on complexes 3P9W, 5FV2, and 6BFT and generated 10'000 CDR3 sequences (multiple point mutations) for each, among which 197, 185, and 5811 were unique for 3P9W, 5FV2, and 6BFT, respectively. We further filtered the mutations on 6BFT to include sequences generated more than 6 times by ProteinMPNN to reduce computational costs. This filtering further reduced the number of mutants to 229. The greater diversity of mutations observed for the 6BFT antibody complex may result from an increased number of possible mutation combinations, facilitated by the presence of both heavy and light chain CDR3, rather than simply one CDR3. Within the unique mutations generated, 68, 57, and 142 mutations were favorable for 3P9W, 5FV2, and 6BFT complexes, respectively.

We ran the structure-adapted maturation process on the CDR3 of 5FV2 (positions 97 to 109 of VHH) ten independent times (**Fig. 3A** and **S6**). For ease of visualization, we selected 6 processes out of 10, from which 5 converged to favorable mutants (**Fig. 3A**). For all generated mutant sequences, the corresponding melting temperature is always lower than the wildtype (71.5 °C).

Similarly, we ran the iterative process on the CDR3 of complex 5FV2 ten independent times. Each process has 7 iterations. Almost all resulting sequences are favorable (**Fig. 3B** and **S7**). We also
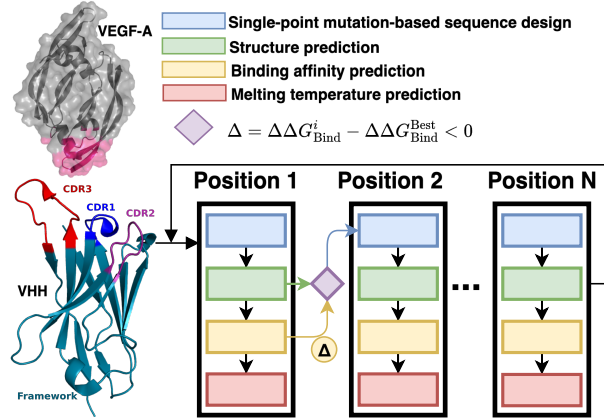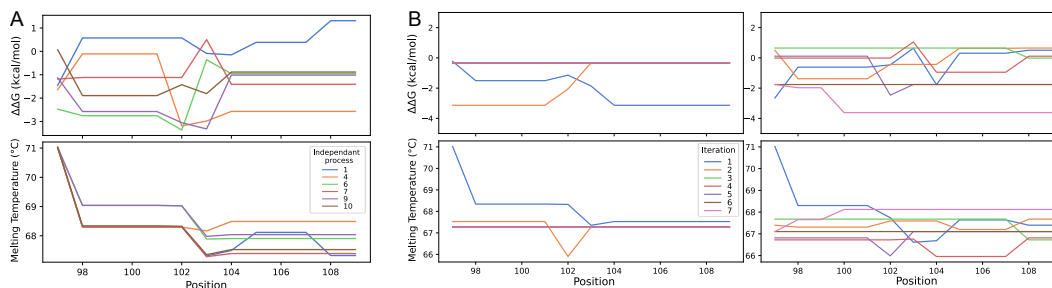
Figure 3: Evolution of the predicted $\Delta\Delta G_{Bind}$ (top row) and T_m (bottom row) through mutant positions within the CDR3 of the complex 5FV2. **(A)** Each curve corresponds to a selected independent design process (see **Fig. S6** for all results). **(B)** Evolution over 7 iterations for two independent iterative processes (left and right). On the left, while iteration 1 (blue) reduces $\Delta\Delta G_{Bind}$, iteration 2 (orange) increases it back to close to 0 (almost no improvement), and subsequently, the sequence and structure remain the same for all following iterations. On the right, $\Delta\Delta G_{Bind}$ fluctuates around -1 $kcal/mol$, throughout iterations, and at the end of the last iteration, we reach the most negative $\Delta\Delta G_{Bind}$ value, around -4 $kcal/mol$. The X-axis corresponds to the CDR3 positions.

observed that the final mutant obtained may not necessarily be the most favorable during the iterative process. This could be due to divergences between mutations judged as favorable by ProteinMPNN and the ones predicted as favorable by DDgPredictor. We observed lower T_m for mutants compared to the wild-type. All along the iterative process, T_m fluctuates within 3°C.

We explored the space of single-point mutations for the two VHH-VEGF complexes. We observed that certain wild-type positions support multiple favorable amino acid substitutions (**Fig. 4A**). To assess whether ProteinMPNN performance aligns with the predicted mutational landscape associated with exhaustive single-point mutations, we ran ProteinMPNN on each position 10 independent times and averaged the output probability vectors. There are differences in the choice of amino acid substitutions between ProteinMPNN and the $\Delta\Delta G_{Bind}$ predictions from exhaustive single-mutational scanning (**Fig. 4A**). The charge distributions on the interface of wild-type and mutant VHH-VEGF structures align with the DDgPredictor predictions. Building on this observation, we decided to run the guided maturation process on the CDR3 of 5FV2 ten independent times, by evaluating the predicted $\Delta\Delta G_{Bind}$ after each mutation to converge to the highest complex stability. We observed a progressive decrease of $\Delta\Delta G_{Bind}$, reaching -4.92 $kcal/mol$ at the end of affinity maturation for complex 5FV2 (**Fig. 4B**). The most favorable mutants, reaching $\Delta\Delta G_{Bind}$ values smaller than -4 kcal/mol (processes 2, 3, 5, 9, 10), also reach higher T_m. The least favorable mutant (process 4) reaches the lowest T_m while the most favorable one (process 3) reaches the highest one.

Since ProteinMPNN and DDgPredictor perform differently in identifying favorable mutations, we sought to determine whether AbMPNN can align with DDgPredictor and generate diverse mutations. We calculated the mean of probability vectors of amino acid substitutions (computed at Softmax temperature of T=0.1) over ten runs per position and compared the results with the exhaustive mutational analysis obtained from DDgPredictor (**Fig. S8**). For 5FV2, we observed that AbMPNN predictions often stick to the wild-type amino acid and other substitutions are mostly favorable according to DDgPredictor. For 3P9W, AbMPNN frequently predicts wild-type amino acids while its alternative amino acid substitutions are often unfavorable (according to DDgPredictor) and tend to be limited to a small set of choices. For instance, Tyrosine is predicted at 6 of the 13 positions. At the given Softmax temperature, AbMPNN tends to have less variability, resulting in a smaller exploration of the mutational landscape, and may miss potentially beneficial mutations. We used predictions from DDgPredictor to guide the mutations introduced by AbMPNN, either preserving the mutation or reverting to the wild-type amino acid. This combination enabled us to achieve energy reductions as low as -4.5 $kcal/mol$ (**Fig. S9**).

All mutants exhibit a T_m within a range of 3.7°C and 5.3°C, while the wild-type T_m are 72.49°C and 71.54°C for VHH in complexes 3P9W and 5FV2, respectively (**Fig. 5AB**). We observed that specific CDR3 positions favor higher T_m (3rd for 5FV2 and 4th for 3P9W) while others are very sensitive to mutations and tend to reduce it (2nd for 5FV2 and 5th and 11th for 3P9W).

Our fine-tuned model for predicting melting temperature outperforms the baseline standard model, reaching lower Root Mean Square Error (RMSE) (6.0°C vs 18.89°C), higher R-square (0.59 vs -0.01), and improved Pearson correlation coefficient (0.772 vs 0.005) (**Fig. 5C**). The predicted vs
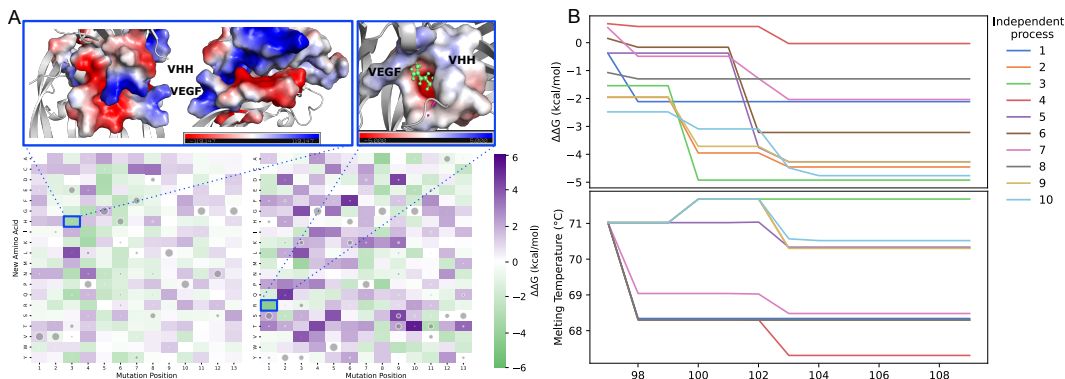
4

Figure 4: **(A)** Predicted $\Delta\Delta G_{Bind}$ for exhaustive single point mutation on the CDR3 of 5FV2 (left) and 3P9W (right). Gray circles represent the mean probability of substitution, generated at Softmax temperature 0.25 for 5FV2 and 0.4 for 3P9W, obtained over 10 runs by ProteinMPNN for each position. In the structure of complexes, the mutant amino acid is shown in green, and the electrostatic charge distribution goes from negative charge (red) to positive charge (blue). **(B)** Evolution of the predicted $\Delta\Delta G_{Bind}$ (top) and T_m (bottom) across CDR3 mutations of VHH in 5FV2 during guided maturation over 10 independent processes. All processes reach lower $\Delta\Delta G_{Bind}$ values, and process 3 gives the lowest value. All T_m predictions are lower compared to wild-type at the end of each process, except for process 3. The X-axis corresponds to the CDR3 positions.

experimental curve shows a closer alignment with the ground truth, highlighting the fine-tuned model's ability to capture the effects of mutations on T_m more effectively than the standard model. See **Fig. S2** for model improvement across different epochs, and **Fig. S3-S5** for the ablation study to find the best architecture and hyperparameters.
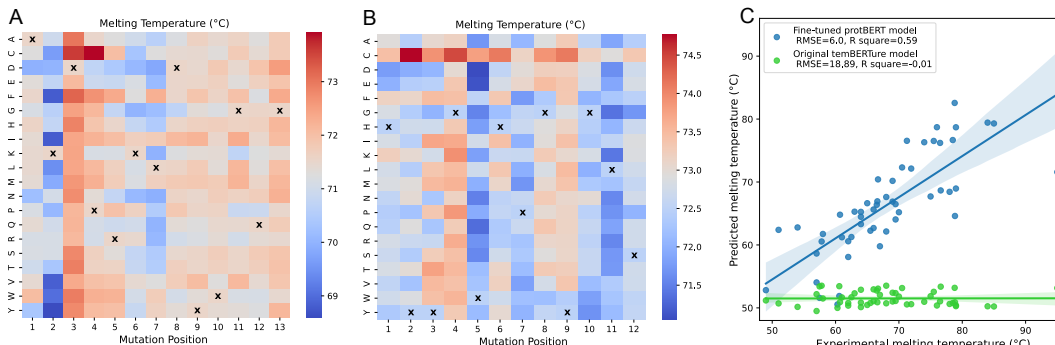

Figure 5: **(A-B)** T_m prediction for exhaustive point mutation performed on CDR3 of VHH in 5FV2 **(A)** and 3P9W **(B)**. Crosses indicate the wild-type amino acid. Higher temperatures mean higher thermostability. **(C)** Comparison between the melting temperature (T_m) determined experimentally and the one predicted by the temBERTure and fine-tuned models. RMSE=Root Mean Square Error.

4 Discussion

In this work, we developed a deep-learning framework for VHH CDR3 design via affinity maturation, modeling how mutations affect backbone conformation, VEGF binding affinity, and VHH thermostability. To enhance the reliability of our $\Delta\Delta G_{Bind}$ predictions, future improvements will incorporate molecular dynamics simulations to better capture protein complex conformational variability. Moreover, the accuracy of mutant protein structure models remains an active area of research. Studies suggest AlphaFold2 [12] is not reliable for predicting mutation-induced changes in protein stability or function, as it can model folded structures for mutations that experimentally yield unfolded proteins [28, 20]. While our VHH T_m model outperforms the state-of-the-art, its sensitivity must be improved to fully capture the effects of single-point mutations on VHH stability. Finally, the model was trained exclusively on folded VHHs due to the lack of experimentally measured T_m for non-folding variants. Availability: The source code and model weights are available from the corresponding author upon request.

5

Acknowledgements

References

[1] Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J. Ballard, Joshua Bambrick, Sebastian W. Bodenstein, David A. Evans, Chia-Chun Hung, Michael O'Neill, David Reiman, Kathryn Tunyasuvunakool, Zachary Wu, Akvilė Žemgulytė, Eirini Arvaniti, Charles Beattie, Ottavia Bertolli, Alex Bridgland, Alexey Cherepanov, Miles Congreve, Alexander I. Cowen-Rivers, Andrew Cowie, Michael Figurnov, Fabian B. Fuchs, Hannah Gladman, Rishub Jain, Yousuf A. Khan, Caroline M. R. Low, Kuba Perlin, Anna Potapenko, Pascal Savy, Sukhdeep Singh, Adrian Stecula, Ashok Thillaisundaram, Catherine Tong, Sergei Yakneen, Ellen D. Zhong, Michal Zielinski, Augustin Žídek, Victor Bapst, Pushmeet Kohli, Max Jaderberg, Demis Hassabis, and John M. Jumper. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, pp. 1–3, May 2024. Publisher: Nature Publishing Group.

[2] Alain Beck, Liliane Goetsch, Charles Dumontet, and Nathalie Corvaïa. Strategies and challenges for the next generation of antibody–drug conjugates. *Nature Reviews Drug Discovery*, 16(5):315–337, May 2017. Publisher: Nature Publishing Group.

[3] J. Dauparas, I. Anishchenko, N. Bennett, H. Bai, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, A. Courbet, R. J. de Haas, N. Bethel, P. J. Y. Leung, T. F. Huddy, S. Pellock, D. Tischer, F. Chan, B. Koepnick, H. Nguyen, A. Kang, B. Sankaran, A. K. Bera, N. P. King, and D. Baker. Robust deep learning–based protein sequence design using ProteinMPNN. *Science*, 378(6615):49–56, October 2022. Publisher: American Association for the Advancement of Science.

[4] Frédéric A. Dreyer, Daniel Cutting, Constantin Schneider, Henry Kenlay, and Charlotte M. Deane. Inverse folding for antibody sequence design using deep learning, October 2023. arXiv:2310.19513 [q-bio].

[5] Walead Ebrahimizadeh, Seyed Latif Mousavi Mousavi Gargari, Zahra Javidan, and Masoumeh Rajabibazl. Production of Novel VHH Nanobody Inhibiting Angiogenesis by Targeting Binding Site of VEGF. *Applied Biochemistry and Biotechnology*, 176(7):1985–1995, August 2015.

[6] Nehad S. El Salamouni, Jordan H. Cater, Lisanne M. Spenkelink, and Haibo Yu. Nanobody engineering: computational modelling and design for biomedical and therapeutic applications. *FEBS Open Bio*, 15(2):236–253, 2025. _eprint: https://febs.onlinelibrary.wiley.com/doi/pdf/10.1002/2211-5463.13850.

[7] 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: Towards Cracking the Language of Lifes Code Through Self-Supervised Deep Learning and High Performance Computing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[8] Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Žídek, Russ Bates, Sam Blackwell, Jason Yim, Olaf Ronneberger, Sebastian Bodenstein, Michal Zielinski, Alex Bridgland, Anna Potapenko, Andrew Cowie, Kathryn Tunyasuvunakool, Rishub Jain, Ellen Clancy, Pushmeet Kohli, John Jumper, and Demis Hassabis. Protein complex prediction with AlphaFold-Multimer. Technical report, October 2021. doi: 10.1101/2021.10.04.463034.

[9] Bo-kyung Jin, Steven Odongo, Magdalena Radwanska, and Stefan Magez. NANOBODIES®: A Review of Diagnostic and Therapeutic Applications. *International Journal of Molecular Sciences*, 24(6):5994, January 2023. Publisher: Multidisciplinary Digital Publishing Institute.

[10] Sara Joubbi, Alessio Micheli, Paolo Milazzo, Giuseppe Maccari, Giorgio Ciano, Dario Cardamone, and Duccio Medini. Antibody design using deep learning: from sequence and structure design to affinity maturation. *Briefings in Bioinformatics*, 25(4):bbae307, July 2024.

[11] Ivana Jovčevska and Serge Muyldermans. The Therapeutic Potential of Nanobodies. *BioDrugs*, 34(1):11–26, February 2020.

[12] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, August 2021.

[13] Elmira Karami, Shamsi Naderi, Reyhaneh Roshan, Mahdi Behdani, and Fatemeh Kazemi-Lomedasht. Targeted therapy of angiogenesis using anti-VEGFR2 and anti-NRP-1 nanobodies. *Cancer Chemotherapy and Pharmacology*, 89(2):165–172, February 2022.

[14] Jiaqi Li, Guangbo Kang, Jiewen Wang, Haibin Yuan, Yili Wu, Shuxian Meng, Ping Wang, Miao Zhang, Yuli Wang, Yuanhang Feng, He Huang, and Ario de Marco. Affinity maturation of antibody fragments: A review encompassing the development from random approaches to computational rational optimization. *International Journal of Biological Macromolecules*, 247:125733, August 2023.

[15] Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, Allan dos Santos Costa, Maryam Fazel-Zarandi, Tom Sercu, Salvatore Candido, and Alexander Rives. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, March 2023. Publisher: American Association for the Advancement of Science.

[16] Gianluca Lombardi and Alessandra Carbone. MuLAN: Mutation-driven Light Attention Networks for investigating protein-protein interactions from sequences, August 2024. Pages: 2024.08.24.609515 Section: New Results.

[17] Vitória Meneghetti Minatel, Carlos Roberto Prudencio, Benedito Barraviera, and Rui Seabra Ferreira. Nanobodies: a promising approach to treatment of viral diseases. *Frontiers in Immunology*, 14, January 2024. Publisher: Frontiers.

[18] Milot Mirdita, Konstantin Schütze, Yoshitaka Moriwaki, Lim Heo, Sergey Ovchinnikov, and Martin Steinegger. ColabFold: making protein folding accessible to all. *Nature Methods*, 19(6):679–682, June 2022. Publisher: Nature Publishing Group.

[19] Yasser Mohseni Behbahani, Elodie Laine, and Alessandra Carbone. Deep Local Analysis deconstructs protein–protein interfaces and accurately estimates binding affinity changes upon mutation. *Bioinformatics*, 39(Supplement_1):i544–i552, June 2023.

[20] Marina A. Pak, Karina A. Markhieva, Mariia S. Novikova, Dmitry S. Petrov, Ilya S. Vorobyev, Ekaterina S. Maksimova, Fyodor A. Kondrashov, and Dmitry N. Ivankov. Using AlphaFold to predict the impact of single mutations on protein stability and function. *PLOS ONE*, 18(3):e0282689, March 2023. Publisher: Public Library of Science.

[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. *PNAS*, 118(15), April 2021.

[22] Chiara Rodella, Symela Lazaridi, and Thomas Lemmin. TemBERTure: advancing protein thermostability prediction with deep learning and attention mechanisms. *Bioinformatics Advances*, 4(1):vbae103, January 2024.

[23] Constantin Schneider, Matthew I J Raybould, and Charlotte M Deane. SAbDab in the age of biotherapeutics: updates including SAbDab-nano, the nanobody structure tracker. *Nucleic Acids Research*, 50(D1):D1368–D1372, January 2022.

[24] Sisi Shan, Shitong Luo, Ziqing Yang, Junxian Hong, Yufeng Su, Fan Ding, Lili Fu, Chenyu Li, Peng Chen, Jianzhu Ma, Xuanling Shi, Qi Zhang, Bonnie Berger, Linqi Zhang, and Jian Peng. Deep learning guided optimization of human antibody against SARS-CoV-2 variants with broad neutralization. *Proceedings of the National Academy of Sciences*, 119(11):e2122954119, March 2022. Publisher: Proceedings of the National Academy of Sciences.

[25] Pallab Shaw, Shailendra Kumar Dhar Dwivedi, Resham Bhattacharya, Priyabrata Mukherjee, and Geeta Rao. VEGF signaling: Role in angiogenesis and beyond. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1879(2):189079, March 2024.

[26] Shuyang Sun, Ziqiang Ding, Xiaomei Yang, Xinyue Zhao, Minlong Zhao, Li Gao, Qu Chen, Shenxia Xie, Aiqun Liu, Shihua Yin, Zhiping Xu, and Xiaoling Lu. Nanobody: A Small Antibody with Big Implications for Tumor Therapeutic Strategy. *International Journal of Nanomedicine*, 16:2337–2356, March 2021. Publisher: Dove Medical Press _eprint: https://www.tandfonline.com/doi/pdf/10.2147/IJN.S297631.

[27] Mario S Valdés-Tresanco, Mario E Valdés-Tresanco, Esteban Molina-Abad, and Ernesto Moreno. NbThermo: a new thermostability database for nanobodies. *Database*, 2023:baad021, January 2023.

[28] Lei Wang, Zehua Wen, Shi-Wei Liu, Lihong Zhang, Cierra Finley, Ho-Jin Lee, and Hua-Jun Shawn Fan. Overview of AlphaFold2 and breakthroughs in overcoming its limitations. *Computers in Biology and Medicine*, 176:108620, June 2024.

[29] Qian Zhang, Nan Zhang, Han Xiao, Chen Wang, and Lian He. Small Antibodies with Big Applications: Nanobody-Based Cancer Diagnostics and Therapeutics. *Cancers*, 15(23):5639, January 2023. Publisher: Multidisciplinary Digital Publishing Institute.
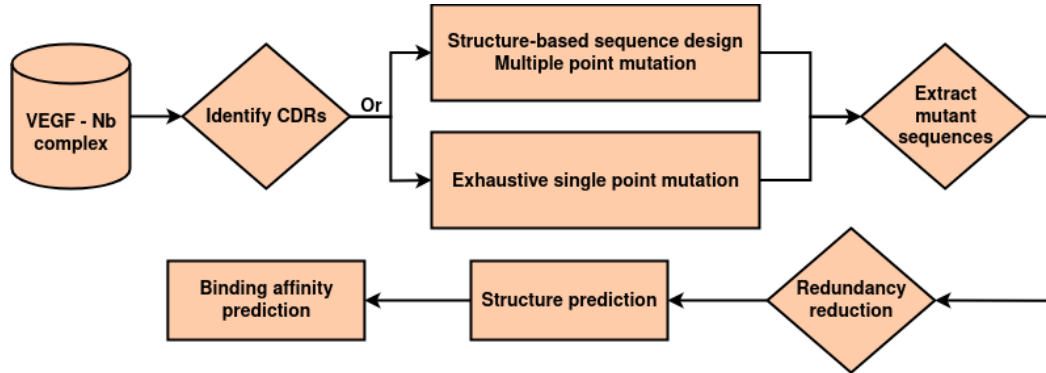
Supplementary Information

Figure S1: Wild-type based maturation pipeline. Multiple point mutations or exhaustive single point mutations are performed on the CDR3. After redundancy reduction, the structures are generated from the mutant sequences, and the binding affinity changes are predicted.
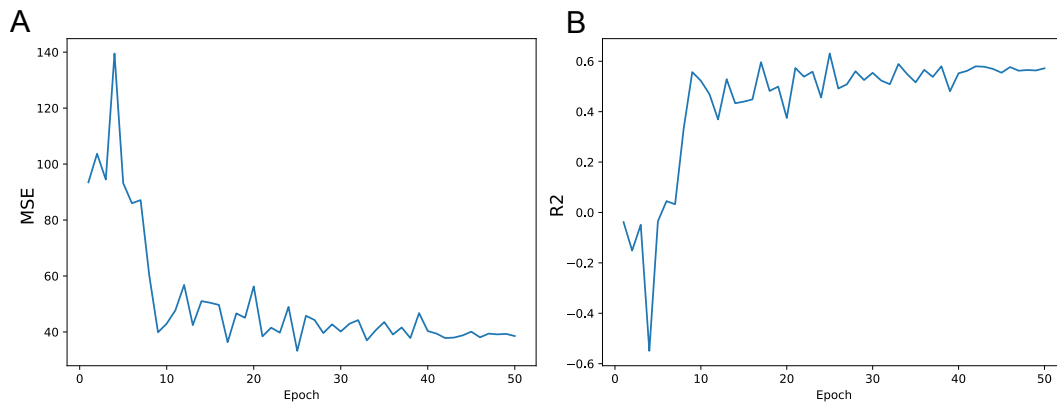

Figure S2: Evolution of (**A**) the Mean Square Error (MSE) and (**B**) R^2 over finetuning protBERT. The X-axis is the epochs and the Y-axis is the MSE or R^2.
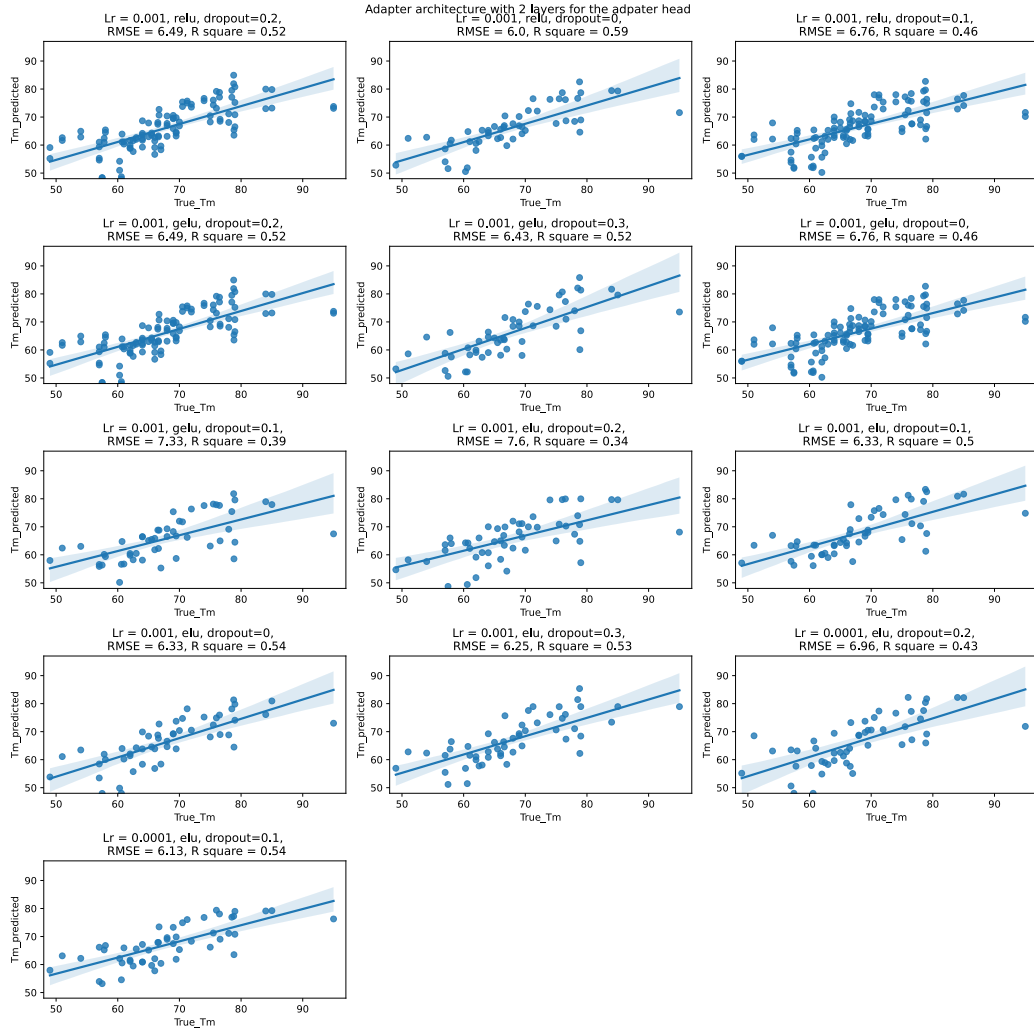
Figure S3: The performance of VHH melting temperature predictor on the test set. The finetuning of the parameters was performed on an adapter architecture on protBERT and with 2 layers in the adapter head. The learning rate (Lr), dropout, and activation function in the adapter heads were changed. RMSE stands for root mean square error.
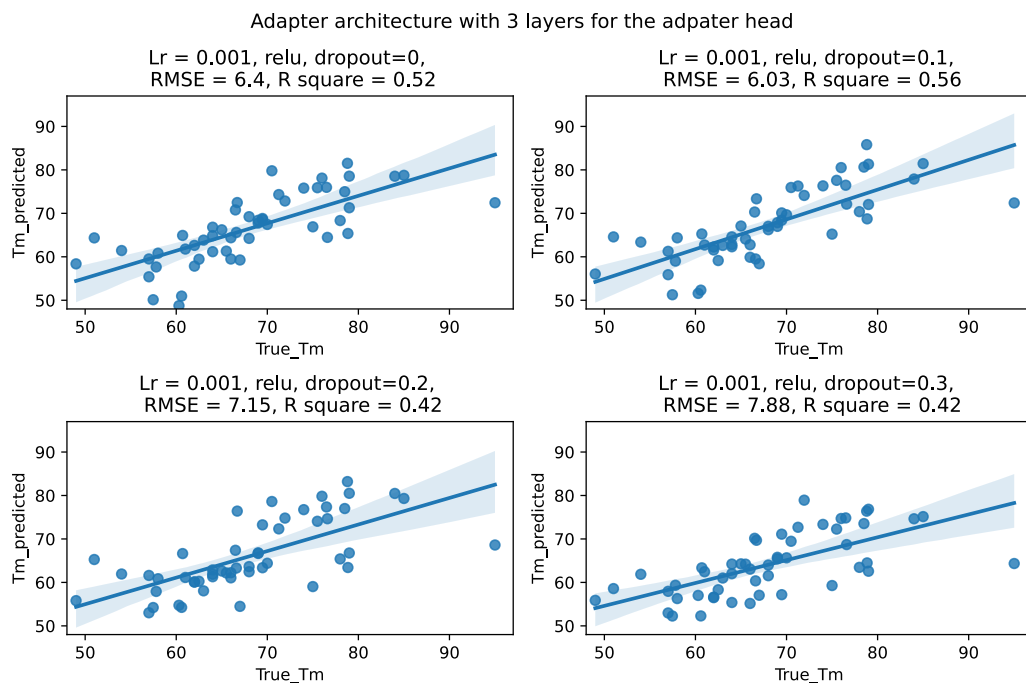
Figure S4: The performance of VHH melting temperature predictor on the test set. The finetuning of the parameters was performed on an adapter architecture on protBERT with 3 layers in the adapter head and a learning rate of 0.001. The dropout was changed. RMSE stands for root mean square error.
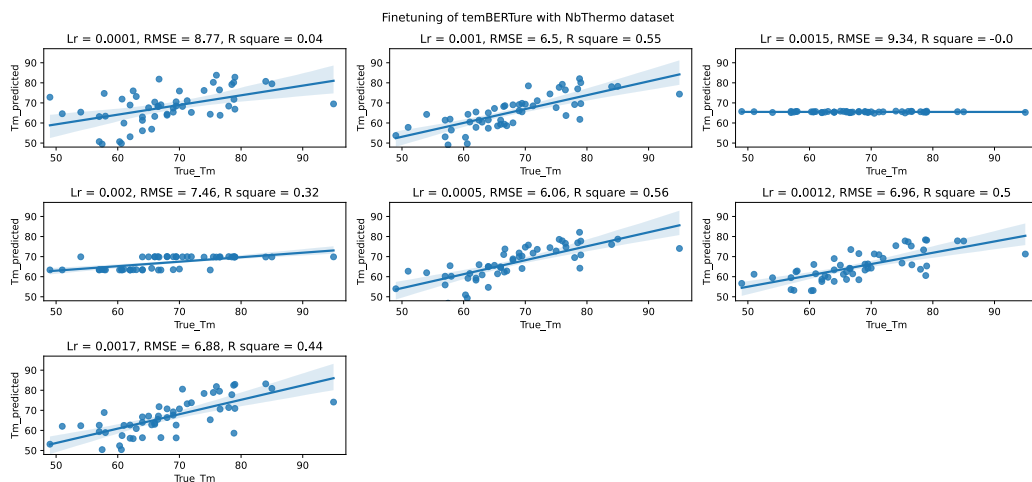

Figure S5: The performance of VHH melting temperature predictor on the test set. The parameters were finetuned on the temBERTure best model (replica 2). The learning rate (lr) was changed. RMSE stands for root mean square error.
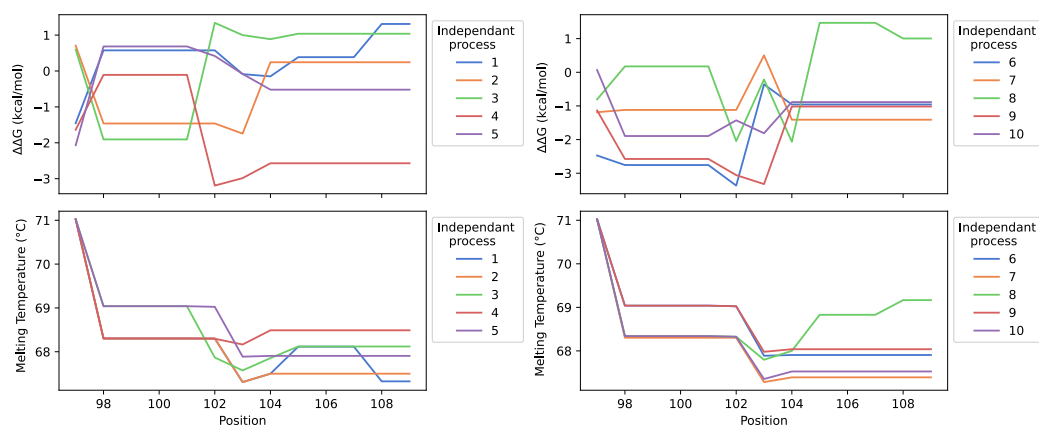
Figure S6: Evolution of the predicted changes in binding affinity (top) and melting temperature (bottom) through mutant positions within the CDR3 of the 5FV2 complex. Each curve corresponds to an independent process. The X-axis position corresponds to the amino acid number with respect to the first amino acid of the nanobody. CDR3 region has been identified to be between amino acids 97 and 109.
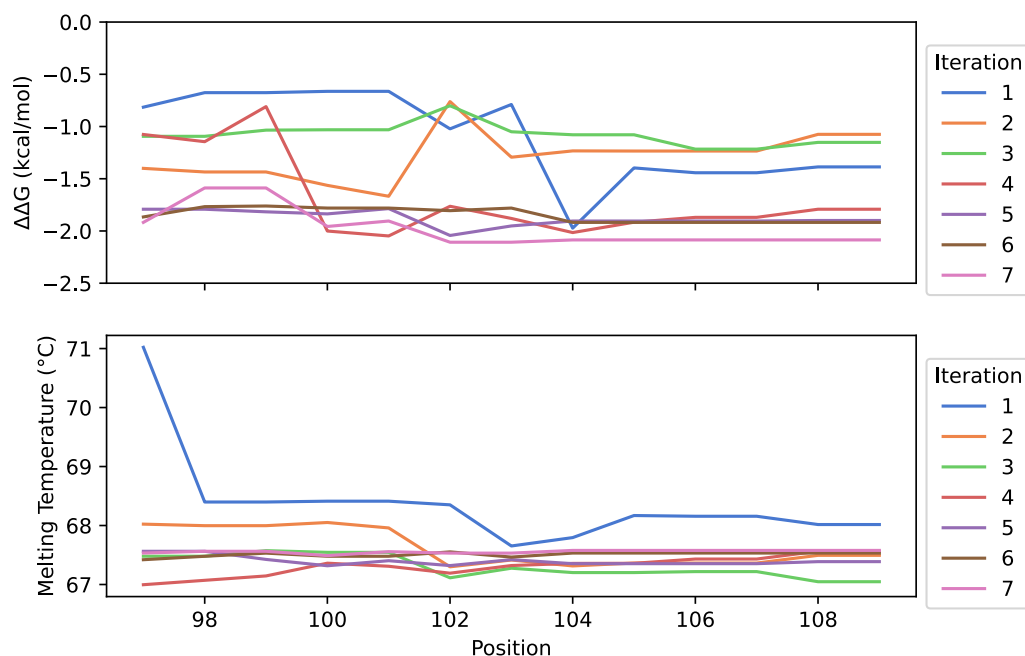

Figure S7: Evolution of the mean of predicted changes in binding affinity (top) and the mean of predicted melting temperatures (bottom) through mutant positions within the CDR3 of the 5FV2 complex for 7 iterations. The mean is calculated over the 10 independent runs. The mean of changes in binding affinity (top) reaches a lower value at the end of the last iteration. The mean melting temperature (bottom) remains in a 2°C range for all iterations except the first one, which reaches slightly higher values.
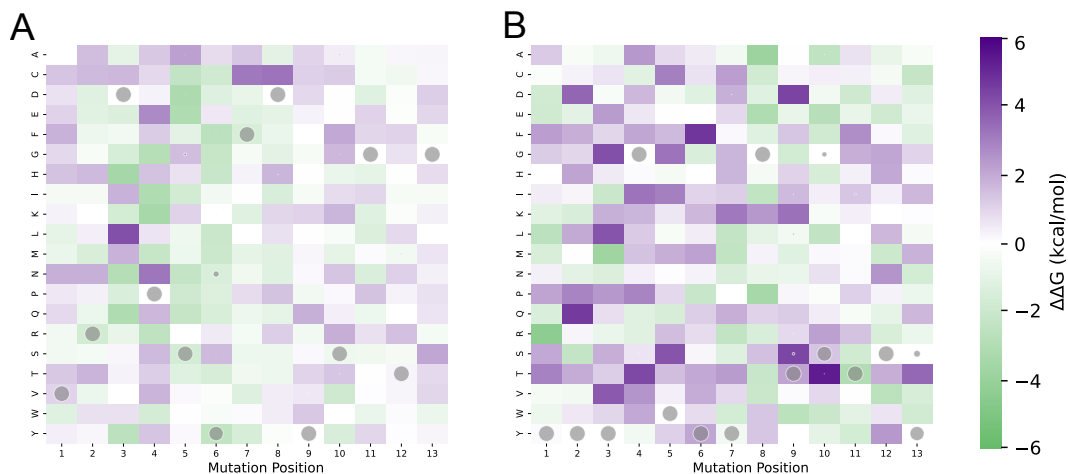
4

Figure S8: Predicted $\Delta\Delta G_{Bind}$ for exhaustive single point mutation on the CDR3 of VHH in 5FV2 (**A**) and in 3P9W (**B**). Gray circles represent the mean probability of amino acid substitution over 10 independent runs by AbMPNN for each amino acid position.
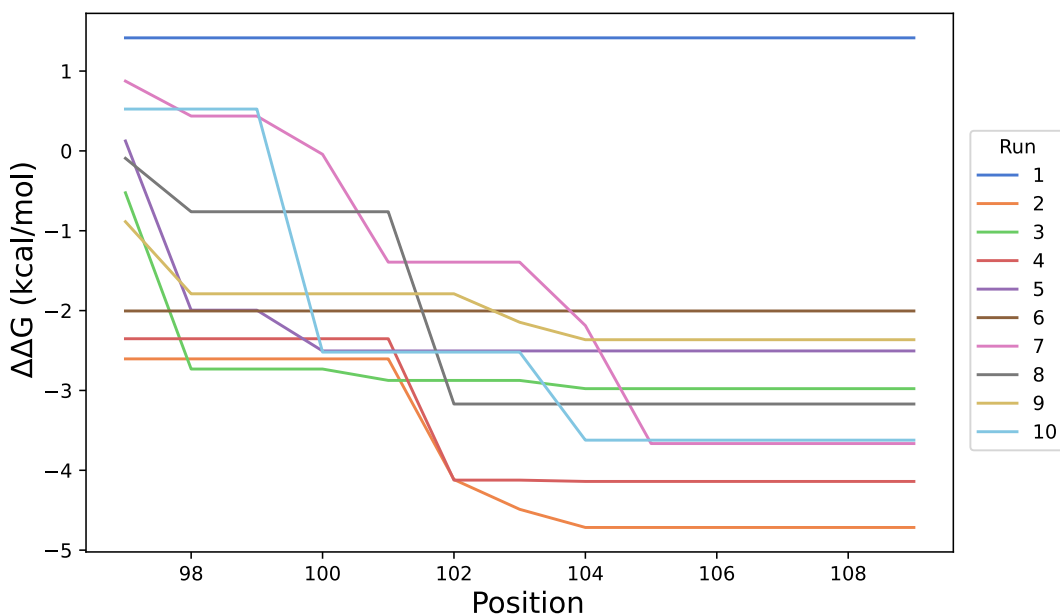

Figure S9: Evolution of the predicted $\Delta\Delta G_{Bind}$ through CDR3 mutant positions of VHH in 5FV2 complex for the guided mutation process over 10 independent runs. AbMPNN was used instead of ProteinMPNN. The affinity maturation stops in the first amino acid position for processes 1 and 6, while for all the other processes, more favorable mutants are reached at later positions. The X-axis is the position corresponding to the amino acid number with respect to the first amino acid of VHH. CDR3 region has been identified to be between amino acids 97 and 109.