JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

Application Note

Cite This: J. Chem. Inf. Model. 2018, 58, 2189−2192

pubs.acs.org/jcim

CRISPro: An Automated Pipeline for Protein Conformation Stabilization by Proline

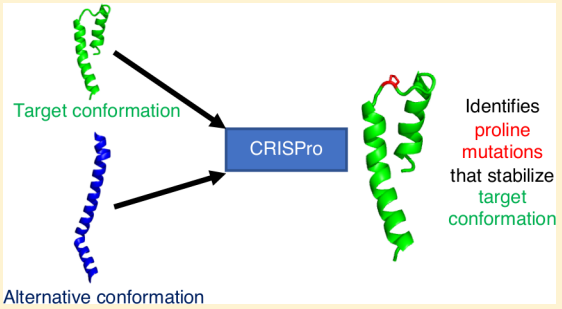
Reda Rawi,[†] Chen-Hsiang Shen,[†] Peter D. Kwong,[†,‡] and Gwo-Yu Chuang*[,†]

[†]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

[‡]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, United States

Supporting Information

ABSTRACT: Recent studies have shown that the yield, antigenicity, and immunogenicity of an immunogen can be enhanced by stabilizing it into a specific conformation. Such stabilization often involves the engineering of proline mutations at residue positions where a proline is structurally compatible with the target conformation but not with an alternative conformation. However, there is no publicly available tool that can design proline mutations for this purpose automatically. Here we implemented an automated tool, CRISPro, that inputs structural coordinates of the target conformation and/or an alternative conformation and outputs a list of residue positions where proline mutations are predicted to stabilize the target conformation based on compatibility of phi−psi angles, secondary structure, and steric constraints. Thus, CRISPro can be used to engineer immunogens into specific conformation and to design serologic probes, capable of isolating antibodies that recognize a target shape.



INTRODUCTION

In recent years, a number of subunit-based vaccine candidates have been developed. Compared to traditional vaccines such as live-attenuated vaccines or inactivated vaccines, subunit-based vaccines are safe and include only the antigen or antigens that best stimulate the desired immune response. Many candidate viral antigens, such as the HIV-1 envelope (Env) trimer or the paramyxovirus fusion (F) glycoprotein, undergo conformational change. To develop these antigens into effective immunogens, these antigens should be stabilized in the conformation that is best recognized by the immune system to provide an effective neutralizing response.[1,2]

There are multiple ways to stabilize a protein in a specific conformation, including engineered disulfides, cavity-filling mutations, and proline designs.[1,2] An engineered disulfide can rigidly lock two domains if designed properly, but it may disrupt folding of the protein as the engineered cysteines can also form disulfides with native cysteines. Cavity-filling mutations can stabilize conformation by filling holes needed for structural rearrangement and by strengthening the interaction between flexible regions; however, these designs may require time-intensive energetic calculations.[3]

Proline is unique among the standard 20 amino acids as its side chain is attached to the backbone nitrogen, which leads to more restricted backbone flexibility and to an inability to form helical secondary structures. The utility of proline mutations to enhance protein stability has been well documented.[4,5] Proline design has become an attractive way to stabilize protein conformation and has been applied in immunogen design to target diverse pathogens. For instance, Qiao, White, and colleagues have designed a V55P mutation to reduce the fusogenic activity of influenza hemagglutinin (HA).[6] Sanders, Moore, and colleagues have incorporated an I559P mutation to disrupt transitions from the prefusion conformation in their native-like soluble HIV-1 Env trimers.[7] Krarup, Langedijk, and colleagues have designed an S215P mutation to improve the stability of respiratory syncytial virus (RSV) F glycoprotein in the prefusion conformation.[8] Kwon, Gorman, and colleagues have found an A433P mutation to improve the prefusion-closed stability of HIV-1 Env trimers.[2] Pallesen, McLellan, and colleagues have shown that the yield for the spike protein of middle east respiratory syndrome coronavirus (MERS-CoV) can be substantially improved with the introduction of two proline mutations: V1060P and L1061P.[9] More recently, Battles, McLellan, and colleagues have designed an A185P mutation that stabilizes metapneumovirus (MPV) F glycoprotein in the prefusion conformation.[10] However, despite the success of proline-based protein stabilization, there is no publicly available automated tool that performs proline-based design.

Here we present CRISPro (ConfoRmatIon Stabilization by Proline), a software package that identifies residue positions suitable for proline design to stabilize the protein in a target conformation, using structural coordinates of the target conformation and/or an alternative conformation as input.

Received: August 31, 2018
Published: October 5, 2018

Downloaded via JOHNS HOPKINS UNIV on July 6, 2021 at 13:08:26 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications

This article not subject to U.S. Copyright. Published 2018 by the American Chemical Society

DOI: 10.1021/acs.jcim.8b00592
J. Chem. Inf. Model. 2018, 58, 2189−2192

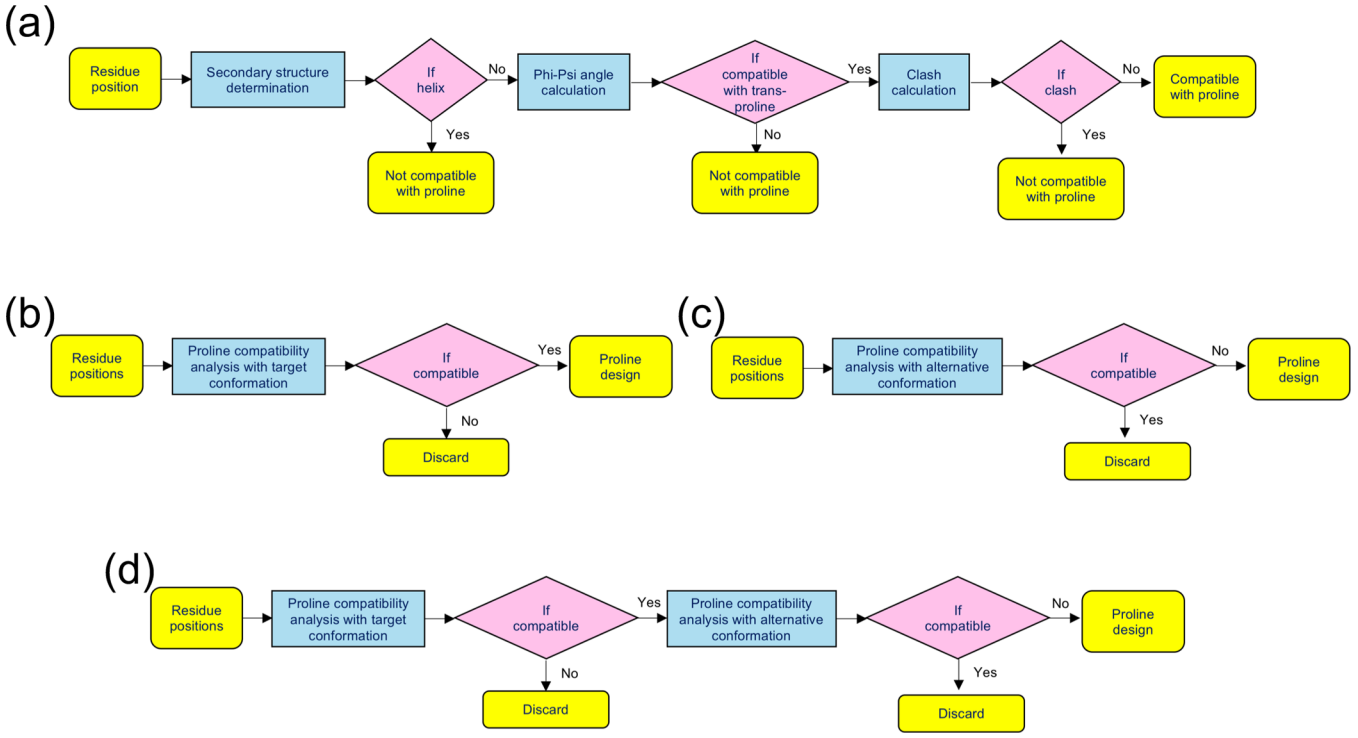


Figure 1. CRISPro workflow. (a) Workflow to determine residue compatibility to proline mutation. (b) Workflow of CRISPro when structural coordinates of a target conformation are supplied. (c) Workflow of CRISPro when structural coordinates of an alternative conformation are supplied. (d) Workflow of CRISPro when structural coordinates of both a target and an alternative conformation are supplied.

CRISPro is easy to install and use, thereby providing researchers with a convenient tool to design proline mutations for the stabilization of antigens in their target conformation.

CRISPro

CRISPro was implemented in R software and is available for download from GITHUB. It requires three programs as dependencies: R software environment and R package bio3d,[11] DSSP for secondary structure assignment,[12] and PyMOL version 1.7.2.1−2.2 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.).[13]

The workflow of CRISPro is illustrated in Figure 1. CRISPro predicts a residue position to be compatible with proline mutation if (1) the residue position is not part of a helical secondary structure, with the exception of the first three residues of the helix where the introduction of proline does not interrupt intrahelical hydrogen bond, and the last residue of the helix where the introduction of proline does not substantially affect the stability of the rest of the helix, (2) the phi−psi angles are within the allowable region for trans-prolines, and (3) there are no clashes between mutated proline side chains and any heavy atom at another residue position (Figure 1A). The secondary structure for each residue position was determined using DSSP. To determine if the phi−psi angles of a residue backbone were within the allowed angles for trans-prolines, we utilized the Top8000 rotamer reference data for trans-prolines from the Richardson lab (https://github.com/rlabduke/reference_data) with a density cutoff of 0.01. The number of clashes for designed proline were determined first by mutating the residue position to proline using PyMOL, and then by summing the number of instances where the distance between any heavy atom from the mutated proline side chain and any heavy atom from other residues was less than the sum of the radii of the two atoms minus 1.5 Å. Users have the flexibility to change the parameters for criteria 2 (proline rotamer density threshold) and 3 (atom overlap threshold). A higher proline rotamer density threshold results in less residue position with compatible torsional angles, while a higher atom overlap threshold allows more clashes from introduction of the proline side chain.

To use the program, users would need to input structural coordinates (in PDB format) of a target conformation, structural coordinates of an alternative conformation, or both. If structural coordinates of a target conformation were used as input, CRISPro generates a list of residue positions that are compatible with proline mutation and which stabilizes the target conformation (Figure 1B). If the structural coordinates of an alternative conformation were used as input, CRISPro generates a list of residue positions that are not compatible with proline mutation and which destabilizes the alternative conformation (Figure 1C). If the structural coordinate of both the target conformation and alternative conformation were provided, CRISPro generates a list of residue positions where proline mutations are compatible with the target conformation but not with the alternative conformation (Figure 1D).

CRISPro was successfully installed and tested on a Linux Ubuntu 16.04.3 LTS. The runtime of CRISPro, which can utilize multiprocessing, with the RSV F prefusion structure (PDB ID: 4JHW) as the input target conformation and the RSV F postfusion structure (PDB ID: 3RRR) as the input alternative conformation, was ∼78 min using one core and ∼14 min using four cores on a Dell Precision M6700 notebook with Intel Core i7-3920XM processors. We applied CRISPro to T4 lysozyme and B1 immunoglobulin-binding domain of streptococcal protein G using a single target conformation as input and found that proline mutations known to increase

DOI: 10.1021/acs.jcim.8b00592
J. Chem. Inf. Model. 2018, 58, 2189−2192

stability could be successfully predicted by CRISPro (Table S1). We also applied CRISPro to a number of viral antigens, including RSV F, MPV F, HIV-1 Env, and influenza HA, using prefusion and postfusion structures of each of these fusion machines as target and alternative input conformation, respectively, and showed retrospectively that experimentally identified proline mutants that stabilized these antigens in their prefusion states could be identified by CRISPro (Table S2). Altering the proline rotamer density thresholds (0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, and 0.0001) and atom overlap thresholds (1, 1.5, and 2 Å) showed that a number of parameter combinations led to the identification of all of the published prefusion-stabilizing mutations, while the default parameters (proline rotamer density threshold = 0.01, atom overlap threshold = 1.5 Å) identified these mutations with the highest specificity (Figure S1). We have also compared the backbone torsional angles and secondary structures between wild type and the published proline mutant for T4 lysozyme and hRSV F,[5,8] for which the crystal structures of both the wildtype and proline mutant were available (Table S3). We observed that the secondary structure of these two proteins remained the same after introduction of proline mutations, while the torsional angles were also affected minimally, with the exception of the psi angle for hRSV F S215P mutation. These analyses demonstrate the general applicability of CRISPro to the design of proline mutations that stabilize select target conformations of viral antigens in real-life settings.

CONCLUSIONS

In this study, we developed CRISPro, a software tool that predicts residue positions most suitable for proline mutations designed to stabilize proteins in a target conformation. Notably, we found that CRISPro was able to identify retrospectively—in several viral settings—multiple proline mutations that have been shown experimentally to improve the stability or the yield of an immunogen in the desired target conformation. CRISPro will thus be a useful tool to stabilize immunogens or probes in specified target conformation.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.8b00592.

> Retrospective analyses on proline mutations that improved the stability of T4 lysozyme and B1 immunoglobulin-binding domain of streptococcal protein G using CRISPro with default parameters and single structure as input (Table S1); retrospective analyses on proline mutations that improved the stability of prefusion conformation of hRSV F, hMPV F, HIV-1 Env, and influenza hemagglutinin using CRISPro with default parameters and both prefusion and postfusion structures as inputs (Table S2); comparison of phi−psi angles and secondary structure before and after introduction of published proline stabilizing mutations for T4 lysozyme and hRSV F (Table S3); impact of proline rotamer density and atom overlap thresholds on CRISPro prediction accuracy and specificity (Figure S1) (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: gwo-yu.chuang@nih.gov.

ORCID

Gwo-Yu Chuang: 0000-0003-0375-1013

Funding

This work has been supported by the Intramural Research Program (National Institute of Allergy and Infectious Diseases, National Institutes of Health, USA).

Notes

The authors declare no competing financial interest.

CRISPro is available for download free of charge at https://github.com/RedaRawi/CRISPro.

ACKNOWLEDGMENTS

We thank members of the Structural Biology Section, Vaccine Research Center, for helpful comments. This work utilized the computational resources of the NIH HPC Biowulf cluster (NIH HPC). This study used the Office of Cyber Infrastructure and Computational Biology (OCICB) High Performance Computing (HPC) cluster at the National Institute of Allergy and Infectious Diseases (NIAID), Bethesda, MD.

REFERENCES

(1) McLellan, J. S.; Chen, M.; Joyce, M. G.; Sastry, M.; Stewart-Jones, G. B.; Yang, Y.; Zhang, B.; Chen, L.; Srivatsan, S.; Zheng, A.; Zhou, T.; Graepel, K. W.; Kumar, A.; Moin, S.; Boyington, J. C.; Chuang, G. Y.; Soto, C.; Baxa, U.; Bakker, A. Q.; Spits, H.; Beaumont, T.; Zheng, Z.; Xia, N.; Ko, S. Y.; Todd, J. P.; Rao, S.; Graham, B. S.; Kwong, P. D. Structure-Based Design of a Fusion Glycoprotein Vaccine for Respiratory Syncytial Virus. *Science* **2013**, *342*, 592−598.

(2) Kwon, Y. D.; Pancera, M.; Acharya, P.; Georgiev, I. S.; Crooks, E. T.; Gorman, J.; Joyce, M. G.; Guttman, M.; Ma, X.; Narpala, S.; Soto, C.; Terry, D. S.; Yang, Y.; Zhou, T.; Ahlsen, G.; Bailer, R. T.; Chambers, M.; Chuang, G. Y.; Doria-Rose, N. A.; Druz, A.; Hallen, M. A.; Harned, A.; Kirys, T.; Louder, M. K.; O'Dell, S.; Ofek, G.; Osawa, K.; Prabhakaran, M.; Sastry, M.; Stewart-Jones, G. B.; Stuckey, J.; Thomas, P. V.; Tittley, T.; Williams, C.; Zhang, B.; Zhao, H.; Zhou, Z.; Donald, B. R.; Lee, L. K.; Zolla-Pazner, S.; Baxa, U.; Schon, A.; Freire, E.; Shapiro, L.; Lee, K. K.; Arthos, J.; Munro, J. B.; Blanchard, S. C.; Mothes, W.; Binley, J. M.; McDermott, A. B.; Mascola, J. R.; Kwong, P. D. Crystal Structure, Conformational Fixation and Entry-Related Interactions of Mature Ligand-Free Hiv-1 Env. *Nat. Struct. Mol. Biol.* **2015**, *22*, 522−531.

(3) Magliery, T. J. Protein Stability: Computation, Sequence Statistics, and New Experimental Methods. *Curr. Opin. Struct. Biol.* **2015**, *33*, 161−168.

(4) Choi, E. J.; Mayo, S. L. Generation and Analysis of Proline Mutants in Protein G. *Protein Eng., Des. Sel.* **2006**, *19*, 285−289.

(5) Matthews, B. W.; Nicholson, H.; Becktel, W. J. Enhanced Protein Thermostability from Site-Directed Mutations That Decrease the Entropy of Unfolding. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 6663−6667.

(6) Qiao, H.; Pelletier, S. L.; Hoffman, L.; Hacker, J.; Armstrong, R. T.; White, J. M. Specific Single or Double Proline Substitutions in the "Spring-Loaded" Coiled-Coil Region of the Influenza Hemagglutinin Impair or Abolish Membrane Fusion Activity. *J. Cell Biol.* **1998**, *141*, 1335−1347.

(7) Sanders, R. W.; Vesanen, M.; Schuelke, N.; Master, A.; Schiffner, L.; Kalyanaraman, R.; Paluch, M.; Berkhout, B.; Maddon, P. J.; Olson, W. C.; Lu, M.; Moore, J. P. Stabilization of the Soluble, Cleaved, Trimeric Form of the Envelope Glycoprotein Complex of Human Immunodeficiency Virus Type 1. *J. Virol* **2002**, *76*, 8875−8889.

(8) Krarup, A.; Truan, D.; Furmanova-Hollenstein, P.; Bogaert, L.; Bouchier, P.; Bisschop, I. J.; Widjojoatmodjo, M. N.; Zahn, R.;

DOI: 10.1021/acs.jcim.8b00592
J. Chem. Inf. Model. 2018, 58, 2189−2192

Schuitemaker, H.; McLellan, J. S.; Langedijk, J. P. A Highly Stable Prefusion Rsv F Vaccine Derived from Structural Analysis of the Fusion Mechanism. *Nat. Commun.* **2015**, *6*, 8143.
(9) Pallesen, J.; Wang, N.; Corbett, K. S.; Wrapp, D.; Kirchdoerfer, R. N.; Turner, H. L.; Cottrell, C. A.; Becker, M. M.; Wang, L.; Shi, W.; Kong, W. P.; Andres, E. L.; Kettenbach, A. N.; Denison, M. R.; Chappell, J. D.; Graham, B. S.; Ward, A. B.; McLellan, J. S. Immunogenicity and Structures of a Rationally Designed Prefusion Mers-Cov Spike Antigen. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E7348−E7357.
(10) Battles, M. B.; Mas, V.; Olmedillas, E.; Cano, O.; Vazquez, M.; Rodriguez, L.; Melero, J. A.; McLellan, J. S. Structure and Immunogenicity of Pre-Fusion-Stabilized Human Metapneumovirus F Glycoprotein. *Nat. Commun.* **2017**, *8*, 1528.
(11) Grant, B. J.; Rodrigues, A. P.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. Bio3d: An R Package for the Comparative Analysis of Protein Structures. *Bioinformatics* **2006**, *22*, 2695−2696.
(12) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22*, 2577−2637.
(13) Frankel, S. S.; Steinman, R. M.; Michael, N. L.; Kim, S. R.; Bhardwaj, N.; Pope, M.; Louder, M. K.; Ehrenberg, P. K.; Parren, P. W.; Burton, D. R.; Katinger, H.; VanCott, T. C.; Robb, M. L.; Birx, D. L.; Mascola, J. R. Neutralizing Monoclonal Antibodies Block Human Immunodeficiency Virus Type 1 Infection of Dendritic Cells and Transmission to T Cells. *J. Virol* **1998**, *72*, 9788−9794.

DOI: 10.1021/acs.jcim.8b00592
J. Chem. Inf. Model. 2018, 58, 2189−2192