Predicting Binding Energy of Staphylocoagulase

Erik Nebel, Jonathan Sheehan, David Nannemann, and Jens Meiler

KEYWORDS. Rosetta, Staphylocoagulase-Prothrombin Complex, nonproteolytic protein activation, computational mutations

BRIEFS. Computational mutations of protein-protein interfaces.

ABSTRACT. Thrombin is a protein involved in the process of blood clotting. It is found in the blood plasma of mammals as an inactive precursor called prothrombin (PR). PR is activated by proteolytic cleavage. In this process the first 15 residues of the N-terminus are cleaved. The enzyme staphylocoagulase (SC), produced by Staphylococcus aureus, is able to activate PR without the usual proteolytic cleavage, thereby directly initiating blood clotting. This works as follows: the N-terminal tail of SC binds into a "pocket" of PR resulting in a conformational rearrangement and activation of PR. To understand the underlying mechanism of this process, a library of peptides was computationally docked into PR. The predicted binding affinity was compared with experimental studies to verify the predictive power of the model.

INTRODUCTION.

Prethrombin is a protein involved in the process of blood clotting. Normally it gets activated by cleavage of the single Arg15-Ile16 peptide bond [1].

Staphylocoagulase (SC) is a toxin secreted by Staphylococcus aureus, which leads to many diseases like sepsis or endocarditis because it is able to activate prethrombin (PR) [1]. SC belongs to the nonproteolytic protein activators. It activates PR without the proteolytic cleavage to directly initiate blood clotting. The process is initiated by SC binding to PR and inserting its N-terminal peptide into the activation pocket of PR (Figure 1). The exact mechanism of this process is not completely understood. It was hypothesized that the N-Terminal residues of SC imitate the Ile-Val-Gly/Asn sequences of trypsin and chymotrypsin. Recently performed crystallographic studies on the PR/SC complex support this theory [1].

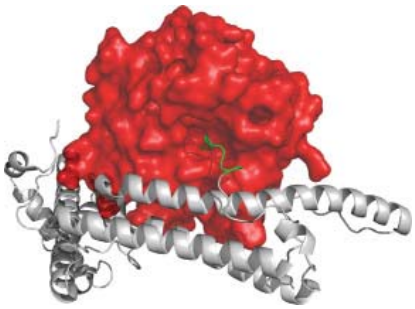

Figure 1. SC activates PR by inserting it's N-terminal residues (green) into the activation "pocket" of PR. One effort to understand this process better involves mutating the N-Terminal residues of SC and study the effect on PR activation.

When considering just the three most N-terminal residues of SC, considering the 20 genetically encoded amino acids this procedure yields up to 8,000 peptide variation–a number that is experimentally difficult to handle. The focus of the present study is to leverage computational power to generate these mutations in silico to predict the activity of SC mutants thereby focusing the experimental studies on the most interesting peptides.

METHODS.

The protein modeling suite Rosetta was used to computationally generate mutations of the peptide. Rosetta predicts protein structures and intermolecular interactions through evaluation of van der Waal's packing, solvation, hydrogen bonding and electrostatic scoring terms, among others. The generated protein models were analyzed for lowest binding energies indicating tight binding, The most promising candidates will be experimentally validated.

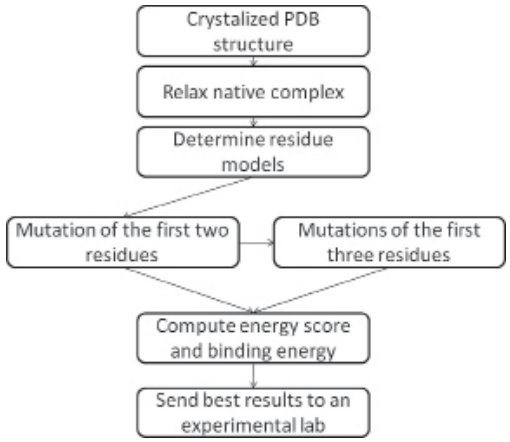

Figure 2. The graphic summarizes the Rosetta protocol.

The Rosetta Scrips framework was utilized to provide access to confromational sampling and scoring functions within Rosetta. Briefly, the experimental coordinates of the PR/SC complex were obtained from the PDB (code: 1nu9) and relaxed in the force field. Afterwards, mutations were introduced using design functionalities, the SC N-terminus was docked and the complex energy minimized. Finally, the complex was scored to determine the binding affinities.

The two N-terminal positions were exhaustively screened with combinations of the 20 genetically encoded amino acids yielding a total of 400 peptides. The 25 double mutants predicted to bind most tightly were retained and merged with a list of 7 double mutants of particular interest to our experimental collaborators. All possible triple mutants were created for these 32 peptides yielding a total of 640 sequences.

RESULTS.

To validate our computational protocol we correlated predicted with known experimental binding affinities for 30 peptides (r^2=0.43). The correlation between the experimental results and Rosetta verifies the computational mutations.

The triple mutants have stronger binding affinities than the double mutants. This result suggests that the third N-Terminal residue is involved in binding. The third amino acid sequence is dominated by large polar amino acids like Tyrosine or Arginine. The most realistic results are shown in the Table 1.

Table 1. The table shows the triple mutations with the strongest binding affinities.

File Name	Average_Score	StDev	ddg	StDev
ALA-1_TRP-2_TYR-3	-679.354	0.491	-21.314	0.177
ILE-1_ARG-2_TRP-3	-678.806	0.621	-22.774	0.217
ILE-1_HIS-2_TRP-3	-681.557	1.223	-22.036	0.294
ILE-1_MET-2_TYR-3	-681.579	0.203	-21.077	0.017
ILE-1-SER-2_ARG-2	-681.857	0.138	-19.714	0.059
ILE-1_TRP-2_HIS-3	-681.116	0.346	-22.440	0.103
LEU-1_GLN-2_HIS-3	-681.499	0.429	-19.863	0.189
MET-1_ARG-2_ARG-3	-682.703	0.45	-21.018	0.077
PRO-1_TRP-2_LYS-3	-679.538	0.961	-21.638	0.336
THR-1_HIS-2_TYR-3	-682.427	0.777	-21.492	0.114

CONCLUSIONS.

The predicted triple mutations of Rosetta are dominated by Isoleucine in the first position. The second position is dominated by large residues like Tryptophan, Histidine, or Arginine. Rosetta places large polar amino acids in the third position.

One limitation of the present analysis is that it does not predict whether PR is activated after binding of the peptide. Therefore, peptides predicted to have the highest binding affinity must be experimentally validated. Interestingly, Rosetta predicts a few mutations which fit better into the activation pocket than the wild type peptide. If this finding is experimentally verified it would be most interesting because of the question why it is not used by nature. One hypothesis is that these peptides might activate PR too fast killing the host before the bacterium can replicate.

Our results will help to understand the process of PR activation by SC. This understanding might yield opportunities to develop drugs against endocarditis or sepsis that prevent PR activation.

ACKNOWLEDGMENTS. The author hereby acknowledges the Bock Group, who supported us with their experimental data, the Meiler Group and the Research Experience for High School Students (REHSS) at Vanderbilt University.

REFERENCES.

1. R. Friedrich, *et al., Nature.* 425, 535-539 (2003).
2. T. Kortemme, *et al., Sci STKE.* 219 (2004).
3. S. Fleishman, *et al., PLoS One.* 6,. e20161 (2011).

Erik Nebel is a student at Leibnizschule in Wiesbaden, Germany, and participated in the Research Experience for High School Students Program (REHSS).