Supplementary Materials: A Novel Strategy for Molecular Interfaces Optimization: the case of Ferritin-Transferrin Receptor Interaction

Lorenzo Di Rienzo[a,*], Edoardo Milanetti[a,b], Claudia Testi[a], Linda Celeste Montemiglio[c], Paola Baiocco[a,d], Alberto Boffi[d], Giancarlo Ruocco[a,b]

[a] *Center for Life Nanoscience, Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161, Rome, Italy*
[b] *Department of Physics, Sapienza University, Piazzale Aldo Moro 5, 00185, Rome, Italy*
[c] *Institute of Molecular Biology and Pathology, National Research Council, Rome, Italy*
[d] *Department of Biochemical Sciences "A. Rossi Fanelli" Sapienza University, Piazzale Aldo Moro 5, 00185, Rome, Italy*

Abstract

Protein-protein interactions regulate almost all cellular functions and rely on a fine tune of surface amino acids properties involved on both molecular partners. The disruption of a molecular association can be caused even by a single residue mutation, often leading to a pathological modification of a biochemical pathway. Therefore the evaluation of the effects of amino acid substitutions on binding, and the *ad hoc* design of protein-protein interfaces, is one of the biggest challenges in computational biology. Here, we present a novel strategy for computational mutation and optimization of protein-protein interfaces. Modeling the interaction surface properties using the Zernike polynomials, we describe the shape and electrostatics of binding sites with an ordered set of descriptors, making possible the evaluation of complementarity between interacting surfaces. With a Monte Carlo approach, we obtain protein mutants with controlled molecular complementarities. Applying this strategy to the relevant case of the interaction between Ferritin and Transferrin Receptor, we obtain a set of Ferritin mutants with increased or decreased complementarity. The extensive molecular dynamics validation of the method results confirms its efficacy, showing that this strategy represents a very promising approach in designing correct molecular interfaces.

*Corresponding author
Email address: lorenzo.dirienzo@iit.it (Lorenzo Di Rienzo)

Preprint submitted to Computational and Structural Biotechnology Journal September 10, 2020

1. Analysis of the obtained Ferritin mutants

In Supplementary Table 1 we summarize the complementarity values obtained for all the 20 mutants produced. In particular, we report the shape and electrostatic complementarity and the percentage of difference each mutant shows in this values with respect to the wild type and the arithmetic mean of this quantities. We used the mean complementarity ro rank the mutants within the H-C set and L-C set.

As the complementarity increases when the numerical values decrease (and vice-versa), it is evident that the difference in complementarity between the wild type and both the categories of mutants is notable.

We next analyze the changes in three important chemico-physical properties between the mutants and the wild type Ferritin. We selected for the analysis the variation in charge [1], side chain volume [2] and hydropathy [3]. We calculated the properties X for each residue of the wild type and mutant proteins, and then we sum the values obtained separately for each protein. We finally obtained the changes due to the mutations with $X_{wt} - X_{mut}$.

From Supplementary Table 2 it emerges that when the Monte Carlo aims to increase the binding complementarity some negatively charged groups are inserted, preferring also small and hydrophobic residues. Expectably, when on the contrary the complementarity has to be reduced, a high increase in positive, voluminous, and hydrophilic residues is observed.

	Shape Complementarity	Electrostatic Complementarity	Mean Complementarity
WT	0.137 (0%)	1.071 (0%)	0.604
H-C A	0.068 (- 50%)	0.516 (- 52%)	0.292 (4^{th})
H-C B	0.077 (- 44%)	0.527 (- 51%)	0.302 (8^{th})
H-C C	0.073 (- 47%)	0.525 (- 51%)	0.299 (6^{th})
H-C D	0.072 (- 47%)	0.508 (- 53%)	0.290 (2^{nd})
H-C E	0.078 (- 43%)	0.503 (- 53%)	0.291 (3^{rd})
H-C F	0.079 (- 42%)	0.461 (- 57%)	0.270 (1^{st})
H-C G	0.072 (- 47%)	0.522 (- 51%)	0.297 (5^{th})
H-C H	0.069 (- 50%)	0.570 (- 47%)	0.320 (9^{th})
H-C I	0.088 (- 36%)	0.670 (- 37%)	0.379 (10^{th})
H-C J	0.076 (- 45%)	0.523 (- 51%)	0.300 (7^{th})
L-C A	0.191 (+ 39%)	2.445 (+ 128%)	1.318 (6^{th})
L-C B	0.159 (+ 16%)	2.043 (+ 90%)	1.101 (10^{th})
L-C C	0.186 (+ 36%)	2.477 (+ 131%)	1.332 (4^{th})
L-C D	0.177 (+ 29%)	2.654 (+ 148%)	1.416 (2^{nd})
L-C E	0.212 (+ 55%)	2.623 (+ 128%)	1.418 (1^{st})
L-C F	0.191 (+ 39%)	2.445 (+ 128%)	1.318 (6^{th})
L-C G	0.209 (+ 53%)	2.072 (+ 93%)	1.141 (9^{th})
L-C H	0.189 (+ 38%)	2.450 (+ 129%)	1.320 (5^{th})
L-C I	0.187 (+ 36%)	2.608 (+ 144%)	1.398 (3^{rd})
L-C J	0.178 (+ 30%)	2.349 (+ 119%)	1.264 (8^{th})

Supplementary Table 1: Variation in binding stability, as calculated by shape and electrostatic Zernike complementarity. In each row we show the values relating to the wild type or to the mutants Ferritin in complex with Transferrin receptor and their percentage of variation compared to the wild type. In the last column we report the arithmetic mean of the complementarities and the rank of each mutant: for the high complementarity mutants the rank goes from the more complementary (lower numerical value) to the less complementary (higher numerical value), while for the low complementary mutants the reverse is true.

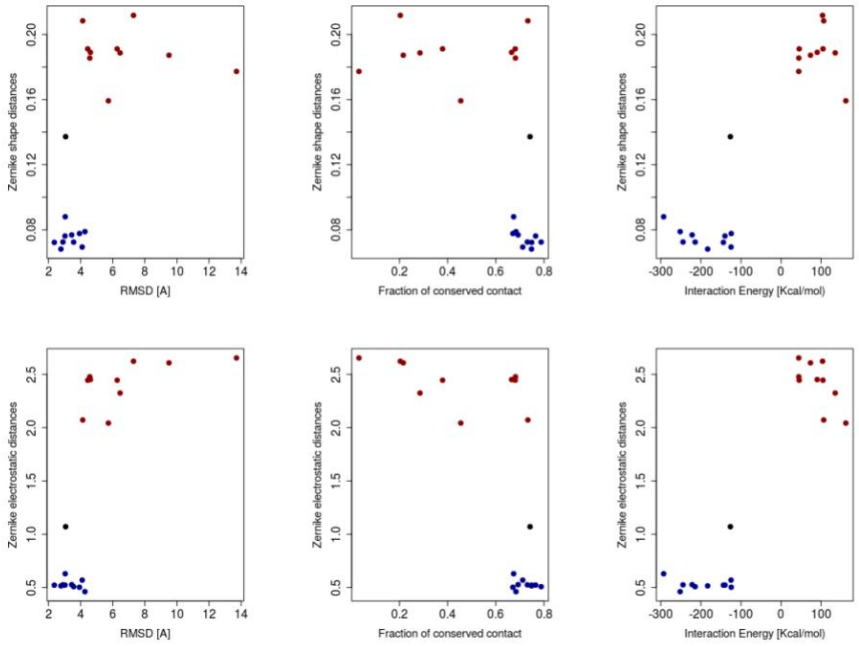
	$\Delta Q[e]$	$\Delta Vol[A^3]$	ΔHph
H-C A	4	213.0	-5.8
H-C B	4	174.7	-9.2
H-C C	4	62.5	-5.2
H-C D	4	191.0	-5.8
H-C E	3	151.9	-34.8
H-C F	4	163.1	-18.5
H-C G	3	121.0	-5.4
H-C H	3	218.2	-16.2
H-C I	2	- 4.1	-10.2
H-C J	4	96.9	-8.1
L-C A	- 8	- 211.0	3.4
L-C B	- 5	- 239.3	14.0
L-C C	- 8	- 206.6	3.2
L-C D	- 8	- 292.3	6.9
L-C E	- 8	- 321.2	12.4
L-C F	- 8	- 211.0	3.4
L-C G	- 6	- 210.8	-0.2
L-C H	- 7	- 245.5	8.7
L-C I	- 8	- 248.8	9.1
L-C J	- 7	- 197.3	4.0

Supplementary Table 2: Changes in charge, side chain volume and Hydropathy between the wild type and the mutants ($X_{wt} - X_{mut}$)

2. Molecular Dynamics Validation

We studied the correlation between changes in Zernike-based E used in the Monte Carlo procedure and the various stability indices deduced from the Molecular Dynamics simulation results. In particular we analyze how the variations in shape and electrostatic complementarities are related with the mean RMSD of the ferritin in respect to the initial configuration, the mean percentage of conserved contacts with respect to the wild type experimental complex, the mean

energy exchanged between the atoms at the interface (Fig. 1).

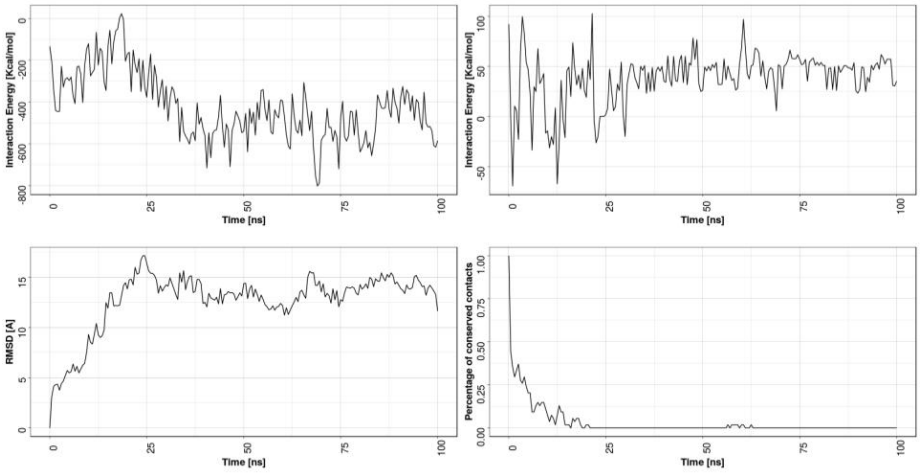


Supplementary Figure 1: Correlations between the changes in Zernike-based Energy during the Monte Carlo and the stability indices obtained from the Molecular Dynamics. On the top row the plots regard the shape complementarity while in the bottom row we face the electrostatic complementarity. **left** On the x axis the mean RMSD of the ferritin in respect with the initial configuration, on the y axis the Zernike complementarity. **center** On the x axis the mean fraction of conserved contact with respect to the wild type experimental complex, on the y axis the Zernike complementarity. **right** On the x axis the mean energy exchanged between the atoms belonging to the experimental binding site residues (Fig. 1), on the y axis the Zernike complementarity. The H-C set, the L-C set and the Wild Type are represented with blue, red and black points respectively.

30 The results of this analysis is shown in Supplementary Figure 1. As evident, these plots suggest that exist a clear distinction between the two categories of mutants.

3. A case of study: Mutant L-C D.

The aim of our work is to develop a computational protocol able to increase the compatibility between 2 molecular interfaces, in order to give some insights for the design of molecules with increased affinity and unaltered functional properties. We choose to validate our protocol with Molecular Dynamics, and our purpose is to verify if the modified interface could be able to bind its molecular partner. We think that it could be explicating to discuss in detail an example, in particular analyzing the time evolution of the simulations regarding the L-C mutants D.



Supplementary Figure 2: **top left)** The mean energy of non-covalent interactions between all the atoms of the ferritin and the receptor. **top right)** The mean energy of non-covalent interactions between all the atoms of the ferritin and the receptor binding site (as defined in Figure 1). **bottom left)** The RMSD with respect to the initial configuration. **bottom right)** the percentage of conserved ferritin-receptor contacts with respect to the wild type complex.

In the case of such mutant, in order to worsen the complementarity, the Monte Carlo protocol selects the following mutations:

- 17 E $\longrightarrow$ K

- 21 N $\longrightarrow$ K
- 78 G $\longrightarrow$ R
- 83 Q $\longrightarrow$ K
- 116 E $\longrightarrow$ R
- 123 D $\longrightarrow$ R

As evident, many charged groups are involved in these substitutions. The electrostatic composition of the Ferritin interface is therefore dramatically modified, with a very high increase of positive charges. As intuitive, the Monte Carlo simulation, in order to create a ferritin binding site unsuitable for its corresponding receptor binding site, lead to a substantial inversion of the charge distribution on the Ferritin surface.

In light of these considerations, it is not surprising that the initial configuration of the molecular dynamics situation is very unstable. As shown in the Supplemenary Figure 2, in the initial 25 ns of simulation the mutated ferritin moves away from its original position, as evidenced by the extremely high values of RMSD (bottom left) and by the lost of all the original contacts (bottom right); indeed the significant change in charge composition make the presence of the ferritin on the original binding site energetically very unfavourable, as shown in Supplementary Figure 2 top panels.

After this initial time, the ferritin mutants find another configuration, apparently stable and with a good interaction.

This new configuration, that is the repositioning of the ferritin in interaction with another epitope on transferrin receptor, for our purposes it is not acceptable for several reasons.

Firstly, the investigation of the goodness of this new interface through Molecular Dynamics would require a very long and costly simulation, since the starting conformation it is not derived by an experimental data. More importantly, the validation of a new ferritin-receptor interface is largely beyond the scope of this work, since our purpose is the optimization of existing interface and not the

identification of new ones. In our framework, if the mutant moves away from the transferrin binding sites means that the original interaction is abolished, indicating the success of the MonteCarlo process in creating low complementarity mutants.

References

[1] P. Klein, M. Kanehisa, C. DeLisi, Prediction of protein function from sequence properties: Discriminant analysis of a data base, Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology 787 (3) (1984) 221–226.

[2] W. Krigbaum, A. Komoriya, Local interactions as a structure determinant for protein molecules: Ii., Biochimica et biophysica acta 576 (1) (1979) 204.

[3] J. Kyte, R. F. Doolittle, A simple method for displaying the hydropathic character of a protein, Journal of molecular biology 157 (1) (1982) 105–132.