Solubility-Aware Protein Binding Peptide Design Using AlphaFold

Takatsugu Kosugi* and Masahito Ohue

Department of Computer Science, School of Computing, Tokyo Institute of Technology kosugi@li.c.titech.ac.jp

Kosugi, T.; Ohue, M. Solubility-Aware Protein Binding Peptide Design Using AlphaFold. Biomedicines 2022, 10, doi:10.3390/biomedicines10071626.

Abstract

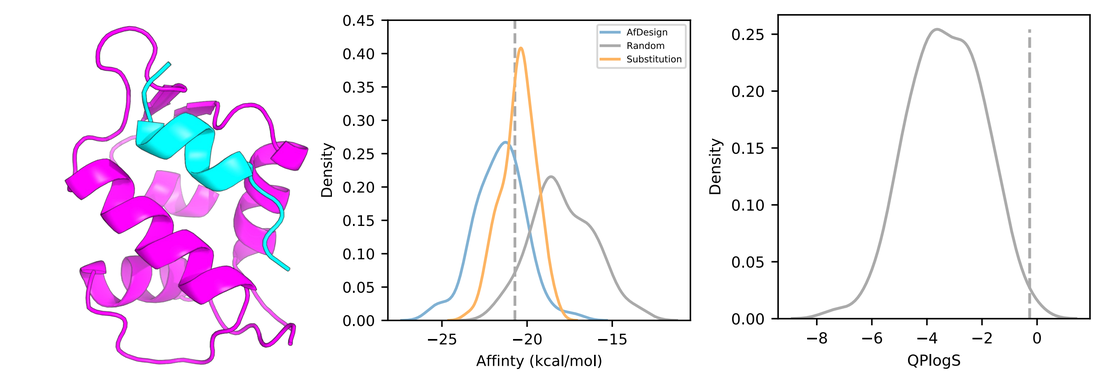
Peptides offer a new modality to target protein protein interactions (PPIs), but designing appropriate peptide sequences by computation is challenging.
We designed peptides likely to have PPI as the target protein using the "binder hallucination" protocol of AfDesign, a de novo protein design method using AlphaFold. However, the solubility of the pes tended to be low. Therefore, we designed a solubility loss function using solubility indices for amino acids and developed a solubility-aware AfDesign binder hallucination protocol. The peptide solubility in sequences designed using the new protocol increased with the weight of the solubility loss function; moreover, they captured the characteristics of the solubility indices. Moreover, the new protocol sequences tended to have higher affinity than native or random sequences when evaluated by docking binding affinity. Our approach shows that it is possible to design peptide sequences that can bind to the interface of PPI while controlling solubility.

Introduction

There are several possible strategies for peptide design to target PPIs considering solubility

- A library of highly solubility peptide sequences is created, and then docking scores and binding affinities are predicted by protein–peptide docking.
- Design peptide sequences that are likely to bind to target proteins using peptide sequence prediction methods such as AfDesign, and then evaluate water solubility and filter out those that exceed water solubility thresholds.

Challenges in Designing Peptides Targeting PPI with AfDesign binder hallucination



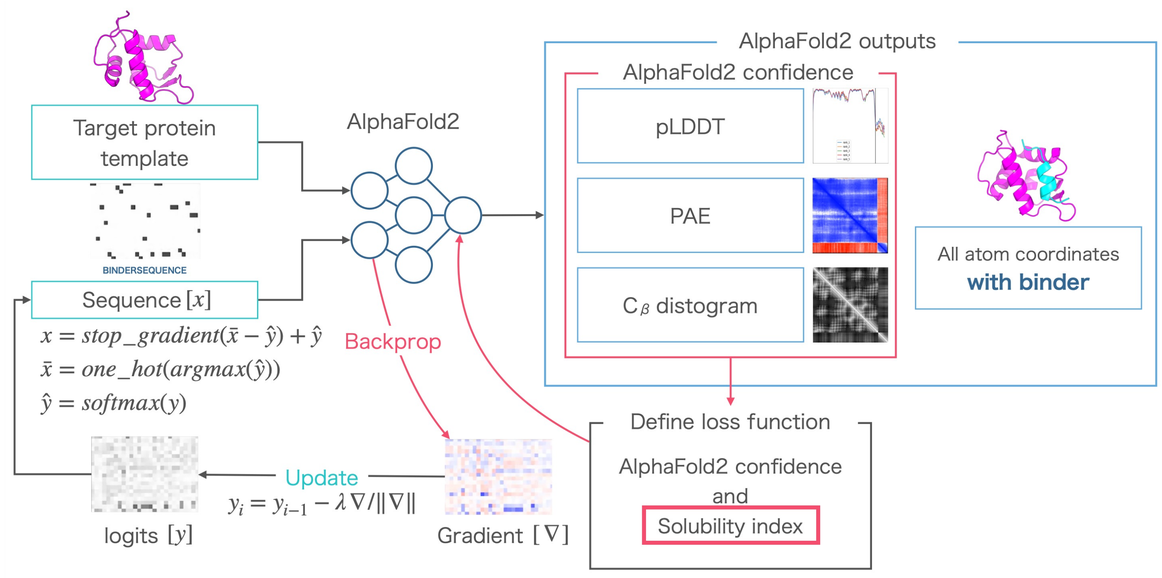
MDM2/p53 Bindig affinity Solubilty (logS)

In this study, we uses AfDesign binder hallucination [1], a method developed by Sergey Ovchinnikov, the developer of RoseTTAFold, was used to "hallucinate" binders to target proteins. The results of this study focused on binding affinity and solubility. The results showed that peptides with low solubility tended to be generated. Sequence logo for sequences designed with the AfDesign binder hallucination protocol

Methods

Schematic of the optimization method of the AfDesign binder hallucination with Solubility index

Sequences and solubility optimization at the same time



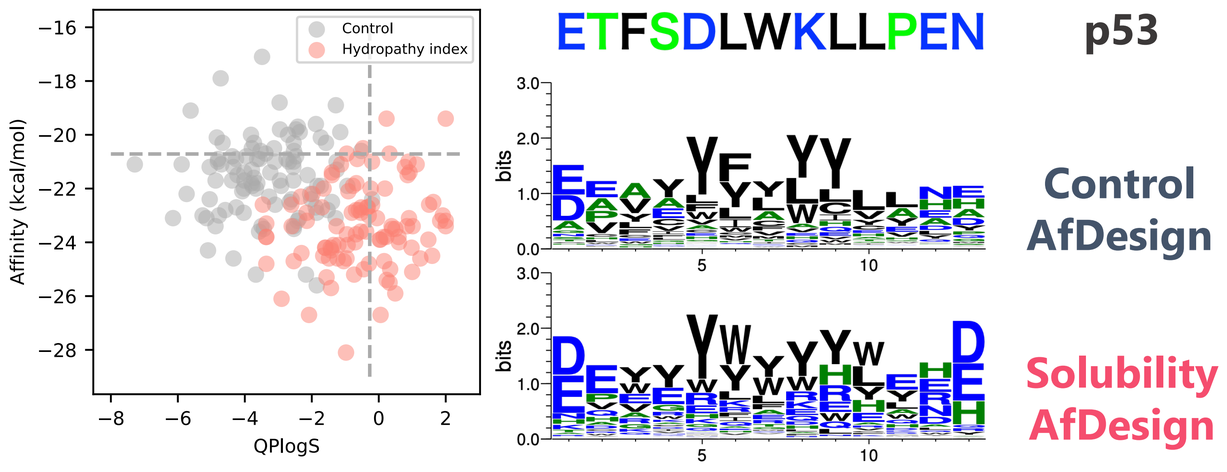
$x = stop\_gradient(\bar{x} - \hat{y}) + \hat{y}$
$\bar{x} = one\_hot(argmax(\hat{y}))$
$\hat{y} = softmax(y)$

$y_i = y_{i-1} - \lambda \nabla / \|\nabla\|$

The output of AlphaFold is an all-atom coordinate, two reliability indices (predicted aligned error (PAE) and predicted local-distance difference test (pLDDT)), and a distogram. These are used to define the loss function, which is back-propagated to compute the gradient for the design sequence and then updated and predicted in a loop for optimization.

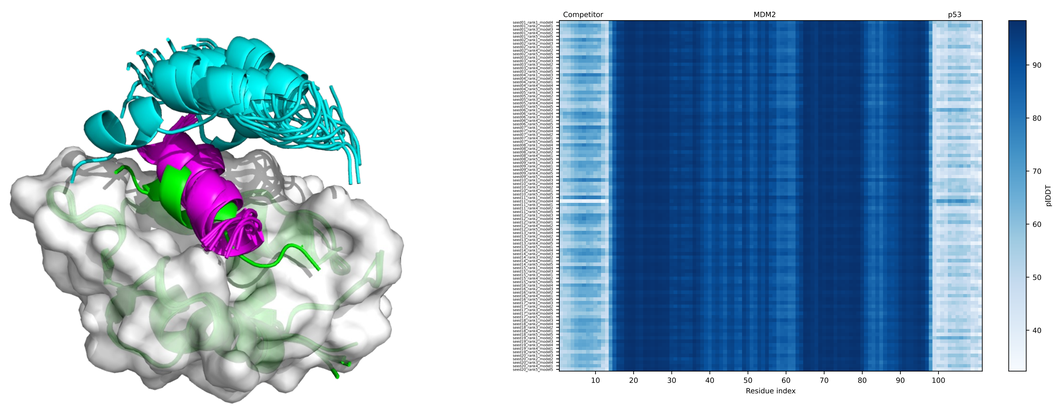
Results

Scatter plots of binding affinity and logS



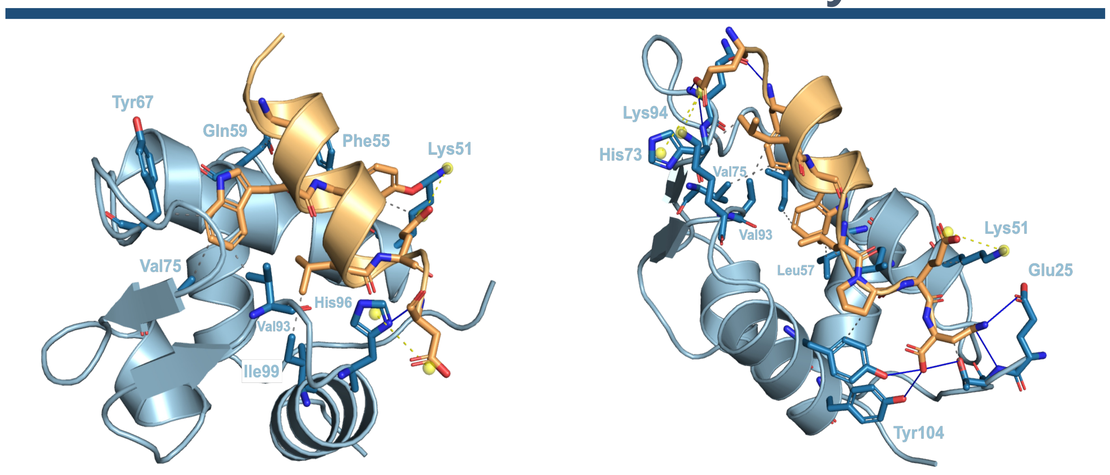
The sequences designed using the Hydropathy Index tended to have higher binding affinity with increasing logS than sequences without solubility loss. Thus, the Hydropathy Index is able to control logS while maintaining or increasing the binding affinity of sequences designed by the AfDesign binder hallucination protocol with solubility loss.

Competitive peptide binding prediction



Left Figure:All top rank models from competitive peptide binding prediction [2] of MDM2 with p53 and our highest affinity peptide consistently indicate the highest affinity peptide as the strong binder. Green indicates native MDM2 and p53 as well as MDM2 in the competitive peptide binding prediction. Cyan indicates p53 in the competitive peptide binding prediction. Magenta indicates the competitor peptide in the competitive peptide binding predictions. Right Figure: Heatmap of all pLDDT values in predictions of competitive peptide binding of the competitor peptide and p53 peptide with MDM2 using ColabFold. Our highest affinity competitor peptide was found at the MDM2 binding interface in all cases, except for one prediction.

Interatomic interaction analysis



Furthermore, PLIP [3] was used to compare the interatomic interaction between MDM2/the designed peptide and the MDM2/p53. As shown in Left figure a hydrogen bond between His96 of MDM2 and Asp13 of the designed peptide, a salt bridge between Lys51 of MDM2 and Glu11 of the designed peptide, and between His96 of MDM2 and Asp13 of the designed peptide were identified. In addition, several hydrophobic bonds were identified, such as that between Val75 of MDM2 and Trp6 and that between Val93 of MDM2 and Trp6 of the designed peptide. In the p53 peptide shown in Right figure, there is also a salt bridge between Lys51 of MDM2 and Glu28 of p53 peptide and hydrophobic binding between Val75 of MDM2 and Phe19 of p53 peptide and between Val93 of MDM2 and Leu22 of p53 peptide,

Code Availability

Solubility_AfDesign

- Solubility-aware AfDesign Code
- Easy-to-use Jupyter Notebook

References

1. @sokrypton, AfDesign. Available online: https://github.com/sokrypton/ColabDesign/tree/main/af
2. Chang, L.; Perez, A. AlphaFold Encodes the Principles to Identify High Affinity Peptide Binders. bioRxiv 2022
3. Adasme, M.F. et al. PLIP 2021: Expanding the Scope of the Protein-Ligand Interaction Profiler to DNA and RNA. Nucleic Acids Res. 2021, 49, W530–W534.

Acknowledgments

科研費 KAKENHI JST 研究成果展開事業 リサーチコンプレックス推進プログラム CREST 戦略的創造研究推進事業 文部科学省 Regional Innovation Ecosystems AMED BINDS