



**HAL**  
open science

## **Cross-Talk Between Intramolecular and Intermolecular Amino Acid Networks Orchestrates the Assembly of the Cholera Toxin B Pentamer via the Residue His94**

Mounia Achoch, Giovanni Feverati, Kave Salamatian, Laurent Vuillon, Claire Lesieur

### ► To cite this version:

Mounia Achoch, Giovanni Feverati, Kave Salamatian, Laurent Vuillon, Claire Lesieur. Cross-Talk Between Intramolecular and Intermolecular Amino Acid Networks Orchestrates the Assembly of the Cholera Toxin B Pentamer via the Residue His94. International Conference on Advanced Intelligent Systems for Sustainable Development ; AI2SD 2022, May 2022, Rabat, Morocco. pp.391 - 404, 10.1007/978-3-031-35248-5\_35 . hal-04278362

**HAL Id: hal-04278362**

**<https://hal.science/hal-04278362>**

Submitted on 12 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Cross-talk between intramolecular and intermolecular amino acid networks orchestrates the assembly of the cholera toxin B pentamer via the residue His94

Mounia Achoch<sup>1\*</sup>, Giovanni Feverati<sup>2</sup>, Kave Salamatian<sup>3</sup>, Laurent vuillon<sup>4</sup> and Claire Lesieur<sup>5</sup>

<sup>1</sup>Systems Engineering Laboratory, Sultan Moulay Slimane University, Beni Mellal, Morocco

<sup>2</sup>Fr3405 Research Federation, Modeling, Simulations, Fundamental Interactions, Annecy-le-Vieux, France

<sup>3</sup>System Computing Laboratory, Information and Knowledge Processing (LISTIC), University of Savoie, Annecy le Vieux, France

<sup>4</sup>Mathematics laboratory (LAMA UMR 5127), Savoie Mont Blanc University, CNRS, Le Bourget du Lac, France;

<sup>5</sup>Univ. Lyon, CNRS, INSA Lyon, Université Claude Bernard Lyon 1, Ecole Centrale de Lyon, Ampère UMR5005, 69622, Villeurbanne, France

\*[achochmounia@gmail.com](mailto:achochmounia@gmail.com)

**Abstract.** Protein assembly is the mechanism of combining two or more protein chains. Living organisms to trigger biological activity often uses this mechanism. The cholera toxin B subunit pentamer (CtxB<sub>5</sub>), the binding moiety of CtxAB<sub>5</sub>, a member of the AB<sub>5</sub> toxin family, is presented here as a model for studying protein assembly. Experimental results showed the importance of histidine residues in CtxB<sub>5</sub> assembly. Nevertheless, the histidines are located outside the pentamer interfaces suggesting an indirect role. Gemini and Spectral-Pro, two network-based models developed in our team, were used in combination with *in silico* mutations produced with Fold X to investigate this question. All the residues of the pentamer interfaces, so-called hotspots, were identified and some appeared to be chemical neighbors of the His 94, making possible a propagation path of conformational changes from the His 94 to the interface to regulate the pentamer assembly. Mutations of the residues along intra-to-inter amino acid paths produce non additive mutational perturbations of the interface energy, indicating influences of His 94 on its hotspot neighbors to regulate the interface. The role of intra-to-inter amino acid paths in regulating the pentamer interface was confirmed by applying similar approach to the heat labile toxin B pentamer (LTB<sub>5</sub>), which share 84% sequence identity, structural and functional similarity with CtxB<sub>5</sub>. Different intra-to-inter amino acid paths were identified for LTB<sub>5</sub> consistently with the different assembly mechanisms followed by both toxin pentamers. These results open up avenues for understanding why these two toxins follow different assembly mechanisms.

**Keywords:** Protein assembly, histidine, network, communication paths, interfaces.

## 1 Introduction

Computational and mathematical modeling of complex physical and biological phenomena is a real challenge for computer scientists and mathematicians. Faced with the increase of biological data, it becomes important to provide models and tools to exploit them and help investigating biological questions.

The function of a large majority of proteins depends on their capacity for self-assembly, either transiently or permanently, into oligomers<sup>1,2</sup>. Understanding protein assembly mechanisms is also important due to the involvement of oligomers in pathologies, from bacterial infection to conformational diseases (e.g. Alzheimer's disease or Parkinson's disease)<sup>3-5</sup>.

Cholera toxin is the most important virulence factor produced by *Vibrio cholerae*. This toxin is composed of two subunits A and B. The B subunit is a pentamer mainly involved in the transport of the A subunit to its target in cells via its binding to GM1 on the cell surface<sup>6,7</sup>. The B subunit can be produced by bacteria, in the absence or presence of the A subunit, which means the pentamer folds and assembles alone<sup>8</sup>.

Heat-labile enterotoxin (LTB<sub>5</sub>) like cholera toxin (CtxB<sub>5</sub>) is a hetero-hexameric AB<sub>5</sub> complex<sup>6</sup>. CtxB<sub>5</sub> and LTB<sub>5</sub> share 82% sequence identity. The crystallographic structures of LTB<sub>5</sub> and CtxB<sub>5</sub> are almost superimposable, circular in shape and each B subunit interacts extensively with its adjacent subunits<sup>9,10</sup>. Therefore, pentamers are very stable, and dissociate only at pH values below 2.0<sup>11,12</sup>. *In vitro* human LTB<sub>5</sub> reassembly is inhibited by pH with a pKa around 7.0 while *in vitro* CtxB<sub>5</sub> reassembly is inhibited at a pKa around 6.0<sup>13,14</sup>. The deprotonation of the LTB<sub>5</sub> N-terminals regulates the folding of the toxin monomer, subsequently allowing association of folded monomers and toxin assembly. The alanine residue in position 1 is one hot spots of the LTB<sub>5</sub>, namely it is at the pentamer interface, in good agreement with the N-terminal involvement in the toxin assembly. For CtxB<sub>5</sub>, deprotonation of histidines is necessary for interface formation and assembly. CtxB<sub>5</sub> has four histidines, His 13, His 18, His 57 and His 94 while human LTB<sub>5</sub> has only His 13 and His 57, which argues on the specific role of His 18 and/or His 94 in CtxB assembly (Fig. 1A). This possibility is supported by the position of His 18 upstream the  $\beta$  strand 23 to 31 and the position of His 94 upstream the  $\beta$  strand 96 to 103, both  $\beta$  strands making up the main toxin interface (Figure 1B).

Nevertheless, none is located at the CtxB<sub>5</sub> interface, which makes their direct involvement in the interface formation and the assembly unlikely and suggests an indirect role through the propagation of large-scale conformational changes from their position to the interface, as in allosteric mechanisms.

Such a propagation mechanism is difficult to determine experimentally because of the use of macroscopic measures when amino acid to amino acid motions are at play. Instead, we propose to use computational tools with the objective to establish a link between histidine residues and the interface of CtxB<sub>5</sub> that would explain the histidine involvement in the pentamer assembly.



effect of mutations on the stability, folding, and dynamics of proteins and nucleic acids<sup>15</sup>. The core functionality of Fold-X is the calculation of the free energy of a macromolecule based on high resolution 3D images of its structure.

### ***Generation of mutant structures***

Mutant structures were generated using Fold-X's <PositionScan> or <BuildModel> position function. During this design procedure, Fold-X tests different rotamers and allows neighboring side chain atoms to move. The program first introduces a mutation in alanine, then mutates in the desired residue (while displacing neighboring residues).

### ***Optimization of models using the Fold-X repair function***

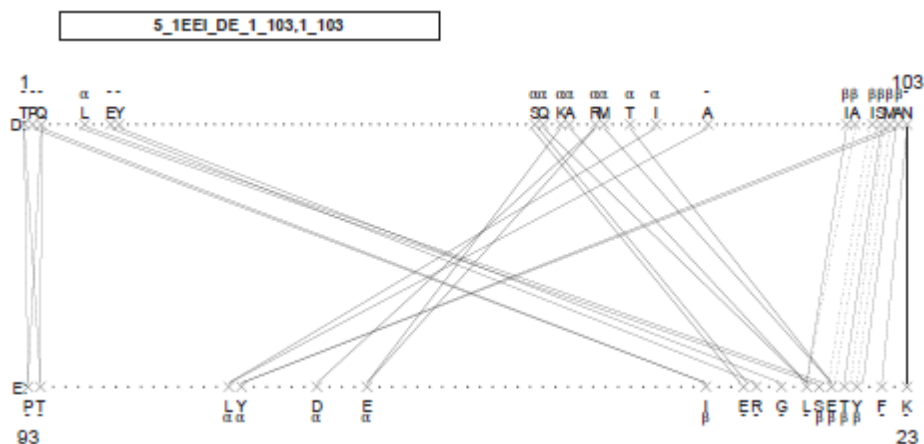
The WT 3D structure of CtxB<sub>5</sub> was taken from the Protein Data Bank (PDB 1EEI), and subjected to an optimization procedure using the Fold-X repair function. During this procedure, Fold-X identifies residues that have low torsion angles, Van Der Waals exposure, or clashing total energies. Fold-X works as follows: first, it mutates the selected position to an alanine residue and annotates the side chain energies of neighboring residues. Then it mutates the alanine to the selected amino acid, and recalculates the side chain energies of the same neighboring residues. These residues which present an energy difference are then mutated by themselves, in order to identify the most favorable energy rotamer. This procedure contains an additional function, where Fold-X quickly eliminates small local clashes, and saves the calculation time.

### ***Energy calculations***

Fold-X aims to describe energetic contributions to protein stability in simple empirical terms that allow easy interpretation by non-specialists. It is therefore suitable for high-throughput structural bioinformatics tasks. It will be used to calculate the energy of the stability of a protein, the energy of the interactions at the interface of an oligomer (eg the pentamer of the cholera toxin) and in wild-type or mutated versions. Energy calculations for mutant proteins were performed with the Fold-X energy function which includes terms considered important for protein stability and the energy of the interactions at the interface. The values of the energies obtained by Fold-X are then converted into more realistic values through the use of a normalization function obtained by fitting the experimental and calculated data.

## **2.2 Gemini**

This program creates a graphical representation of interface regions and their amino acid interactions in the style of graph theory<sup>16</sup>. Here, the vertices of the graph are the amino acids; those which are selected as participants to the interface via weak chemical bonds are symbolized by crosses "X" (Hot spots), the others which do not contribute to an interface are symbolized by dots "." (Figure 2).



**Figure 2.** Example of Gemini graph: Graph shows the hot spots that are in intermolecular interactions in the cholera toxin based on the PDB 1EEI of CtxB<sub>5</sub>.

Gemini proposes a skeleton of the intermolecular interactions between the amino acids of an interface. The details on Gemini can be found here<sup>16</sup>.

### 2.3 Spectral-Pro

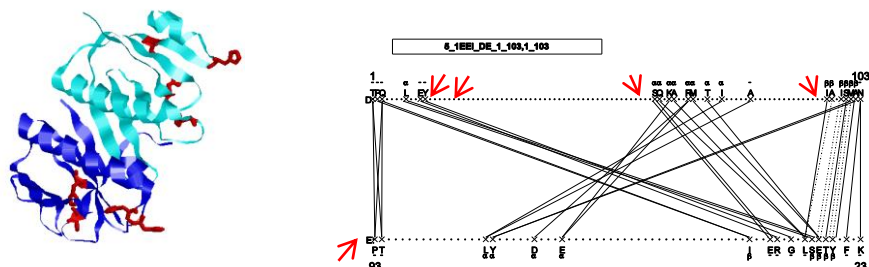
This algorithm, also developed on graph theory, is based on adjacency matrices and the spectral properties of networks to model interfaces in the form of a network of interacting amino acids. The properties of interfaces are inferred from the properties of their networks. Spectral-Pro uses the graph between the atoms of the protein to produce several other graphs. A first graph, which we will call unweighted, connects two amino acids if at least one atom of each amino acid is connected in the initial graph. A second graph, which is called weighted, adds to the preceding graph a weight on each bond representing the exact number of bonds between atoms of each amino acid. The spectral analysis is reproduced on each of these graphs making it possible to group the amino acids of a protein into groups of acids strongly interconnected inside the group and weakly connected with amino acids outside the group. The details of Spectral-Pro can be found here<sup>17</sup>.

## 3 Results

A complete picture of protein assembly, involves intermolecular amino acid interactions (between chains) for the interface formation and intramolecular amino acid interactions (within chains) for the chain folding. The challenge is to understand how both are coordinated<sup>14,18–20</sup>. To address this question, we use approaches and tools derived from computer science and mathematics, interpreted from experimental data on protein assembly. Our case study is the cholera toxin B subunit pentamer (CtxB<sub>5</sub>).

In the laboratory, it was shown experimentally and by simulations that the assembly of the cholera toxin CtxB<sub>5</sub> (1EEI) was inhibited at acidic pH with a pK<sub>a</sub> around 6.0 suggesting the involvement of histidine residues in the regulation of assembly<sup>14</sup>. However, none of the four histidine present in CtxB<sub>5</sub> (His 13, His 18, His 57 and His 94), is located at the interface and accordingly the Gemini program does not identify them as hotspots (Figure 3). This makes their direct involvement in the interface formation and the pentamer assembly unlikely and suggests an indirect role through the propagation of large-scale conformational

changes from their position to the interface. In order to verify that histidines had no direct effect on the interface, the histidine residues were mutated to asparagine and the effect of the mutations was assessed by comparing the interaction energy of the interface calculated with Fold X for the wild type (WT) and the mutants (Table 1).



**Figure 3.** Histidines in CtxB5. **A.** Representation by Rasmol of two chains out of the five of the pentamer, designated by different colors (PDB 1EE1). Histidines are in red. **B.** Gemini graph of the CtxB5 interface. Histidines are not identified as hotspots and their positions are indicated by red arrows.

The Fold-X program is used to generate the mutations and perform the calculations of the interaction energies<sup>15</sup>. The individual histidine mutation has no effect on the interaction energy. Double, triple and quadruple mutations of histidines were also tested and none led to a change in the interaction energy (Table 1).

**Table 1.** Interaction energy of single, double, triple and quadruple mutations of histidines

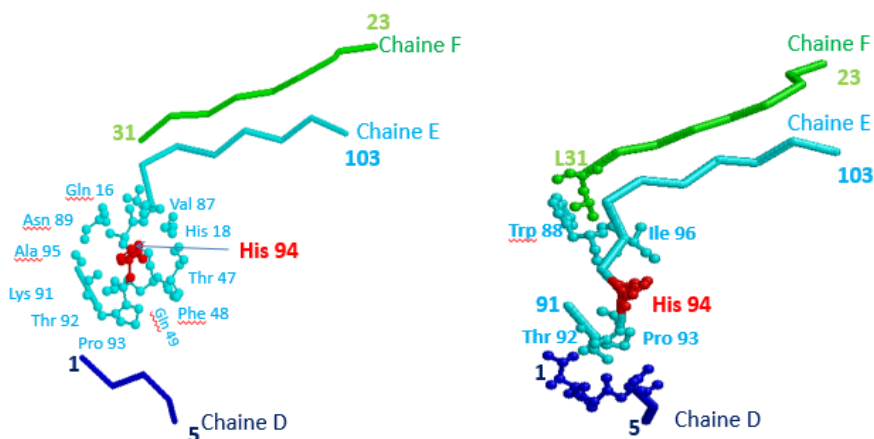
Mutant	Energie d'interaction (Kcal/mol)
WT	-10,9
H13N	-10,9
H18N	-10,8
H57N	-10,2
H94N	-10,8
H13N+H18N	-10,8
H57N+H94N	-10,6
H13N+H57N	-10,5
H13N+H94N	-10,7
H18N+H57N	-10,9
H18N+H94N	-10,7
H13N+H18N+H57N	-10,5
H13N+H57N+H94N	-10,0
H18N+H57N+H94N	-10,6
H13N+H18N+H57N+H94N	-10,5

The individual or combined mutations of the histidines have no effect on the interaction energy of the interface supporting our hypothesis that histidines regulate assembly indirectly. The objective now is to find evidences of an indirect mechanism involving the propagation of atomic motions from the histidine positions to amino acids at the interface. We are looking for the presence of "intermediate" residues close to histidines, i.e. having intramolecular interactions with histidines that are hotspots, i.e. that have intermolecular interactions with adjacent chains.

This hypothesis is based on the possibility of having local perturbations on a histidine

propagating beyond the histidine chemical neighbors to hotspots on adjacent chains. Such a phenomenon is involved in allostery. The notion of network is very relevant to understand the mechanisms of communication from a local scale to a global scale<sup>21,22</sup>.

Half of His 94 chemical intramolecular neighbors interact with amino acids on adjacent chains, i.e. half of His 94 chemical neighbors are hotspots (Table 2). In addition, His 94 hotspot neighbors are involved in two different interface regions located on two different chains. The His 94 on chain E has its neighbors 47, 49, 92 and 93, which interact intermolecularly with residues 1 to 3 on chain D and its neighbors 88 and 96, which interact intermolecularly with residues 23 to 31 on chain F (Figure 4). The residue His 94 is therefore at a strategic position connecting three chains. Moreover, His 94 interacts intramolecularly with His 18, also presents only in CtxB<sub>5</sub>, and located upstream the  $\beta$ -strand 23-31 which participates to the main  $\beta$ -interface with the  $\beta$ -strand 96 to 103 on the adjacent chain<sup>19</sup>. This makes His 94 a key position to coordinate the formation of the N-terminal interface (residues 1 to 3 with residues 88 and 96) and the main  $\beta$ -interface. In comparison, His 13 on chain E has only a fourth of its intramolecular neighbors which are hotspots (residues 11, 12 and 88) interacting intermolecularly with residues 31 to 35 on chain F. His 57 on chain E has a third of its intramolecular neighbors which are hotspots (residues 61, 58, 53 and 65) interacting intermolecularly with residues 31 to 36 on chain F and residue 63 on chain D. Histidine 18 has no intramolecular neighbor hot spots.



**Figure 4.** Neighborhood of histidine 94 (Rasmol). Histidine 94 of the E chain (turquoise color) is indicated in red, these intramolecular neighbors located at a distance of 5 Å are indicated in ball and stick representation. Their intermolecular neighbors are shown in skeleton representation in blue for the interface on the D chain and in green for the interface on the F chain.

LTB<sub>5</sub> also has His 13 and His 57 and hence based on the histidine neighborhoods' analysis, His 94 is the most likely histidine to coordinate CtxB assembly indirectly through intermediate residues 47, 49, 92, 93, 88 and 96. The neighborhood analysis highlights intra-to-inter amino acid interaction paths going from the His 94 to the intermediate residues at the interface and from them to the interface residues on the adjacent chain (Table 2 and Appendix).

To test the role of the intra-to-inter amino acid interaction paths in regulating the toxin pentamer assembly, all the residues along the paths were mutated to asparagine, individually and in combination, using Fold X<sup>15</sup>. The interaction energy of the interface was calculated with Fold-X as well and the effect of the mutations was established by comparing the



interaction energy of the interface of the WT and the mutants. Non-additive effects on the interaction energy suggest dependent positions and support the presence of an interaction path between positions.

**Table 2.** Hotspot intermediate amino acids in intramolecular interaction with histidine 94 (in bold).

<b>His 94</b>	<b>Intramoléculaire</b>	<b>Intermoléculaire</b>
	Gln 16	***
	His 18	***
	<b>Thr 47</b>	Gln 3
	Phe 48	***
	<b>Gln 49</b>	Thr 1
	Val 87	***
	<b>Trp 88</b>	Leu 31, Ala 32
	Asn 89	***
	Lys 91	***
	<b>Thr 92</b>	Thr 1, Gln 3
	<b>Pro 93</b>	Thr 1, Pro 2, Gln 3
	<b>Ile 96</b>	Leu 31

Two examples are given on Table 3 to illustrate on one hand, an intra-to-inter amino acid interaction path identified by non-additive effects of combined mutations (His 94, Gln 49, Thr 1) and on the other a non interaction path, identified by additive effects of combined mutations (His 94, Ile 96 and Leu 31).

**Tables 3.** Single, double and triple mutations of two communication pathways in cholera toxin (CtxB<sub>5</sub>)

<b>CtxB<sub>5</sub></b>		<b>CtxB<sub>5</sub></b>	
<b>Mutants</b>	<b>Energie d'interaction (Kcal/mol)</b>	<b>Mutants</b>	<b>Energie d'interaction (Kcal/mol)</b>
<b>WT</b>	<b>-10,9</b>	<b>WT</b>	<b>-10,9</b>
H94N	-10,9	H94N	-10,9
Q49N	-11,3	I96N	-10,8
H94N+Q49N	-11,4	H94N+I96N	-10,9
T1N	-11,0	L31N	-8,4
H94N+T1N	-11,1	H94N+L31N	-8,8
Q49N+T1N	-13,1	I96N+L31N	-8,8
<b>H94N+Q49N+T1N</b>	<b>-11,7</b>	<b>H94N+I96N+L31N</b>	<b>-8,9</b>

The notion of additivity is a property of measurable quantities used to evaluate mixtures. Interactions or influences between compounds of a mixture are revealed when the effect of a mixture is different from the sum of the effect of the individual compounds, namely when the effects are non-additive. In biology, the influence of amino acids on one another is called epistasis and it occurs when the effect of a mutation changes with the presence of one or more mutation(s) elsewhere.

In our situation, the triple multiple mutations of the His 94, the Ile 96 intermediate hotspot and the Leu 31 hotspot on the adjacent chain (Table 3, right, grey rows) give an energy equal to the energy of the double mutations (without the His 94 mutation), indicating that the His 94 has no influence on the effect of the double mutation. The mutations of the two

hotspots Ile 96 and Leu 31 disturb the energy of interactions of the interface while the mutation of His 94 has not effect on the interface energy. The triple mutation has the same effect than the double showing the independence of the effect of the His 94 mutation and of the double mutations. This indicates that there is no communication between the His 94 and the two positions I96 and L31 and no intra-to-inter amino acid interaction paths from the His 94 to the interface.

On the contrary, the triple multiple mutations of the His 94, the Gln 49 intermediate hotspot and the Thr 1 hotspot on the adjacent chain (Table 4, left, grey rows) has an energy of interactions of the interface different from the energy of interactions of the interface of the double mutations (without the His 94 mutation), indicating that the His 94 mutation influences the effect of the mutations at positions 49 and 1. This establishes an intra-to-inter amino acid interaction path from the His 94 to the intermediate residue 49 and the interface residue 31 on the adjacent chain.

An intra-to-inter amino acid interaction path between His 94 and the main  $\beta$ -interface which involves the  $\beta$ -strand 23 to 31 and the  $\beta$ -strand 96 to 103 on the adjacent chain exists through the intermediate hotspot Trp 88 and Leu 31 as shown by the difference in the energy of interactions of the interface in the triple mutant and the double mutant (without the His 94 mutation) (Appendix and Table 4, grey rows).

**Tables 4.** Single, double and triple mutations of residues H94, L31 and W88 in the communication pathway in cholera toxin (CtxB<sub>5</sub>)

CtxB <sub>5</sub>	
Mutants	Energie d'interaction (Kcal/mol)
H94N	-10,9
W88N	-10,8
L31N	-8,4
H94N+L31N	-8,8
H94N+W88N	-10,8
W88N+L31N	-9,1
H94N+W88N+L31N	-10,8

In conclusion, the results suggest that there are privileged intra-to-inter amino acid interaction paths that allow His 94 to regulate the formation of the N-terminal interface and the main  $\beta$ -interfaces.

Another interesting question is to determine whether the position of residue 94 is strategic or the type at the position matters as well. To answer this question, we could test the mutation of residue 94 by the other 18 amino acid types and do double and triple mutations, but that would be tedious. Instead, we choose to consider the LTB<sub>5</sub> case which has an asparagine residue at position 94. Moreover CtxB<sub>5</sub> and LTB<sub>5</sub> have different assembly mechanisms despite having superimposable 3D structures, so finding different intra-to-inter amino acid interaction paths would provide some understanding on why the two pentamers assemble differently. We applied the same procedure on LTB<sub>5</sub> to determine if the same paths exist and whether the amino acid-to-amino acid influences are due to the position or to the type of residue at this position. The intramolecular neighboring residues of the Asn<sub>94</sub> residue and their intermolecular neighbors are shown in Table 5.

**Table 5.** Intramolecular intermediate residues of Asn 94 (in bold) and their intermolecular neighborhood in the LTB<sub>5</sub> toxin (PDB 1LTR)

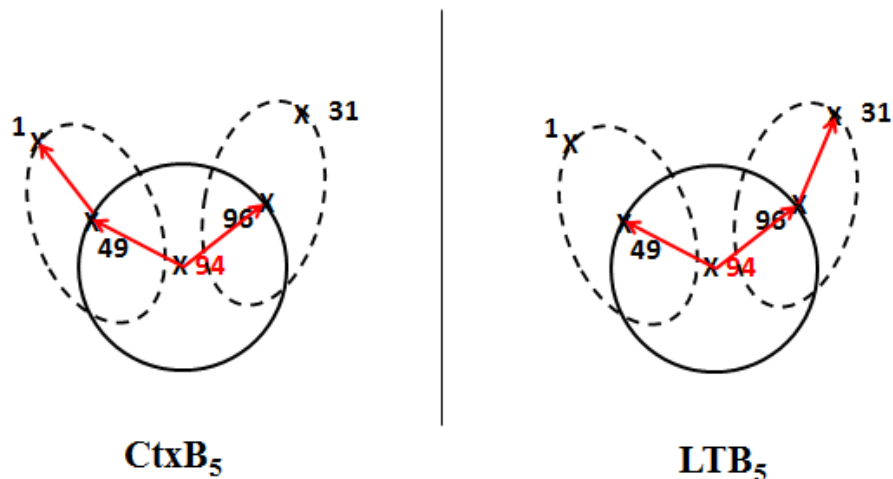
Asn 94	Intramoléculaire	Intermoléculaire
	Tyr 18	***
	<b>Thr 47</b>	Gln 3
	Phe 48	***
	Gln 49	Ala 1
	<b>Trp 88</b>	Met 31, Ala 32
	Asn 89	***
	Asn 90	***
	Lys 91	***
	<b>Thr 92</b>	Ala 1, Gln 3
	<b>Pro 93</b>	Ala, Pro 2, Gln 3
	Ser 95	***
	<b>Ile 96</b>	Met 31

Table 6 shows that an intra-to-inter amino acid interaction path between Asn 94, Gln 49 and Ala 1 exist as well in LTB<sub>5</sub> but the difference in interaction energy between the triple and the double mutants (Table 6, left, grey rows) is smaller than for CtxB<sub>5</sub>, suggesting a less privileged path and lesser influence of position 94 on the position 49 and 1. In contrast to what is observed for CtxB<sub>5</sub>, there is an interaction path between Asn 94, Ile 96 and Met 31 (Table 6, right, grey rows).

The results show that the position of residue 94 alone does not explain the influences between positions and that the type of amino acid at the position 94 is also important. Moreover, the results provide some explanation on why the two toxin pentamers assemble by different mechanisms because they do not involve the same intra-to-inter amino acid paths (Fig. 5). In addition, we notice that the neighbors of the position 94 differs between the two toxin pentamers, for example position 31 is a leucine in CtxB<sub>5</sub> but a methionine in LTB<sub>5</sub>, or position 18 is an histidine in CtxB<sub>5</sub> but a tyrosine in LTB<sub>5</sub>. This indicates that the type of the neighbors also matter to establish influences between residues.

**Tables 6.** Single, double and triple mutations of two communication pathways in LTB<sub>5</sub>

LTB <sub>5</sub>		LTB <sub>5</sub>	
Mutants	Energie d'interaction (Kcal/mol)	Mutants	Energie d'interaction (Kcal/mol)
WT	-6,0	WT	-6,0
N94H	-6,0	N94H	-6,0
Q49N	-5,5	I96N	-5,2
Q49N+N94H	-6,0	N94H+I96N	-6,1
A1N	-7,5	M31N	-3,9
A1N+N94H	-7,3	M31N+N94H	-6,8
Q49N+A1N	-7,1	I96N+M31N	-6,3
<b>Q49N+A1N+N94H</b>	<b>-7,9</b>	<b>N94H+I96N+M31N</b>	<b>-8,4</b>



**Figure 5.** Representation of the difference between the two toxins CtxB<sub>5</sub> and LTB<sub>5</sub> at the level of the communication pathways. The phenomenon of amino acid to amino acid influence is indicated by the red arrows.

## 4 Conclusion

Overall, the results show that the influences between residues that are coordinating protein assembly are not only due to the position of a residue but also to its type at a given position, and to the type of its neighbors. Combining network-based models and *in silico* mutations appear to be a good approach to investigate protein assembly and the impact of mutations on such mechanism.

## Appendix

CtxB<sub>5</sub> communication paths

H94+T47+Q3	
Mutants	intermolecular
H94N	-10,9
T47N	-10,6
Q3N	-11,1
H94N+Q3N	-11,0
H94N+T47N	-10,6
T47N+Q3N	-10,8
H94N+T47N+Q3N	-10,7

H94+W88+L31	
Mutants	intermolecular

H94+W88+A32	
Mutants	intermolecular
H94N	-10,9
W88N	-10,8
A32N	-9,8
H94N+W88N	-10,8
H94N+A32N	-10,0
W88N+A32N	-9,9
H94N+W88N+A32N	-10,2

H94+T92+T1	
Mutants	intermolecular

<b>H94N</b>	-10,9
<b>W88N</b>	-10,8
<b>L31N</b>	-8,4
<b>H94N+L31N</b>	-8,8
<b>H94N+W88N</b>	-10,8
<b>W88N+L31N</b>	-9,1
<b>H94N+W88N+L31N</b>	-10,8

<b>H94N</b>	-10,9
<b>T92N</b>	-11,1
<b>T1N</b>	-11,0
<b>H94N+T92N</b>	-11,7
<b>H94N+T1N</b>	-11,1
<b>T92N+T1N</b>	-13,3
<b>H94N+T92N+T1N</b>	-14,0

<b>H94+P93+P2</b>	
<b>Mutants</b>	<b>intermolecular</b>
<b>H94N</b>	-10,9
<b>P93N</b>	-8,3
<b>P2N</b>	-11,8
<b>H94N+P2N</b>	-12,4
<b>H94N+P93N</b>	-9,0
<b>P93N+P2N</b>	-10,0
<b>H94N+P93N+P2N</b>	-10,2

<b>H94+T92+Q3</b>	
<b>Mutants</b>	<b>intermolecular</b>
<b>H94N</b>	-10,9
<b>T92N</b>	-11,1
<b>Q3N</b>	-11,1
<b>H94N+Q3N</b>	-11,0
<b>H94N+T92N</b>	-11,7
<b>T92N+Q3N</b>	-11,0
<b>H94N+T92N+Q3N</b>	-11,1

<b>H94+Q3+P93</b>	
<b>Mutants</b>	<b>intermolecular</b>
<b>H94N</b>	-10,9
<b>Q3N</b>	-11,1
<b>P93N</b>	-8,3
<b>H94N+Q3N</b>	-11,0
<b>H94N+P93N</b>	-9,0
<b>P93N+Q3N</b>	-8,3
<b>H94N+P93N+Q3N</b>	-8,4

<b>H94+P93+T1</b>	
<b>Mutants</b>	<b>intermolecular</b>
<b>H94N</b>	-10,9
<b>P93N</b>	-8,3
<b>T1N</b>	-11,0
<b>H94N+T1N</b>	-11,1
<b>H94N+P93N</b>	-9,0
<b>P93N+T1N</b>	-10,0
<b>H94N+P93N+T1N</b>	-13,7

## References

- 1 Y. Ofraim and B. Rost, Protein–protein interaction hotspots carved into sequences, *PLoS computational biology*, 2007, **3**, e119.
- 2 S. J. de Vries and A. M. Bonvin, How proteins get in touch: interface prediction in the study of biomolecular complexes, *Curr Protein Pept Sci*, 2008, **9**, 394–406.
- 3 M. D. Kirkitadze and A. Kowalska, Molecular mechanisms initiating amyloid beta-fibril formation in Alzheimer’s disease, *Acta Biochim Pol*, 2005, **52**, 417–23.
- 4 C. Soto and S. Pritzkow, Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases, *Nature neuroscience*, 2018, **21**, 1332–1340.
- 5 M. D. Kirkitadze, G. Bitan and D. B. Teplow, Paradigm shifts in Alzheimer’s disease and other neurodegenerative disorders: the emerging role of oligomeric assemblies, *J Neurosci Res*, 2002, **69**, 567–77.

- 6 T. R. Hirst, *Biogenesis of Cholera toxin and Related oligomeric Enterotoxins*, M. Dekker, New York, 1995, vol. 8.
- 7 Y. Feng, A. P. Jadhav, C. Rodighiero, Y. Fujinaga, T. Kirchhausen and W. I. Lencer, Retrograde transport of cholera toxin from the plasma membrane to the endoplasmic reticulum requires the trans-Golgi network but not the Golgi apparatus in Exo2-treated cells, *EMBO Rep*, 2004, **5**, 596–601.
- 8 T. R. Hirst, S. J. Hardy and L. L. Randall, Assembly in vivo of enterotoxin from *Escherichia coli*: formation of the B subunit oligomer, *J Bacteriol*, 1983, **153**, 21–6.
- 9 E. A. Merritt, S. Sarfaty, F. van den Akker, C. L’Hoir, J. A. Martial and W. G. Hol, Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide, *Protein Sci*, **3**, 166–75.
- 10 T. K. Sixma, S. E. Pronk, K. H. Kalk, E. S. Wartna, B. A. van Zanten, B. Witholt and W. G. Hol, Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli*, *Nature*, 1991, **351**, 371–7.
- 11 M. J. De Wolf, G. A. Van Dessel, A. R. Lagrou, H. J. Hilderson and W. S. Dierick, pH-induced transitions in cholera toxin conformation: a fluorescence study, *Biochemistry*, 1987, **26**, 3799–806.
- 12 L. W. Ruddock, H. M. Webb, S. P. Ruston, C. Cheesman, R. B. Freedman and T. R. Hirst, A pH-dependent conformational change in the B-subunit pentamer of *Escherichia coli* heat-labile enterotoxin: structural basis and possible functional role for a conserved feature of the AB5 toxin family, *Biochemistry*, 1996, **35**, 16069–76.
- 13 L. W. Ruddock, J. J. Coen, C. Cheesman, R. B. Freedman and T. R. Hirst, Assembly of the B subunit pentamer of *Escherichia coli* heat-labile enterotoxin. Kinetics and molecular basis of rate-limiting steps in vitro, *J Biol Chem*, 1996, **271**, 19118–23.
- 14 J. Zrimi, A. Ng Ling, E. Giri-Rachman Arifin, G. Feverati and C. Lesieur, Cholera toxin B subunits assemble into pentamers - proposition of a fly-casting mechanism, *PLoS One*, 2010, **5**, e15347.
- 15 J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau and L. Serrano, The FoldX web server: an online force field, *Nucleic Acids Research*, 2005, **33**, W382–W388.
- 16 G. Feverati and C. Lesieur, Oligomeric interfaces under the lens: gemini, *PloS one*, 2010, **5**, e9897.
- 17 M. Achoch, R. Dorantes-Gilardi, C. Wymant, G. Feverati, K. Salamatian, L. Vuillon and C. Lesieur, Protein structural robustness to mutations: an in silico investigation, *Physical Chemistry Chemical Physics*, 2016, **18**, 13770–13780.
- 18 K. Luke, M. Perham and P. Wittung-Stafshede, Kinetic folding and assembly mechanisms differ for two homologous heptamers, *J Mol Biol*, 2006, **363**, 729–42.
- 19 C. Lesieur, M. J. Cliff, R. Carter, R. F. James, A. R. Clarke and T. R. Hirst, A kinetic model of intermediate formation during assembly of cholera toxin B-subunit pentamers, *J Biol Chem*, 2002, **277**, 16697–704.
- 20 B. A. Shoemaker, J. J. Portman and P. G. Wolynes, Speeding molecular recognition by using the folding funnel: the fly-casting mechanism, *Proc Natl Acad Sci U S A*, 2000, **97**, 8868–73.
- 21 L. Vuillon and C. Lesieur, From local to global changes in proteins: a network view, *Current opinion in structural biology*, 2015, **31**, 1–8.

22 L. Di Paola and A. Giuliani, Protein contact network topology: a natural language for allostery, *Current opinion in structural biology*, 2015, **31**, 43–48.