Studies of TomB/YmoB family of proteins: biofilm regulators in enteric bacteria

Oriol Marimon Garrido

PhD thesis

Studies of TomB/YmoB family of proteins: biofilm regulators in enteric bacteria

Oriol Marimon Garrido

Department of Organic Chemistry

Faculty of Chemistry

Universitat de Barcelona

2013

Memòria presentada per

Oriol Marimon Garrido

Per obtenir el grau de doctor per la Universitat de Barcelona

Revisat per:

Director de Tesi
Prof. Miquel Pons Vallès
Universitat de Barcelona

Codirector de la Tesi
Jesús García Arroyo
Institut de Recerca Biomèdica

Programa de Química Orgànica

2008-2010

1931. Charlie Chaplin to Albert Einstein:

"The people are applauding you because none of them understands you and applauding me because everybody understands me."

ACKNOWLEDGEMENTS

I wander to acknowledge my thesis director Professor Miquel Pons for his support, his toughness, his scientific sense and his fantastic ability to find solutions, sometimes daring and ambitious, to the problems that have arisen during the course of this thesis. Also for his enormous capacity for teaching, always in the easiest way, the intricate and complex concepts related with the NMR, structural biology and related issues. And finally, because he has always been available when I needed. I am also thankful to my thesis codirector Jesús Garcia.

To Tiago Cordeiro and Irene Amata my utmost gratitude. Their address, advice, and the long hours of interpretation and discussion of results, as well as designing experiments, have helped me to train as a scientist and as a person. Thank you very much.

I would like to thank our collaborators, Tomas K. Wood in the Penn State University, and Maxim Maicel and Vladislav Orekhov in the Swedish NMR Centre, Goteborg University, for their help and constant support given.

I thank the support of the NMR LRB (Laboratori de RMN de Barcelona) of the CCIT (Centres Científics I Tecnològics), which belongs to the ICTS (Institución Científico Tecnológica Singular). To the IRB and PCB for the good working conditions, and to the Mass Spectrometry Core Facility at the IRB Barcelona for support with the MS experiments. Particulary to Dr Margarida Gairí and Dr. M. Vilaseca for their time and help.

To the University of Barcelona and to the IRB of Barcelona. Also to the Spanish MICIN (Ministerio de Ciéncia e INnovación) and Fundació Bosch i Gimpera for the financial support.

I finalment a tots els companys i companyes de laboratori, a la família i sobretot, de la manera més especial, a la Helena González, per voler compartir amb mi aquest estrany camí que és la vida.

DECLARATIONS

Good Use Right

"It is strictly prohibited to use, to investigate or to develop, in a direct or indirect way, any of the scientific contributions of the author contained in this work by any army or armed group in the world, for military purposes and for any other use which is against human rights or the environment, unless a written consent of all the persons in the world is obtained"

Clàusula del bon ús

"Queda absolutament prohibida la utilització, investigació y desenvolupament, de manera directa o indirecta, de qualsevol de les aportacions científiques pròpies de l'autor, que es presenten en aquesta memòria, per part de qualsevol exèrcit del món, o per part de qualsevol grup armat, per qualsevol grup militar y i per a qualsevol alter ús que atempti contra els drets humans o contra el medi ambient, excepte permís escrit de totes les persones del món."

Cláusula de buen uso

"Queda absolutamente prohibida la utilización, investigación y desarrollo, de forma directa o indirecta, de cualquiera de las aportaciones científicas propias del autor, que se presentan en esta memoria, por parte de cualquier ejército del mundo o por parte de cualquier grupo armado, para cualquier uso militar y para cualquier otro uso que atente contra los derechos humanos o contra el medio ambiente, salvo permiso por escrito de todas las personas del mundo."

INDEX

Chapter I.

INTRODUCTION & OBJECTIVES

I. 1 BIOFILMS: sessile bacterial communities.

Biofilms are communities of microorganisms attached to a surface by a hydrated Extracellular Polymeric Substance (EPS), which is primarily composed of proteins, nucleic acids and polysaccharides[1]. The ability to form these sessile communities is an ancient and integral characteristic of prokaryotes. In the context of evolution and adaptation, biofilms provided homeostasis in the face of the fluctuating and harsh conditions of the primitive earth (extreme temperatures, pH and exposure to ultraviolet light), thereby facilitating the development of complex interactions between individual cells and providing an optimal environment for the development of signaling pathways, quorum sensing and chemotactic motility[2].

Furthermore, biofilms are observed literally everywhere in nature. In principle they can be found on any solid-liquid interface, as on rocks and pebbles at the bottom of most streams or rivers, on organic surfaces such as plant and body tissues and on inorganic manmade materials like intravenous catheters or the interiors of water pipes. Biofilms grow even in the most inhospitable environments on earth, as the hot, acidic pools at Yellowstone National Park or on glaciers in Antarctica.

Intriguingly, the biofilms growing patterns in diverse environments are strikingly similar, pointing to important convergent survival strategies that are conferred in part by structural specialization. Different types of biofilm structures have been identified according to the environment where biofilms are growing, such as streamers (filamentous biofilm that form in flowing water) or stromatolites ('mushroom' shape biofilm formed in quiescent or low-shear environments).

Within these structures bacteria grow forming micro colonies encased in the EPS and interspersed with open water channels. These channels facilitate efficient nutrient uptake by infusing fluid from the bulk phase into the biofilm, thereby optimizing nutrient and waste-product exchange[3].

Biofilms formation takes places following a series of discrete and well-regulated steps. While molecular mechanisms differ from species to species, the broad stages of biofilm development appear to be conserved among a wide range of genera. These stages generally occur in three phases: adherence, proliferation and dispersion[4] (Fig 1).

Biofilm cycle depends on environmental cues and requires the coordinate expression and simultaneous regulation of many genes. Previous studies done on aerobic or facultative pathogens have shown differential gene expression and protein profiles for biofilm cells compared to planktonic cells, and the results suggest that biofilm formation is a complex and highly regulated process[5]. For example, in vitro comparative transcriptome analysis carried on *Streptococcus* mutans, which are strongly associated with the pathogenesis of

dental diseases, identified using DNA-microarray analyses 12% of genes that are differentially expressed in biofilm compared with free-living cells[6].

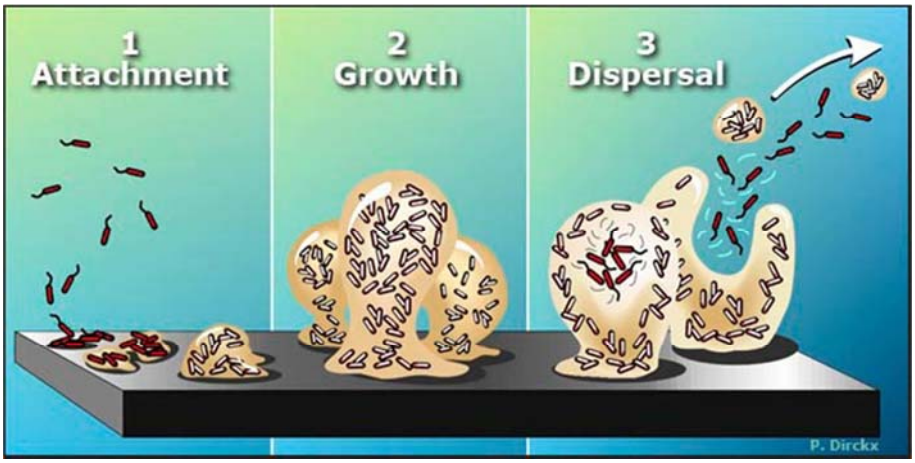


Fig 1: The biofilm cell cycle. **1:** Planktonic bacteria encounter a submerged surface and within minutes can become attached. They begin to produce slimy EPS and to colonize the surface. **2:** EPS production allows the emerging biofilm community to develop a complex, three-dimensional structure that is influenced by a variety of environmental factors. Biofilm communities can develop within hours. **3:** Biofilms can propagate through detachment allowing bacteria to attach to a surface or to a biofilm downstream of the original community. Figure by P. Dirckx, Center for Biofilm Engineering, Montana State University, Bozeman.

Microbial biofilms affect world economy at the level of billions of dollars with regard to equipment damage, product contamination, energy losses and infections[7] (up to 80% of human bacterial chronic inflammatory and infectious diseases are associated to biofilm formation[8]). As conventional methods that would otherwise lead to eradication of planktonic microbes are often ineffective to the microbial populations inside the biofilms, understanding the different mechanisms of biofilm regulation to develop novel strategies to control biofilms becomes of primary interest[9].

I. 2 BIOFILM REGULATION: the Toxin-Antitoxin systems.

Toxin–Antitoxin (TA) systems are small genetic elements found on plasmids or chromosomes of countless bacteria, archaea, and possibly also unicellular fungi[10]. Under normal growth conditions, the stable toxin protein would cause cell death or cell arrest but the labile antitoxin protein or RNA preclude toxicity[11]. Under a variety of stress conditions[12,13], the antitoxin is either proteolysed or its transcription is changed, allowing the toxin to disrupt some essential cellular process[14].

TA systems were initially related to plasmid addiction systems due to their ability to stabilize plasmids by post-segregational killing[15]. But further studies demonstrate that TA systems are also ubiquitous as chromosomal elements. In 126 prokaryotic genomes (16

archaea and 110 bacteria) searched in 2005, 671 TA loci were identified[16]. These number increases as in an analysis of 360 bacterial and 28 archaeal genomes (cumulatively, more than 1.5 million genes) done in 2013, the number of TA systems increased to 400 pairs, organized into 24 families[17].

At present, these 24 families of TA systems are classified in five types. All of them comprise a toxin protein and an antitoxin that can be either a small noncoding RNA (sRNA) in type I and III or a low molecular weight protein in types II, IV, and V[14]. In type I the toxin is usually a hydrophobic peptide whose mRNA transcript is intercepted by the RNA antitoxin. In type III the RNA antitoxin consists of 5.5 repeats, which is processed and inhibits toxicity by directly binding the toxic protein. In type II toxin and antitoxin are proteins, and the toxicity is prevented by formation of a TA complex. In type IV the toxin inhibits a target molecule, whereas the antitoxin counteracts these effects. In type V the mRNA of the small toxin-encoding ORF is cleaved by the antitoxin[10].

The most prevalent kinds of TA systems are type II, IV and V systems, where both toxin and antitoxin are proteins[17]. The two genes, which reside on the same operon, code for small proteins, and inhibition of the toxin is carried out by antitoxin protein. As a rule, the toxin is a stable protein, and the antitoxin is unstable and is degraded rapidly[18]. Although these five types of TA systems have been described, to date the extent and the roles of such systems in bacterial genomes are far from fully understood.

The prevalent existence of TA systems on chromosomes suggests that TA modules play pivotal roles in prokaryotic cellular biology, including programmed cell death[19], stress response[20], generation of persister cells[21], biofilm formation[22], and phage defense via abortive infection[23].

A clear evidence of the role of TA systems in biofilm formation was obtained by studying a strain that had five of the most-studied TA systems deleted (MazF/MazE, RelE/RelB, YoeB/YefM, YafQ/DinJ and ChpB). Upon deletion of these five TA systems, biofilm formation decreased after 8h and increased after 24h in rich medium at 37°C due to a reduction in dispersal[22,24].

Biofilm formation and external stresses are related. In some cases, under stress conditions bacteria activate the General Stress Response (GSR), leading to a change in the gene regulation pattern which switches bacteria from planktonic state to biofilm state[25]. The GSR has been directly implicated in chronic infections due to biofilm formation. Major reasons for the reduced susceptibility of biofilms to antibiotics are the increased dormancy cells and the dramatically reduced growth rates produced in biofilm systems under external stress[26].

There is evidence that TA systems mediate the GSR, help direct cells toward the formation of biofilm and enhance biofilm resistance by increasing persister cells within biofilm[21]. Persistence, in which bacteria enter a dormant, metabolically silent state, is a major obstacle in effective antibiotic treatment against biofilm. It is speculated that different TA systems allow the cell to respond to a specific stress or group of stresses in a highly regulated, elegant fashion[21].

One of the most studied TA systems, which is related to GSR, biofilm formation and persistence, is the MqsR/MqsA pair of *E. coli*[27]. It was the first TA system related to biofilms, since mqsR was identified as differentially regulated in biofilms than in planktonic cells[28], and *mqs*RA deletion reduces biofilm formation[29]. The three-dimensional structure of MqsR:MqsA complex shows that toxin MqsR is an RNAse and that antitoxin MqsA binds DNA via its helix-turn-helix motif in its C-terminal domain while binding MqsR at its N-terminal domain[30].

MqsR 3D structure together with the study performed by Kim et al[31] confirmed that MqsA affected cell dormancy through its DNA-binding motif by regulating other cellular systems, including other toxins. Further studies demonstrate that MqsA blocks the GSR by repression of rpoS[32] (RpoS is a positive regulator of catalase activity via katG and katE[33]) leading to a 850-fold reduction in oxidative stress resistance via repression of catalase activity.

Therefore, MqsA is the first antitoxin shown to regulate more than its own synthesis; this creates a new paradigm where antitoxins of TA systems may be viewed as regulators.

On the other hand toxin MqsR is induced in biofilms[34] and mediates the ability of autoinducer 2 to increase *E. coli* biofilm regulating motility and quorum sensing that is directly associated with biofilm development[27]. MqsR also influences persister formation since deletion of mqsRA reduces persister cell formation and MqsR overexpression increases persister cell formation. Hha and CspD are also related to persister cell formation with MqsR. Hha and CspD increase the production of persister cells and deletion of hha and cspD decreases persistence significantly by diminishing the ability of MqsR to cause persistence. Also, hha and cspD are induced by stress and by the production of MqsR[35]. The mechanism of persister production via MqsR is shown in Fig 2.

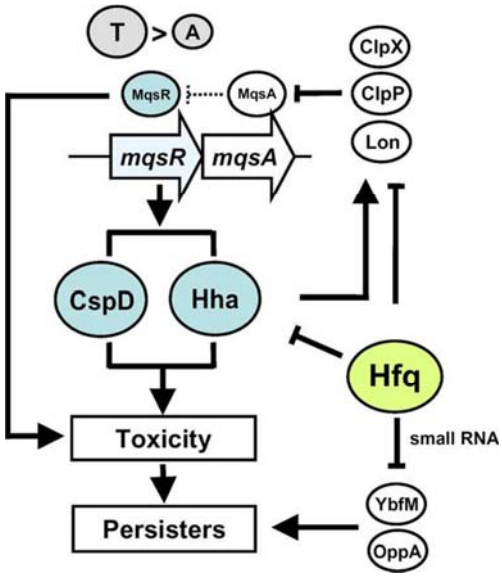


Fig 2: Schematic representation of the mechanism of persister production via MqsR. Hha and CspD are directly involved in persister formation via MqsR, and Hfq negatively regulates transport and outer membrane proteins. → indicates induction, and ⊥ indicates repression. Figure extracted from Kim et al.[35].

I. 3 Hha/TomB FAMILY OF PROTEINS: biofilm regulators as TA system?

Hha is a global transcriptional regulator of *Escherichia coli*[36], and a putative toxin with TomB as a proposed antitoxin[37]. Hha (High hemolysin activity, 72 aa) belongs to the Hha-YmoA family of low-molecular mass proteins (about 8 kDa) that regulate many genes in Gram-negative planktonic bacteria[36]; Hha was identified as reducing hemolysin production by repressing the *hly* operon of *Escherichia coli*[38]. However, Hha does not bind specific DNA sequences; instead, Hha has a protein partner, H-NS, and the Hha-HNS complex binds specific sequences of the *hly* regulatory operon[39] and affected the expression of at least 162 genes controlled by the Hha-H-NS complex[40]. Hha also represses the pathogenicity Locus of Enterocyte Effacement (LEE) set of operons in enterohemorrhagic *E. coli* by repressing transcription of *ler,* which encodes the activator of LEE. Hence, Hha is a global transcriptional regulator in modulating cell physiology[41].

Contreras et al. demonstrated that Hha expression is toxic for the cell, while TomB diminishes although does not counteract its toxicity (Fig 3)[42]. These data suggest that Hha and TomB may form a toxin-antitoxin module in planktonic cells with Hha the toxin (Fig 3 A) and the poorly-characterized TomB the antitoxin (Fig 3 B). This hypothesis is supported by the following observations, which are characteristics of a toxin-antitoxin pair[43]; a) hha and tomB are in the same operon, b) the separation between the genes is short (25bp), c) both proteins are small, d) Hha exerts negative transcriptional self-regulation, and e) overexpression of Hha is toxic.

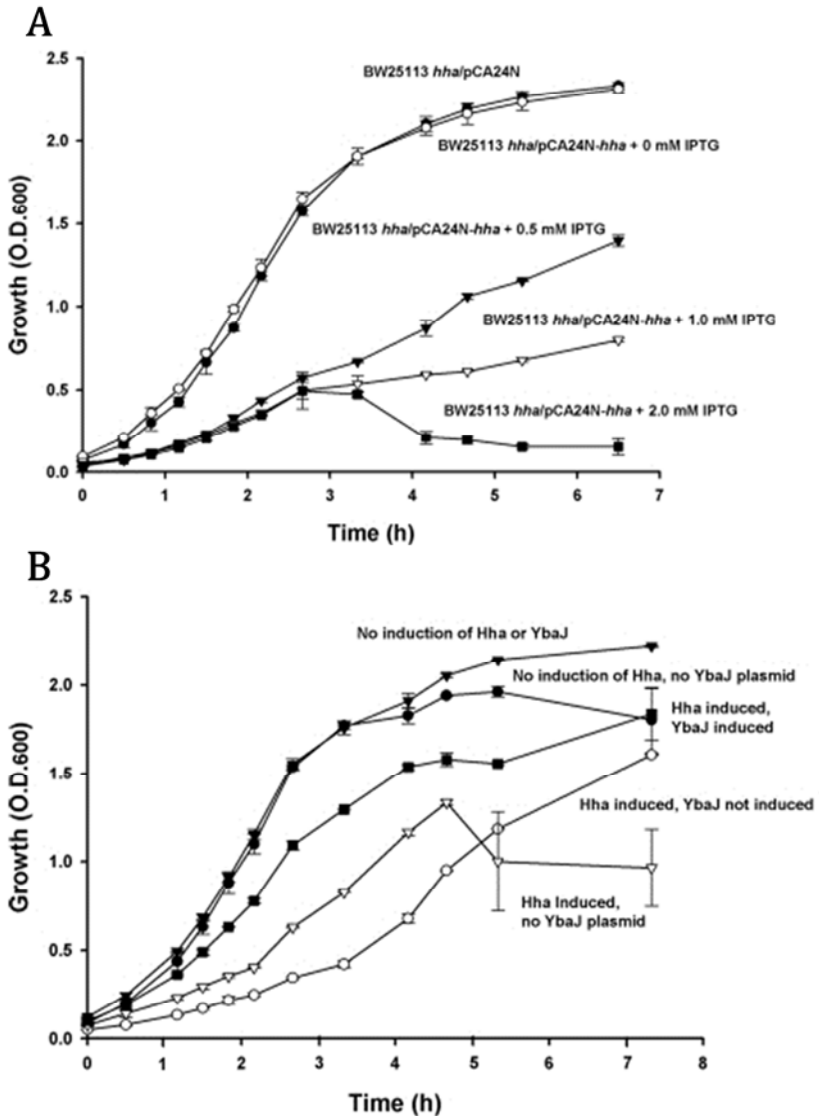


Fig 3: A: Growth of BW25113 hha/pCA24N-hha at 37°C in LB. Hha was induced by the addition of increasing concentrations of IPTG **B:** Effect of TomB overexpression from pBAD-TomB in *E. coli* K-12 BW25113 hha for Hha mediated toxicity from pVLT31-hha in LB at 37°C; Hha was induced by the addition of 2mM IPTG and TomB was induced by the addition of 0.1% L-arabinose. The experiments were performed in triplicate, and one standard deviation is shown. Figure from Garcia-Contreras et al.[42].

Hha is induced 30 fold in *E. coli* biofilms relative to planktonic cells[44] and overexpression of Hha decreases total biofilm formation (Fig 4). Garcia-Contreras et al. demonstrated that Hha influences biofilm formation by repressing the transcription of rare codon tRNAs[42].

Hha is toxic and leads to cell lysis and biofilm dispersal due to activation of prophage lytic genes *appY* and *rzpD* of DLP12 and *alpA* and *yfjZ* of CP4-57, and due to the induction of protease ClpXP[42]. In addition, Hha induces excision of prophages CP4-57 and DLP-12 of *E. coli*[45].

Hha also plays an important role in biofilms by repressing transcription of type I fimbrial genes. A critical step in transitioning from a planktonic state to biofilm formation is

cessation or reduction in the expression of the flagellar system in order to reduce bacterial motility and to promote expression of adhesive fimbriae, such as curli fimbriae, and other extracellular adhesins that allow adherence of bacterial cells to solid surfaces[46]. Hha represses biofilm formation in enterohemorrhagic *Escherichia coli* O157:H7 by induction of flagellar and repression of curli gene expression. *In vitro* gel shift analysis revealed positive interactions of Hha with the promoter sequences of *flhDC* and *csgDEFG*, two global regulators of the flagellar and curli gene regulons, respectively[47].

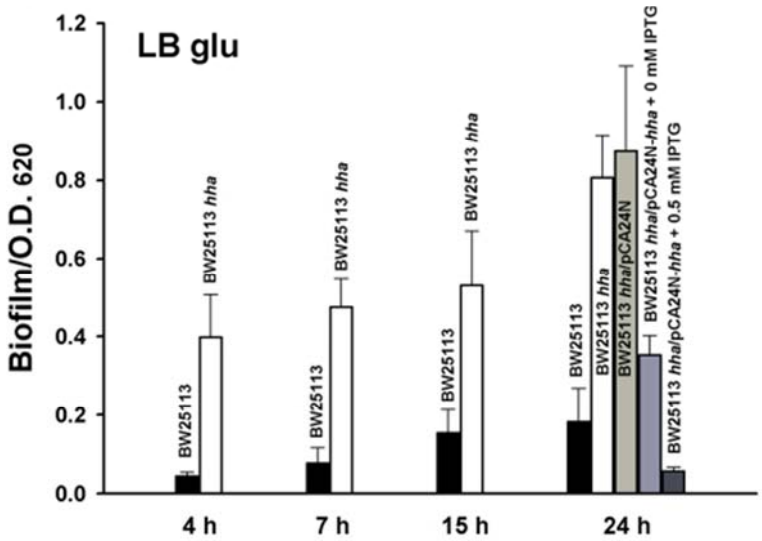


Fig 4: Biofilm formation normalized by cell growth shown after 4 h, 7 h, 15 h, and 24 h at 37°C for the wild-type BW25113 and BW25113 hha in LB glu medium. Each experiment was performed in triplicate, and one standard deviation is shown. Figure from Garcia-Contreras et al.[42].

The final stage of the biofilm cycle is the dispersal of cells from the biofilm into the environment[48]. Like other stages of biofilm development, biofilm dispersal is a highly-regulated process involving many sensory circuits[49]. To date, little is known about the intracellular molecular mechanisms of bacterial biofilm dispersal. Cell death in biofilms is an important mechanism of dispersal since it leads to the creation of voids inside the biofilm which facilitates dispersal of the surviving cells[50]. Hence, programmed cell death is a social activity that in biofilms is mediated by prophage and autolysis proteins[51]. Since in the EPS biofilm matrix structural proteins are present, protease activity is necessary for biofilm dispersal. Hha disperses *E. coli* biofilms by inducing protease ClpXP that activates toxins by degrading antitoxins[42]. Hence, induction of protease activity is a key mechanism to trigger biofilm dispersal by causing cell death and so by disrupting biofilm structure[41].

In contrast to the Hha, which has been extensively studied in planktonic systems and in biofilm systems, the Hha-TomB pair and the protein TomB have been poorly characterized. TomB shows no homology sequence with any other family of proteins, and

no 3D structure of any TomB family members is available. Gonzalez Barrios and coworkers demonstrated that Hha/TomB promotes conjugation, which increase biofilm formation by favoring cell aggregation. When they deleted hha/tomB genes, they observed less conjugation, less aggregation, more motility, and less biofilm[52].

Also Garcia-Contreras and coworkers published that TomB may bind to 21 genes and two intergenic regions. Of these putative targets, the more interesting include the TomB gene itself, indicating that TomB may regulate hha and its own expression. Also, TomB binds sfmH, which encodes a predicted Salmonella fimbrial-like adhesion protein that is only one gene away from fimZ and argU (Hha binds these sites), which indicates Hha and TomB bind nearby loci and may compete for the regulation of genes near the binding sites. The most recent model that summarizes the Hha/TomB E. coli biofilm and cell death regulation is shown in Fig 5.

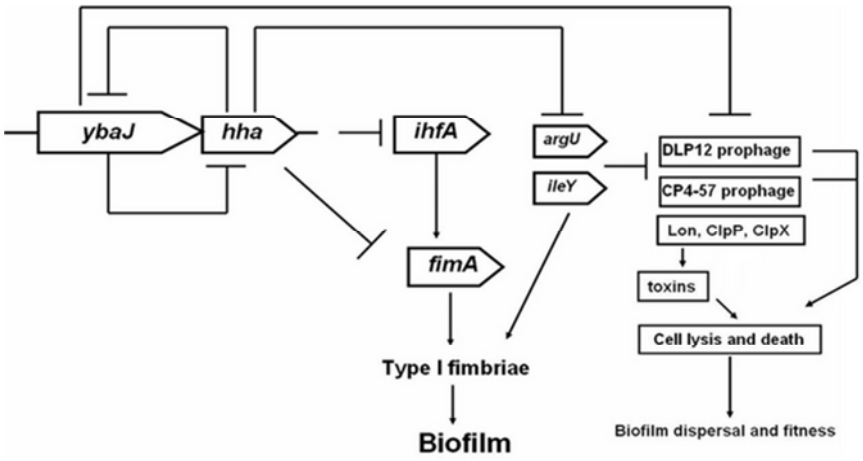


Fig 5: Model for Hha regulation of *E. coli* biofilm and cell death. → indicates induction, and ⊥ indicates repression. tomB is shown as ybaJ, its previous name. Figure from Garcia-Contreras et al.[42].

Hence, Hha and TomB may form in biofilm a TA pair similar to MqsR/MqsA, where the antitoxin regulates more than their own synthesis. Understanding the different mechanisms of biofilm regulation and in particular clarifying the biological activity of TomB family of proteins as antitoxins of Hha in biofilms becomes of primary interest. The aim of this thesis is to study the TomB family of proteins, being our first goal the elucidation of the 3D structure of a TomB family member using Nuclear Magnetic Resonance (NMR).

We hypothesized that the analysis of the 3D structure of a member of TomB family of proteins could shed light on their function and regulation in biofilm system. Indeed, in this work we describe the singular properties of YmoB, the *Yersinia* homologue of TomB, and their potential functional relevance. The richness of TomB/YmoB reactivity associated to

its peculiar structural features made the apparently trivial task of determining the 3D structure of a small soluble protein a major endeavor. In Chapter II.1 we describe our effort to identify a construct suitable for structural studies, in Chapter II.2 we show that the construct is biologically functional, in Chapter II.3 we report the use of novel acquisition methods to solve the 3D structure of our suitable construct and finally, in Chapter II.4 we used the elucidated 3D structure to drive the discovery of the functional properties of YmoB.

I. 4 OBJECTIVES

The three main objectives of this PhD thesis project are:

- The determination of the 3D structure in solution of a member of TomB family of proteins
- The confirmation of the role of YmoB as an antitoxin capable of complement the antitoxin role of TomB in *E. coli.*
- The use of the structural information to identify chemical transformations potentially underlying the transition between active and inactive forms of the TomB/YmoB antitoxins.

I. 5 BIBLIOGRAPHY

[1] Sauer, K., Rickard, A. H., & Davies, D. G. (2007). Biofilms and Biocomplexity. *Microbe, 2*(7), 347–353.

[2] Hall-Stoodley, L., Costerton, J. W., & Stoodley, P. (2004). Bacterial biofilms: from the natural environment to infectious diseases. *Nature reviews. Microbiology*, *2*(2), 95–108.

[3] Wilking, J. N., Zaburdaev, V., Volder, M. De, Losick, R., & Brenner, M. P. (2012). Liquid transport facilitated by channels in Bacillus subtilis biofilms. *PNAS, 110(3),* 848–852.

[4] Toole, G. O., Kaplan, H. B., & Kolter, R. (2000). Biofilm Formation As Microbial Development. *Annu. Rev. Microbiol. 54,* 49–79.

[5] Clark, M. E., He, Z., Redding, A. M., Joachimiak, M. P., Keasling, J. D., Zhou, J. Z., Arkin, A. P., et al. (2012). Transcriptomic and proteomic analyses of Desulfovibrio vulgaris biofilms: carbon and energy flow contribute to the distinct biofilm growth state. *BMC genomics*, *13*(1), 138.

[6] Shemesh, M., Tam, A., & Steinberg, D. (2007). Differential gene expression profiling of Streptococcus mutans cultured under biofilm and planktonic conditions. *Microbiology (Reading, England)*, *153*(Pt 5), 1307–17.

[7] Yang, L., Liu, Y., Wu, H., Song, Z., Høiby, N., Molin, S., & Givskov, M. (2012). Combating biofilms. *FEMS immunology and medical microbiology*, *65*(2), 146–57.

[8] Thomas K. Wood, Seok Hoon Hong, & Qun Ma (2012). Engineering biofilm formation and dispersal *Trends Biotechnol*, *29*(2), 87–94.

[9] Römling, U., & Balsalobre, C. (2012). Biofilm infections, their resilience to therapy and innovative treatment strategies. *Journal of internal medicine*, *272*(6), 541–61.

[10] Ghafourian, S., Raftari, M., Sadeghifard, N., & Sekawi, Z. (2013). Toxin-antitoxin Systems: Classification, Biological Function and Application in Biotechnology. *Current issues in molecular biology*, *16*(1), 9–14.

[11] Gerdes, K., Christensen, S. K., & Løbner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nature reviews. Microbiology*, *3*(5), 371–82.

[12] Amato, S. M., Orman, M. a, & Brynildsen, M. P. (2013). Metabolic Control of Persister Formation in Escherichia coli. *Molecular cell*, *50*(4), 475–487.

[13] Wang, X., Kim, Y., Hong, S. H., Ma, Q., Brown, B. L., Pu, M., Tarone, A. M., et al. (2011). Antitoxin MqsA Helps Mediate the Bacterial General Stress Response. *Nat Chem Biol. 7*(6), 359–366.

[14] Schuster, C. F., & Bertram, R. (2013). Toxin-antitoxin systems are ubiquitous and versatile modulators of prokaryotic cell fate. *FEMS microbiology letters*, *340*(2), 73–85.

[15] Ogura, T., & Hiraga, S. (1983). Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(15), 4784–8.

[16] Pandey, D. P., & Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic acids research*, *33*(3), 966–76.

[17] Sberro, H., Leavitt, A., Kiro, R., Koh, E., Peleg, Y., Qimron, U., & Sorek, R. (2013). Discovery of functional toxin/antitoxin systems in bacteria by shotgun cloning. *Molecular cell*, *50*(1), 136–48.

[18] Masuda, H., Tan, Q., Awano, N., Wu, K.-P., & Inouye, M. (2012). YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in Escherichia coli. *Molecular microbiology*, *84*(5), 979–89.

[19] Hazan, R., & Engelberg-Kulka, H. (2004). Escherichia coli mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1. *Molecular genetics and genomics : MGG*, *272*(2), 227–34.

[20] Norton, J. P., & Mulvey, M. a. (2012). Toxin-Antitoxin Systems Are Important for Niche-Specific Colonization and Stress Resistance of Uropathogenic Escherichia coli. *PLoS pathogens*, *8*(10), e1002954. doi:10.1371/journal.ppat.1002954

[21] Wang, X., & Wood, T. K. (2011). Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Applied and environmental microbiology*, *77*(16), 5577–83.

[22] Kim, Y., Wang, X., Ma, Q., Zhang, X.-S., & Wood, T. K. (2009). Toxin-antitoxin systems in *Escherichia coli* influence biofilm formation through YjgK (TabA) and fimbriae. *Journal of bacteriology*, *191*(4), 1258–67.

[23] Koga, M., Otsuka, Y., Lemire, S., & Yonesaki, T. (2011). Escherichia coli rnlA and rnlB compose a novel toxin-antitoxin system. *Genetics*, *187*(1), 123–30.

[24] Kolodkin-Gal, I., Verdiger, R., Shlosberg-Fedida, A., & Engelberg-Kulka, H. (2009). A differential effect of *E. coli* toxin-antitoxin systems on cell death in liquid media and biofilm formation.

[25] Pesavento, C., Becker, G., Sommerfeldt, N., Possling, A., Tschowri, N., Mehlis, A., & Hengge, R. (2008). Inverse regulatory coordination of motility and curli-mediated adhesion in Escherichia coli. *Genes & development*, *22*(17), 2434–46.

[26] Brown, M. R. W., Collier, P. J., & Gilbert, P. (1990). Influence of Growth Rate on Susceptibility to Antimicrobial Agents : Modification of the Cell Envelope and Batch and Continuous Culture Studies. *Antimicrob. Agents Chemother. 1990, 34*(9):1623.

[27] Zuo, R., Hashimoto, Y., Yang, L., Bentley, W. E., & Wood, T. K. (2006). Autoinducer 2 Controls Biofilm Formation in Escherichia coli through a Novel Motility Quorum-Sensing Regulator (MqsR , B3022), *188*(1), 305–316.

[28] Ren, D., Bedzyk, L. a, Thomas, S. M., Ye, R. W., & Wood, T. K. (2004). Gene expression in *Escherichia coli* biofilms. *Applied microbiology and biotechnology*, *64*(4), 515–24.

[29] Kasari, V., Kurg, K., Margus, T., Tenson, T., & Kaldalu, N. (2010). The Escherichia coli mqsR and ygiT genes encode a new toxin-antitoxin pair. *Journal of bacteriology*, *192*(11), 2908–19.

[30] Brown, B. L., Grigoriu, S., Kim, Y., Arruda, J. M., Davenport, A., Wood, T. K., Peti, W., et al. (2009). Three dimensional structure of the MqsR:MqsA complex: a novel TA pair comprised of a toxin homologous to RelE and an antitoxin with unique properties. *PLoS pathogens*, *5*(12), e1000706.

[31] Kim, Y., Wang, X., Zhang, X.-S., Grigoriu, S., Page, R., Peti, W., & Wood, T. K. (2010). Escherichia coli toxin/antitoxin pair MqsR/MqsA regulate toxin CspD. *Environmental microbiology*, *12*(5), 1105–21.

[32] Wang, X., Kim, Y., Hong, S. H., Ma, Q., Brown, B. L., Pu, M., Tarone, A. M., et al. (2011). Antitoxin MqsA helps mediate the bacterial general stress response. *Nature Chemical Biology*, *7*(6), 359–366.

[33] Lacour, S., Landini, P., Lacour, S., & Landini, P. (2004). σ S-Dependent Gene Expression at the Onset of Stationary Phase in Escherichia coli : Function of σ S-Dependent Genes and Identification of Their Promoter Sequences. *J. Bacteriol. 2004, 186*(21):7186.

[34] Ren, D., Bedzyk, L. a, Thomas, S. M., Ye, R. W., & Wood, T. K. (2004). Gene expression in Escherichia coli biofilms. *Applied microbiology and biotechnology*, *64*(4), 515–24.

[35] Kim, Y., & Wood, T. K. (2010). Toxins Hha and CspD and small RNA regulator Hfq are involved in persister cell formation through MqsR in Escherichia coli. *Biochemical and biophysical research communications*, *391*(1), 209–13.

[36] Madrid, C., Balsalobre, C., García, J., & Juárez, A. (2007). The novel Hha/YmoA family of nucleoid-associated proteins: use of structural mimicry to modulate the activity of the H-NS family of proteins. *Molecular microbiology*, *63*(1), 7–14.

[37] García-Contreras, R., Zhang, X.-S., Kim, Y., & Wood, T. K. (2008). Protein translation and cell death: the role of rare tRNAs in biofilm formation and in activating dormant phage killer genes. *PloS one*, *3*(6), e2394.

[38] Godessart, N., Muñoa, F. J., Regue, M., & Juárez, a. (1988). Chromosomal mutations that increase the production of a plasmid-encoded haemolysin in Escherichia coli. *Journal of general microbiology*, *134*(10), 2779–87.

[39] Madrid, C., Nieto, J. M., Paytubi, S., Gualerzi, C. O., Juárez, A., Madrid, C., Nieto, M., et al. (2002). Regulation of a Plasmid-Encoded Virulence Operon Expressing Escherichia coli Hemolysin Temperature- and H-NS-Dependent Regulation of a Plasmid-Encoded Virulence Operon Expressing Escherichia coli Hemolysin.

[40] Baños, R. C., Vivero, A., Aznar, S., García, J., Pons, M., Madrid, C., & Juárez, A. (2009). Differential regulation of horizontally acquired and core genome genes by the bacterial modulator H-NS. *PLoS genetics*, *5*(6), e1000513.

[41] Seok Hoon Hong, Jintae Lee, and Thomas K. Wood (2011). Engineering Global Regulator Hha of *Escherichia coli* to Control Biofilm Dispersal . *Microb Biotechnol, 3*(6), 717–728.

[42] García-Contreras, R., Zhang, X.-S., Kim, Y., & Wood, T. K. (2008). Protein translation and cell death: the role of rare tRNAs in biofilm formation and in activating dormant phage killer genes. *PloS one*, *3*(6), e2394.

[43] Magnuson, R. D. (2007). Hypothetical functions of toxin-antitoxin systems. *Journal of bacteriology*, *189*(17), 6089–92.

[44] Ren, D., Bedzyk, L. a, Thomas, S. M., Ye, R. W., & Wood, T. K. (2004). Gene expression in Escherichia coli biofilms. *Applied microbiology and biotechnology*, *64*(4), 515–24.

[45] Wang, X., Kim, Y., & Wood, T. K. (2009). Control and benefits of CP4-57 prophage excision in Escherichia coli biofilms. *The ISME journal*, *3*(10), 1164–79.

[46] Barnhart, M. M., & Chapman, M. R. (2006). Curli biogenesis and function. *Annual review of microbiology*, *60*, 131–47.

[47] Sharma, V. K., & Bearson, B. L. (2013). Hha controls Escherichia coli O157:H7 biofilm formation by differential regulation of global transcriptional regulators FlhDC and CsgD. *Applied and environmental microbiology*, *79*(7), 2384–96.

[48] Kaplan, J. B. (2010). Biofilm dispersal: mechanisms, clinical implications, and potential therapeutic uses. *Journal of dental research*, *89*(3), 205–18.

[49] Karatan, E., & Watnick, P. (2009). Signals, regulatory networks, and materials that build and break bacterial biofilms. *Microbiology and molecular biology reviews : MMBR*, *73*(2), 310–47.

[50] Webb, J. S., Thompson, L. S., James, S., Tolker-nielsen, T., Koch, B., Givskov, M., Kjelleberg, S., et al. (2003). Cell Death in Pseudomonas aeruginosa Biofilm Development. *J. Bacteriol, 185*(15):4585.

[51] Mai-Prochnow, A., Webb, J. S., Ferrari, B. C., & Kjelleberg, S. (2006). Ecological advantages of autolysis during the development and dispersal of Pseudoalteromonas tunicata biofilms. *Applied and environmental microbiology*, *72*(8), 5414–20.

[52] Zuo, R., Ren, D., & Wood, T. K. (2005). Hha, YbaJ, and OmpA Regulate Escherichia coli K12 Biofilm Formation and Conjugation Plasmids Abolish Motility. *Published online: Wiley InterScience* (www.interscience.wiley.com).

Chapter II.

RESULTS & DISCUSSION

II.1 Optimization of NMR sample parameters to elucidate the 3D structure of a member of TomB/YmoB family of proteins by NMR

II.1.1. INTRODUCTION

Compared to other spectroscopic techniques NMR is a relatively insensitive method. To determine the structure of proteins multidimensional experiments require long acquisition times and protein concentrations at the millimolar (mM) range[53]. Optimization of the protein constructs to achieve the required solubility and stability to obtain good quality NMR spectra is the first, non-trivial step for 3D protein structure elucidation.

In order to identify the most suitable construct of the TomB family of proteins for NMR structure determination we employed a rapid ''screening'' of ^{15}N-labeled constructs, concentrated at the mM range, by ^{1}H-^{15}N-HSQC spectroscopy[54]. The HSQC spectrum provides a diagnostic fingerprint of a protein. One peak is expected for each nonproline residue. ^{1}H-^{15}N-HSQC spectra of soluble proteins could be classified as good, promising, or poor. The ''good'' spectra showed dispersion of peaks with roughly equal intensity and in the number expected from the sequence of the protein. These spectra indicate that the protein was amenable to determine its structure by NMR. ''Promising'' spectra showed well dispersed peaks but were either too few or too many in number and were often of differing intensities. The HSQC spectra classified as "poor" were characterized by a group of broad peaks with little dispersion in the center of the spectrum.

Once a suitable construct was selected, we optimized the conditions for stability, solubility and NMR sensitivity of the sample.

II.1.2. PROTEIN SELECTION

E. coli TomB was our starting point as biofilm been most studied in this system. We overexpressed TomB fused to GST trough a thrombin sensitive connector, but when we digested it with thrombin we obtained two species. The less abundant species corresponded to a truncated form of TomB lacking the short peptide -NATKENPASLSC located at the C-Terminus. The sensitivity to proteases suggests that this part of the protein is not structured and is highly exposed.

We decided to generate two new constructs, H_6-TomB[i] and TomB-Stag[ii] in pET14b commercial vectors. The short tags do not interfere with structure determination and can be retained in the final construct, avoiding the use of proteases. The two constructs were

[i] H_6-TomB corresponds to TomB protein with a purification tag of six histidines at N-Terminal.
[ii] TomB-Stag corresponds to TomB protein with a purification Strep tag at C-Terminal.

soluble at µM range but when we tried to concentrate them to the mM range they precipitated, so they were discarded.

The presence of free cysteines may result in disulfide bond formation, leading to irreversible oligomers. TomB has four cysteines; at least one of them located at the C-terminus is very exposed. Although reducing conditions were used throughout, we tried to minimize possible cysteine related problems. We mutated the C-terminal cysteine to serine producing the mutant C124S H_6-TomB in order to reduce the intermolecular disulfide-bridge formation. The new construct could be concentrated at 0.5mM without protein precipitation.

The ^{1}H-^{15}N-HSQC of [C124S] H_6-TomB is shown in Fig 6A. We classified this spectrum as "poor". It shows groups of broad peaks with strong overlapping, it is not possible to identify all the expected peaks, and there are peaks with different intensities.

In spite that reducing the number of cysteines resulted in an improvement of protein solubility, the spectra were still of poor quality. We decided to focus in the homologous protein YmoB from *Yersinia spp*. This protein has only two cysteines, instead of the four that are present at TomB. With this construct we expected a further reduction in the oligomerization equilibrium and an increase in the quality of the spectra. We generated a H_6-YmoB construct and performed a "screening" ^{1}H-^{15}N-HSQC (Fig 6B).

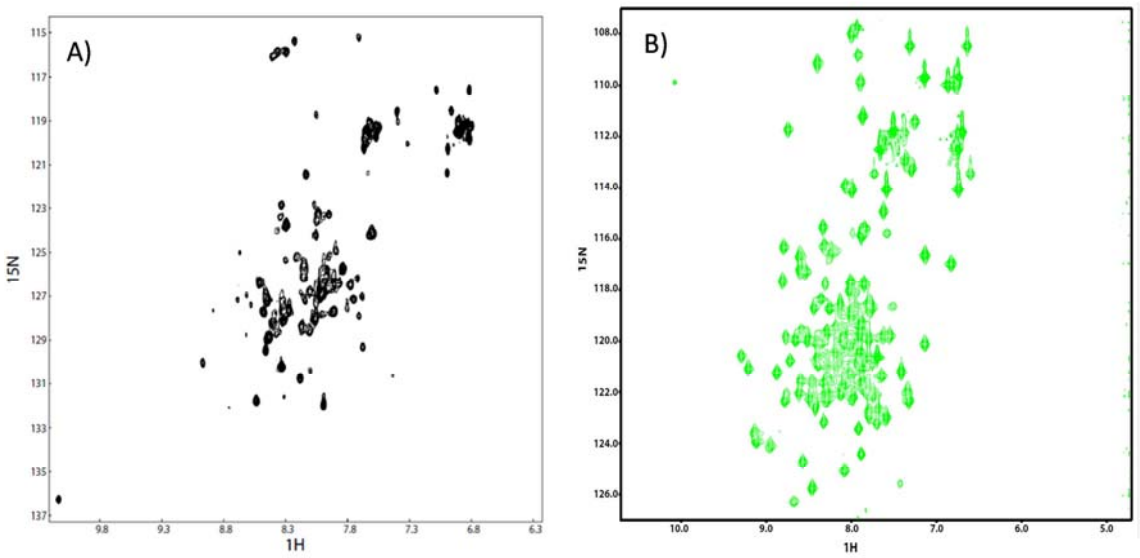


Fig 6: A: ^{1}H-^{15}N-HSQC spectrum of ^{15}N-[C124S]-H_6-TomB. We can observe a high degree of overlapping in the central zone of the spectra, two set of peak intensities and a significant loss of signals. **B:** ^{1}H-^{15}N-HSQC spectrum of ^{15}N-H_6-YmoB. The intensities of the signals are homogenous, with low overlap allowing the observation of all the expected signals. Spectrum was recorded at pH=7.00 phosphate buffered, 150mM NaCl, 0,1mM EDTA[iii], 1mM TCEP[iv], 0,1%NaN_3, 37°C and 0.5mM of protein.

[iii] EDTA: Ethylenediaminetetraacetic acid.

We obtained a "good" spectrum where it was possible to count all the expected signals of the construct, well dispersed and with roughly similar intensity of the peaks. So we decided to use this construct to optimize the solvent conditions to improve the efficiency of the NMR structural experiments.

II.1.3. NMR SAMPLE OPTIMIZATION

We optimized the pH of the sample, the type and concentration of salts used to minimize the electrostatic interactions and the maximum concentration of protein that could be reached.

In NMR, the optimal pH to minimize the loss of signals produced due to the chemical exchange between NH protons from protein and water is around 5.00[55]. But to improve the solubility of the proteins is crucial to work at a pH far from the isoelectric point of the protein. For H_6-YmoB the theoretical isoelectric point is 5.20, so we evaluated the solubility of samples at 50μM H_6-YmoB from pH=5.00 to pH=7.50. Below pH=6.50 the protein precipitates but at pH equal or higher than 7.00 becomes totally soluble. We fixed the pH of the sample at 7.00.

Often at mM concentrations proteins have oligomerization problems due to intermolecular electrostatic interactions. Adding high salt concentrations can improve solubility. However, the presence of salts increases the ionic conductivity of the sample, and results in a significant reduction in the sensitivity of cryoprobes[56].

We tested two types of salts; NaCl (phosphate buffered) because is an aprotic salt that would produce no interference in future NMR experiments, and HEPES[v], that is auto buffered at our working pH and has a low ionic conductivity[57]. We calculated the melting temperatures (T_m) in samples at 50μM H_6-YmoB, with increasing salt concentrations (from 0mM to 200mM) using differential scanning fluorimetry[58]. 100 and 150mM of NaCl gave the highest thermal stability (giving 43.9±0.2°C and 44.2±0.5°C T_m respectively). The highest T_m in HEPES (42.1±0.4°C) was observed at concentrations equal or higher than 100mM.

Sample stability at long times was studied at high protein concentration (1mM) at 100mM NaCl, 150mM NaCl and 100mM HEPES. We kept the samples to 25°C during 10 days. After that we checked the presence of precipitate and the final concentration of the protein. As soluble protein oligomers scatter light causing an increase in the absorption measurement at 340nm[59] we also measured the absorbance of the sample at this wavelength.

[iv] TCEP: Tris (2-carboxyethyl) phosphine hydrochloride.

[v] HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

We discarded the sample at 100mM NaCl because more than 20% of H_6-YmoB precipitated and the absorption at 340nm saturated the optical reader. However, at 150mM NaCl and 100mM HEPES there was not protein precipitation, the final protein concentration was reduced by less than 10% and the absorbance at 340nm remained stable for each case.

We recorded "screening" ^{1}H-^{15}N-HSQCs of ^{15}N-H_6-YmoB 1mM at 150nM NaCl and 100mM HEPES. In both cases we obtained "good" spectra, but for NaCl, the spectrum (Fig 7, green) showed better homogeneity of intensities and peak dispersion than for HEPES (Fig 7, blue).

As 150mM NaCl gives no interference and produces "good" ^{1}H-^{15}N-HSQC, we decided to use 150mM NaCl buffered with phosphate at pH=7.00.

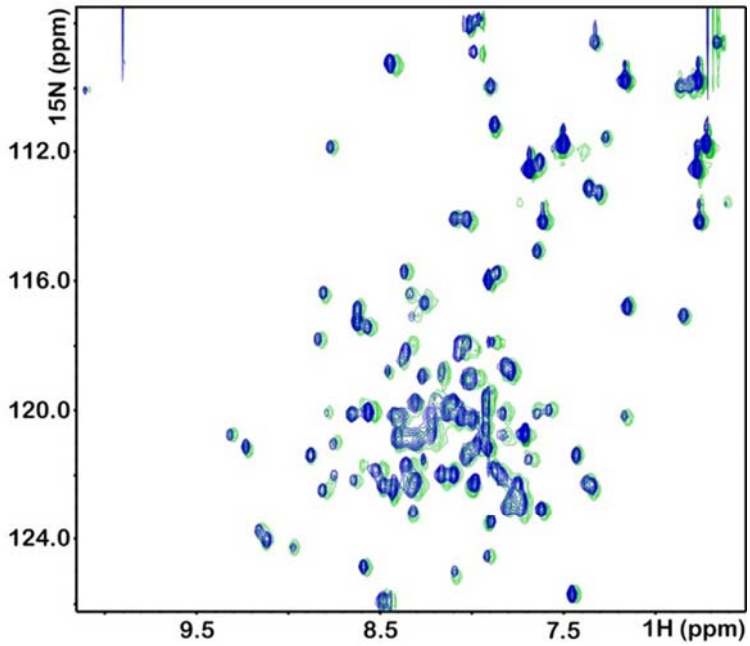


Fig 7: Superposition of ^{1}H-^{15}N-HSQC of ^{15}N-H_6-YmoB 1mM, pH=7.0, 1mM TCEP and 0.01% NaN_3. In blue sample with 100mM HEPS and in green sample with 150mM NaCl.

We next determined the maximum concentration for H_6-YmoB. We recorded 1D ^{1}H-WG-NMR[vi,60] experiments of samples at increasing concentrations of H_6-YmoB (Fig 8A). We also recorded the absorbance of the samples at 340_{nm} (Fig 8B). Until 1mM of H_6-YmoB the intensity increased in accordance with the increasing of the protein concentration. However between 1 and 1.4mM of H_6-YmoB the NMR intensity dropped at the same time that absorbance at 340nm increased.

[vi] ^{1}H-WG-NMR is a one dimensional ^{1}H-NMR experiment the water signal is eliminated using watergate sequence.

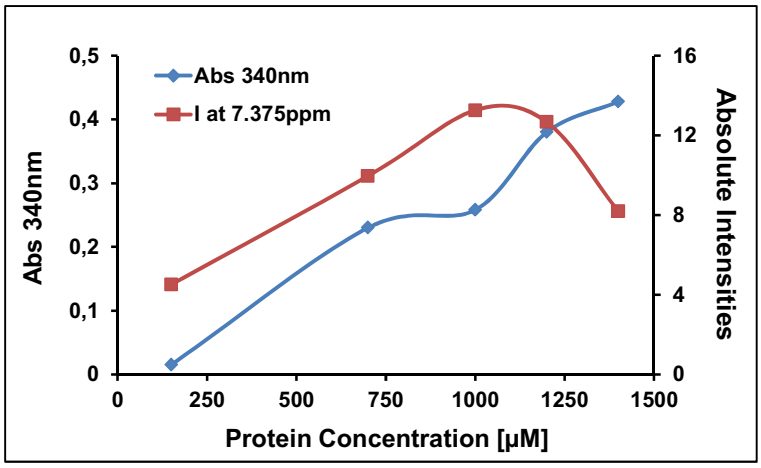


Fig 8: For samples of H_6-YmoB at growing concentrations we show in blue the absorbance at 340nm and in red the intensity of the peaks from ^{1}H-WG-NMR spectra at 7.375ppm.

The profile of the ^{1}H-WG-NMR also changed with the increasing of the protein concentration. Below 0.7mM H_6-YmoB we identified signals changing with concentration: peaks that disappeared, new peaks that appeared and peaks that changed their chemical shift (Fig 9, red, black and blue dashed lines respectively).

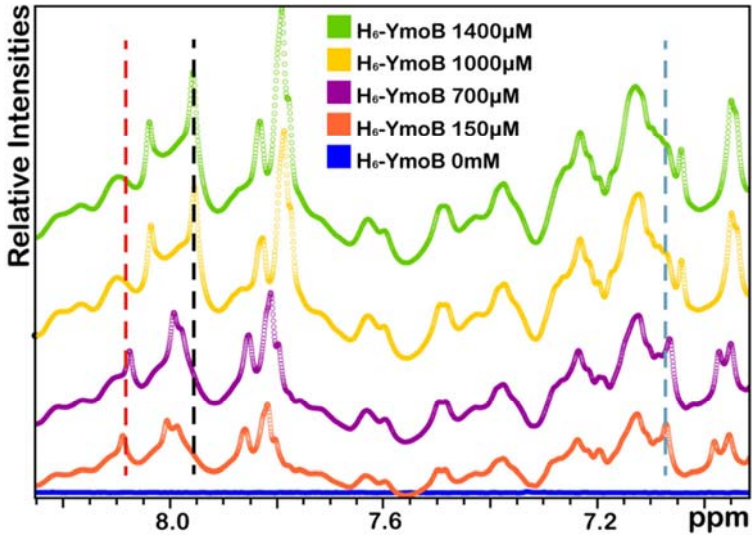


Fig 9: Expansions of ^{1}H-WG-NMR spectra corresponding to a portion of the amida and aromatic proton region of H_6-YmoB at increasing protein concentration. Dashed lines highlight changes at some signals; red for a peak that disappears upon concentration, black for a peak that appears only in concentrated conditions and blue for a peak that changes its chemical shift.

We performed size exclusion chromatography of 1.0mM H_6-YmoB NMR sample (Fig 10, Blue) and we identified two well resolved peaks, one corresponding to a species identified by mass spectrometry as a covalent dimer (Fig 10, A: D) and a second one corresponding to the monomeric form of the construct(Fig 10, B: M). No higher order oligomers were observed. Surprisingly, covalent dimer species were former in spite of the presence of

TCEP. Furthermore, the absence of larger oligomers suggests that only one cysteine was involved in the dimerization process. We hypothesized that the two cysteines of YmoB had very different reactivity and that mutating only one of them may prevent the formation of the dimeric species.

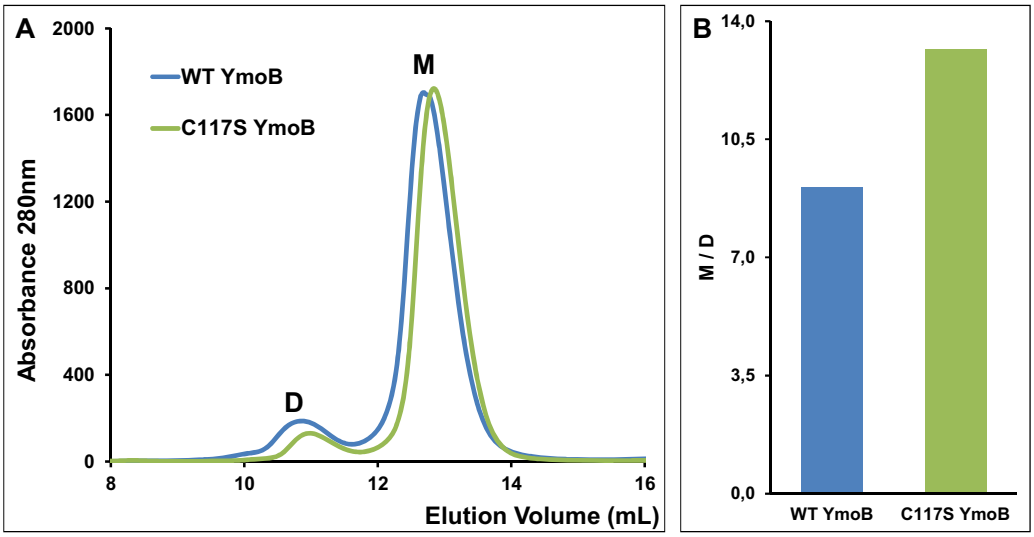


Fig 10: A: Elution profile for size exclusion chromatography of samples at 1mM WT YmoB (blue) and [C117S] YmoB (green). D peak corresponds to the covalent-dimmer form. M peak corresponds to the monomeric form. **B:** The ratio between the absorbance of the constructs in the monomeric state (peak 2 at Fig 10) and in the covalent-dimmer state (peak 1 at Fig 10)

We performed a thiol alkylation of the cysteines of H_6-YmoB with sodium iodoacetate[61]. By mass spectrometry we checked that only one cysteine was alkylated. YmoB cysteines are at positions 18 and 117 of the sequence, in a total of 122 amino acids. We hypothesized that C117 was the alkylated cysteine and that C18 remained deeply protected by the tertiary structure of the protein. We obtained the [C117S] H_6-YmoB (C117S) construct using standard mutagenic techniques. To check our hypothesis we repeated the thiol alkylation with the [C117S] H_6-YmoB construct. By mass spectrometry we checked that there was not alkylation. Surprisingly, after size exclusion chromatography of [C117S] H_6-YmoB at 1.0mM we still obtained protein covalent dimer (Fig 10, green), but in a lower amount than in the case of the WT.

Furthermore, in contrast to WT YmoB, the C117S mutant could be concentrated until 1.4mM without significant changes in the ^{1}H-WG-NMR profile (Fig 11).

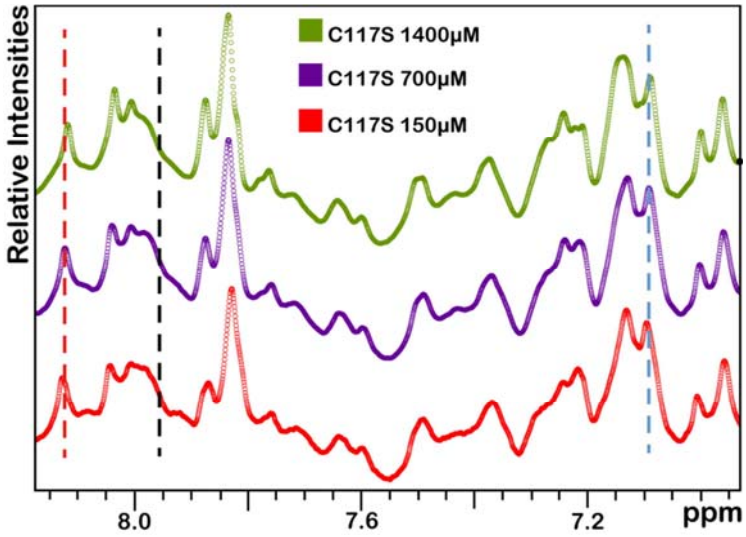


Fig 11: Expansions of ^{1}H-WG-NMR spectra corresponding to a portion of the amidic and aromatic proton region of C117S H_6-YmoB at increasing protein concentration. Dashed lines are located at the same chemical shifts than in Fig 9. Same optimal buffer conditions determined for the WT were used.

We recorded ^{1}H-^{15}N-HSQCs of 1.4mM ^{15}N-H_6-YmoB-C117S and we compared with ^{1}H-^{15}N-HSQC of 1.4mM H_6-YmoB (Fig 12). The mutant presents narrow peaks that enhance their intensity and improve the resolution of the spectra. In the case of the WT, peaks were broader, less intense, and had less dispersion than in the case of C117S H_6-YmoB.

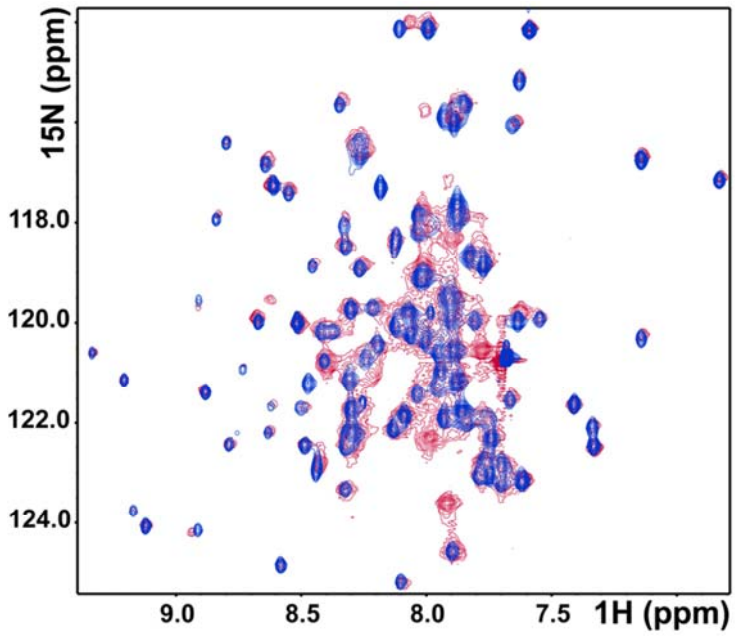


Fig 12: Superposition of ^{1}H-^{15}N-HSQCs of ^{15}N-H_6-YmoB at 1mM in red and ^{15}N-H_6-YmoB-C117S at 1.4mM in blue.

II.1.4. CONCLUSIONS

In TomB and YmoB proteins a reduction in the number of cysteines generates constructs that can be concentrated up to mM range and reduces the formation of soluble high molecular weight oligomers.

H_6-YmoB-C117S can be concentrated up to 1.4mM, giving "good" ^{1}H-^{15}N-HSQC spectra and being an optimal sample for 3D structure elucidation by NMR.

Slow formation of a dimer species still remained a problem with the C117S YmoB mutant but could be eventually solved by using fast and parallel acquisition protocols as described in Chapter II.3.

The final conditions of the NMR sample to obtain the 3D-structure of a member from TomB family of proteins were: H_6-YmoB-C117S at 1.4mM, pH=7.00 buffered with 20mM [NaH_2PO_4 + Na_2HPO_4], 150mM NaCl, 1mM TCEP, 0.1mM EDTA and 0.1% NaN_3.

II.1.5. BIBLIOGRAPHY

[53] Yee, A., Gutmanas, A., & Arrowsmith, C. H. (2006). Solution NMR in structural genomics. *Current opinion in structural biology*, *16*(5), 611–7.

[54] Yee, A., Chang, X., Pineda-Lucena, A., Wu, B., Semesi, A., Le, B., Ramelot, T., et al. (2002). An NMR approach to structural proteomics. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(4), 1825–30.

[55] Wüthrich, K. Nagayama, K. and Ernst, R. Trends Biocem. Science. 1979

[56] Kelly, A. E., Ou, H. D., Withers, R., Do, V., & Corp, B. B. (2002). Low-Conductivity Buffers for High-Sensitivity NMR Measurements copy is a relatively insensitive method , requiring concentrations, (14), 12013–12019.

[57] Hautbergue, G. M., & Golovanov, A. P. (2008). Increasing the sensitivity of cryoprobe protein NMR experiments by using the sole low-conductivity arginine glutamate salt. *Journal of magnetic resonance (San Diego, Calif. : 1997)*, *191*(2), 335–9.

[58] Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocol. 2(9), 2212-21.*

[59] Canaan, P., Vincentelli, R., Valencia, C., Maurin, D., Frassinetti, F. R. E. D. E. R. I. C., Scappucini-calvo, N. A., Bourne, Y., et al. (2004). High-throughput automated refolding screening of inclusion bodies, 2782–2792.

[60] Piotto, M., Saudek, V., & Sklenár, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of biomolecular NMR*, *2*(6), 661–5.

[61] Trivedi, M., Laurence, J. S., Williams, T. D., Middaugh, C. R., & Siahaan, T. J. (2012). Improving the stability of the EC1 domain of E-cadherin by thiol alkylation of the cysteine residue. *International journal of pharmaceutics*, *431*(1-2), 16–25.

II.2. Functional studies of TomB family of proteins

II.1.1. INTRODUCTION

Up to 80% of human bacterial chronic inflammatory and infectious diseases are associated to biofilm formation. Biofilms are surface-attached resistant communities, which protect bacteria from adverse environmental conditions and host immune response[62]. Biofilm development is a complex process that generally occurs in three phases: adherence, proliferation and dispersion[63]. Biofilm cycle depends on environmental cues and requires the coordinate expression and simultaneous regulation of many genes.

Hha is a well-known environmental regulator of *E. coli* genes in planktonic cells. Hha represses the transcription of rare tRNAs, which serves to inhibit fimbriae production, leads to cell lysis by activation on lytic genes and activate toxins by the induction of the proteases that degrade the antitoxins. In planktonic cells, Hha acts as a toxin while co-transcribed TomB seems to attenuate Hha toxicity[64]. These data suggested that Hha and TomB may form a toxin-antitoxin (TA) pair.

In *E. coli* biofilms Hha overexpression decreases biofilm formation, but the effect of TomB is still unknown[64]. As other TA systems are known to regulate the formation of biofilms[65], understanding the different mechanisms of biofilm regulation and in particular clarifying the biological activity of TomB family of proteins as antitoxins of Hha in biofilms becomes of primary interest.

For technical reasons the structural study of TomB could not be performed and the efforts were directed to the Yersinia homologue of TomB, YmoB and its C117S mutant. In order to validate the YmoB constructs used, we determined the effect of YmoB and C117S YmoB in the well-studied biofilm system formed by *E. coli,* by looking at the following topics:

1. The effects of TomB family of proteins in independent cells: the TA role of Hha-TomB pair and the function of YmoB and C117S YmoB mutant.
2. The effects of TomB family of proteins in biofilms: the TA role of Hha-TomB in biofilms and the function of YmoB and C117S YmoB mutant.
3. The effects of overexpression of Hha and YmoB in the biofilm morphology: the characterization of the biofilm in presence of the TA Hha-YmoB pair.

An important part of this work was carried out during a stay in the laboratory of Dr. T. K. Wood in the Pennsylvania State University, at the Frenske Laboratory, in the department of Chemical Engineering & Biochemistry and Molecular Biology (USA).

The strategy used was based on the transformation of *E. coli* K12MG1655 (Δhha) strain with two plasmids: pCA24Nhha, which expresses Hha under the control of a T7 promoter,

induced by IPTG, and pBAD30x* (x*= tomb, ymob or C117S ymob), which expresses the indicated protein under the control of an araC promoter, and is induced by arabinose. Table 1 shows the abbreviations referring to the strains used in this study.

ABBREVIATION	GENOTYPE	INDUCER	OVEREXPRESSED PROTEIN
WT	WT	-	-
Δhha	Δhha	-	-
Hha(-) X*(-)	Δhha	-	-
Hha(+)X*(-)	Δhha	IPTG	Hha
Hha(+)X*(+)	Δhha	IPTG + Arabinose	Hha + X*
Hha(-) X* (+)	Δhha	Arabinose	X*

Table 1: List of abbreviations used to design the strain *E coli* K12MG1655 carrying the pCA24Nhha and pBAD30x* plasmids. X*= TomB, YmoB or C117S YmoB proteins. x*= tomb, ymob or C117S ymob. IPTG and arabinose were used at final concentrations of 1mM and 1% respectively. We will use (-) to point out no overexpression and (+) to point out overexpression of proteins.

II.1.2. EFFECTS OF TomB FAMILY OF PROTEINS IN PLANKTONIC CELLS

In *E. coli* planktonic cells, the overexpression of Hha is toxic for the cells and the co-expression with TomB attenuates the toxic effect of Hha by reducing cell mortality[64]. These results have been previously obtained in the T.K. Wood laboratory by measuring bacterial growth curves in LB-Luria medium of *E coli* BL21 (DE3) strains transformed with pET28-hha and pBAD-tomb. However, the presence of sugars in the LB-Luria medium could disturb the expression of protein under the control of the araC promoter in pBAD. This is one of the possible explanations of why TomB did not totally counteract the toxic effect of Hha and only attenuates its toxicity.

In order to fix this problem we used tryptone minimal medium for measuring the growth curves of the strains listed in Table 1. Cells were pre-cultured over-night at 37°C from single colonies with shaking (250rpm) and inoculated in 25mL of tryptone medium (with the corresponding antibiotics and inducers) until final cell turbidity of 0.05 at 600nm. Growth curves were measured using two independent cultures at 37°C and 250rpm.

No differences were observed at the growth curves between WT and Δhha. In the presence of the two plasmids but no inducers, growth curves were comparable with respect to the WT. Overexpression of Hha (Fig 13, dark green) strongly decreased cells growth rate. The effect was totally suppressed by the simultaneous overexpression of TomB (Fig 13, light red). Therefore, we demonstrated the TA action of these pair of proteins in planktònic cells.

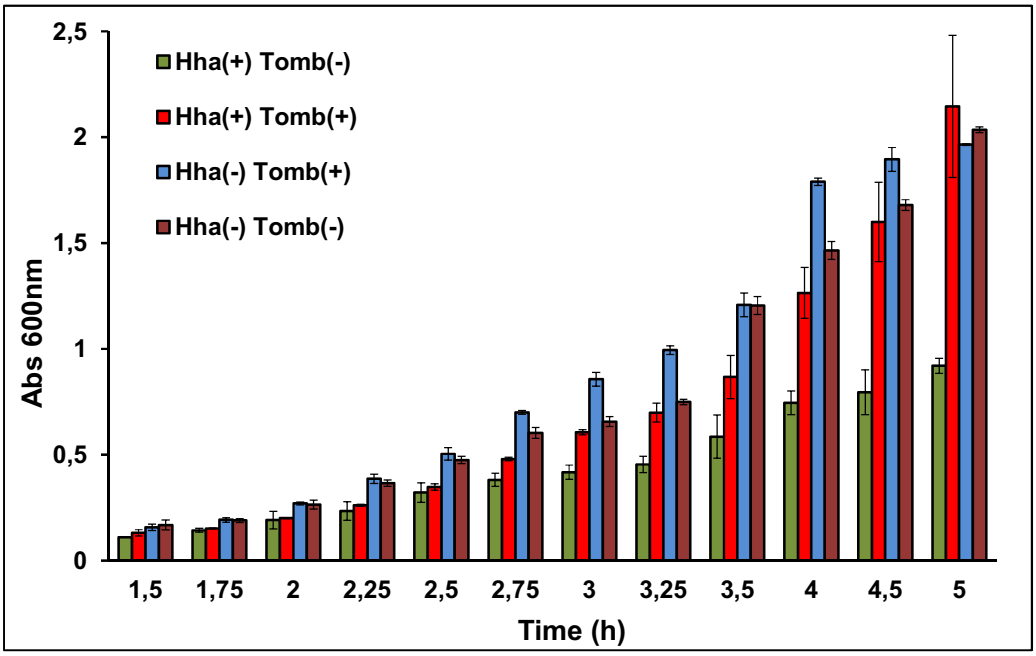


Fig 13: Growth of *Ecoli* K12MG1655 Δhha. Colored bars refers to which proteins were overexpressed; dark green for Hha, light red for co-expression of Hha and TomB, light blue for TomB and dark red for a control in which no proteins were overexpressed. Experiments were performed in duplicate, and one standard deviation is shown.

From the growth curves we calculated the Specific Growth Rates (SGRs)[66], which is a measure of how fast cells are diving. Fig 14 shows that the co-expression of the TA pair (Fig 14, red) allows cells to grow at the same rate with respect to the control (Fig 14, purple). When we overexpressed TomB the total growth of the cells (Fig 13, purple) was comparable with the control (Fig 13, purple), but showed different growing rates (Fig 14, light blue). The increase on the growth velocity may be explained by the addition of arabinose to induce TomB overexpression. Arabinose is a sugar that cells can use to grow faster, producing the observed increase in SGRs.

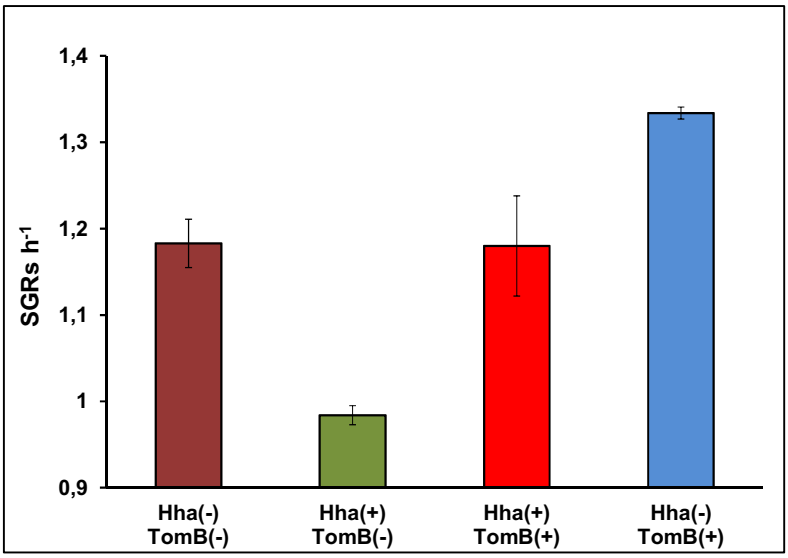


Fig 14: SGRs of *Ecoli* K12MG1655 Δhha were derived from the slope of $\ln OD_{600}$ vs. time (h) of the ODs, considering the exponential phase of the growth curves (OD_{600}=0.05~0.7). Experiments were performed in duplicate, and one standard deviation is shown.

We next compared the effect of the *Yersinia* variants YmoB and C117S YmoB in planktonic cells. Overexpression of TomB, YmoB and [C117S] YmoB in *E. coli* cells overexpressing Hha causes similar growth curves profiles (Fig 15).

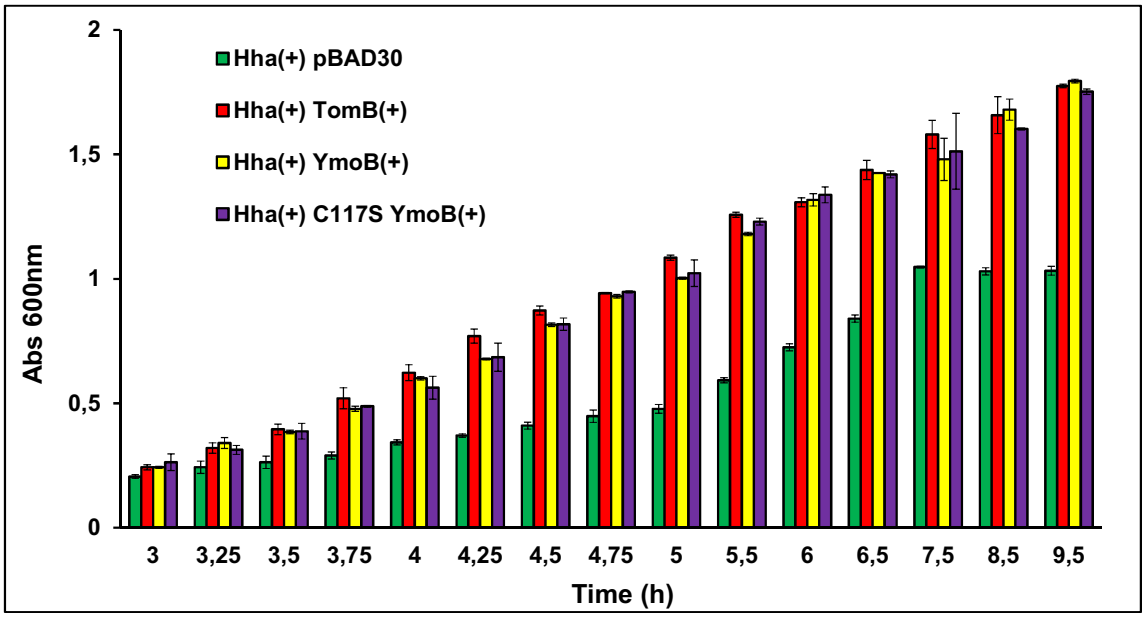


Fig 15: Growth of *Ecoli* K12MG1655 Δhha. Colored bars refers to which proteins were overexpressed; light green for Hha, red for co-expression of Hha and TomB, yellow for co-expression with YmoB and lilac for co-expression with [C117S] YmoB. Experiments were performed in duplicate, and one standard deviation is shown.

YmoB and [C117S] YmoB can replace TomB producing comparable cell growth rates. The analysis of SGRs (Fig 16) shows that cells overexpressing Hha and YmoB grow a little more slowly than those overexpressing Hha and TomB (0.917±0.015 h^{-1} for TomB and 0,862±0.010 h^{-1}). YmoB WT and then [C117S] mutant did not present significant differences (0,855±0.024 h^{-1} for [C117S] YmoB).

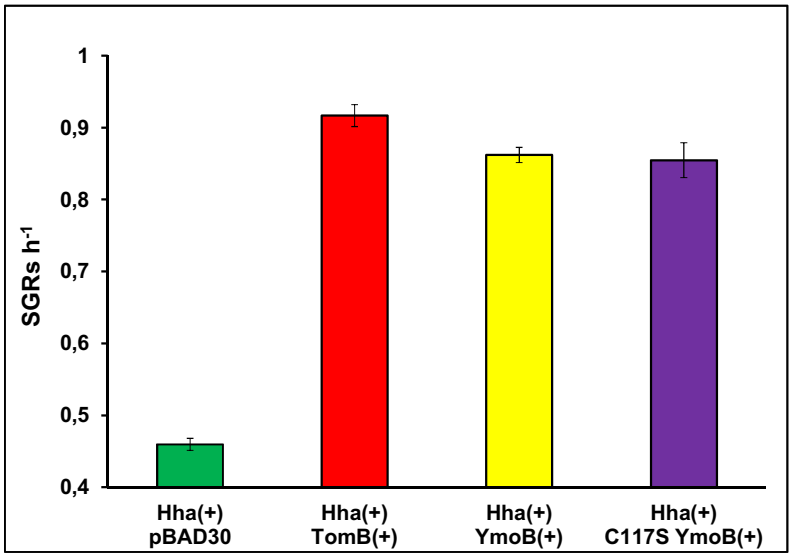


Fig 16: SGRs of *Ecoli* K12MG1655 Δhha overexpressing Hha (light green), Hha and TomB (red), Hha and YmoB (yellow) and Hha and [C117S] YmoB (lilac). Experiments were performed in duplicate, and one standard deviation is shown.

II.1.3. EFFECTS OF TomB FAMILY OF PROTEINS IN BIOFILMS

Biofilm assays were conducted in 96-well microtitre plates. We inoculated bacteria in tryptone minimal medium (with the corresponding antibiotics and inducers) at an initial turbidity of 0.05 (at 600nm) and were incubated for 15h without shaking. Biofilms were visualized at 520nm by staining with 0.5% crystal violet and standardized with the total growth (620nm)[67]. Each data point was averaged from twenty replicate wells from two independent cultures of each strain.

Overexpression of Hha caused a 2.6-fold decrease in standard biofilm formation in comparison with the Δhha strain (Fig 17). A smaller reduction in biofilm formation was also observed in Δhha cells transformed with the plasmid encoding Hha even in the absence of IPTG. This effect is probably a consequence of basal expression of the pCA24Nhha plasmid.

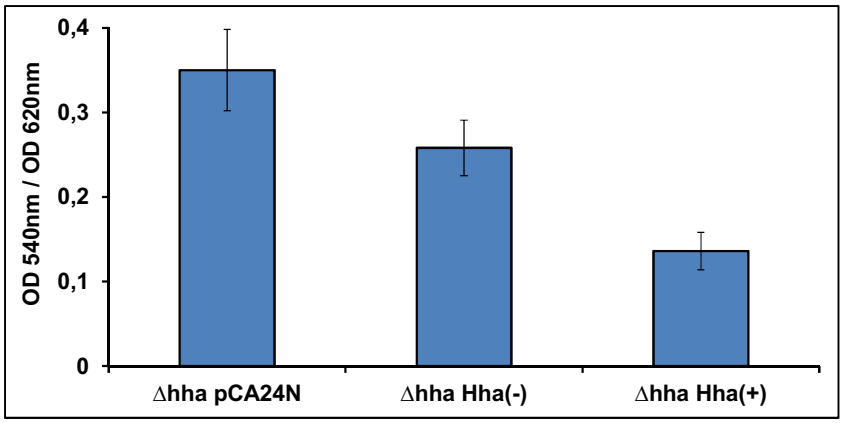


Fig 17: Biofilm formation normalized by total cell growth. The overexpression of Hha decreases total biofilm formation. The experiments were performed in duplicate, and one standard deviation is shown.

Overexpression of TomB and YmoB caused a similar increase in the amount of standard biofilm (around 4-fold) formed by cells overexpressing Hha (Fig 18, blue). In the absence of Hha (Fig 18, red), TomB and YmoB had a small effect on standard biofilm formation. As in the case of SGRs, this effect may be related to the addition of arabinose in the tryptone minimal medium for the induction of the antitoxins. [C117S] YmoB mutant was also able to enhance standard biofilm formation in *E. coli* overexpressing Hha, although its activity was around 75% with respect to wild type YmoB or TomB.

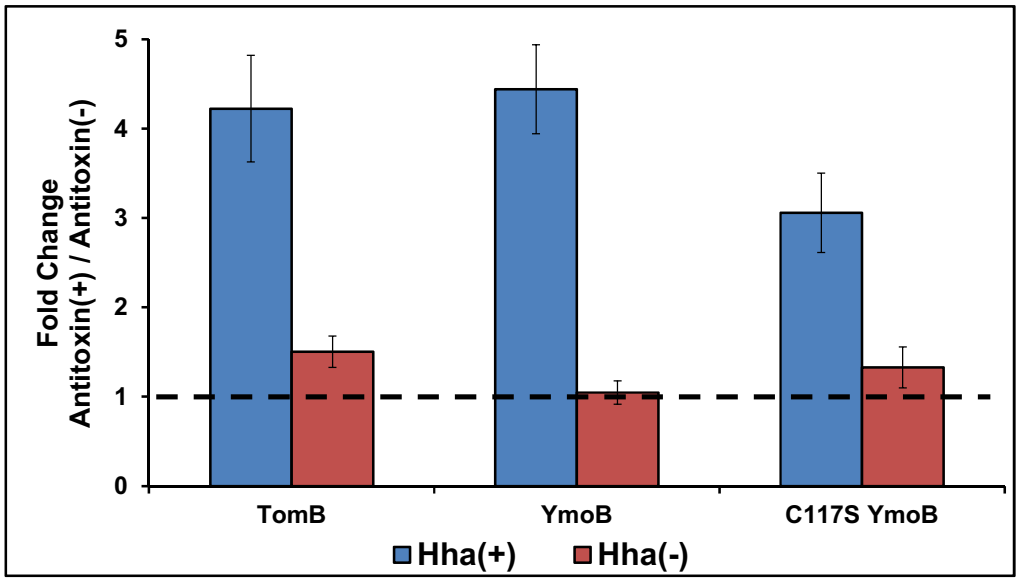


Fig 18: Effect of TomB/YmoB/ [C117S] YmoB antitoxins in biofilms. We show the fold change between the biofilm formation with and without (highlighted by the dashed line) overexpression of the antitoxins. The fold change is represented in blue when Hha is overexpressed and in red when the toxin is not present. The experiments were performed in duplicate, and one standard deviation is shown.

Summarizing, the experiments described show that a) Hha and TomB act as a TA module in *E. coli* biofilms, b) YmoB can replace TomB in biofilm formation and c) C117S mutation in YmoB does not prevent the effect of YmoB in biofilm formation, although the activity is slightly reduced.

II.1.4. EFFECTS OF Hha AND YmoB OVEREXPRESSION IN THE BIOFILM MORPHOLOGY

To study the biofilm morphology we performed flow cell biofilm assays and image analysis with and without overexpression of YmoB in *E. coli* cells overexpressing Hha. The flow cells were inoculated with tryptone minimal medium cultures at an initial turbidity of 0.05 (at 600nm) at a flow rate of 10 ml h-1 during 3 hours at 37°C. Culture medium was replaced with fresh tryptone minimal medium containing the corresponding antibiotics and inducers and then fed to the flow cell at the same rate for 15h at 37°C. Flow cell biofilms were stained using LIVE/DEAD *Bac*Light Bacterial Viability Kit, giving two wavelengths, one at 500nm (green) for living cells and the other at 635nm (red) for dead cells. Flow cell biofilm images from eight random positions were taken using confocal microscopy[68] and analyzed with Imaris confocal software (Bitplane, Zurich, Switzerland). Fig 19 shows the most representative images for each experiment, and all the spots are shown at Suppl. Fig. 1

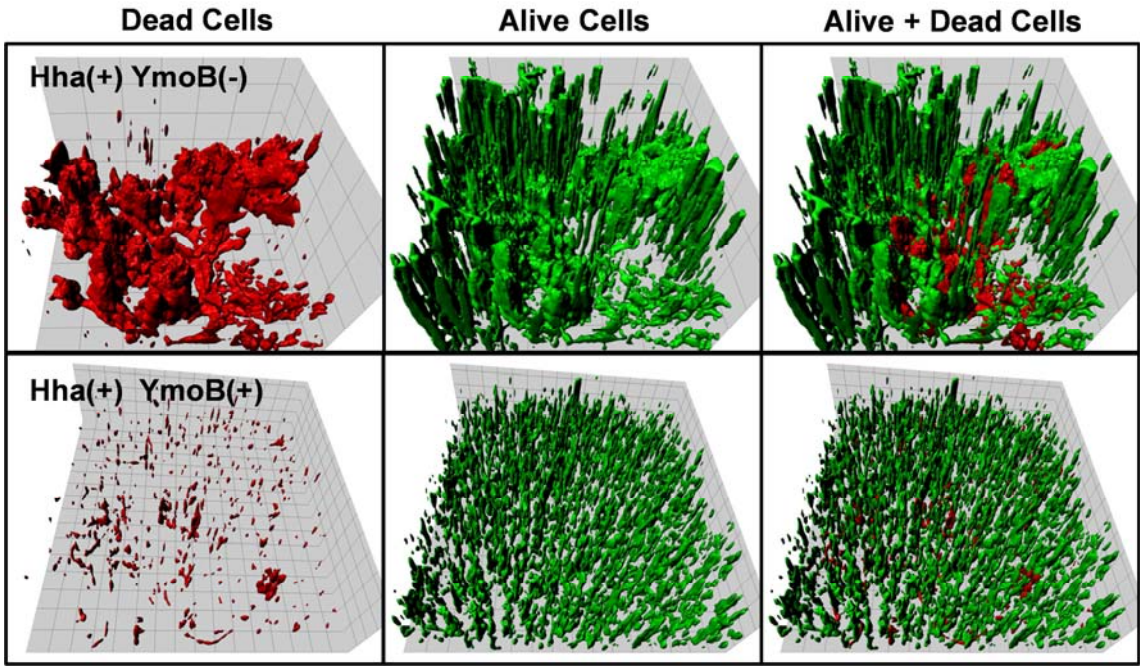


Fig 19: Imaris images of flow cell biofilm of *Ecoli*K12MG1655 Δhha pCA24Nhha pBAD30ymob. We show dead cells in red and alive cells in green. Representative images for Hha(+) YmoB(-) (up) and Hha(+) YmoB(+) (down) are shown.

In cells overexpressing Hha, amorphous biofilms with a high presence of dead cells were formed. Instead, when YmoB was also expressed, the amount of dead cells was significantly reduced and the biofilm was composed by micro colonies uniformly spread in the entire surface.

The distribution of Live and Dead (L-D) cells was characterized a) by counting the relative biomass of the two groups of cells, b) by the relative surface covered by L-D cells and c) by the mean thickness of distribution of L-D cells (Fig 20). These parameters have been shown to fully describe the characteristics of biofilm systems and were obtained using COMSTAT[69] confocal software.

The overexpression of YmoB resulted in a significant decrease in the number of dead cells and a corresponding increase of the L/D cells ratio up to 35.4 and 36.9 for biomass and for substratum coverage, respectively. In cells overexpressing Hha, L/D ratio for biomass and substratum coverage dropped down to 5.6 and 3.2, respectively. The mean thickness ratio between live and dead cells was not affected by overexpression of YmoB.

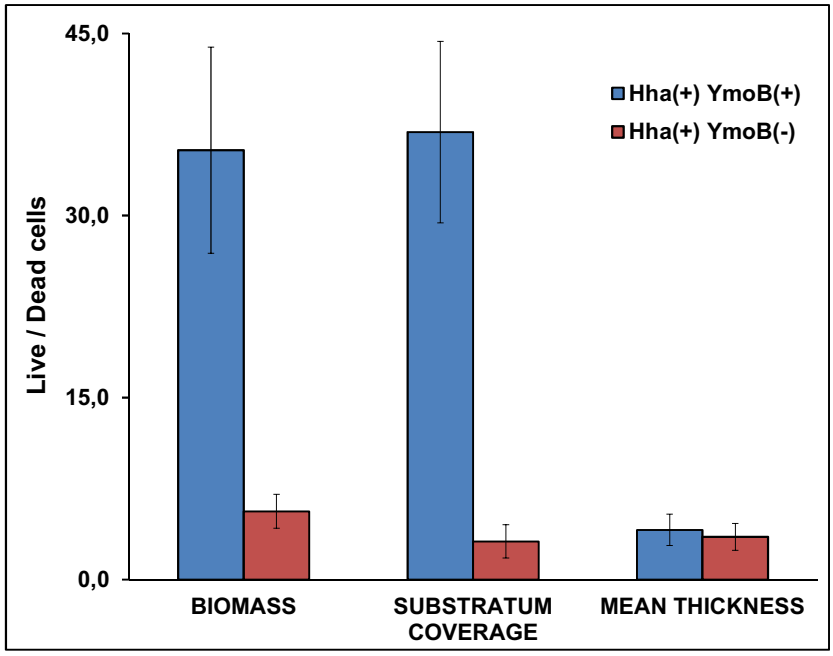


Fig 20: Statistical analysis of flow cell biofilms. For the biomass ($\mu m^3/\mu m^2$), the substratum coverage (%) and the mean thickness (μm) we calculated the ratio between Live and dead cells (L/D). Blue bars refer to co-expression of Hha and YmoB, and red to Hha overexpression.

II.1.5. CONCLUSIONS

Hha-TomB proteins form a toxin-antitoxin system with Hha the toxin and TomB the antitoxin, in planktonic cells as well as in biofilms.

In *E. coli* planktonic and biofilm system, we demonstrated that the biological activity of TomB proteins is always related to Hha.

We showed that TomB and YmoB are functionally interchangeable in planktonic cells and in biofilms, demonstrating that this TA module is functional in enteric bacteria.

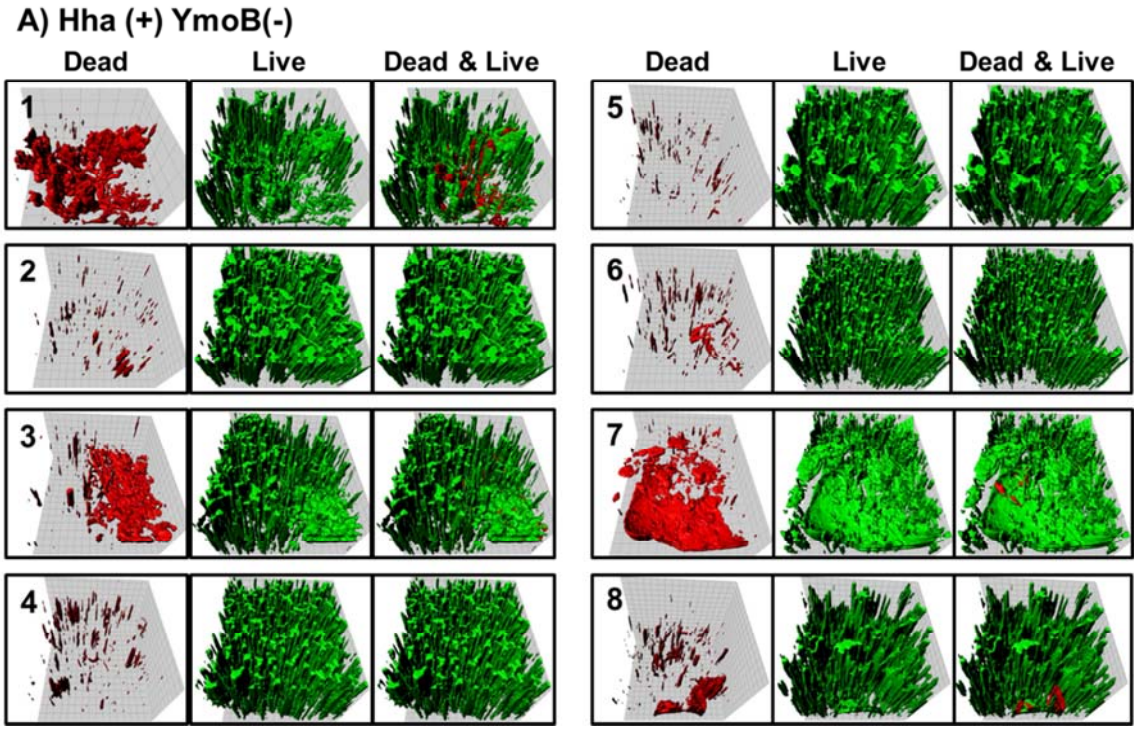
Moreover, we confirmed that the [C117S] YmoB mutant, our optimal construct for NMR studies, is a biologically active antitoxin of Hha in *E. coli* planktonic and biofilm systems.

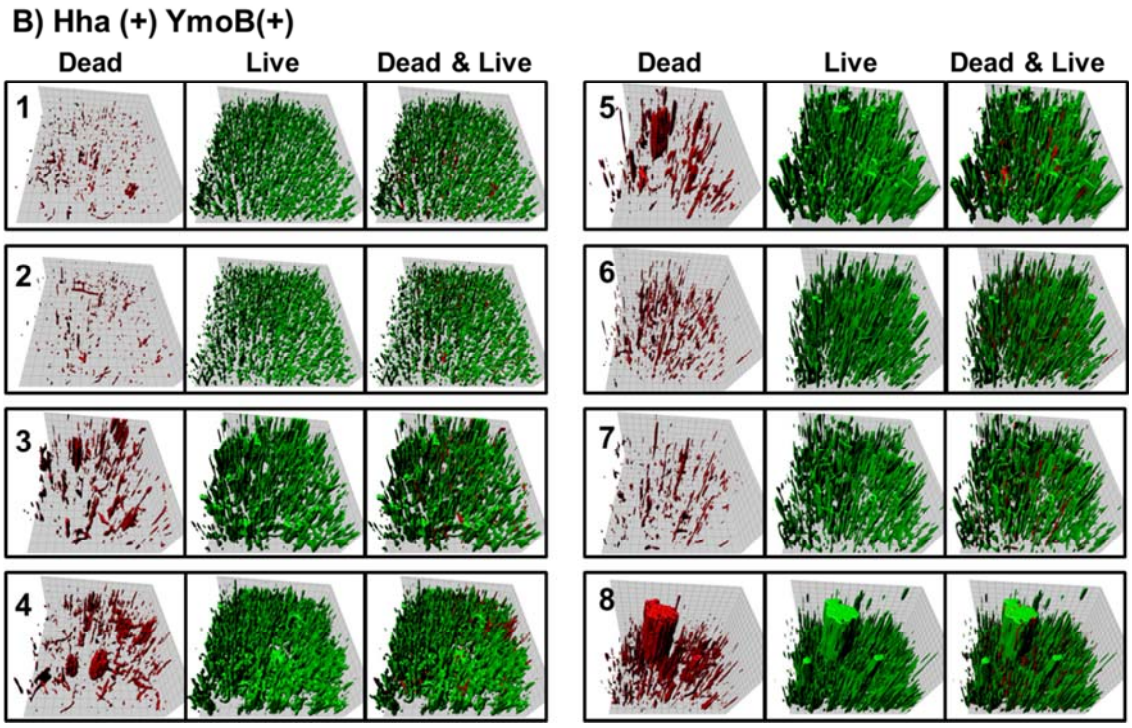
When the TA system of Hha-TomB family of proteins is misfunctioning, the deregulation produced in the biofilm gives rise to a dramatic increase of dead cells with the core of the biofilm.

Finally, when Hha and YmoB are simultaneously overexpressed, the biofilm is organized in micro colonies, which are mainly composed by living cells uniformly spread over the entire surface.

II.1.6. SUPPLEMENTARY MATERIAL

Suppl. Fig. 1: Imaris images of flow cell biofilm of *Ecoli* K12MG1655 Δhha pCA24Nhha pBAD30ymob. **A:** Cells overexpressing Hha. **B:** Cells overexpressing Hha and YmoB. Dead cells are red and live cells green.





II.1.7. BIBLIOGRAPHY

[62] Römling, U., & Balsalobre, C. (2012). Biofilm infections, their resilience to therapy and innovative treatment strategies. *Journal of internal medicine*, *272*(6), 541–61.

[63] Toole, G. O., Kaplan, H. B., & Kolter, R. (2000). Biofilm Formation As Microbial Development. *Annu. Rev. Microbiol. 54:*49–79.

[64] García-Contreras, R., Zhang, X.-S., Kim, Y., & Wood, T. K. (2008). Protein translation and cell death: the role of rare tRNAs in biofilm formation and in activating dormant phage killer genes. *PloS one*, *3*(6), e2394.

[65] Wang, X., & Wood, T. K. (2011). Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Applied and environmental microbiology*, *77*(16), 5577–83.

[66] Berney, M., Weilenmann, H., Ihssen, J., Bassin, C., Egli, T., Berney, M., Weilenmann, H., et al. (2006). Specific Growth Rate Determines the Sensitivity of Escherichia coli to Thermal , UVA , and Solar Disinfection. *Appl. Environ. Microbiol. 72*(4):2586-2593.

[67] Wakimoto, N., Nishi, J., Sheikh, J., Nataro, J. P., Sarantuya, J. A. V, Iwashita, M., Manago, K., et al. (2004). Quantitative biofilm assay using microtiter plate to screen for enteroaggregative Escherichia coli. *71*(5), 687–690.

[68] Ma, Q., Yang, Z., Pu, M., Peti, W., & Wood, T. K. (2011). Engineering a novel c-di-GMP-binding protein for biofilm dispersal. *Environmental microbiology*, *13*(3), 631–42.

[69] Heydorn, a, Nielsen, a T., Hentzer, M., Sternberg, C., Givskov, M., Ersbøll, B. K., & Molin, S. (2000). Quantification of biofilm structures by the novel computer program COMSTAT. *Microbiology (Reading, England)*, *146 (Pt 10)*, 2395–407.

II.3. 3D Structure determination of [C117S] YmoB by NMR

II.3.1. INTRODUCTION

The NMR structure determination of a protein requires the preparation of (uniformly $^{13}C/^{15}N$-labeled) soluble protein, the acquisition of a set of 2D and 3D NMR experiments, NMR data processing, chemical shift assignment, NOE assignment and collection of conformational restraints, structure calculation, refinement and validation[70].

Although a variety of NMR parameters contain structural information (as hydrogen bonds or torsion angles about chemical bonds), the key conformational data are upper distance limits derived from NOEs[71]. NOEs result from cross-relaxation due to the dipole–dipole interactions between nearby pairs of nuclear spins in a molecule undergoing Brownian motion, and in two-dimensional (2D) or higher-dimensional heteronuclear-resolved [^{1}H,^{1}H]-NOESY spectra they are manifested by cross-peaks. In order to extract informative distance constraints from a NOESY spectrum, its cross-peaks have to be assigned, i.e. the pairs of hydrogen atoms that give rise to the observed cross-peaks need to be identified. These NOESY assignments are based on ^{1}H chemical shift values that result from previous sequence-specific resonance assignments. Resonance assignment is the process of identifying the exact resonance frequency of each ^{1}H, ^{13}C and ^{15}N nucleus of the protein amino acids. To obtain the complete assignment of [C117S] YmoB, collect the structural restraints and obtain the 3D structure of the protein, we organized this work in four main blocks: backbone assignment, side chain assignment, collection of structural restraints and structure calculation, refinement & validation.

II.3.2. BACKBONE ASSIGNMENT

To assign the H-N backbone atoms we acquired a set of 2D and 3D NMR experiments listed in Table 2. We assigned the H-N atoms to their corresponding resonances using ^{1}H-^{15}N-HSQC spectrum. With the HNCO/HNCACO experiments we assigned C' atoms and linked them to its neighbor H-N resonances.

Experiment	Dimensionality	Observed Spins
^{1}H-^{15}N-HSQC	2D: ^{1}H-^{15}N	H_i-N_i
HNCO	3D: ^{1}H-^{15}N-^{13}C	H_i-N_i-$C'O_{-i}$
HNCACO	3D: ^{1}H-^{15}N-^{13}C	H_i-N_i-$C'O_{-i}$-$C'O_i$
CBCAcoNH	3D: ^{1}H-^{15}N-^{13}C	H_i-N_i-C^{α}_{-i}-C^{β}_{-i}
CBCANH	3D: ^{1}H-^{15}N-^{13}C	H_i-N_i-C^{α}_i-C^{β}_i-C^{α}_{-i}-$C\beta_{-i}$

Table 2: NMR experiments recorded in order to assign the backbone atoms H-N-C' and the side chain atoms C^{α}/C^{β}.

With CBCAcoNH and CBCANH spectra we assigned the $C^{\alpha}_{-i}/C^{\beta}_{-i}$ and $C^{\alpha}_{i}/C^{\beta}_{i}$ resonances. We connected the backbone resonances with its side chain resonances and we performed an initial identification for the type of residue that a particular spin system belongs to (Fig 21). In order to identify each spin we compared the determined H-N-C'-C^{α}-C^{β} resonance positions with chemical shift tables listing the random coil chemical shifts of all 20 proteogenic amino acids[72].

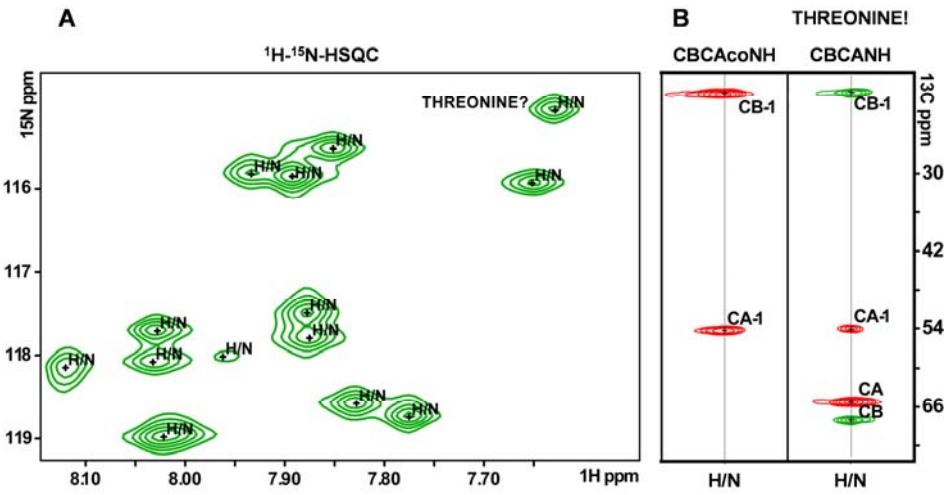


Fig 21: A: Expansion of ^{1}H-^{15}N-HSQC of 1.4mM [C117S] YmoB. All peaks were identified. **B:** CBCAcoNH (left) and CBCANH (right) strips corresponded to the H-N peak from A) labeled as THREONINE?. C^{α} and C^{β} chemical shifts allowed us to preliminary assign this spin system as a threonine by comparing them with tabulated random coil chemical shifts.

Typical, the range of chemical shifts for different amino acids show considerable overlapping. However, some of them present characteristic values, allowing their unambiguous identification to a specific type. Threonines (C^{α}~62ppm & C^{β}~69ppm), serines (C^{α}~58ppm & C^{β}~64ppm), alanines (C^{α}~53ppm & C^{β}~19ppm) and glycines (only C^{α}~45ppm) are the most suitable spin systems to be unequivocally identified at this first step.

To elucidate the sequential connection between neighbor spin systems we aligned CBCANH strips at forward and reverse directions using both the $C^{\alpha}_{-i}/C^{\beta}_{-i}$ and $C^{\alpha}_{i}/C^{\beta}_{i}$ resonances. Starting from the threonine assigned in Fig 21 we identified a short peptide formed by G-I-A-T-L (Fig 22). Only one combination of these amino acids matches with the [C117S] YmoB sequence, allowing their unambiguous assignment to the G25-I26-A27-T28-L29 tract. Following this strategy we assigned all backbone atoms (Fig 23) and also the side chain atoms C^{β}.

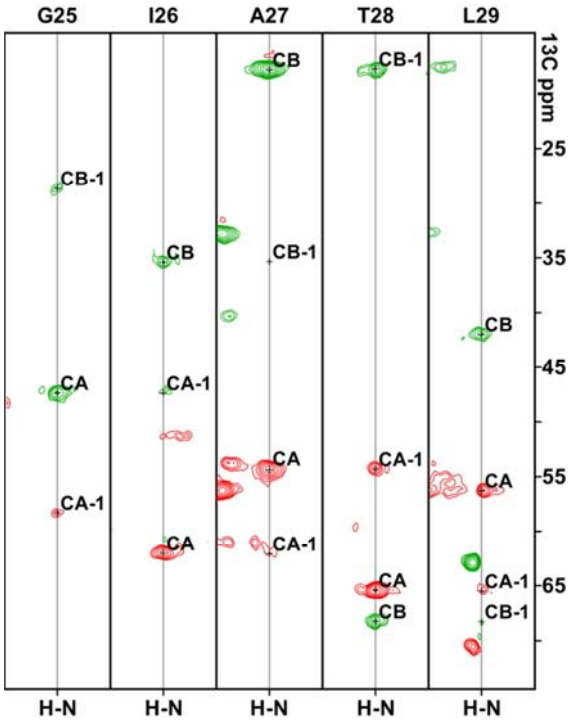


Fig 22: Alignment of CBCANH strips by matching $C^{\alpha}_{-i}/C^{\beta}_{-i}$ with $C^{\alpha}_{i}/C^{\beta}_{i}$ resonances. $C^{\alpha}_{-i}/C^{\beta}_{-i}$ peaks not observed at CBCANH strips were identified using CBCAcoNH experiment.

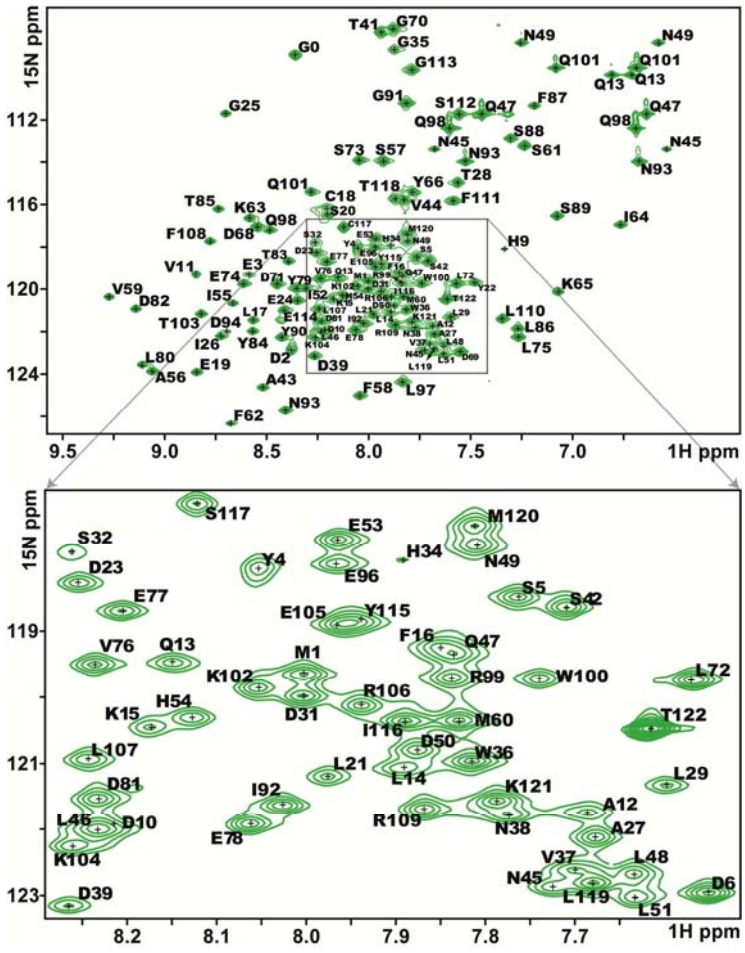


Fig 23: ^{1}H-^{15}N-HSQC of 1.4mM [C117S] YmoB mutant with the H-N peaks assigned.

II.3.3. SIDE CHAIN ASSIGNMENT

II.3.3.1. Assignment Strategy

We first assigned the H^{α}/H^{β} using HBHAcbcacoNH. With the C^{α}-H^{α}/C^{β}-H^{β} assigned we could connect the ^{1}H-^{13}C-HSQC resonances with the 3D HccH-&-HCCh-TOCSY NMR experiments (Fig 24). TOCSY experiments yield strips in which all side-chain hydrogen (HccH) or carbon (HCCh) resonances are visible. An example of the complete assignment of leucine 119 side chain atoms is shown in Fig 24.

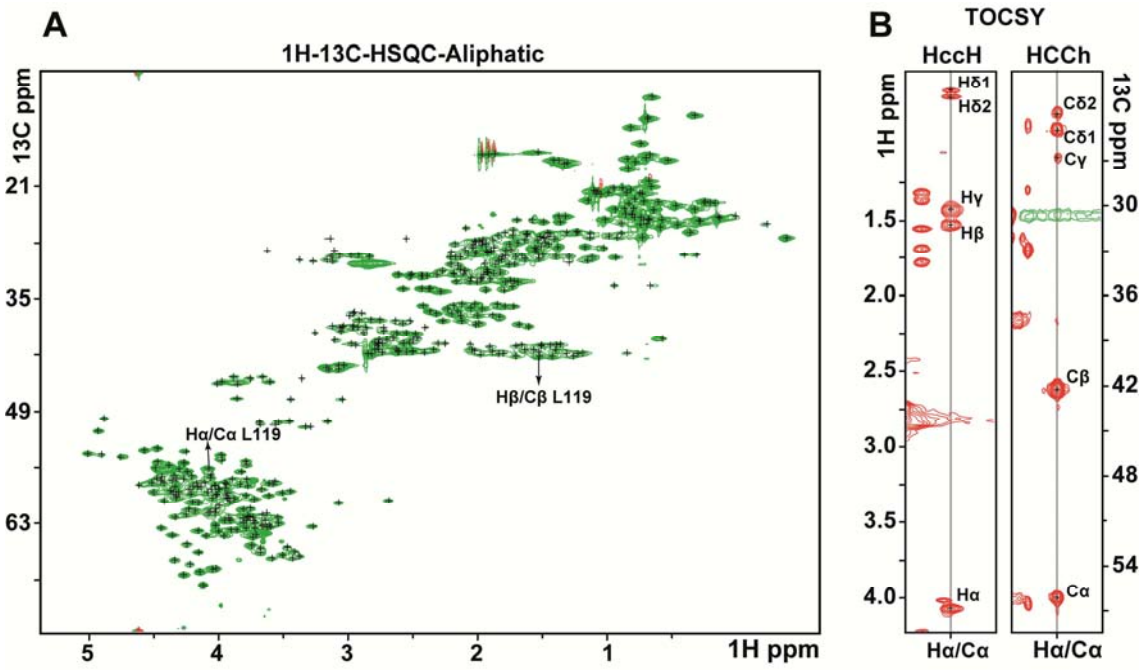


Fig 24: A: ^{1}H-^{13}C-HSQC of aliphatic proton region. C^{α}-H^{α}/C^{β}-H^{β} of L119 were identified using HBHAcbcacoNH and used to anchor the 2D ^{1}H-^{13}C-HSQC to the 3D-TOCSYs. **B:** Strips of HccH (left) and HCCh (right) TOCSY 3D experiments related to the C^{α}-H^{α} of L119 were used to the complete assignment of the side chain of this spin system.

Stereospecific assignment of the pro-chiral methyl groups of Val and Leu residues was achieved using a fractionally labeled sample[73]. In the 2D ^{1}H-^{13}C-HSQC-CT experiment the sign of a cross-peak depends on the number of directly coupled carbon nuclei (-1 for an even number of neighbors, +1 for an odd number). Due to the biosynthetic pathway in *E. coli,* protein expression systems grown on minimal media containing 10% [$^{13}C_6$] 90% [$^{12}C_6$]-glucose are always coupled to ^{13}C nucleus. As consequence, the pro-R methyl groups (γ^1 and δ^1 for V and L) lead to a negative sign of the associated cross-peak[74]. In Fig 25 we show the stereospecific assignment of the $C^{\delta 1}$-$H^{\delta 1}/C^{\delta 2}$-$H^{\delta 2}$ corresponding to the L119.

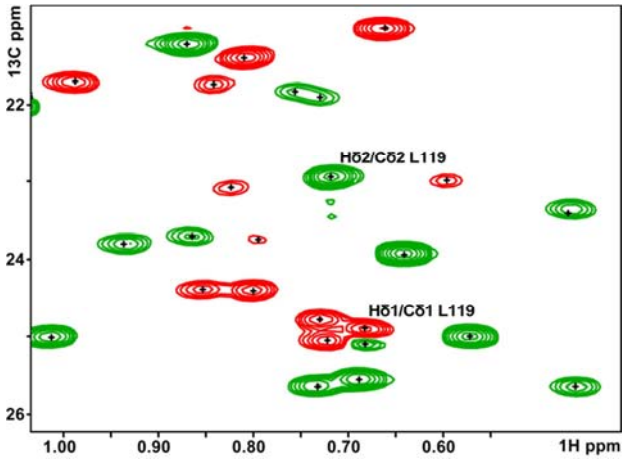


Fig 25: ^{1}H-^{13}C-HSQC-CT expansion of 10%-^{13}C labeled 1.4mM [C117S] YmoB sample. $C^{\delta1}$-$H^{\delta1}$/$C^{\delta2}$-$H^{\delta2}$ resonance positions for L119 were assigned using TOCSY experiments (Fig 4). To assign them stereospecifically, L119 $C^{\delta1}$-$H^{\delta1}$ was assigned to the peak at positive phase (red) and L119 $C^{\delta2}$-$H^{\delta2}$ to the peak at negative phase (green). Same strategy was followed for the stereospecific assignment of all the Val and Leu pro-chiral methyl groups.

Aromatic protons were assigned taking advantage of the characteristic pattern of the resonances that the different aromatic spin systems have. We used the 2D ^{1}H-^{13}C-HSQC-Aromatic and the 3D-^{1}H-^{13}C-HSQC-NOESY-Aromatic NMR experiment and the sequential connection of the spin systems previously elucidated.

II.3.3.2. NMR experiments acquisition using INTERLIVED-NUS[vii]-MDD[viii].

In spite of the improvement achieved by using the [C117S] mutant, YmoB samples still show some residual intermolecular interactions and slow evolution with time, probably associated to the formation of disulfide bonded dimers, even though reducing conditions used.

As around one week of recording time is needed for each 3D NMR experiment and the sample evolves with time, for each NMR experiment we had to prepare new fresh samples. Experiments acquired with independent samples shows slight variations in the observed chemical shifts (Fig 26A). Furthermore, the spectral evolution of the sample with time (Fig 26B) is incorporated in long 3D NMR experiments increasing the variations observed in the chemical shifts.

While these variations were not critical for the backbone assignment, they introduced inconsistences between side chain assignment and NOESY experiments, leading to a dramatic reduction on the measurable upper distance limits derived from NOEs.

[vii] NUS: Non-Uniform Sampling.

[viii] MDD: Multi-Dimensional Decomposition.

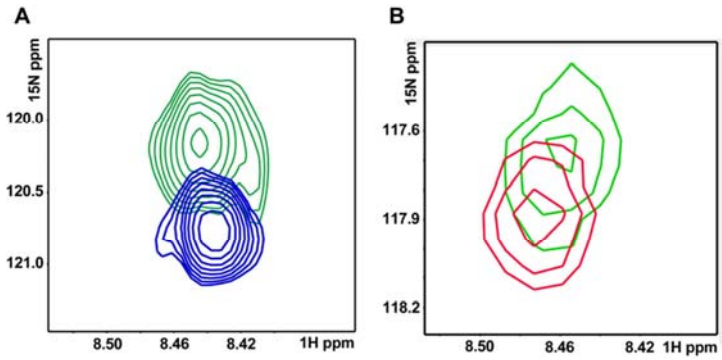


Fig 26: A: Superposition of the ^{1}H-^{15}N 2D projections from 3D CBCANH (green) and CBCAcoNH (blue) at 57.0ppm of ^{13}C. Peaks correspond to the same $^{13}C\alpha$ atom, which should present the same resonance, but there is a clear difference between the two independent experiments. **B:** Superposition of 1.2mM ^{1}H-^{15}N-HSQC [C117S] YmoB at time 0 (green) and after 15 days (red).

To avoid these problems and to obtain correct and coherent data valid for the structure calculation, we started a collaboration with the group of Vladislav Orekhov (Swedish NMR Centre, Goteborg University). They have developed a very efficient strategy for nD spectral acquisition that is especially suited to solve sample instability problems.

The basis of the strategy is the interleaved acquisition of a complete set of spectra, allowing the assignment and the acquisition of NOE data in the same sample and, effectively, at the same time. The interleave concept is applied not only to the acquisition of various experiments but also to the acquisition of the increments, that form the indirect detected dimensions.

Thus, instead of the classical approach in which the evolution times are sequentially increased by a fixed amount of time (the Dwell time), in the interleaved approach the various evolution times are acquired in an arbitrary order. Therefore, any slow evolution of the sample with time is not mistakenly incorporated as an additionally spectral evolution. The arbitrary acquisition of the evolution times is done using NUS[75], where the data is collected according to a user-specified sampling schedule[76] and also for only a subset of all incremented evolution periods, leading to an important saving in the acquisition time. To reconstruct the multidimensional NMR spectrum and obtain complete NMR data sets we applied MDD[77].

Using Interleaved-NUS-MDD we were able to acquire the NMR experiments listed at Table 3 using only one sample of [C117S] YmoB at lower concentration (1.0mM), in only seven days and alternating not only the different experiments acquired but also the increments of the indirect dimensions.

Following this strategy we assigned 92.4% of the non-labile protons (aromatic and aliphatic). Taking in account all the protons, we reached the 94.1% of the assignment.

Experiment	Solvent	Dimensionality	Observed Spins
^{1}H-^{15}N-HSQC	H_2O & D_2O	2D: ^{1}H-^{15}N	H_i-N_i
^{1}H-^{13}C-HSQC Aliphatic	H_20 & D_2O	2D: ^{1}H-^{13}C	H_i-C_i aliphatic
^{1}H-^{13}C-HSQC Aromatic	D_2O	2D: ^{1}H-^{13}C	H_i-C_i aromatic
^{1}H-^{13}C-HSQC-CT 10% [^{13}C]-glucose*	D_2O	2D: ^{1}H-^{13}C	H_i-C_i aliphatic
HBHAcbcacoNH*	H_20	3D: ^{1}H-^{15}N-^{1}H	H_i-N_i-$H\alpha_i$-$H\beta_i$-$H\alpha_{-i}$-$H\beta_{-i}$
HccH-TOCSY	D_2O	3D: ^{1}H-^{13}C-^{1}H	H_i-$H\alpha_i$-$H\beta_i$-$H\gamma$-$H\delta$-$H\varepsilon$
HCCh-TOCSY	D_2O	3D: ^{1}H-^{13}C-^{13}C	$C\alpha_i$-$C\beta_i$-$C\gamma$-$C\delta$-$C\varepsilon$
^{1}H-^{15}N-NOESY-HSQC	H_20	3D: ^{1}H-^{15}N-^{1}H	H_i-$H\alpha_i$-$H\beta_i$-$H\gamma$-$H\delta$-$H\varepsilon$
^{1}H-^{13}C-NOESY-HSQC Aliphatic	D_2O	3D: ^{1}H-^{13}C-^{1}H	H_i-$H\alpha_i$-$H\beta_i$-$H\gamma$-$H\delta$-$H\varepsilon$
^{1}H-^{13}C-NOESY-HSQC Aromatic	D_2O	3D: ^{1}H-^{15}N-^{1}H	H_i-$H\alpha_i$-$H\beta_i$-$H\gamma$-$H\delta$-$H\varepsilon$-HZ-HH

Table 3: NMR experiments recorded in order to assign side chains and obtain the structural restrains. Interleaved-NUS-MDD acquisition technology allowed us to record all the experiments listed (except those marked by *) using one [C117S] YmoB sample at 1.0mM in seven days of acquisition time.

II.3.4. COLLECTION OF STRUCTURAL RESTRAINS

In this study we used three types of structural restrains, that we obtained following this strategy; a) inter proton distances derived from NOEs measurements, b) torsion angles φ and ψ from ^{13}C, ^{15}N and ^{1}H chemical shifts c) hydrogen bonds determined experimentally and D) inter proton distance refinement.

II.3.4.1. Inter proton distances derived from NOEs measurements

To obtain the list of inter proton distances we used UNIO'10 software. This software combines the use of ATNOS[78] for the automatic NOE peak picking, CANDID[79] for the automatic NOE assignment and CYANA[80] for the structure calculation.

The input for ATNOS consisted of the amino acid sequence of the [C117S] YmoB, chemical shift lists from the sequence-specific resonance assignment, and the 3D NOESY spectra listed at Table 2. ATNOS performs 7 cycles of NOE peak identification in concert with automated NOE assignment with the software CANDID and protein structure calculation with the program CYANA. In the second and subsequent cycles, the intermediate protein structures are used as an additional guide for the interpretation of the NOESY spectra and for the validation of the NOEs identified with ATNOS and assigned with CANDID. The

output is a set of 20 preliminary models of [C117S] YmoB and a preliminary list of inter proton distance restrains. The whole process is schematically represented in Fig 27.

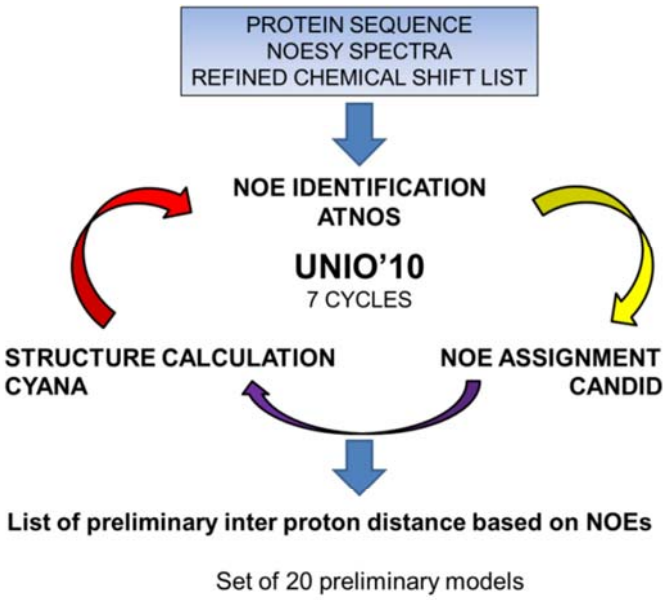


Fig 27: Scheme of the strategy followed to obtain the structural restraints for the calculation of the structure of [C117S] YmoB.

II.3.4.2. Torsion angles φ and ψ from ^{13}C, ^{15}N and ^{1}H chemical shifts

NMR chemical shifts in proteins strongly depend on local structure. The program TALOS+ establishes an empirical relation between ^{13}C, ^{15}N and ^{1}H chemical shifts and backbone torsion angles φ and ψ[81]. As inputs for TALOS+, we used the chemical shift values obtained from the assignment of [C117S] YmoB together with a preliminary set of 20 lowest energy models of the structure calculated by using only inter proton distances as restrains. We obtained 198 structural restrains, 99 φ and 99 ψ torsion angles.

II.3.4.3. Experimental determination of hydrogen bonds

In α-helix, protons from the backbone HN_i can form hydrogen bonds with the oxygen of the carbonyl O_{i-3} or O_{i-4}. H/D exchange experiments and 2D ^{1}H-^{15}N-HSQC of partially exchanged samples permit to identify protected protons that are involved in hydrogen bonds. We lyophilized a sample of 1.0mM [C117S] YmoB and we suspended it in D_2O until reaching the same concentration. We acquired a ^{1}H-^{15}N-HSQC spectrum after 4 minutes of sample preparation (Fig 28), and we identified 43 signals that did not disappeared.

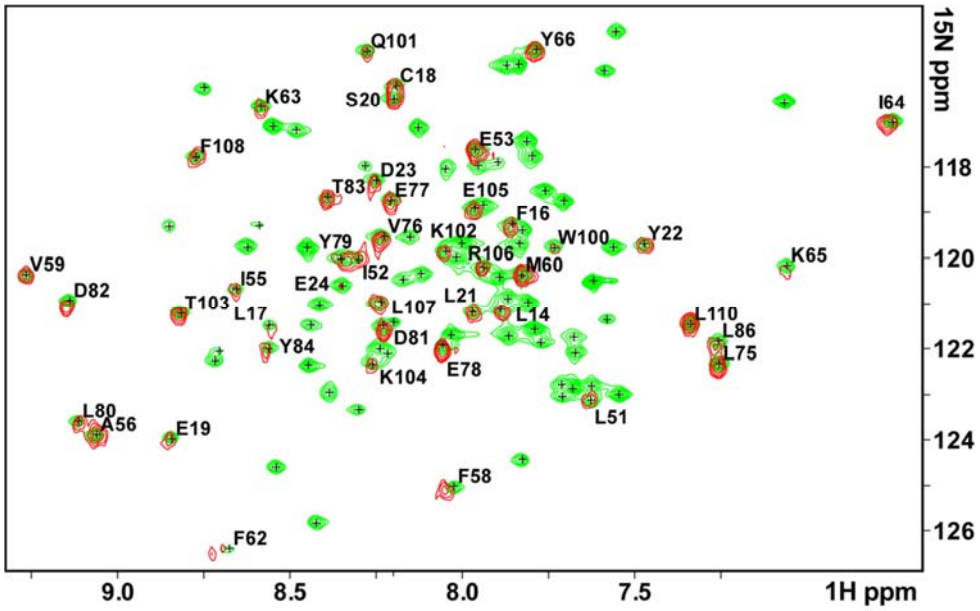


Fig 28: Superposition of ^{1}H-^{15}N-HSQC of [C117S] YmoB in water (green) and after lyophilizing and being suspended in D_2O, acquired after 4 minutes of sample preparation (red).

After assigning those signals to the corresponding amino acids, we used the lowest energy model from the 20 preliminary lowest energy models to identify the hydrogen bond partner (Fig 29). In the cases where the oxygen-donor amino acid was ambiguous the hydrogen bond was not taken into account. We obtained a final set of 36 hydrogen bond restrains.

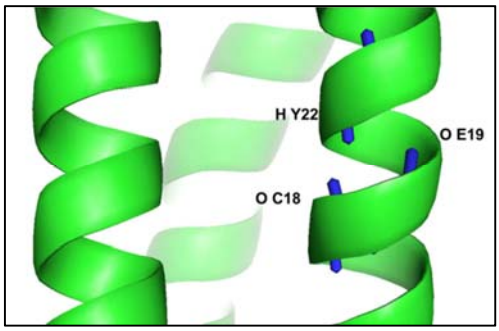


Fig 29: Identification of hydrogen bond partner of Y22. We highlight the amidic hydrogen of Y22 and the oxygen from carbonyl (O) of $E19_{i-3}$ and $C18_{i-4}$. The oxygen presenting a conformation compatible with hydrogen bond formation was O-C18.

II.3.4.4. Inter proton distance refinement

To improve the NOE assignment procedure necessary for obtaining the inter proton distances, torsion angles and hydrogen bonds can be added as structural restrains to UNIO'10. As the identification and assignment of NOEs is automatic, to validate the data and increase its robustness, we run 10 independent calculations with UNIO'10 changing only the seed generator number, *i.e.* the starting point that CYANA uses to generate the

structure calculation. If the assignment of the chemical shifts and the automated NOEs assignment are correct, all the calculations will rise to the same result, independently of the starting point that we choose (Fig 30).

For each independent calculation we obtained as ensemble of 20 models of the protein [C117S] YmoB. For each ensemble, we calculate the mean structure using the program MOLMOL[82]. We fitted the 10 mean structures and we calculated the root mean square deviation (RMSD). All 10 models presented exactly the same secondary structure. Taking in account all the atoms of the protein, the RMSD of the fitting was 2.23Å. This value decreased to 1.33Å when we took into account only the atoms forming the α-helices.

Also, for each independent calculation we obtained an inter-proton distances list. We retained the restraints from the list with the fewest restrains (1480) that were reproduced in at least seven of the all other nine lists to produce the final inter-proton distances list, which had a total of 1349 restrains.

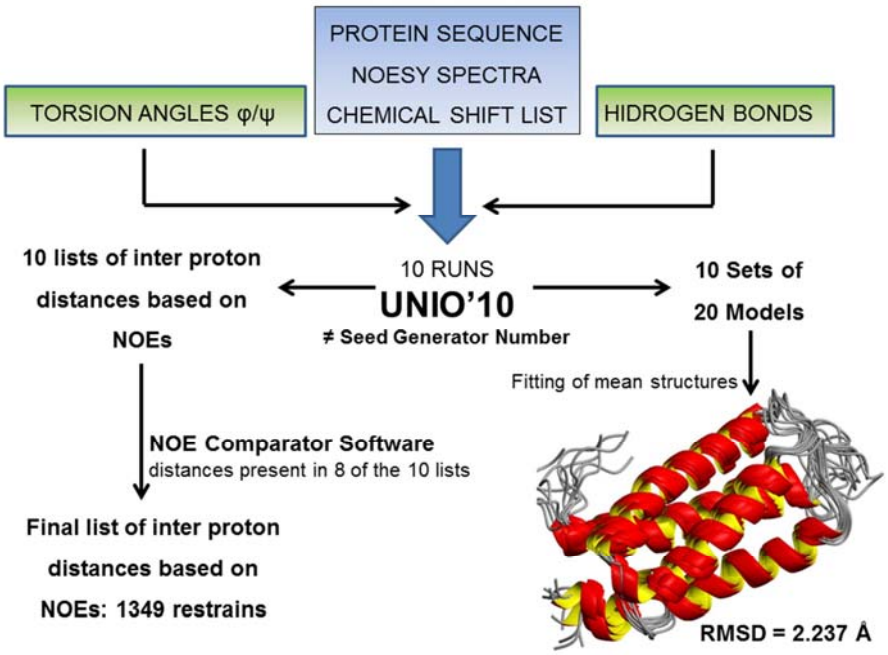


Fig 30: Scheme of the strategy followed for NOE-derived inter proton distances determination.

Summarizing, we obtained a final list of 198 torsion angles, 36 hydrogen bonds and 1349 inter proton distances as restrains for the structure calculation of [C117S] YmoB.

II.3.5. STRUCTURE CALCULATION, REFINEMENT & VALIDATION

The structure of [C117S] YmoB was determined by simulated annealing[83] using the torsion angle dynamics simulation program CYANA 2.1[80]. Further water refinement of [C117S] YmoB was performed using CNS 1.2.1[84]. An ensemble of 100 structures was

generated and the 20 lowest energy conformers were refined with explicit solvent. Models were validated using PROCHECK. The restraint and structural statistics are shown in Table 4. None of the structures contained distance or dihedral angle violations >0.5 Å or 5º, respectively.

NMR Distance and Dihedral Constrains	
NOE Restrains	**1349**
• Intra-residue	541
• Inter-residue	808
• Sequential	353
• Non-sequential	455
• Medium range	241
• Long range	214
Hydrogen bonds	**36**
Torsion Angle Restraints	**198**
• φ	99
• ψ	99
Total Restraints per Residue	13
Total Restraints per Residue in Structured Residues*	16
Structure Calculation Statistics	
Violations	
• Distance Constraints (Å)	0.0288 ± 0.0024
• Dihedral Angle Constrains (°)	0.2968 ± 0.0453
Average pairwise RMSD (Å) for All Atoms	
• Heavy atoms (Å)	2.62 ± 0.44
• Backbone atoms (Å)	1.89 ± 0.45
Average pairwise RMSD (Å) for Atoms in Structured Residues*	
• Heavy atoms (Å)	1.49 ± 0.12
• Backbone atoms (Å)	0.87 ± 0.14

Ramachandran Analysis	
Most favored regions	93.1%
Additional allowed regions	6.3%
Generously allowed regions	0.5%
Disallowed regions	0.2%

Table 4: NMR calculation and refinement statistics. * Structured Residues were considered the amino acids forming α-helices; K7-G30, A43-K65, G70-F87, D94-F111, I116-L119, a total of 87 amino acids from 122.

II.3.6. DESCRIPTION OF THE STRUCTURE

[C117S] YmoB is a monomer formed by four long α-helices (α1 K7-G30, α2 A43-K65, α3 G70-F87, and α4 D94-F111) and one short α-helix (S-α5 I116-L119) at the C-terminus (Fig 31 A). Helices are connected by four loops (L1 D31-S42, L2 Y66-D69, L3 S88-N93 and L4 S112-Y115). In Fig 31 B we represented the 20 lowest energy models of the [C117S] YmoB 3D structure by fitting the heavy atoms of the Structured Residues (Table 4). YmoB forms a four helix bundle arranged in an antiparallel manner (Fig 31, C) and stabilized by hydrophobic contacts as inferred by aromatic NOEs. The five helices fit with an RMSD under 1.5 Å (Table 4).

Loops L2, L3 and L4 are short (4 amino acids for L2, 6 amino acids for L3 and 4 amino acids for L4) and well defined. They fit with RMSDs of 1.97 $\pm$ 0.68Å, 1.41 $\pm$ 0.39Å and 1.02 $\pm$ 0.25Å, respectively. However, loop L1 is longer (12 amino acids) and heavily disordered, with a RMSD of 3.30 $\pm$ 1.22Å.

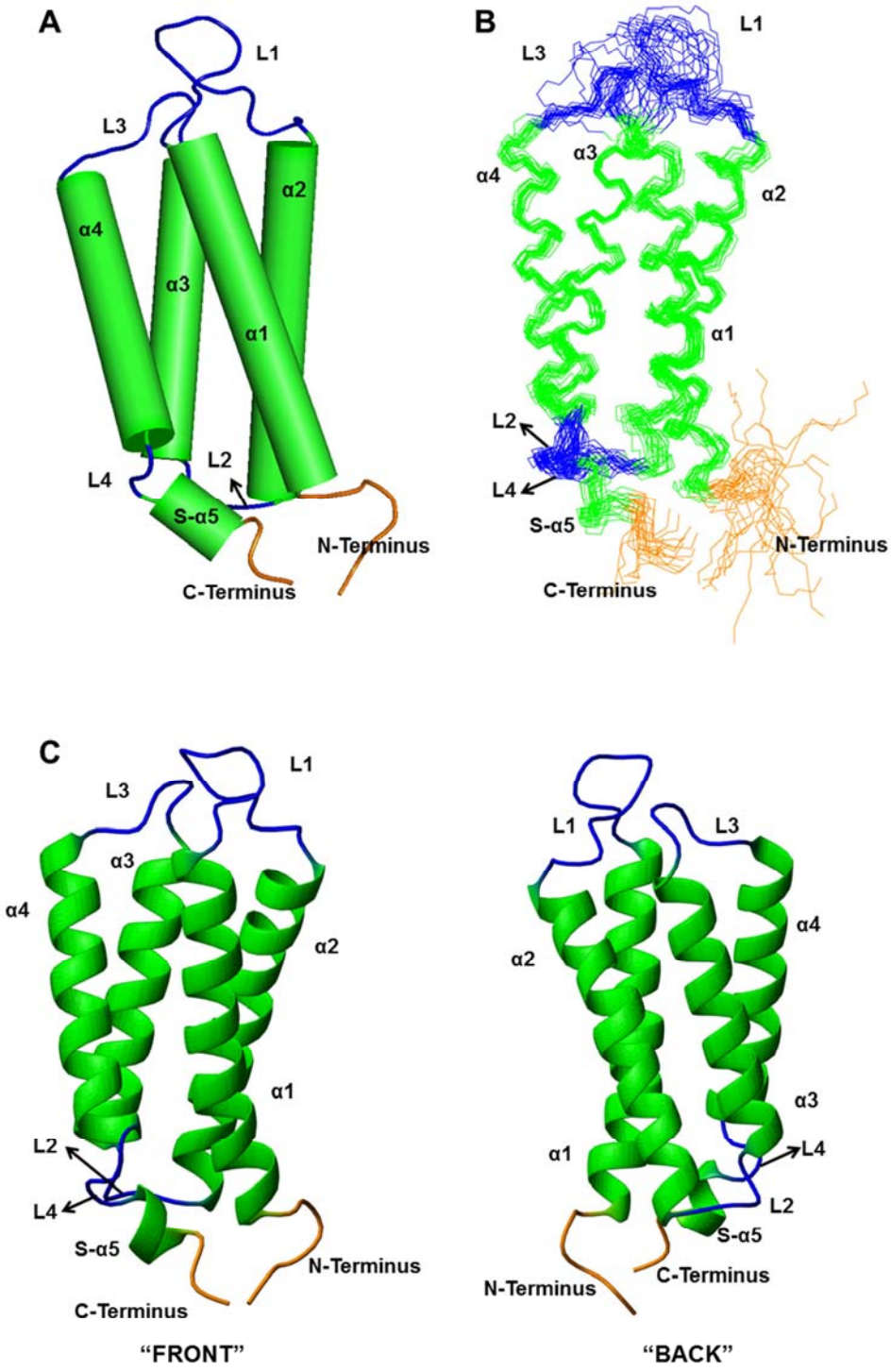


Fig 31: Different views of the 3D structure of [C117S] YmoB mutant. α-helices are represented in green, loops in blue and N- & C-terminus in orange. **A:** Cylindrical helix representation of lowest energy model. **B:** Fitting of the heavy atoms of the Structured Residues for the 20 lowest energy models. **C:** Representation of the secondary structure for the lowest energy model.

We calculated the electrostatic potential at the surface of [C117S] YmoB. Fig 32 shows that the electrostatic surface of the mutant is mainly negative. However we identified two positive patches: P1 (residues K7, R8, H9, K15 and k121) and P2 (R99, K102, R106, R109).

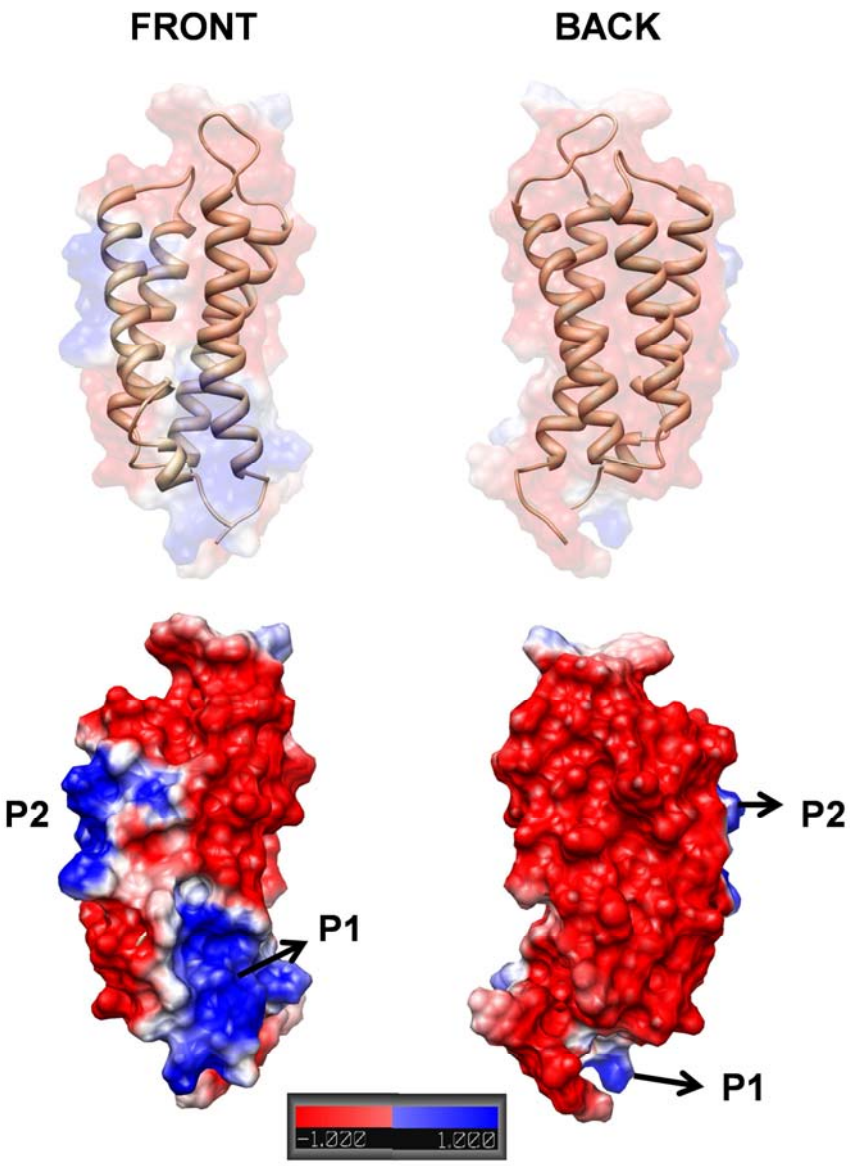


Fig 32: Electrostatic potential at the surface of [C117S] YmoB lowest energy model for "front" and "back" views.

II.3.7. CONCLUSIONS

[C117S] YmoB structure consists in a four helix bundle. We obtained a high resolution structure, in which the 20 lowest energy models present an average pairwise RMSD under 1.5Å for the heavy atoms of residues forming α-helices.

The protein is mainly negatively charged, but present two positives patches.

Cysteine 18 is totally buried into the structure. This amino acid does not contribute to the protein surface, as using PROCHECK[85] the calculation of the exposition of C18 to the solvent was 0%. These observations are in total agreement with our previous results on the alkylation of the cysteines, where only [C117S] was accessible to the alkylation reaction.

II.3.8. BIBLIOGRAPHY

[70] Wider, G. (2000). Structure Determination of Biological Macromolecules in Solution Using NMR spectroscopy, *1294*, 1278–1294.

[71] Herrmann, T., Güntert, P., & Wüthrich, K. (2002). Protein NMR Structure Determination with Automated NOE Assignment Using the New Software CANDID and the Torsion Angle Dynamics Algorithm DYANA. *Journal of Molecular Biology*, *319*(1), 209–227.

[72] Wuthrich, K. 1986. NMR of proteins and nucleic acids. John Wiley & Sons, Inc., New York.

[73]. Dario Neri, Thomas Szyperski, Gottfried Otting, Hans Senn, and Kurt Wiithrich. (1989). Stereospecific Nuclear Magnetic Resonance Assignments of the Methyl Groups. *Biophys, E. E. A. B*, 7510–7516.

[74] Klaus, W., Gsell, B., Labhardt, a M., Wipf, B., & Senn, H. (1997). The three-dimensional high resolution structure of human interferon alpha-2a determined by heteronuclear NMR spectroscopy in solution. *Journal of molecular biology, 274(4)*, 661–75.

[75] Sven G. Hyberts, Haribabu Arthanari, and Gerhard Wagner (2012). Applications of non-uniform sampling and processing. *Top Curr Chem., 316,* 125–148.

[76] Isabella C. Felli, Bernhard Brutscher (2009). Recent Advances in Solution NMR: Fast Methods and Heteronuclear Direct Detection. *ChemPhysChem 10,* 1356 – 1368.

[77]. Orekhov, V. Y., & Jaravine, V. a. (2011). Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Progress in nuclear magnetic resonance spectroscopy*, *59*(3), 271–92.

[78] Herrmann, T., Güntert, P., & Wüthrich, K. (2002). Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *Journal of biomolecular NMR*, *24*(3), 171–89.

[79] Herrmann, T., Güntert, P., & Wüthrich, K. (2002). Protein NMR Structure Determination with Automated NOE Assignment Using the New Software CANDID and the Torsion Angle Dynamics Algorithm DYANA. *Journal of Molecular Biology*, *319*(1), 209–227.

[80] Güntert, P. (2004). Automated NMR structure calculation with CYANA. *Methods in molecular biology (Clifton, N.J.)*, *278*, 353–78.

[81] Yang Shen, Frank Delaglio, Gabriel Cornilescu, and Ad Bax. (2009). TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of Biomolecular NMR, 44(4): 213–223.*

[82] Koradi R, Billeter M, Wüthrich K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph. 14(1): 51-5, 29-32.*

[83] Nilges, M., Clore, G. M., & Gronenborn, a M. (1988). Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. Circumventing problems associated with folding. *FEBS letters*, *239*(1), 129–36.

[84] Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-kunstleve, R. W., Jiang, J., Kuszewski, J., et al. (1998). Crystallography & NMR System : A New Software Suite for Macromolecular Structure Determination. *Acta Cryst. D54: 905–921.*

[85] Laskowski R A, Rullmannn J A, MacArthur M W, Kaptein R, Thornton J M (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*, 8, 477-486.

II.4 Oxidation and oligomerization of YmoB: a structure-driven hypothesis for YmoB function

II.4.1. INTRODUCTION

Functional studies presented in Chapter II.2 demonstrate that YmoB acts as an antitoxin, compensating the toxic effect of Hha overexpression. Antitoxins tightly regulate the expression of the lethal toxin genes at both transcriptional and post-transcriptional levels[86]. Under a variety of stress conditions, modifications of the antitoxin impair its repressor capacity, releasing the toxin[87]. During the process to obtain a stable sample needed to solve the structure of YmoB we observed that YmoB and its *E. coli* homologue TomB are indeed labile proteins. TomB showed spontaneous proteolysis of a C-terminal fragment and both proteins formed oligomers involving cysteine oxidation. A mutant of YmoB containing a single cysteine residue at position 18 was relatively stable; nevertheless this construct required a special fast NMR data acquisition protocol to obtain a consistent set of experiments needed to determine its three dimensional structure, suggesting that this construct was still undergoing slow modifications. The remaining C18 and its immediate environment are conserved between YmoB and TomB. This residue is buried in the hydrophobic core of [C117S] YmoB (Fig 33).

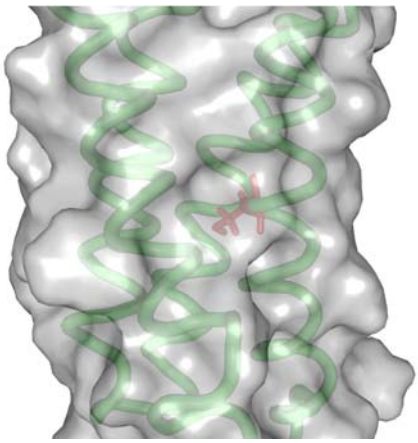


Fig 33: Expansion of [C117S] YmoB 3D structure. Cysteine 18, highlighted in red, is completely oriented to the center of the 4-helix bundle and totally protected by the tertiary structure of the protein.

In Chapter II.1 we described that only one of the two cysteine residues of WT YmoB could be modified by treatment with sodium iodoacetate (Fig 34A). The lack of reactivity of Cys18 towards sodium iodoacetate was confirmed in the [C117S] mutant (Fig 34B).

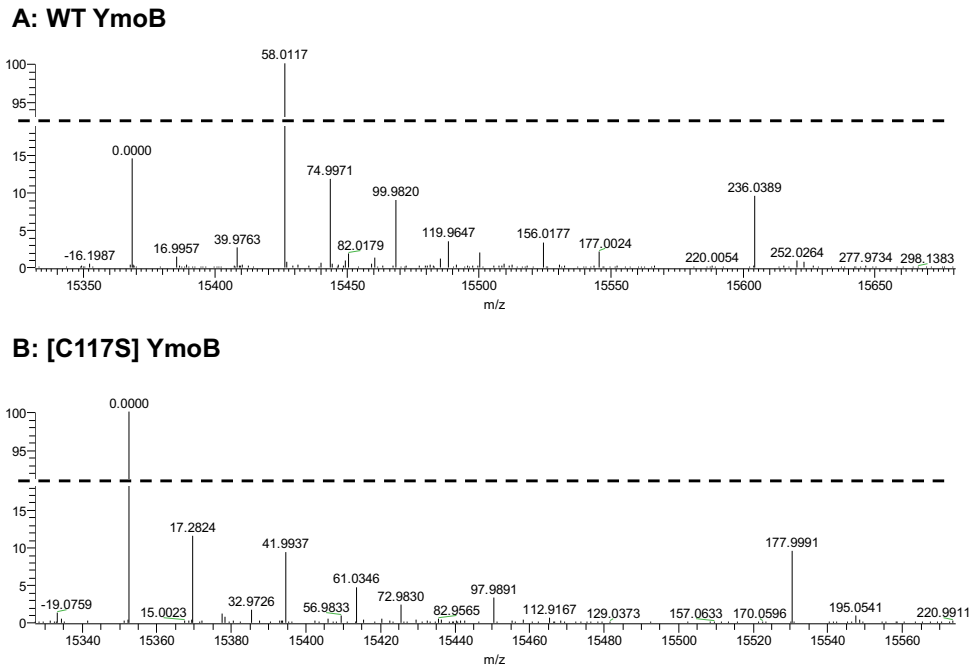


Fig 34: Mass spectra of YmoB constructs after alkylation reaction with sodium iodoacetate (58 atomic mass unit). Peaks are labeled with the mass increase with respect to the unmodified protein. **A:** The 58 atomic mass unit increment in WT YmoB shows that only one of the two cysteine residues was alkylated. **B:** The [C117S] mutant, containing only C18 was not alkylated.

We hypothesized that the conserved, buried cysteine may be functionally relevant and we decided to investigate in more detail the reactivity of this residue and the oligomerization of YmoB.

II.4.2. OXIDATION AND OLIGOMERIZATION STUDIES

The formation of oligomers by YmoB and [C117S] mutant was initially studied by native PolyAcrylamide Gel Electrophoresis (native PAGE) for diluted samples (μM range) and Size Exclusion Chromatography (SEC) for concentrated ones (mM range). Fig 35 shows native PAGE of YmoB and the [C117S] mutant incubated for 10 days in TCEP-free buffer (Fig 35 A) and in the presence of 1mM TCEP (Fig 35 B).

After incubation in the absence of reducing agents, WT YmoB was mainly converted to dimer and higher molecular weight oligomers even at the lowest protein concentration tested (10μM). Dimers were also detected in the presence of 1mM TCEP in a protein-concentration dependent manner. However, the presence of reducing agents was very efficient in preventing the formation of larger oligomers.

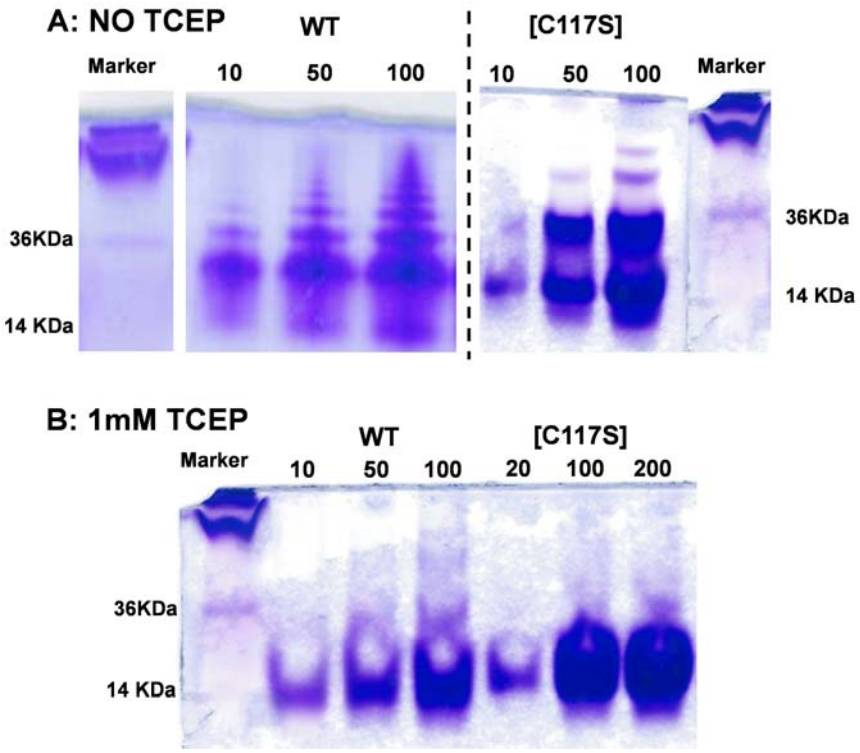


Fig 35: Native-Page of WT and [C117S] YmoB samples incubated for 10 days, at increasing concentrations (μM range) of protein. **A:** In absence of reducing agents WT formed a series of oligomers. [C117S] formed predominantly dimers, although higher molecular weight species were also observed at higher concentrations. **B:** Dimer species were observed even in the presence of 1mM TCEP, especially at high concentrations.

The [C117S] mutant required much higher concentrations to form small amounts of dimer in the presence of reducing agents. Under non-reducing conditions the relative amounts of monomer and dimer remained similar between 50 and 100 μM. The formation of dimer species decreased in the presence of reducing agents, confirming that oligomer formation involves disulfide bonds. In the case of [C117S], covalent dimers have to be formed via cysteine 18, which is protected by the tertiary structure of the protein. Thus, a conformational change has to occur to expose C18 and allow the disulfide bond formation. A small fraction of higher molecular weight species was observed at high concentrations of [C117S] YmoB under non-reducing conditions. Those species may be formed by additional interactions involving the covalent dimer, as they are not observed in the presence of TCEP even at higher protein concentrations.

Concentrated (1.5mM) WT and [C117S] YmoB samples were incubated at room temperature for 10 days in the presence and in the absence of a reducing agent (TCEP) and were analyzed by size exclusion chromatography (Fig 36).

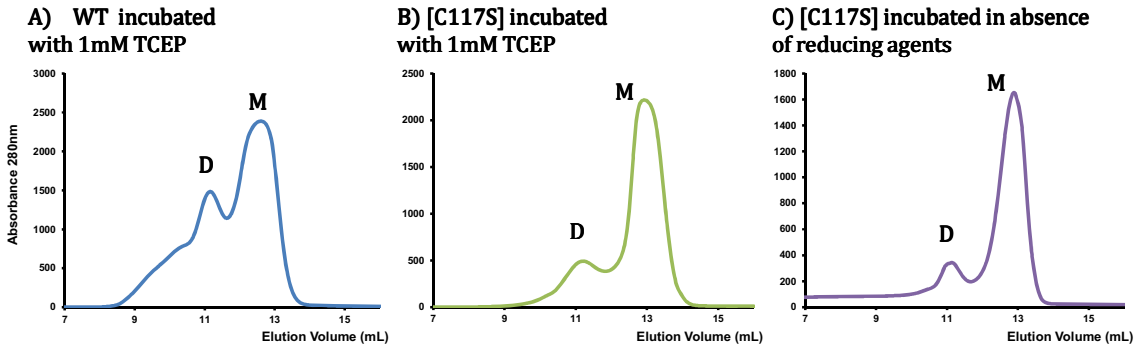


Fig 36: SEC profiles corresponding to the separation of 1.5mM samples incubated for 10 days at 25ºC in the presence 1mM TCEP for WT (**A**) and [C117S] (**B**) and in the absence of TCEP for [C117S] (**C**). D corresponds to dimeric form and M to monomeric form of YmoB constructs.

WT YmoB SEC profile reflects the formation of dimer species but also the formation of a faster eluting species, which appears between 8.60mL and 10.30mL of elution volume (Fig 36A). Probably, this higher weight species involve the two cysteines present, as they could not be detected by SEC in the single cysteine mutant (Fig 36B). We hypothesize that they could correspond to alternative dimer forms with different apparent volumes or to higher oligomers. For the [C117S] construct only a single additional species was observed, consistent with the formation of a dimer involving the single cysteine present in this construct. The covalent nature of the WT and [C117S] dimer species was confirmed by mass-spectrometry (see below). TCEP is a very efficient reducing agent used to prevent the formation of disulfide bonds. However, for YmoB constructs it failed to prevent the formation of dimers or even, for WT, higher molecular weight species.

In the absence of reducing agents, the relative integrals of dimer and monomer peaks for [C117S] construct were only slightly larger than those observed in the presence of TCEP. Instead, in the absence of reducing agents WT YmoB formed large oligomers, as seen by native PAGE (Fig 35A).

These results uncover a very unusual oxidation behavior for WT and [C117S] YmoB constructs: dimers are formed in similar amounts in the presence and in the absence of TCEP. The similar reactivity of the mutant and WT species in the presence of TCEP suggests that the dimers formed by WT YmoB under these conditions involve C18.

These data indicate that the participation of C18 in disulfide bond formation is the rate-limiting step and is mostly independent from the presence of TCEP. The limiting role of C18 is indeed suggested by its location within the protein interior (directly observed in the case of [C117S] mutant) and its lack of reactivity towards iodoacetate (in both constructs, WT and [C117S]).

In order to separate the dimer species from the monomer we concentrated and reinjected the material collected from the dimer peak. The process was repeated twice and the

corresponding SEC chromatograms are presented in Fig 37. Comparison of WT and [C117S] mutant in the presence of TCEP shows that a much higher proportion of TCEP-resistant dimer is formed in the WT species. Dimer to monomer ratio (D/M) at 3th SEC was 3.7 for WT YmoB and 0.94 for the mutant. Since the retention times of dimer and monomer are the same for WT and [C117S] YmoB, the difference in the relative (D/M) proportions cannot be explained by differences in the purification process. It must reflect the fact that the C18-C18 dimer is more labile in the presence of TCEP than the dimer(s) formed by WT YmoB.

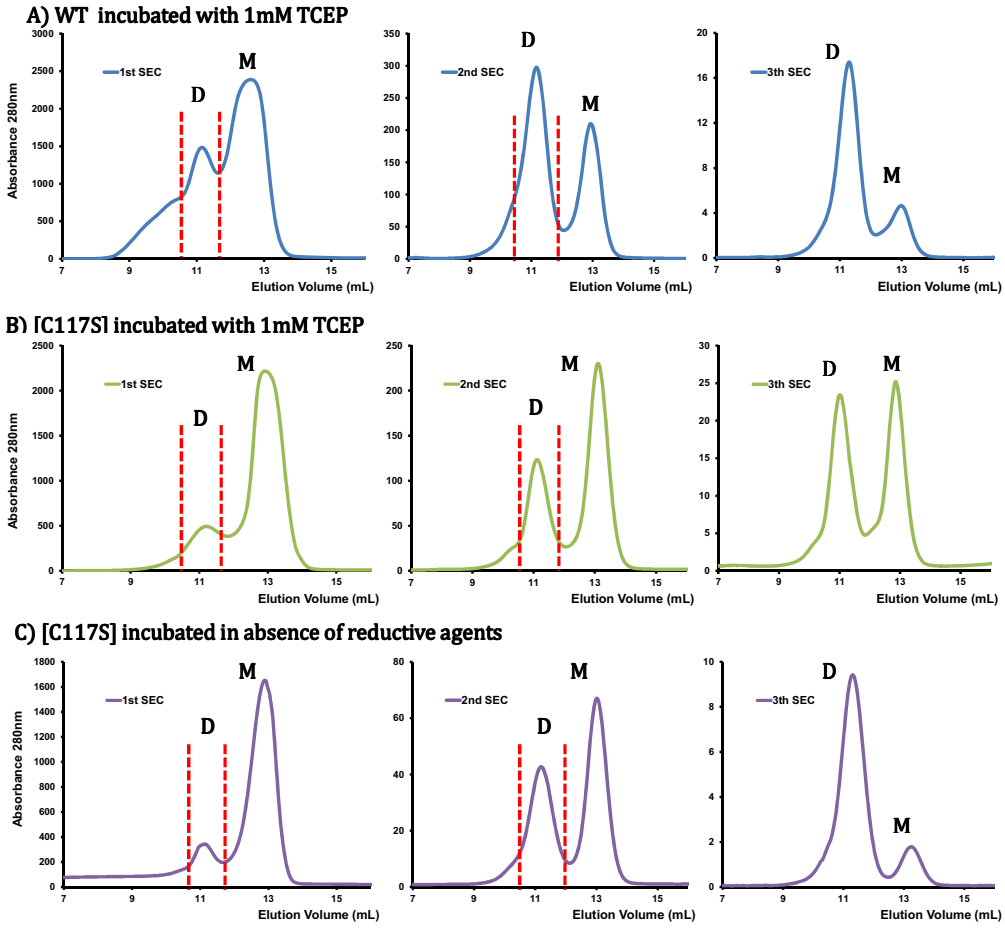


Fig 37: Profiles of consecutive size exclusion chromatography (SEC) assays. 1.5mM samples were used at the first injection. Fractions highlighted by red dashed bars were injected at the next SEC step. **A:** WT YmoB sample incubated and purified in presence of 1mM TCEP. **B:** [C117S] YmoB sample incubated and purified in presence of 1mM TCEP. **C:** [C117S] YmoB sample incubated and purified in absence of reducing agents.

Following this argument, the major dimer species formed by WT YmoB is probably the C18-C117 heterodimer, although the presence of other dimer species cannot be ruled out. Unfortunately, even in the absence of TCEP, no pure dimer could be isolated from the [C117S] YmoB sample. Although the dimer species is clearly more stable under non-

reducing conditions, the observation of a persistent monomer peak suggests that the C18-C18 dimer could also be broken under oxidative conditions (Fig 37 C). Thus, disulfide bonds involving C18 could break, in a TCEP-independent manner, leading to an equilibrium mixture between monomer and dimer.

In order to confirm the covalent structure of the proteins eluting as monomer and dimer, we studied the two peaks resolved in SEC by Liquid Chromatography Mass Spectrometry (LC-MS). The mass of the major products from the two peaks corresponded to the theoretical mass of the monomer (M) and the disulfide bonded dimer (2M-2). Interestingly, we also observed peaks with a 16 atomic mass unit (amu) increase, suggesting sulfur oxidation in species eluting as monomer or dimer species (Table 5).

Sample	Mw Measured	Modification
Monomeric fraction WT	15368,4	M
	15384,4	M+16amu
Dimeric fraction WT	30733,9	2M-2uma
	30750,5	2M(-2 +16)amu
	30762,9	2M(-4 +32)amu
Monomeric fraction [C117S]	15352,4	M
	15368,4	M+16aum
Dimeric fraction [C117S]	30781,1	2M(-2 +16)amu

Table 5: Mw measured in monomer and dimer fractions of 1.5mM WT YmoB and [C117S] YmoB mutant incubated during 10 days at 25°C and resolved by SEC. In the covalent dimer species 2amu corresponds to the loss of 2 protons during the disulfide bond formation. +16amu corresponds to the incorporation of an oxygen atom typical of sulfur oxidation. M represents the atomic mass of the monomer.

The SEC dimer fraction of WT YmoB contains a species with a molecular weight of 2M +28 amu. This species may be explained by the formation of two disulfide bonds and the incorporation of two oxygen atoms. Experimental data is shown in Suppl. Fig. 2.

WT YmoB contains 4 methionine and 2 cysteine residues. Methionine oxidation to sulfoxide is commonly observed. Instead, oxidation of cysteine to sulphenic acid (-SOH) in the absence of strong oxidants is far less common. Nevertheless, cysteine oxidation has been previously observed in other proteins in the presence of Reactive Oxygen Species[88] (ROS) and related to regulatory processes[89]. To identify the oxidized specie(s) in WT and [C117S], we analyzed the SEC fractions corresponding to dimer species using

endoproteinase GluC digestion followed by nanoHPLC-ESI-MS/MS (nano high performance liquid chromatography coupled to tandem mass spectrometry with nanoelectrospray ionization). The detected peptides are listed in Suppl. Table 1. Some of these peptides contained oxidized residues indicating that the monomer peak observed by SEC contains a mixture of species (Table 6). In addition to whose peaks indicating the presence of oxidized methionines, a fragment corresponding to VAQLKFLC(ox)E was unequivocally detected in WT YmoB and in the [C117S] mutant, confirming that C18 was, at least partially, oxidized to sulphenic acid. No oxidation of C117 was detected, indicating an unusual capacity of C18 to be sulphenilated in air.

Sample	Peptide Sequence	Modifications	MH+ [Da]
WT & [C117S] YmoB	SHHHHHHSMGMDE	M1(Oxidation)	1594,60085
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGD	M60(Oxidation)	2085,99808
WT & [C117S] YmoB	HIASFVM	M60(Oxidation)	820,40221
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGDLSE	M60(Oxidation)	2415,15224
WT & [C117S] YmoB	HIASFVMSF	M60(Oxidation)	1054,50260
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGDLSELVEE	M60(Oxidation)	2885,39540
WT & [C117S] YmoB	YICTLMKT	M120(Oxidation)	988,48466
WT & [C117S] YmoB	**VAQLKFLCE**	**C18(Oxidation)**	**1066,56046**

Table 6: Summary of oxidized peptides obtained after endoproteinase GluC digestion followed by nanoHPLC-ESI-MS/MS.

We searched for disulfide-bond cross linked peptides in the samples digested with endoproteinase GluC and analyzed by nanoHPLC-ESI-MS/MS. Searches were run for WT and [C117S] protein constructs taking into account the possible oxidation at cysteines and methionines, using two different cross link possibilities: disulfide bond and cysteine/oxidized-cysteine cross link. The relevant detected peptides are summarized in Table 7, while the complete list of cross-linked peptides is given as Supplementary Material (Suppl. Table 2).

For the [C117S] construct a single disulfide bond is possible and the detected C18-C18 peptide confirmed that the dimer observed by Native-Page and SEC is indeed a covalent dimer. For WT YmoB, peptides with C-C cross-links between C18-C18, C18-C117 and C117-C117 were detected. Thus, dimer peak resolved by SEC (Fig 37 A) corresponds to a mixture of species. Especially noteworthy is the observation of a cross-linked peptide with an unusual thiosulphinate cross link C-S(O)S-C between C18 and C117. The oxidation was always detected on C18 sulphur. No oxidized dimers involving equivalent cysteines (C18-C18 or C117-C117) were observed (Suppl. Table 2). Peptides containing thiosulphinate always presented methionine 120 oxidized. The C18-C117 crosslink was also observed with ordinary disulphide bonds and with peptides containing 120M oxidized or reduced.

Sample	Peptide 1	Peptide 2	Cross link	MW Measured	MW Calculated
[C117S]	VAQLKFLCE	VAQLKFLCE	C18-C18	2098,1107	2098,1075
WT	VAQLKFLCE	VAQLKFLCE	C18-C18	2098,1107	2098,1075
WT	VAQLKFLCE	YICTLMKT	C18-C117	2020,0184	2020,0316
WT	YICTLMKT	YICTLMKT	C117-C117	1941,9524	1941,9556
WT	**VAQLKFLC(ox)E**	**YICTLM(ox)KT**	**C18(O)-C117**	**2052,0164**	**2052,0214**
WT	VAQLKFLCE	YICTLM(ox)KT	C18-C117	2036,0174	2036,0265
WT	YICTLMKT	YICTLM(ox)KT	C117-C117	1957,9422	1957,9506
WT	YICTLM(ox)KT	YICTLM(ox)KT	C117-C117	1973,9384	1973,9455

Table 7: Summary of cross linked peptides obtained after GluC digestion followed by nanoHPLC-ESI-MS/MS using MassMatrix and Stavrox. Only cross-link identifications found by both search engines were considered as positive identifications.

The observation of the sulphenic acid form of C18 in the monomer and the thiosulphinate bond involving an additional oxidation level for C18, confirms the unusual reactivity of this buried cysteine residue. Since this residue is not accessible in the 3D structure of [C117S] YmoB, we hypothesized that non-covalent interactions may be involved in order to expose the buried cysteine side chain and/or bring together the molecules before a disulfide bond is formed.

To study the possible conformational changes of YmoB arising from intermolecular interactions, we compared ^{1}H-^{15}N-HSQC NMR spectra of two [C117S] samples prepared from the same protein batch but at different concentrations: 200µM and 1.2mM. We expected that self-association would result in shifts or changes in intensity of peaks coming from residues close to the dimer interface.

A plot of chemical shift differences and intensity ratios between concentrated and diluted samples along the sequence are shown in Fig 38 A & B. Intensity ratios were normalized to give an average value of 1. Residues with chemical shift changes larger than one standard deviation are M1, Y4, H9, V11, Q13, L17, E19, G25, D31, S32, H34, V37, N38, N45, N49, D50, V59, F62, Y66, D69, S73, E74, N93, D94, Q98, Q101, E105, I116, T118, M120, K121 and T122, which are mapped on a secondary structure representation (Fig 38 C) and on 3D structure models of [C117S] YmoB (Fig 38D). 15 of the 32 affected residues are located at helices α1, α2 and α4. Only two of the affected peaks are at helix α3, whereas the remaining 15 altered residues are located at the most flexible parts of the protein; L1, L3 and S-α5.

The location of residues affected by changes in concentration is consistent with the expected reposition of helix α1 required to allow a C18-C18 disulfide bond formation. This helix is connected to α2 by the long loop L1, which provides α1 with the flexibility needed for the conformational change, and as our data shows, is highly affected by the increasing

of the protein concentration. Reposition of helix α1 in the helix bundle is expected to affect the chemical shifts of the neighbor helices α2 and α4, also observed in our NMR data. Instead, helix α3 which is the one most distant from α1 should be the less affected, as observed experimentally, given that only two residues of helix α3 are affected by changes in the protein concentration.

Thus, reversible concentration dependence of the NMR spectra of [C117S] YmoB is an indication of non-covalent oligomerization. Instead, analysis of ^{1}H-^{15}N-HSQC NMR spectra of samples incubated during days at 25°C in a concentrated (1.2mM) and a diluted (300μM) state revealed slow irreversible changes, which were also dependent on the total protein concentration.

The analysis of the 1.2mM [C117S] YmoB ^{1}H-^{15}N-HSQC at time 0 and after 10 days (Fig 39) shows a number of changes in peaks corresponding to residues that are located in the proximity of C18 in the 3D structure (K15, Y22, H54, I55, V59, F111) or in flexible loops L1 (S32, D39) and L4 (E114). Only K65 is located in helix α2, but at the end of the helix next to L2. We did not observe apparent changes in C18 resonance, as it is overlapped with S20, which is not affected.

Residues K15, showing a frequency shift, and Y22, broadened beyond detection, form hydrogen bonds with the amidic proton and the oxygen from carbonyl of C18 (respectively). Residues I55 (shifted), V59 and F111 (broadened beyond detection), have atoms at less than 5Å from C18 (Fig 40). H54, which is also broadened beyond detection, is not directly close in the space to C18, but its neighbor residue, I55, is highly affected.

N-H signal of S32 is broadened in the concentrated sample at time zero but after incubation for 10 days it recovers the intensity observed in the diluted sample and D39 is broadened beyond detection. Both residues are located in the same long loop L1, which connect helix α1 to the rest of the molecule. In the opposite end of the protein, after long incubation K65 broadens beyond detection and E114 shows a chemical shit in one of the duplicated peaks. These residues are located close to M120, which was found to be oxidized in several proteolytic peptides detected by mass-spectrometry.

C18 is known to be partially oxidized to sulphenic acid under the conditions tested. Sulphenilated monomers may show spectral changes, either in residues that are experiencing a direct effect or by the perturbation of the hydrophobic core around C18 when the polarity of this residue is increased. The accumulation of affected residues in the proximity of this cysteine suggests that these residues (K15, Y22, H54, I55, V59 and F111) may be reporting on the effects of this oxidation process.

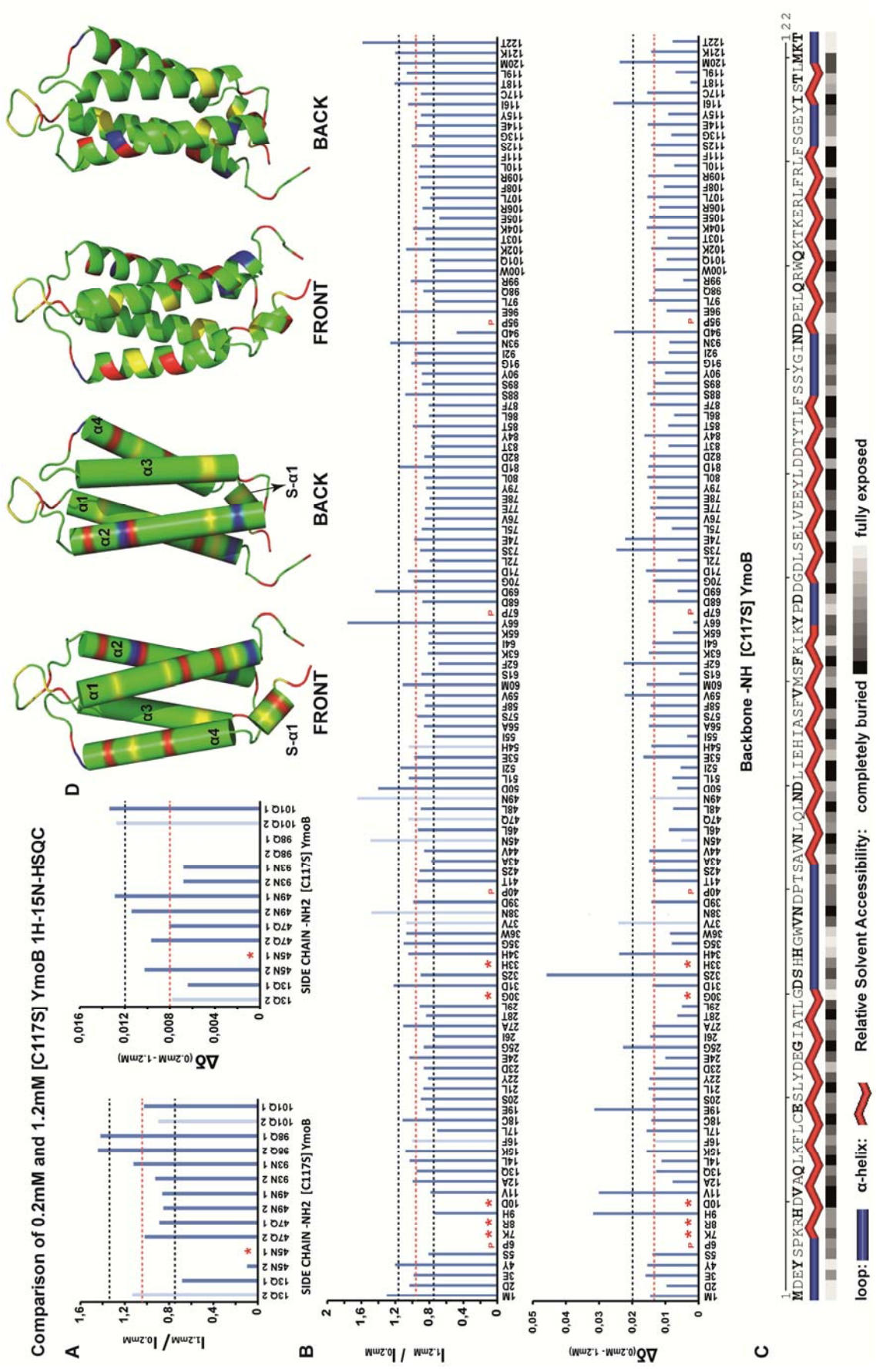

Comparison of 0.2mM and 1.2mM [C117S] YmoB 1H-15N-HSQC
A
B
C
D
SIDE CHAIN -NH2 [C117S] YmoB
Backbone -NH [C117S] YmoB
FRONT
BACK
S-α1
α1
α2
α3
α4
loop:
α-helix:
Relative Solvent Accessibility:
completely buried
fully exposed

Fig 38: Quantitative comparison of ^{1}H-^{15}N-HSQC spectra for ^{15}N-[C117S] YmoB samples at 200µM and 1200µM. **A:** Analysis of normalized intensity ratio (left) and chemical shift variation (right) for NH_2 side chain groups. **B:** Analysis of normalized intensity ratio (up) and chemical shift variation (down) for NH backbone groups. Means corresponding to intensity ratios and chemical shift variations are indicated by red dashed lines and black dashed lines show one standard deviation. Light blue bars correspond to signals with strong overlapping in the ^{1}H-^{15}N-HSQC spectra and asterisks show lack of information. **C:** Secondary structure representation of [C117S] YmoB. Residues with affected intensities or chemical shifts are shown in bold. **D:** 3D models of [C117S] YmoB structure where residues highlighted in red, yellow and blue indicate affected intensities, chemical shifts variations or both, respectively.

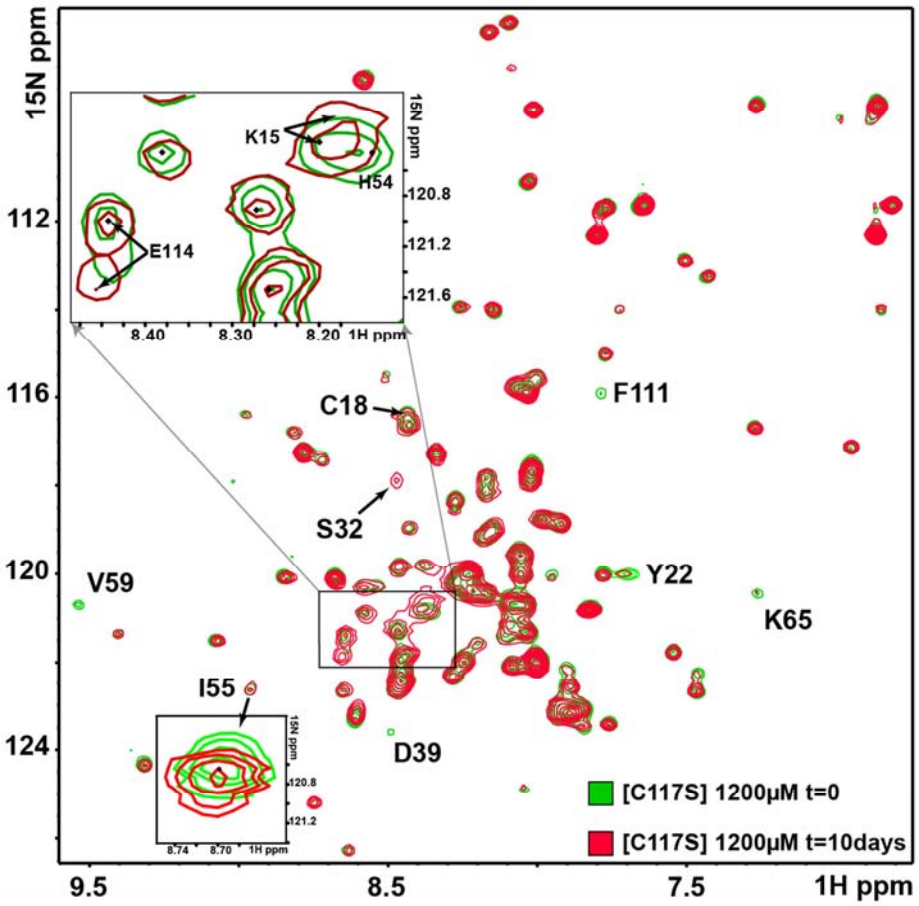


Fig 39: Superposition of ^{1}H-^{15}N-HSQCs of 1200µM ^{15}N-[C117S] YmoB. The same sample, at time 0 and after 10 days (incubated at 25°C) is shown in green and in red, respectively. Peaks K15, Y22, I55, V59 and F111 correspond to amino acids which are affected by C18 oxidation. S32, D39 and H54 are also highlighted as after 10 days they present chemical shift and/or intensity perturbations.

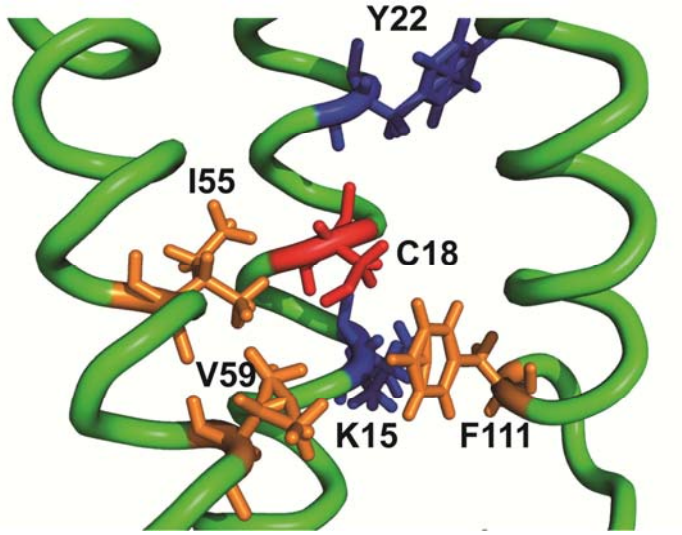


Fig 40: Expansion of C18 (red) amino acid and the affected neighbors due to cysteine oxidation. K15 and Y21 (blue) form hydrogen bonds with C18. I54, V58 and F111 (gold) have atoms at less than 5Å from C18.

The spectral changes observed after incubation of a concentrated sample were mostly absent when a four-fold diluted sample was subjected to the same conditions. Only S32 and D39 show a slight shift while the rest of the peaks remain unaltered after 10 days of incubation (Fig 41).

Thus, we concluded that the oxidation of C18 is a concentration dependent phenomenon. The possibility that the observed spectroscopic changes were due to the formation of dimer can be discarded by the small quantity of dimer detected in similar samples by SEC and by similar line width of the duplicated peaks. The experimental evidences suggest that the spectroscopic changes come from species with similar correlation times, probably C18 sulphenilated YmoB monomers.

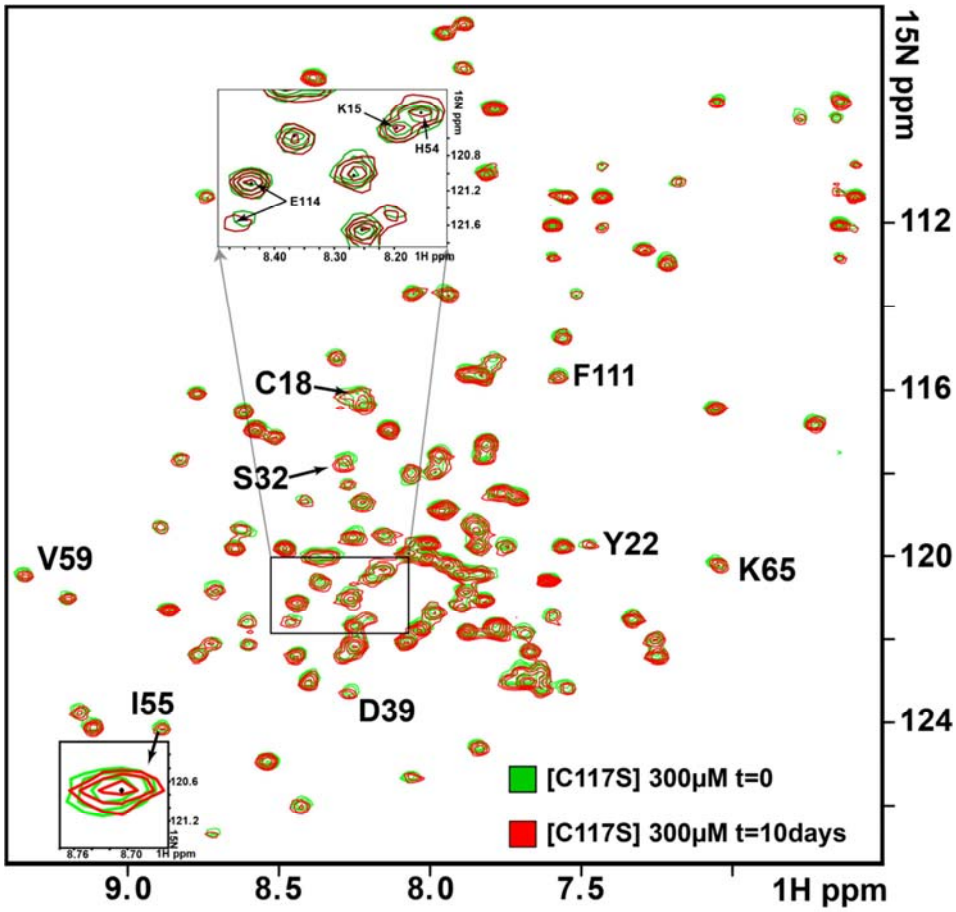


Fig 41: Superposition of ^{1}H-^{15}N-HSQCs of 300µM ^{15}N-[C117S] YmoB. The same sample, at time 0 and after 10 days (incubated at 25°C) is shown in green and in red, respectively. Only S32 and D39 resonances are slightly perturbated.

II.4.3. DISCUSSION

The results presented highlight the special reactivity of Cysteine 18 in spite of the fact that this residue is deeply buried in the tertiary structure of the protein. While this residue is not accessible to a small alkylating agent, such as iodoacetate, it can be oxidized to sulphenic acid (-SOH) and can form disulfide bonds with a second YmoB molecule. In WT YmoB disulfide bonds are formed with the two available cysteine residues while in the [C117S] mutant a C18-C18 bond is formed. The C18-C18 disulfide bond is more easily reduced by TCEP than the C18-C117 bond observed in WT YmoB. Another striking peculiarity involving C18 is the observation of a thiosulphinate bond C18-S(O)-S-C117.

Sulphenic acid formation is known to occur in a number of proteins and constitutes a regulatory mechanism[88], as previously shown in the case of phosphatases and other signaling molecules[90]. The oxidation of cysteine to the sulphenic acid form is usually associated to Reactive Oxygen Species (ROS)[91] but is unexpected for a buried cysteine in an isolated protein in buffer where the only oxidizing species is molecular oxygen. By comparing the signature of C18 oxidation in the NMR spectra after long incubation, we

observed that it occurs preferentially in concentrated samples, suggesting that it requires the presence of more than one protein molecule.

At short incubation times, the spectra of concentrated and diluted samples show differences, which indicate that in concentrated samples non-covalent interactions take place. About half of the residues showing concentration dependent changes in ^{1}H-^{15}N-HSQC spectra are distributed along helices α1, α2 and α4, i.e. the helix containing C18 and its two flanking helices. Other 15 affected residues are located in the flexible loops L1 and L3, on the top of the molecule as well as in the short flexible region near the C-terminus, where the one-turn helix S-α5 is located. The observation of NOEs between Y4/H33, S5/H33, R7/L118, D30/L118, S32/K120 and W36/L119, which are clearly incompatible with the structure of the [C117S] YmoB monomer, could be easily explained by the formation of transient antiparallel dimers.

The observed concentration dependence of the formation of C18 sulphenic acid could be explained either by non-covalent interactions modifying the environment of C18 and facilitating its oxidation or by the formation of covalent, disulfide-bonded, intermediate species. The observation of the unusual thiosulphinate bond between C18 and C117 in WT YmoB suggests that a covalent intermediate may be involved.

Thiosulphinate bonds can be produced by oxidation of disulfide bonds in the laboratory usually by the reaction with peroxides[92,93] and they are also present in natural products such as allicin, isolated from *Allium* vegetables[94]. Conversion of simple disulfides to thiosulphinates results in a considerable weakening of the S–S bond from about 70 to 34.5 kcal mol^{-1} (16.7 to 8.25 kJ mol^{-1}) for the S-S bond in PhS(O)SPh[95]. Thus, thiosulphinate bonds are weaker than disulfide bonds, and can split into monomers even in the absence of reducing agents, leading to sulphenilated cysteines[96].

A number of described reactions connect dimers bound by thiosulfinate or disulfide bonds and monomers containing cysteines with a reduced thiol or an oxidized sulphenic acid. Examples are shown in Fig 44. Most of these reactions have been characterized in small organic compounds. Disulfide bond and sulphenic acid formation are well known reactions in proteins[97]. Disulfide bonds are formed by oxidation with molecular oxygen while sulphenic acid formation is usually associated to the presence of strong oxidizing agents[98]. Thiosulphinates have been detected in oxidation processes involving glutathione[99], in oxidative stress response[100] and in the inhibition of platelet aggregation[94].

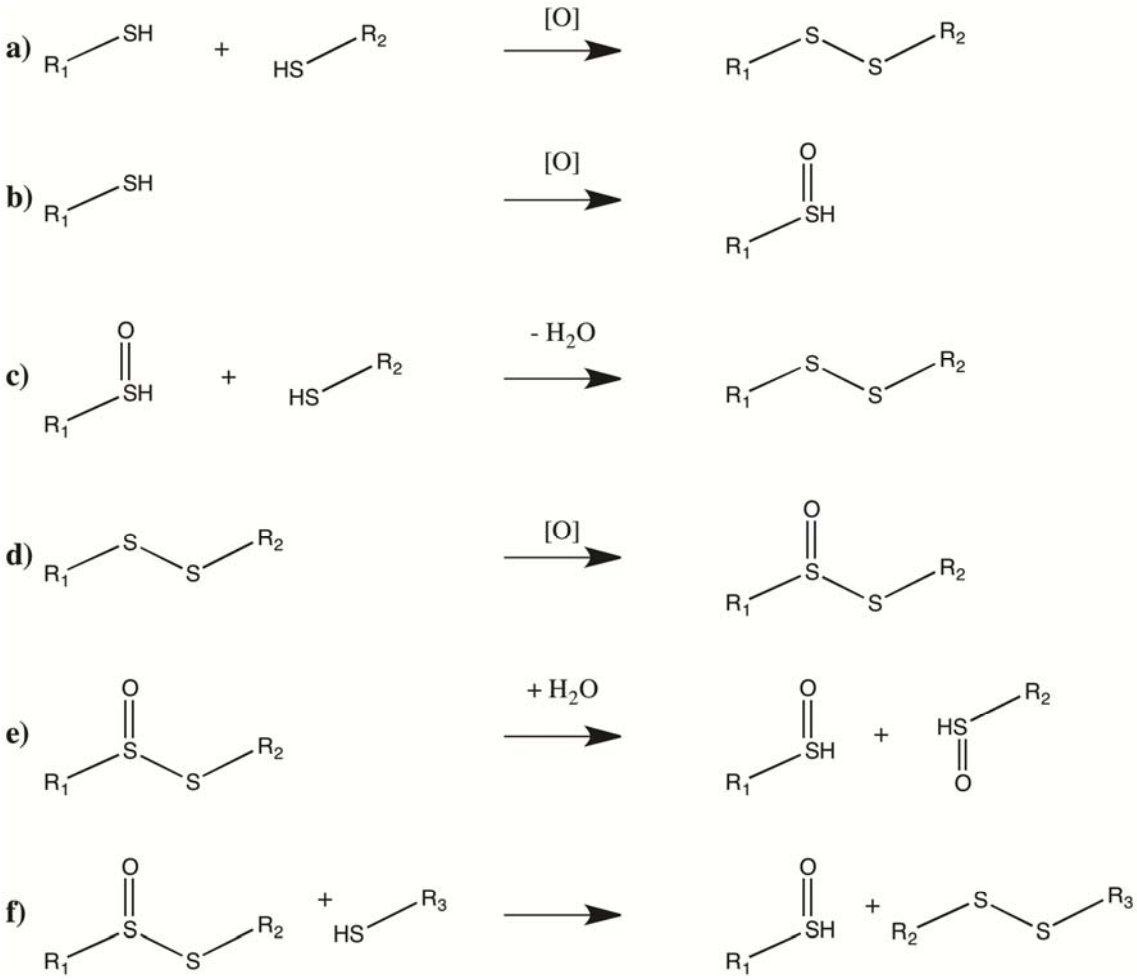
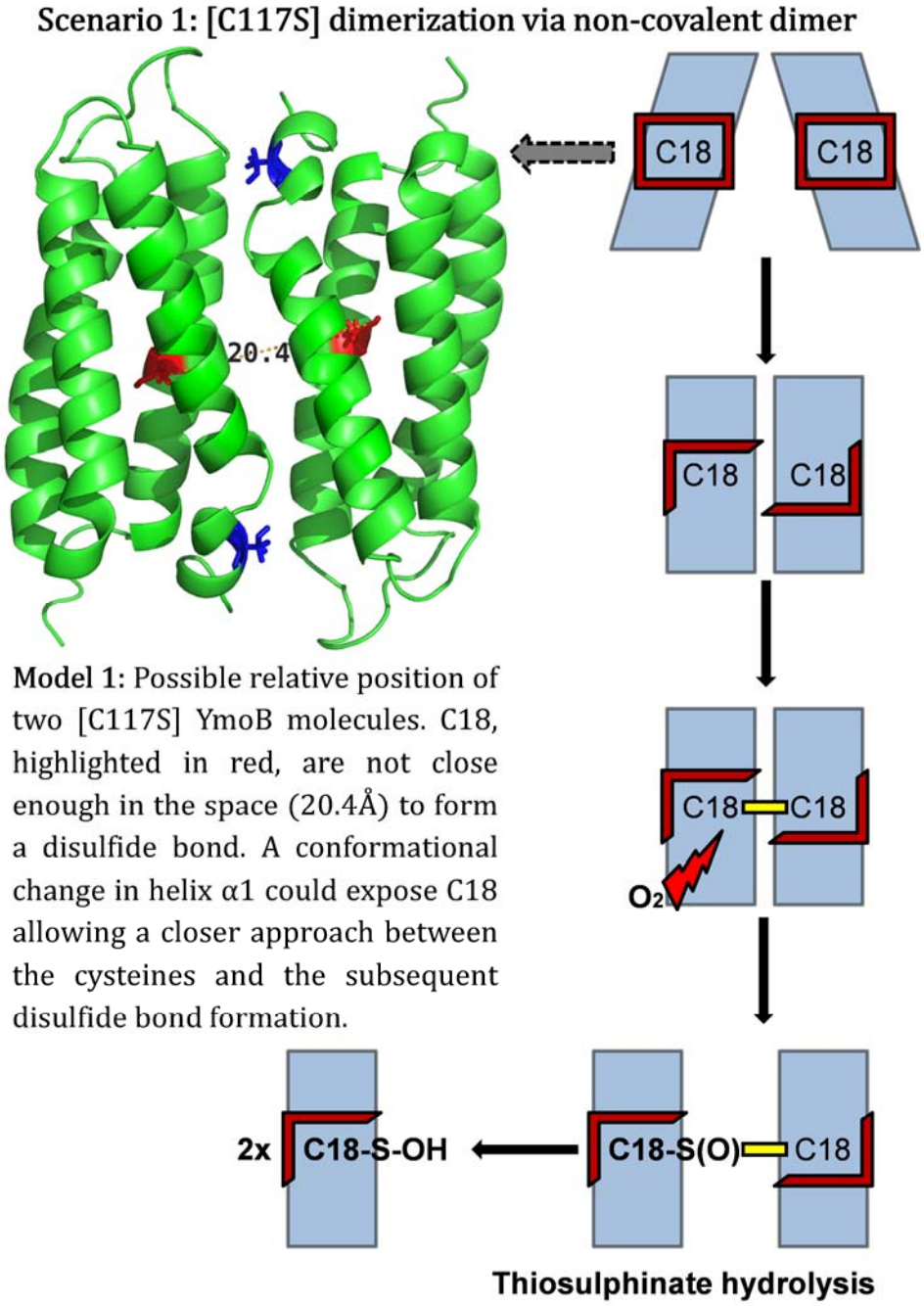


Fig 42: Chemistry of sulphur reactions. **A:** disulfide bond formation. **B:** sulphenic acid formation[101]. **C:** Disulphide bond formation from a sulphenic group and a sulfhydryl group[88]. **D:** Disulphide bond oxidation to thiosulphinate[102]. **E:** Hydrolysis of a thiosulphinate. **F:** Disulphide bond formation from a thiosulphinate and a thiol[96].

Considering the known reactivity of oxidized cysteine species, at least two mechanistic scenarios can be conceived for YmoB sulphenilation. The first one, called scenario 1 (Fig 43), starts with an initial non-covalent intermolecular interaction (non-covalent dimerization). This interaction leads to a distortion of the monomer structure, exposing C18 and, at the same time, bringing a second cysteine into its proximity, and facilitating the formation of the S-S bond. The disulfide bond formed between two conformational distorted molecules could be of high energy, e.g. because of steric repulsion forcing the S-S bond away from the minimum energy 90º C-S-S-C dihedral angle geometry. The distorted S-S bond could be oxidized to an C-S(O)S-C thiosulphinate group and eventually this bond could hydrolyze/break, leading to the formation of two sulphenic acids.

No thiosulfinate C18-S(O)-S-C18 bonds were observed in WT YmoB or in the [C117S] mutant. Thiosulfinate bonds are much weaker than disulfide bonds and the C18-C18 bond

appears to be weaker than the C18-C117 bond. Therefore, the no observation of thiosulphinate bonds in the mutant with a single cysteine could be the result of the instability of the C18-S(O)-S-C18 bond. According to this scenario, accumulation of YmoB monomers containing C18 oxidized to sulphenic acid would be the end result of a series of dimerization-oxidation-dissociation events.



Model 1: Possible relative position of two [C117S] YmoB molecules. C18, highlighted in red, are not close enough in the space (20.4Å) to form a disulfide bond. A conformational change in helix α1 could expose C18 allowing a closer approach between the cysteines and the subsequent disulfide bond formation.

Fig 43: Schematic representation of scenario 1 to explain the C18 reactivity. Model 1 shows the lowest energy model of two [C117S] YmoB 3D structures with C18s highlighted in red an S117s highlighted in blue.

An alternative scenario, scenario 2, would involve the initial oxidation of C18 to sulphenic acid. The increased polarity of the oxidized cysteine would destabilize the hydrophobic core[88] facilitating the conformational transition that would make the sulphenic acid accessible to an approaching thiol from a second YmoB molecule. Interestingly, analysis of the surface of the [C117S] mutant structure reveals two narrow channel connecting position C18 with the protein surface. An oxygen molecule could access to C18 through one of those channels (Fig 44 Model 2).

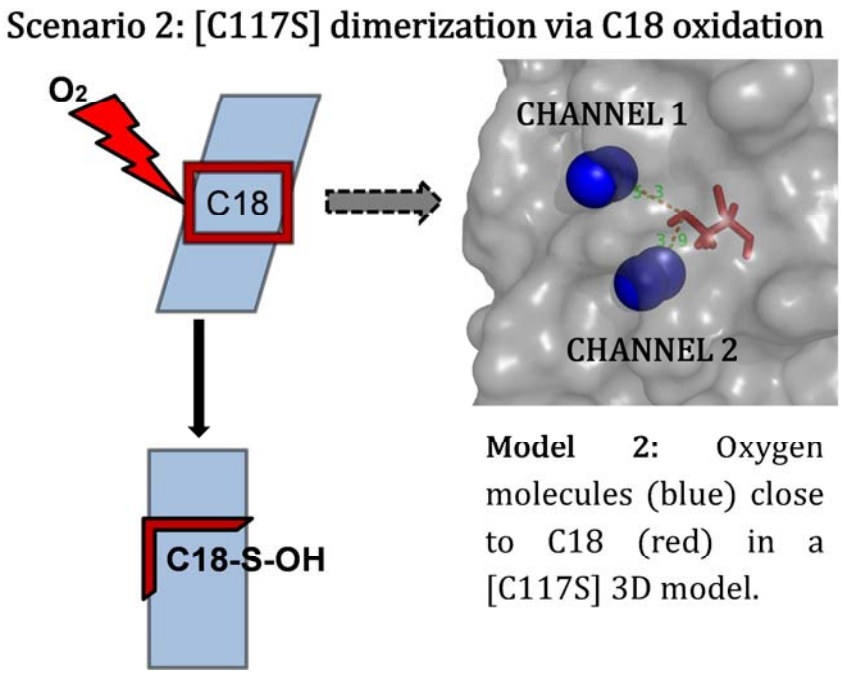


Fig 44: Schematic representation of scenario 2 to explain the C18 reactivity. In Model 2 two oxygen molecules are shown accommodated into the two channels at 5.3Å (channel 1) and 3.9Å (channel 2) to the sulphur atom of C18.

However, the concentration dependence of the sulphenilation reaction suggests that conformational distortion of the monomer induced by non-covalent interactions is needed to facilitate C18 oxidation. Thus, an intermediate situation may exist, where in initial step oxygen molecules may be accommodated in the channels showed at Fig 44 Model 2. The monomer with the oxygen molecules already accommodated could interact with a second YmoB molecule, leading to the conformational change that expose C18, but this time with oxygen molecules in the proximity of C18s, which could then oxidize the sulphur group to sulphenic acid. This situation is schematically presented in Fig 45 as scenario 3.

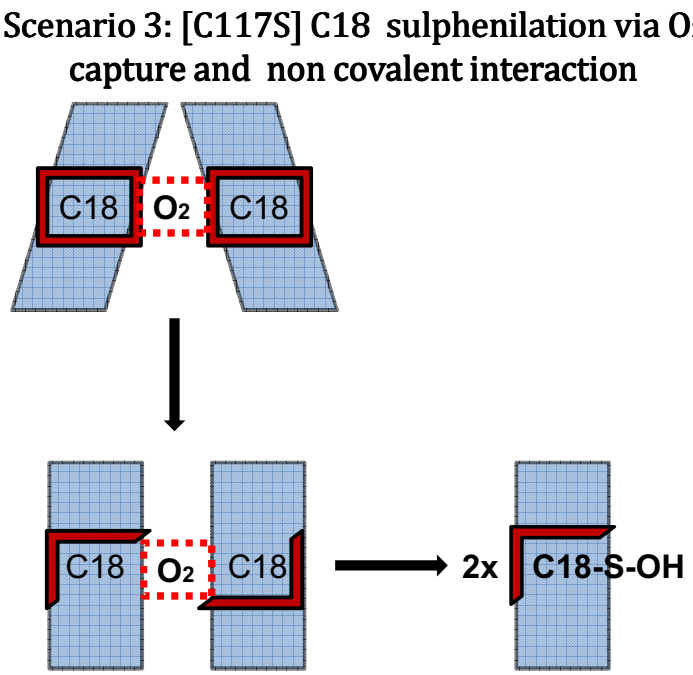


Fig 45: Schematic representation of scenario 3 to explain C18 reactivity. The encounter of two YmoB molecules is produced with oxygen molecules accommodated at the surface channels of the proteins. The conformational change exposes C18 with oxygen atoms at the proximities, producing the sulphenilation of the cysteine.

The channels 1 and 2 in the surface of the [C117S] construct are rigidly formed by two aromatic amino acids, Y22 and F108. The channels are also outflanked by two dipoles, one at the left side of channel 1 formed by R109 and E105, and the second at the right side of channel 2 formed by L15 and Glu19. Both dipoles point at the same direction (Fig 46A). Thus, these channels are rigid enough to accommodate oxygen molecule and present directionality given by the dipoles, which could drive the approximation of two YmoB monomers, not only in an antiparallel way, but also pointing the C18s one close to the other.

We identified a conformation where L103 situates its positive charge at 3.5Å from C18 sulphur. The angle formed by the side chain $-NH_2$ group of K103, the Cβ of C18 and the sulfur of C18 is 109°. Thus, a tetrahedral conformation may be adopted and favor the oxidation of the C18 sulfhydryl to sulphenic acid, as the oxygen situated at channel 2 is at 3.9Å (Fig 46B).

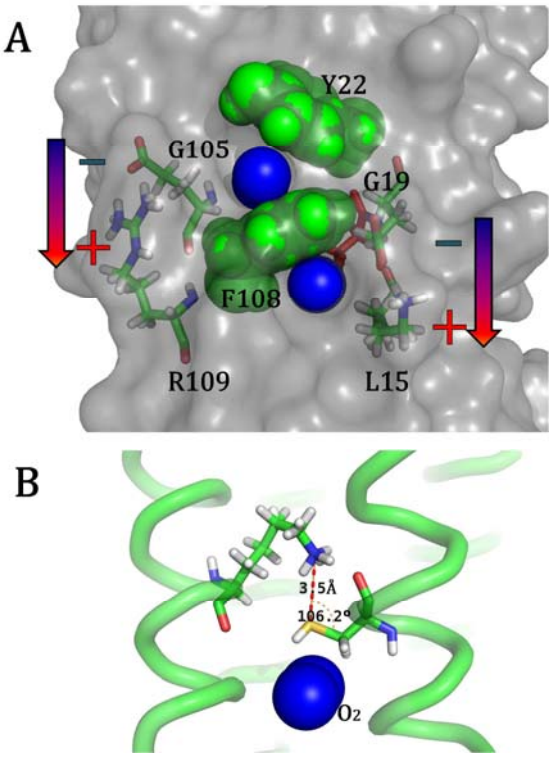


Fig 46: A: Side chain distribution around the oxygen channels 1 and 2 at the [C117S] YmoB surface. Aromatic side chains for Y22 and F108 are shown as green spheres. Positive charges of R109 and L15 are highlighted with a red cross and negative charges of G105 and G19 with a blue minus. The arrows show the direction of the dipoles. B: Representation of the oxygen approximation to C18 across channel 2 when the $-NH_2$ group of K103 is positioned at 3.5Å of the sulfur atom from C18

For all three scenarios described, sulphenic acid is the final product and is accumulated, as suggest our NMR data. On the other hand, the thiosulphinate C18-S(O)-S-C117 observed by mass spectrometry could not be interpreted with any of these scenarios. As reported in the literature, sulphenic acids are reactive species that easily form disulphide bonds by reacting with free thiol[88-91,93,94,101,97,98,100]. In the case of WT YmoB, a sulphenilated C18 could react with the free thiol of a second YmoB molecule. In an antiparallel non-covalent approach, as our NMR data suggest, one or two C18-C117 disulfide bonds could stabilize the dimer (Fig 47).

In this case, the formation of the C18-S(O)-S-C117 thiosulfinate would be produced by oxidation of the C18-C117 disulfide bond, facilitated by the environment of C18 and maybe by the oxidation of neighbor M120, as when thiosulphenilated C18-S(O)-S-C117 is detected, always M120 appears oxidized (Suppl. Table 2). This approach would be mediated by the high flexible single turn helix S-α5 involving residues 116-119 (Fig 48).

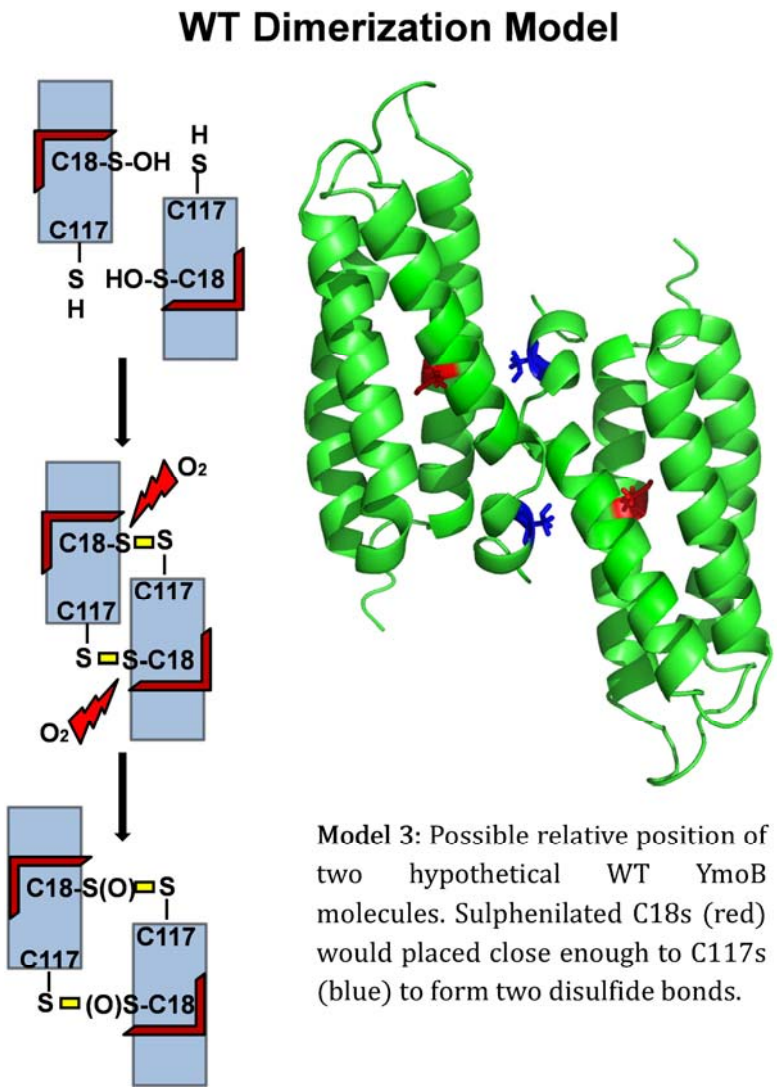


Fig 47: Model for WT YmoB dimerization. With this model the presence of thiosulphenilated C18-S(O)-S-C117 would be explained by the oxidation of the C18-C117 disulfide bond.

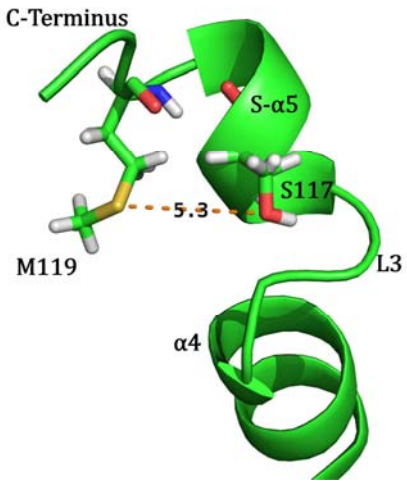


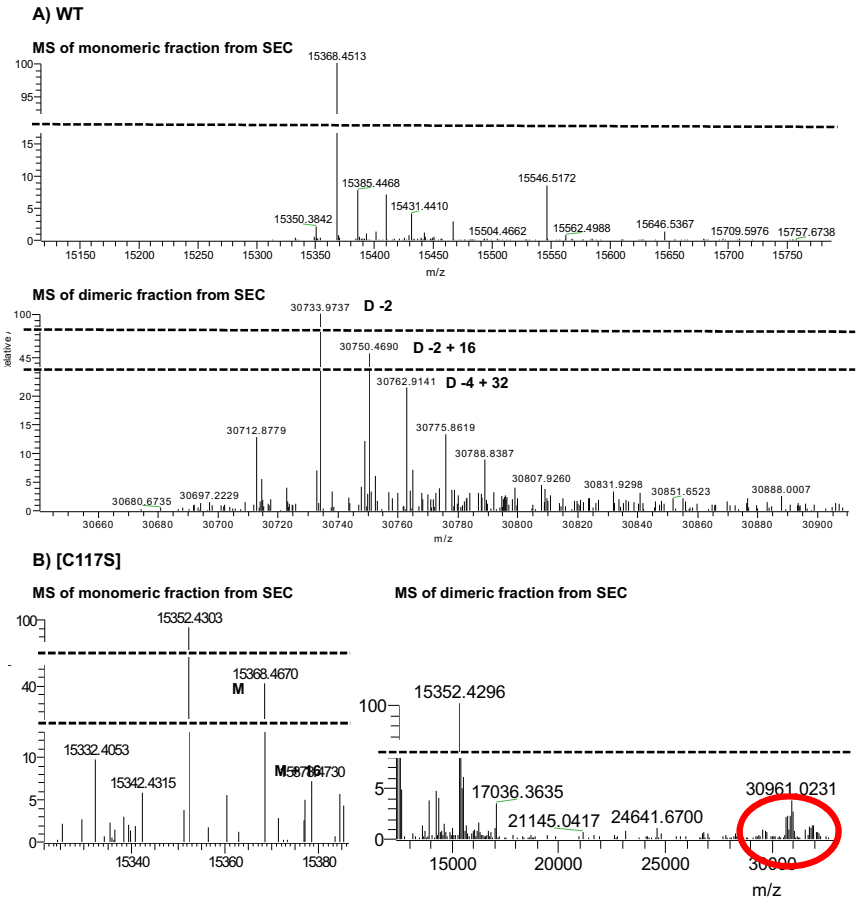
Fig 48: Relative orientation of S117 and M120 in a [C117S] YmoB model. M119 is at 5.3Å of S117. Because of the proximity of these two amino acids, we hypothesize that in a WT model the oxidized M120 could play a role in the thiosulphenilation of the C18-C117S disulphide bone.

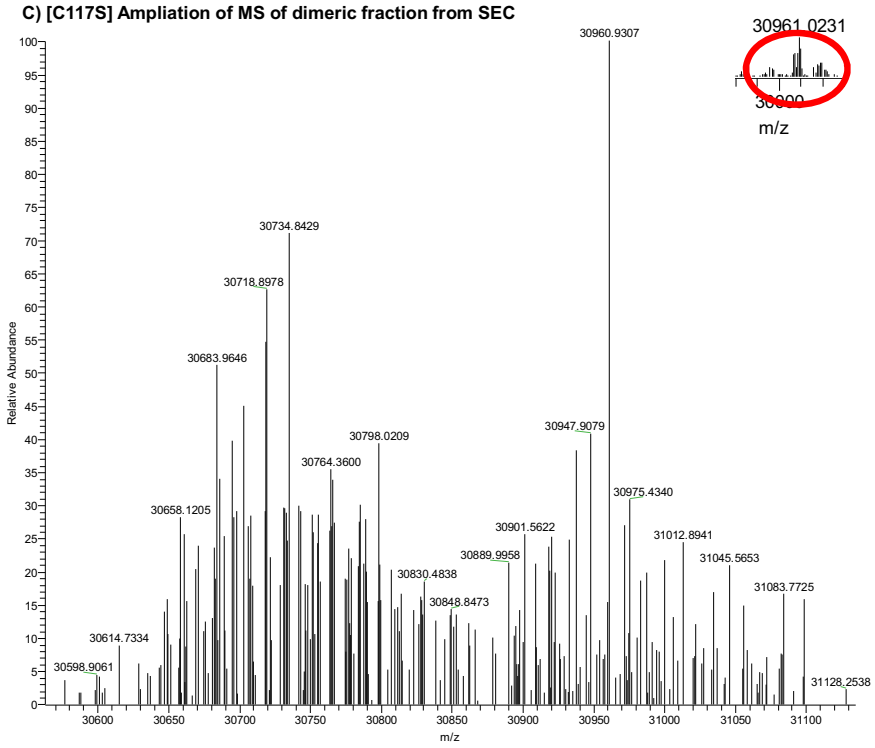
II.4.4. CONCLUSIONS

On the basis of the results obtained, at least three possible YmoB entities could act as active antitoxin: a) the reduced YmoB monomer, b) YmoB disulphide dimers (or higher oligomers), and c) C18 oxidized YmoB monomer is the antitoxin. At Chapter III FINAL CONCLUSIONS & OUTLOOK we address each of these possibilities, taking into account all the results obtained in this PhD thesis project.

II.4.5. SUPPLEMENTARY MATERIAL

Suppl. Fig. 2: A: WT YmoB MS spectra correspond to the monomeric (up) and to the dimeric (down) fractions. **B:** [C117S] YmoB MS spectra correspond to the monomeric (left) and to the dimeric (right) fractions. The highlighted zone at the MS of the dimeric fraction is amplified and showed at **C**: In the covalent dimer species 2amu corresponds to the loss of 2 protons during the disulfide bond formation. +16amu corresponds to the incorporation of an oxygen atom typical of sulfur oxidation.





Suppl. Table 1: Analysis of peptides obtained after endoproteinase GluC digestion followed by nanoHPLC-ESI-MS/MS (nano high performance liquid chromatography coupled to tandem mass spectrometry with nanoelectrospray ionization).

Sample	Peptide Sequence	Modifications	MH+ [Da]
WT & [C117S] YmoB	SHHHHHHSMGMDE	M9(Oxidation)	1594,60085
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGD	M7(Oxidation)	2085,99808
WT & [C117S] YmoB	HIASFVM	M7(Oxidation)	820,40221
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGDLSE	M7(Oxidation)	2415,15224
WT & [C117S] YmoB	HIASFVMSF	M7(Oxidation)	1054,50260
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGDLSELVEE	M7(Oxidation)	2885,39540
WT & [C117S] YmoB	YICTLMKT	M6(Oxidation)	988,48466
WT & [C117S] YmoB	**VAQLKFLCE**	C8(Oxidation)	1066,56046
WT & [C117S] YmoB	SHHHHHHSMGMDE		1578,60700
WT & [C117S] YmoB	LQRWQKTKE		1216,68005
WT & [C117S] YmoB	GIATLGDSHHGWVND		1578,73037
WT & [C117S] YmoB	YSPKRHDVAQLKFLCE		1933,99967
WT & [C117S] YmoB	YICTLMKT		972,48976
WT & [C117S] YmoB	RLFRLFSGE		1124,62113
WT & [C117S] YmoB	GIATLGDSHHGWVNDPTSAVNLQLND		2731,30818
WT & [C117S] YmoB	VAQLKFLCE		1050,56619
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGD		2070,00227
WT & [C117S] YmoB	SHHGWVNDPTSAVNLQLNDLIE		2459,19572
WT & [C117S] YmoB	GIATLGDSHHG		1064,50566
WT & [C117S] YmoB	RLFSGE		708,36771
WT & [C117S] YmoB	FKIKYPDDGD		1197,57878
WT & [C117S] YmoB	GIATLGDSHHGWVN		1463,70301

WT & [C117S] YmoB	FKIKYPDDGDLSE		1526,73693
WT & [C117S] YmoB	LKFLCE		752,40052
WT & [C117S] YmoB	LFSSYGINDPE		1241,56878
WT & [C117S] YmoB	HIASFVMSFK		1166,60282
WT & [C117S] YmoB	FRLFSGE		855,43622
WT & [C117S] YmoB	FKIKYPDDGDLSELVEE		1996,97856
WT & [C117S] YmoB	SLYDEGIATLGDSHHGWVNDPTSAVNLQ		2996,40014
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGDLSE		2399,15569
WT & [C117S] YmoB	LFRLFSGE		968,52081
WT & [C117S] YmoB	YLDDTYTLFSSYGINDPELQRWQKTKE		3310,59677
WT & [C117S] YmoB	HIASFVMSFKIKYPDDGDLSELVEE		2869,40420
WT & [C117S] YmoB	GIATLGDSHHGWVNDPTSAVNLQLNDLIE		3086,52651
WT & [C117S] YmoB	VNDPTSAVNLQLNDLIE		1854,94836
WT & [C117S] YmoB	SLYDEGIATLGDSHHGWVNDPTSAVNLQLNDLIE		3693,77124
WT & [C117S] YmoB	YLDDTYTLFSSYGINDPE		2112,93075
WT & [C117S] YmoB	YLDDTYTLFSSY		1487,65837

Suppl. Table 2: Analysis of cross linked peptides obtained after endoproteinase GluC digestion followed by nanoHPLC-ESI-MS/MS (nano high performance liquid chromatography coupled to tandem mass spectrometry with nanoelectrospray ionization) using two software; MassMatrix and Stavrox. Searches were run twice using two different cross link possibilities (disulfide bridge and cysteine - oxidized cysteine cross link). Only cross-link identifications found by both search engines were considered as positive identifications.

Sample	Peptide 1	Peptide 2	Cross link	MW Measured	charge	ppm
[C117S]	VAQLKFLCE	VAQLKFLCE	C28-C28	2098,1107	3	1,5
[C117S]	VAQLKFLCE	VAQLKFLCE	C28-C28	2098,1107	4	1,5
[C117S]	VAQLKFLCE	VAQLKFLCE	C28-C28	2098,1107	2	1,5
WT	VAQLKFLC(ox)E	YICTLM(ox)KT	C28-C127	2052,0164	3	-2,4
WT	VAQLKFLC(ox)E	YICTLM(ox)KT	C28-C127	2052,0164	3	-2,4
WT	VAQLKFLC(ox)E	YICTLM(ox)KT	C28-C127	2052,0164	3	-2,4
WT	VAQLKFLCE	YICTLM(ox)KT	C28-C127	2036,0174	3	-4,4
WT	VAQLKFLCE	YICTLM(ox)KT	C28-C127	2036,0174	4	-4,4
WT	VAQLKFLCE	YICTLM(ox)KT	C28-C127	2036,0174	3	-4,4
WT	VAQLKFLCE	YICTLM(ox)KT	C28-C127	2036,0174	4	-4,4
WT	VAQLKFLCE	YICTLM(ox)KT	C28-C127	2036,0174	3	-4,4
WT	VAQLKFLCE	YICTLMKT	C28-C127	2020,0184	3	-6,5
WT	VAQLKFLCE	YICTLMKT	C28-C127	2020,0184	2	-6,5
WT	VAQLKFLCE	YICTLMKT	C28-C127	2020,0184	4	-6,5
WT	VAQLKFLCE	YICTLMKT	C28-C127	2020,0184	3	-6,5
WT	VAQLKFLCE	VAQLKFLCE	C28-C28	2098,1107	2	1,5
WT	VAQLKFLCE	VAQLKFLCE	C28-C28	2098,1107	4	1,5
WT	VAQLKFLCE	VAQLKFLCE	C28-C28	2098,1107	3	1,5
WT	VAQLKFLCE	VAQLKFLCE	C28-C28	2098,1107	3	1,5
WT	YICTLM(ox)KT	YICTLM(ox)KT	C127-C127	1973,9384	3	-3,6

WT	YICTLMKT	YICTLM(ox)KT	C127-C127	1957,9422	3	-4,3
WT	YICTLMKT	YICTLMKT	C127-C127	1941,9524	3	-1,6
WT	YSPKRHDVAQLKFLCE	YICTLM(ox)KT	C28-C127	2919,4382	4	-6,3
WT	YSPKRHDVAQLKFLCE	YICTLM(ox)KT	C28-C127	2919,4382	4	-6,3
WT	YSPKRHDVAQLKFLCE	YICTLMKT	C28-C127	2903,4462	4	-5,3
WT	YSPKRHDVAQLKFLCE	YICTLMKT	C28-C127	2903,4462	5	-5,3
WT	YSPKRHDVAQLKFLCE	VAQLKFLCE	C28-C28	2981,5182	4	-6,5
WT	YSPKRHDVAQLKFLCE	VAQLKFLCE	C28-C28	2981,5182	5	-6,5
WT	YSPKRHDVAQLKFLCE	VAQLKFLCE	C28-C28	2981,5182	3	-6,5

II.4.6. BIBLIOGRAPHY

[86] Masuda, H., Tan, Q., Awano, N., Wu, K.-P., & Inouye, M. (2012). YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in Escherichia coli. *Molecular microbiology*, *84*(5), 979–89.

[87] Schuster, C. F., & Bertram, R. (2013). Toxin-antitoxin systems are ubiquitous and versatile modulators of prokaryotic cell fate. *FEMS microbiology letters*, *340*(2), 73–85.

[88] Ryu, S. E. (2012). Structural mechanism of disulphide bond-mediated redox switches. *Journal of biochemistry*, *151*(6), 579–88.

[89] Kulathu, Y., Garcia, F. J., Mevissen, T. E. T., Busch, M., Arnaudo, N., Carroll, K. S., Barford, D., et al. (2013). Regulation of A20 and other OTU deubiquitinases by reversible oxidation. *Nature communications*, *4*, 1569.

[90] Bond, I. D., Buhrman, G., Parker, B., Sohn, J., Rudolph, J., & Mattos, C. (2005). Structural Mechanism of Oxidative Regulation of the Phosphatase Cdc25B via an Intramolecular Disulfide Bond. *Biochemistry,* (6), 5307–5316.

[91] Liu, X., Sun, X., Wu, Y., Xie, C., Zhang, W., Wang, D., Chen, X., et al. (2013). Oxidation-sensing regulator AbfR regulates oxidative stress responses, bacterial aggregation, and biofilm formation in Staphylococcus epidermidis. *The Journal of biological chemistry*, *288*(6), 3739–52.

[92] Takata T, Endo T. (1990). In: Patai S, editor. The chemistry of sulphinic acids, esters and their derivatives. New York: Widely, P. 527-575.

[93] Fukushima, D., Kim, Y. H., Iyanagi, T., & Oae, S. (1978). Enzymatic Oxidation of Disulfides and Thiosulfinates by both Rabbit Liver Microsomes and a Reconstituted System with Purified Cytochrome P-450. *J. Biochem*. *83*(4), 1019–1027.

[94] Badol, P., David-Dufilho, M., Auger, J., Whiteheart, S. W., & Rendu, F. (2007). Thiosulfinates modulate platelet activation by reaction with surface free sulfhydryls and internal thiol-containing proteins. *Platelets*, *18*(7), 481–90.

[95] Koch, P., Ciuffarin, E., & Favalb, A. (1968). The Thermal Disproportionation of Aryl Arenethiolsulfinates. Kinetics and Mechanism'", *417*(2), 5971–5977.

[96] Huang, K.-P., & Huang, F. L. (2002). Glutathionylation of proteins by glutathione disulfide S-oxide. *Biochemical pharmacology*, *64*(5-6), 1049–56.

[97] Benoit, R., & Auer, M. (2011). A direct way of redox sensing. *RNA Biology*, *8*(1), 18–23.

[98] Hu, W., Tedesco, S., McDonagh, B., & Sheehan, D. (2010). Shotgun redox proteomics in sub-proteomes trapped on functionalised beads: Identification of proteins targeted by oxidative stress. *Marine environmental research*, *69 Suppl*, S25–7.

[99] Li, J., Huang, F. L., & Huang, K. P. (2001). Glutathiolation of proteins by glutathione disulfide S-oxide derived from S-nitrosoglutathione. Modifications of rat brain neurogranin/RC3 and neuromodulin/GAP-43. *The Journal of biological chemistry*, *276*(5), 3098–105.

[100] Jönsson, T. J., Tsang, A. W., Lowther, W. T., & Furdui, C. M. (2008). Identification of intact protein thiosulfinate intermediate in the reduction of cysteine sulfinic acid in peroxiredoxin by human sulfiredoxin. *The Journal of biological chemistry*, *283*(34), 22890–4.

[101] Denu, J. M., & Tanner, K. G. (1998). Specific and Reversible Inactivation of Protein Tyrosine Phosphatases by Hydrogen Peroxide : Evidence for a Sulfenic Acid Intermediate and Implications for Redox, *2960*(16), 5633–5642.

[102] Khiar, N., Mallouk, S., Valdivia, V., Bougrin, K., Soufiaoui, M., & Fernández, I. (2007). Enantioselective organocatalytic oxidation of functionalized sterically hindered disulfides. *Organic letters*, *9*(7), 1255–8.

Chapter III.

FINAL CONCLUSIONS & OUTLOOK

Possible biological relevance of YmoB modifications

CONCLUSION 1: *Hha-TomB proteins form a toxin-antitoxin system with Hha the toxin and TomB the antitoxin, in planktonic cells as well as in biofilms. In both systems we demonstrated that the biological activity of TomB proteins is always related to Hha. We also proved that TomB and YmoB are functionally interchangeable in planktonic cells and in biofilms, demonstrating that this TA module is functional in enteric bacteria.*

When the antitoxin YmoB is not active, overexpression of Hha gives rise to a dramatic increase of dead cells with the core of the biofilm. Instead, when Hha and YmoB are simultaneously overexpressed, the biofilm is organized in micro colonies, which are mainly composed by living cells uniformly spread over the entire surface.
Moreover, we confirmed that the [C117S] YmoB mutant, our optimal construct for NMR studies, is a biologically active antitoxin of Hha in *E. coli* planktonic cells. Although in biofilm systems [C117S] YmoB is still functional, its activity as Hha-antitoxin is lower than WT.

CONCLUSION 2: *The [C117S] YmoB structure consists in a four helix bundle and a short one turn helix at the C-terminal. The electrostatic surface of the protein is mainly negatively charged, but presents two well defined positives patches.*

During the studies carried out to obtain a suitable construct for elucidate the [C117S] YmoB 3D structure by NMR we observed that TomB and YmoB proteins show slow oligomerization processes in solution. A reduction in the number of cysteines generates constructs that presents important diminution in the formation of soluble high molecular weight oligomers. This points to the fact that cysteines, in this family of proteins, may play an important role.

CONCLUSION 3: *[C117S] YmoB construct, a mutant with a single cysteine residue deeply protected by the tertiary structure of the protein, still shows slow formation of a dimer species, revealing the importance of the C18 residue.*

CONCLUSION 4: *Our studies uncover a very unusual oxidation behavior of C18. It can form normal disulphide bonds with C18 or C117, it can be sulphenilated and it can also form thiosulphenilated C18-S(O)-S-C117 bonds.*

We hypothesized that oxidation of YmoB is related to its antitoxin activity and we suggest at least three possible YmoB entities as active antitoxin:

a) The reduced YmoB monomer is the antitoxin.

In order to test this hypothesis, some preliminary experiments were carried out to investigate the possible interaction of YmoB with several putative partners in vitro using NMR. YmoB and YmoA do not interact in vitro. Garcia-Contreras and coworkers suggested that DNA could be one of YmoB targets, reducing the expression of YmoA. However, no effect could be observed when ^{15}N-labelled YmoB was treated with *Yersinia* DNA or with soluble *Yersinia* cytoplasmic extracts.

b) YmoB dimers (or higher oligomers) are the antitoxin species.

Oxidized YmoB may exist in the form of oxidized monomer or as dimer or larger oligomers. Interestingly, the structure of the [C117S] YmoB monomer revealed the presence of two positively charged patches. A rough model generated by approaching individual monomers so that C18-C117 distances would be minimized shows that the positively charged patches from neighbor molecules coalesce into a continuous positive strip suggesting that the possibility to interact with nucleic acids (DNA or RNA) would be enhanced by YmoB dimerization (Fig 49).

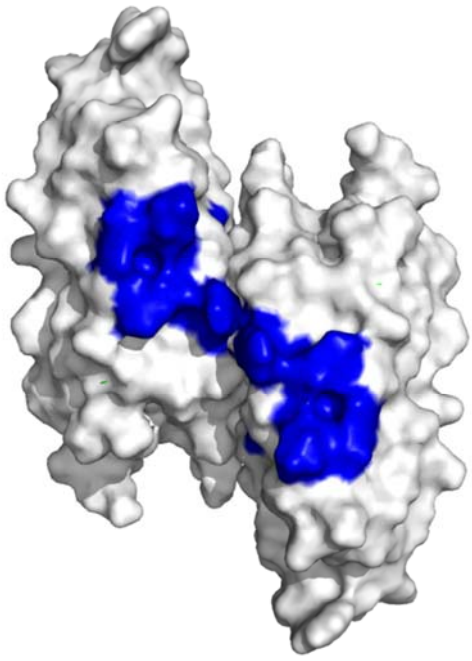


Fig 49: Rough model of YmoB dimer with two C18-C117 close enough to form disulfide bond. The surface of each monomer is represented and the amino acids R99, K102, R106 and R109, which generates the positive patch 2, are highlighted in blue.

This model, however, would predict that the activity of the mutant and WT forms of YmoB would be strongly affected by the disulfide bond connectivity, in contrast to the equivalent activity detected in planktonic cells.

c) C18 Oxidized YmoB monomer is the antitoxin.

The possibility that an oxidized form of YmoB is the active antitoxin is appealing. The conservation of the environment of C18 between TomB and YmoB and the singular reactivity of this residue suggests that it plays a functional role. The easy sulphenilation of C18 even by molecular oxygen suggests that the antitoxin activity of YmoB could be associated to an environmental response to oxygen levels. Within a biofilm, the oxygen pressure varies strongly between the surface and inner regions. The possibility that YmoB controls the survival/death of bacteria in the various regions of the biofilm could be part of the dispersion mechanism. Cells at the base of the biofilm may be exposed to lower oxygen levels, preventing the activation of the antitoxin and leading to cell death. This would result in the dispersion of the biofilm. The fact that oxidation of YmoB is controlled by the protein concentration would add an additional regulation level, by gauging the oxygen levels triggering a response by varying the total YmoB expression levels.

The antitoxin activity of WT and [C117S] mutant of YmoB in *E. coli* is equivalent in well oxygenated planktonic cell. However, the mutant has a lower activity than the WT form in biofilms. This is consistent with the hypothesis that oxidation is required to generate the active antitoxin form from YmoB.

Up to now, neither the relevant YmoB species active as an antitoxin has not been demonstrated nor the mechanism by which the toxic activity of Hha/YmoA is reduced when YmoB is expressed.

Further experiments will be necessary in order to determine which YmoB entity is the active antitoxin, to elucidate the mechanism of action of TomB/YmoB proteins as antitoxins and to define the exact role of C18 and its capacity to be oxidized into sulphenic acid and thiosulphinate in biofilm regulation.

Chapter IV.

MATERIAL & METHODS

IV.1. CLONING AND MUTAGENESIS

TomB was amplified using *Escherichia coli* Dh5α (Life Technologies™) lysates. It was subsequently cloned into pGEX-KG and pET14b vectors. One construct was fused to GST with a thrombin sensitive connector at the N-terminus using the pGEX-KG plasmid. In plasmid pET14b, H_6-TomB construct was obtained using XhoI/BamHI sites and TomB-Stag was obtained using NcoI/XhoI sites. Strep-Tag was fused using appropriate prymers. YmoB was amplified using *Yersinia enterocolitica* W22711 lysates and it was subsequently cloned into pHAT2 using NcoI and XhoI sites.

Cysteine to serine mutations were introduced into TomB and YmoB using the QuickChange™ site directed mutagenesis Kit (Stratagene®) with the appropriate complementary mutagenic primers. We confirmed the constructs by DNA sequencing and Mass Spectrometry.

IV.2. PROTEIN EXPRESSION

The transformation of the expression vector into the host cell was done with the heat shock method using competent BL21(DE3) *E. coli* cells (Novagen, Darmstadt, Germany). Cells were grown for 60 minutes at 37ºC, centrifuged and dispersed on an agar plate with antibiotic for selection. The agar plate was incubated at 37°C overnight (O/N) in order to obtain bacterial colonies. For all constructs glycerolates were prepared for posterior protein expression using 80% of glycerol and 20% of O/N grown culture of BL21(DE3) *E. coli* with the corresponding expression vector.

Unlabeled, ^{15}N-labeled, ^{13}C and ^{15}N were expressed in *E. coli* BL21 (DE3) in LB medium or minimal medium using H_2O enriched with $^{15}NH_4Cl$ and/or D-[^{13}C] glucose as the sole sources of carbon and nitrogen respectively. *E. coli* extracts were lysed by sonication. Soluble fusion proteins were purified by; a) His-Tag affinity chromatography (Ni-NTA agarose resine QIAGEN) and samples were eluted using buffer A (20mM Trizma® hydrochloride, 800mM NaCl, 400mM imidazole, 5mM DTT and pH=8.00), b) Strep-Tag affinity chromatography (Strep-Tacting® sepharose® iba) and samples were eluted using buffer B (20mM Trizma® hydrochloride, 800mM NaCl, 2.5mM desthiobiotin, 5mM DTT and pH=8.00). Proteins and constructs were further purified by size exclusion chromatography (SEC) using HiLoad™ Superdex 75 prepgrade columns (GE Healthcare) and buffer C (20mM [NaH_2PO_4 + Na_2HPO_4], 150mM NaCl, 0.2mM EDTA, 1mM TCEP, 0.01% NaN_3 and pH 7.00). Fractions containing the purified proteins were concentrated to the mentioned concentrations for NMR experiments.

IV.3. THIOL ALKYLATION

thiol alkylation of the cysteines of WT YmoB and [C117S] YmoB was performed with sodium iodoacetate in a final volume of 20mL, in a reaction buffer with; 20mM [NaH_2PO_4 +

Na_2HPO_4], 0.3mg/mL of protein, 0.25mg/mL of sodium iodoacetate and pH=7.5 for 2 hours under darkness and agitation. Reaction was quenched with 2mL of β-mercaptoethanol and sample was recovered using SEC.

IV.4. BACTERIAL GROWTH CURVES

One colony was precultured O/N at 37ºC and 250 rpm in 25mL of Tryptone Minimal Medium (TMM) (10g tryptone and 2.5g NaCl) with the corresponding antibiotic in a flask of 250mL. Two cultures from independent colonies for each strain were grown. 25mL of fresh TMM with the corresponding antibiotic and with the corresponding chemical inducer were inoculated with 200µL of the precultur. OD_{600} was measured every 15 minutes between OD_{600}= 0.05~0.7, every 30 minutes between OD_{600}= 0.7~2 and every 60 minutes until 10 hours. OD_{600} was measured for the two independent cultures and the average with the standard deviation was reported. Growth curves were done for all the strains and conditions listed at Table 8.

IV.5. SPECIFIC GROWTH RATES (SGR)

For SGR we used the exponential part of the bacterial growth curves (OD = 0.05 ~ 0.7). The SGR were calculated as the slope of the $\ln OD_{600}$ vs. time.

IV.6. BIOFILM ASSAYS

Single colonies of all strains were inoculated into 25mL TMM medium (with the corresponding antibiotic) and incubated overnight at 37°C and 250 rpm shaking speed. Two independent cultures were prepared for each strain. O/N cultures for all strains were inoculated respectively into TMM medium (with the corresponding antibiotic) to form a solution with OD_{600} below 0.05. 300µL of culture were added into each well of a 96 well plate. All strains took 20 wells (10 wells for each independent culture) and pure TMM medium took wells all around the plate (Blanc). The 96-well plate was incubated at 37 °C without shaking, for 15 hours. To record the total growth we measured OD_{620}. Supernatant was discarded and the plate was washed three times with room temp MilliQ water by dipping the plate into a 1 L solution of water. 300µL of 0.1% crystal violet were added in each well and the plate was incubated for 20 minutes at room temperature. Staining solution was discarded and the plate was washed three times with water by dipping the plate into a 1 L solution of water. 300µL of 95% ethanol was added to each well and was incubated for 5 minutes. We measured the total biofilm reading OD_{540} after mixing of 50 seconds. Standard biofilm formation was calculated as OD_{540} / [OD_{620} - OD_{620} Blanc]. Biofilm assays were perfomed for all the strains and conditions listed at Table 1. OD measures were done using Infinite® M1000 PRO.

IV.7. FLOW CELL BIOFILM

Strains *E coli* K12MG1655 Δhha pCA24Nhha pBAD30ymob were streaked on LB agar plates with the corresponding antibiotics. One-day old single colonies were inoculated in 25mL of TMM O/N at 37°C at 250 rpm. Overnight cultures were diluted to reach OD_{600} below 0.05 in 150mL TMM with the corresponding antibiotics. Diluted cultures were pumped to the flow cell at 10 mL/h for 3h with stirring at 37°C to facilitate cell attachment to the glass slide surface. Culture medium was replaced with 200mL of fresh TMM containing the corresponding antibiotics and inducers and then fed to the flow cell at the same rate for 15h. After incubation, the flow cell unit was disconnected and blocked (never allowing it to dry). It was placed at room temperature and the SYTO9 mix was added using a 5mL syringe. We incubated the flow cell with SYTO9 was incubated in the dark for 20 min at room temperature. Eight images in random positions were taken with an OLYMPUS FV100 confocal Microscope using a Long Working Distance 40x Olympus LWDCDPLAN 40LP lens. 3D images of each spot were obtained using IMARIS software, and quantitative analysis of biofilm was performed using COMSTAT software.

For the SYTO9 Stain preparation the LIVE/DEAD *Bac*Light Bacterial Viability Kit L7007 was used. 3mL of dye was prepared in a 30-mL glass beaker mixing; 7.5µL of COMPONENT A (SYTO 9 dye, 1.67mM / Propidium iodide, 1.67mM), 7.5µL of COMPONENT B (SYTO 9 dye, 1.67mM / Propidium iodide, 18.3mM) and 3mL of TMM.

IV.8. NMR BUFFER OPTIMIZATION

Protein constructs behavior regarding to pH was tested in samples at 50µM H_6-YmoB from pH=5.00 to pH=7.50. Salts tested were NaCl (phosphate buffered using NaH_2PO_4 and Na_2HPO_4 mixtures[103]) and HEPES[ix]. Melting temperatures (T_m) were calculated for samples at 50µM H_6-YmoB, with increasing salt concentrations (from 0mM to 200mM) using differential scanning fluorimetry in a Photal FluoDia T70 PTI® and following the methodology published by Niesen et al[104].

Sample stability at long times was studied at high protein concentration (1mM) at 100mM NaCl, 150mM NaCl and 100mM HEPES. Samples were incubated at 25°C during 10 days and the presence of precipitate and the final concentration of the protein were checked.

IV.9. NMR SPECTROSCOPY

NMR spectra listed at Table 9 were acquired at 25°C on 500, 600 or 800 MHz Bruker spectrometers in the LRB (Laboratori de RMN de Barcelona) of the CCIT (Centres Científics I Tecnològics), which belongs to the ICTS (Institución Científico Tecnológica Singular). NMR spectra listed at Table 10 were acquired at 25°C on 800 and 900 MHz Varian spectrometers using Interlived-NUS-MDD fast acquisition protocols in the Swedish NMR center. Data processing

[ix]HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

and analysis were carried out with NMRPipe[105], NMRViewJ[106], and CARA[107]. NMR spectra for optimization of NMR sample parameters and for functional studies were recorded at the cited concentrations. NMR spectra for structure determination were recorded at: a) 1.4mM protein concentration for side chain assignment (Table 9) and b) 1.0mM protein concentration for side chain assignment and for collecting conformational restraints (Table 10).

Stereospecific assignments of valine and leucine methyl groups were obtained from a constant time ^{1}H-^{13}C-HSQC on a 10% ^{13}Clabeled protein sample. Proton chemical shifts were referenced using 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as an internal standard, whereas ^{15}N and ^{13}C chemical shifts were indirectly referenced. Protein distance restraints were obtained from 3D ^{1}H-^{15}N-edited NOESY-HSQC, 3D ^{1}H-^{13}C-edited NOESY (aromatic optimized in D2O), and 3D ^{1}H- ^{13}C-edited NOESY-HSQC in D2O experiments with a mixing time of 120ms using the Interlived-NUS-MDD acquisition method. Data were automatically assigned and the NOE distance restraints were obtained iteratively using the Unio'10/CYANA 2.1 suite program[108,109] and manually inspected. Protein backbone dihedral angle restraints were derived using TALOS+[110]. Hydrogen bonds were obtained from H/D exchange ^{1}H-^{15}N HSQC experiments acquired using NUS.

IV.10. STRUCTURAL CALCULATION AND REFINEMENT

The structure of [C117S] YmoB was determined by simulated annealing using the torsion angle dynamic simulation program CYANA 2.1[111], and further water refinement with CNS 1.2.1[112,113]. Protein structure calculation was based on Unio'10/CYANA-generated upper distances using only unambiguously assigned restraints derived from NOESY experiments, TALOS+-driven dihedral angle restraints. Based on H/D exchange experiments, backbone NOE pattern and $^{13}C\alpha/^{13}C\beta$ chemical shifts, hydrogen bond restraints were also used in the structure calculation. An ensemble of 100 protein structures was generated and the 20 lowest energy conformers were checked using PROCHECK-NMR[114]. Molecular images were generated using PyMOL[115] and MolMol[116].

IV.11. MASS SPECTROMETRY

Protein masses were determined on a LTQ-FT Ultra (Thermo Scientific) mass spectrometer after liquid chromatography and processed and analyzed with Xcalibur software (vs. 2.02 SR2). Ion deconvolution to zero charged monoisotopic masses was performed using Xtract algorithm in Xcalibur software.

Peptides were obtained and analyzed using endoproteinase GluC digestion followed by nanoHPLC-ESI-MS/MS (nano high performance liquid chromatography coupled to tandem mass spectrometry with nanoelectrospray ionization) in a LTQ-FT Ultra (Thermo Scientific) mass spectrometer. Cross-Linked peptides were analyzed using MassMatrix and Stavrox and

only cross-link identifications found by both search engines were considered as positive identifications.

IV.12. TABLES

ABBREVIATION	GENOTYPE	INDUCER	OVEREXPRESSED PROTEIN
WT	WT	-	-
Δhha	Δhha	-	-
Hha(-) X*(-)	Δhha	-	-
Hha(+)X*(-)	Δhha	IPTG	Hha
Hha(+)X*(+)	Δhha	IPTG + Arabinose	Hha + X*
Hha(-) X* (+)	Δhha	Arabinose	X*

Table 8: List of abbreviations used to design the strain *E coli* K12MG1655 carrying the pCA24Nhha and pBAD30x* plasmids. X*= TomB, YmoB or C117S YmoB proteins. x*= tomb, ymob or C117S ymob. IPTG and arabinose were used at final concentrations of 1mM and 1% respectively. Strains transformed with pCA24N were growth with 50mg/mL of ampicillin and strains transformed with pBAD30 were growth with 30mg/mL of chloramphenicol. We will use (-) to point out no overexpression and (+) to point out overexpression of proteins.

Experiment	Dimensionality	Observed Spins
^{1}H-WG	1D: ^{1}H	H_i
^{1}H-^{15}N-HSQC	2D: ^{1}H-^{15}N	H_i-N_i
HNCO	3D: ^{1}H-^{15}N-^{13}C	H_i-N_i-$C'O_{-i}$
HNCACO	3D: ^{1}H-^{15}N-^{13}C	H_i-N_i-$C'O_{-i}$-$C'O_i$
CBCAcoNH	3D: ^{1}H-^{15}N-^{13}C	H_i-N_i-$C\alpha_{-i}$-$C\beta_{-i}$
^{1}H-^{13}C-HSQC-CT 10% [^{13}C]-glucose	2D: ^{1}H-^{13}C	H_i-C_i aliphatic
HBHAcbcacoNH	3D: ^{1}H-^{15}N-^{1}H	H_i-N_i-$H\alpha_i$-$H\beta_i$-$H\alpha_{-i}$-$H\beta_{-i}$
CBCANH	3D: ^{1}H-^{15}N-^{13}C	H_i-N_i-$C\alpha_i$-$C\beta_i$-$C\alpha_{-i}$-$C\beta_{-i}$

Table 9: NMR experiments recorded in the LRB (Laboratori de RMN de Barcelona) of the CCIT (Centres Científics I Tecnològics), which belongs to the ICTS (Institución Científico Tecnológica Singular).

Experiment	Solvent	Dimensionality	Observed Spins
^{1}H-^{15}N-HSQC	H_2O & D_2O	2D: ^{1}H-^{15}N	H_i-N_i
^{1}H-^{13}C-HSQC Aliphatic	H_20 & D_2O	2D: ^{1}H-^{13}C	H_i-C_i aliphatic
^{1}H-^{13}C-HSQC Aromatic	D_2O	2D: ^{1}H-^{13}C	H_i-C_i aromatic
HccH-TOCSY	D_2O	3D: ^{1}H-^{13}C-^{1}H	H_i-$H\alpha_i$-$H\beta_i$-$H\gamma$-$H\delta$-$H\varepsilon$
HCCh-TOCSY	D_2O	3D: ^{1}H-^{13}C-^{13}C	$C\alpha_i$-$C\beta_i$-$C\gamma$-$C\delta$-$C\varepsilon$
^{1}H-^{15}N-NOESY-HSQC	H_20	3D: ^{1}H-^{15}N-^{1}H	H_i-$H\alpha_i$-$H\beta_i$-$H\gamma$-$H\delta$-$H\varepsilon$
^{1}H-^{13}C-NOESY-HSQC Aliphatic	D_2O	3D: ^{1}H-^{13}C-^{1}H	H_i-$H\alpha_i$-$H\beta_i$-$H\gamma$-$H\delta$-$H\varepsilon$
^{1}H-^{13}C-NOESY-HSQC Aromatic	D_2O	3D: ^{1}H-^{15}N-^{1}H	H_i-$H\alpha_i$-$H\beta_i$-$H\gamma$-$H\delta$-$H\varepsilon$-H^{Z}-H^{H}

Table 10: NMR experiments recorded with Interlived-NUS-MDD fast acquisition protocols in the Swedih NMR center

IV.13. BIBLIOGRAPHY

[103] Vincent S., Blanchard J. (1990). Methods in enzymology. *Academic Press Elsevier. 182 (4), 24-37.*

[104] Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocol. 2(9), 2212-21.*

[105] Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR, 6(3):277-93.*

[106] Johnson, B. A. (2004). Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods In Molecular Biology Clifton Nj*, *278*, 313-352.

[107] Keller RLJ (2004). The Computer Aided Resonance Assignment Tutorial. *Goldau CANTINA.*

[108] Herrmann, T., Güntert, P., & Wüthrich, K. (2002). Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *Journal of Biomolecular NMR*, *24*(3), 171-189.

[109] Herrmann, T., Güntert, P., & Wüthrich, K. (2002). Protein NMR Structure Determination with Automated NOE Assignment Using the New Software CANDID and the Torsion Angle Dynamics Algorithm DYANA. *Journal of Molecular Biology*, *319*(1), 209-227.

[110] Yang Shen, Frank Delaglio, Gabriel Cornilescu, and Ad Bax. (2009). TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of Biomolecular NMR, 44(4): 213–223.*

[111] Güntert, P. (2004). Automated NMR structure calculation with CYANA. *Methods In Molecular Biology Clifton Nj*, *278*(4), 353-378. Springer.

[112] Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica Section D Biological Crystallography*, *54*(Pt 5), 905-921.

[113] Nederveen, A. J., Doreleijers, J. F., Vranken, W., Miller, Z., Spronk, C. A. E. M., Nabuurs, S. B., Güntert, P., et al. (2005). RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins*, *59*(4), 662-72.

[114] Laskowski R A, Rullmannn J A, MacArthur M W, Kaptein R, Thornton J M (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*, 8, 477-486.

[115] DeLano, W.L. (2002). The PyMOL Molecular Graphics System

[116] Koradi R, Billeter M, Wüthrich K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph. 14(1): 51-5, 29-32.*

Chapter V.

RESUM EN CATALÀ

V. Estudis de les proteïnes de la família TomB/YmoB: reguladors del biofilm en bacteris entèrics.

V.1. Introducció

Seguint els treballs previs realitzats al laboratori del professor Miquel Pons, aquest projecte s'ha focalitzat en els estudis funcionals i estructurals de la proteïna YmoB. Aquesta proteïna de *Yersinia spp.* comparteix una elevada homologia de seqüència amb la proteïna TomB de *Escherichia coli*. Homòlegs de la família de proteïnes TomB/YmoB estan presents en els bacteris entèrics sense que existeixi homologia de seqüència obvia amb altres famílies de proteïnes ni que es conegui l'estructura tridimensional de cap d'elles.

La funció de les proteïnes de la família TomB/YmoB és mol poc coneguda. El grup d'investigació del Dr. T. K. Wood va publicar que Hha-TomB podrien actuar com a parell toxina antitoxina (TA) en la regulació de biofilms, essent Hha la toxina i TomB l'antitoxina (Fig 3)[117]. Els parells TA consisteixen en una toxina estable i una antitoxina làbil. Solen estar involucrats en el control de les tasses de creixement, en la mort cel·lular programada i en la formació i dispersió dels biofilms, entre d'altres[118].

La fisiologia del biofilm permet als bacteris una major supervivència en ambients hostils ja que els dota, entre d'altres característiques, d'una major tolerància als antibiòtics i als atacs dels sistemes immunitaris de les cèl·lules hoste. Això provoca una major persistència i difícil eradicació d'un gran nombre de malalties infeccioses[119].

Així doncs, aquesta tesi doctoral s'ha focalitzat en la obtenció de l'estructura tridimensional mitjançant Ressonància Magnètica Nuclear (RMN) de la proteïna YmoB, en la realització d'estudis funcionals de les proteïnes TomB/YmoB i la validació funcional de la construcció utilitzada en els experiments de RMN y finalment en l'estudi dels processos d'oxidació cisteica i oligomeritzacio que semblen regir la funcionalitat biològica de la família de proteïnes TomB/YmoB.

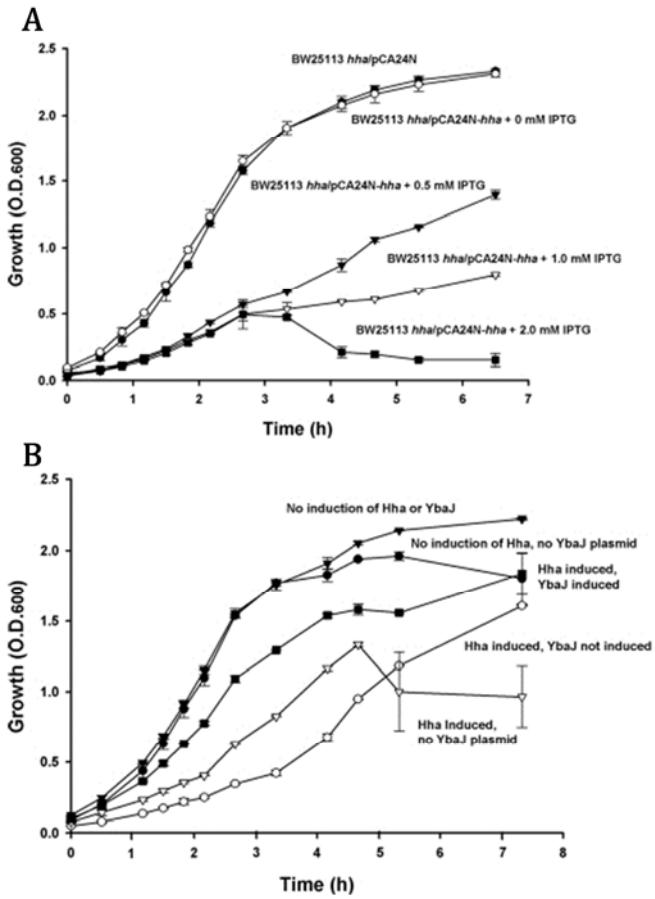


Fig 50: A: Creixement de *E. coli* BW25113 hha/pCA24N-hha a 37°C en LB. Hha es va induir amb concentracions creixents de IPTG **B:** Efecte de la sobre expressió de TomB en *E. coli* K-12 BW25113 hha utilitzant el plasmidi pBAD-TomB per a la sobre expressió de TomB i el plasmidi pVLT31-hha per a la sobre expressió de Hha en LB a 37°C; Hha es va induir addicionant IPTG fins a una concentració final de 2mM i TomB addicionant una concentració final de 0.1% de L-arabinosa. Els experiments es van fer per duplicat I es mostra la desviació estàndard. La figura va ser publicada per Garcia-Contreras i col·laboradors[123].

V.2. Optimització de la mostra de RMN per a la elucidació de l'estructura 3D d'un membre de la família de proteïnes TomB mitjançant RMN.

En la determinació estructural de proteïnes mitjançant RMN són necessaris experiments multidimensionals de llarga durada i concentracions proteiques en el rang mili molar (mM) [120]. La optimització de la construcció proteica per aconseguir la solubilitat i estabilitat requerida per a la obtenció d'espectres de RMN d'elevada qualitat ha estat el nostre primer pas per a la elucidació de l'estructura 3D de la proteïna. Hem utilitzat un mètode ràpid de comprovació de construccions proteiques concentrades en el rang mM i uniformement marcades amb ^{15}N utilitzant experiments de tipus ^{1}H-^{15}N-HSQC[121].

Un cop obtinguda la construcció més adient, hem optimitzat el pH de la mostra, el tipus i la concentració de sals utilitzades i la concentració màxima de proteïna que es podia assolir. Les condicions finals de la mostra de RMN per a l'obtenció de l'estructura 3D d'un membre

de la família de proteïnes TomB/YmoB que hem obtingut són: H_6-YmoB-C117S a 1.4mM, pH=7.00 tamponat amb 20mM [NaH_2PO_4 + Na_2HPO_4], 150mM NaCl, 1mM TCEP, 0.1mM EDTA i 0.1% NaN_3.

Durant l'obtenció de la mostra per a la realització dels estudis de RMN hem observat la formació lenta d'oligòmers. Aquest procés ha estat minimitzat amb la utilització del mutant [C117S] YmoB, però no ha estat completament eliminat. La utilització de protocols d'adquisició ràpida i en paral·lel dels experiments de RMN utilitzant "Interlived-NUS-MDD[122]", ens va permetre resoldre aquest problema.

V.3. Estudis funcionals de les proteïnes de la família TomB i validació biològica de la construcció per a RMN [C117S] YmoB.

En *E. Coli* la sobre expressió de Hha redueix la formació de biofilm, però l'efecte de TomB encara roman desconegut[123]. Degut a que diversos parells TA regulen la formació de biofilms[124], entendre els mecanismes presents en la regulació d'aquests, i en particular, clarificar la funció biològica dels membres de la família TomB com a antitoxines de Hha esdevé un tema de gran interès.

Hem centrat els nostres estudis funcionals en la realització de corves de creixemnt bacterià i en assajos de formació de biofilm en micro plats de 96 pouets[125]. L'estratègia seguida s'ha basat en la transformació de bacteris *E. coli* K12MG1655 (Δhha) amb dos plasmidis: pCA24Nhha que permet la sobre expressió de Hha sota el control del promotor T7, induïble amb IPTG, i el vector pBAD30x (x=tomb, ymob o C117S ymob), que permet la sobre expressió de la proteïna indicada sota el control del promotor araC, induïble amb arabinosa.

Les conclusions a les que hem arribat són:

- Les proteïnes Hha-TomB formen un sistema TA amb Hha la toxina i TomB l'antitoxina, tant per a cèl·lules planctòniques (Fig 13) com per a biofilms (Fig 18). També hem demostrat que la funció de TomB està sempre relacionada amb Hha.
- Hem comprovat que TomB i YmoB són funcionalment intercanviables tant en cèl·lules planctòniques (Fig 15) com en biofilms (Fig 18), demostrant que aquest mòdul TA es funcional en els bacteris entèrics.
- A més a més, hem confirmat que el mutant [C117S] YmoB la nostra construcció òptima per a la realització dels estudis estructurals de RMN, és biològicament activa com a antitoxina de Hha, tant en cèl·lules planctòniques (Fig 15) com en biofilms (Fig 18), tot i que en aquest últim cas, no és tant eficient com el WT.

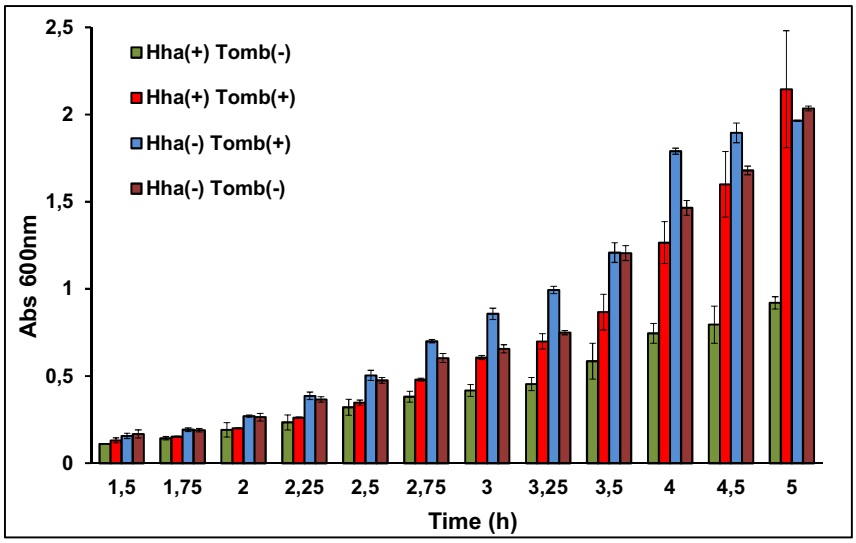


Fig 51: Corbes de creixement de *Ecoli* K12MG1655 Δhha. Les barres calorejades corresponen a les proteïnes que han estat sobre expressades; verd fosc per a Hha, vermell per a la co-expressió de Hha i TomB, blau clar per a TomB i vermell fosc per al control en el qual cap proteïna va ser sobre expressada. Els experiments es van fer per duplicat, i es mostra la desviació estàndard.

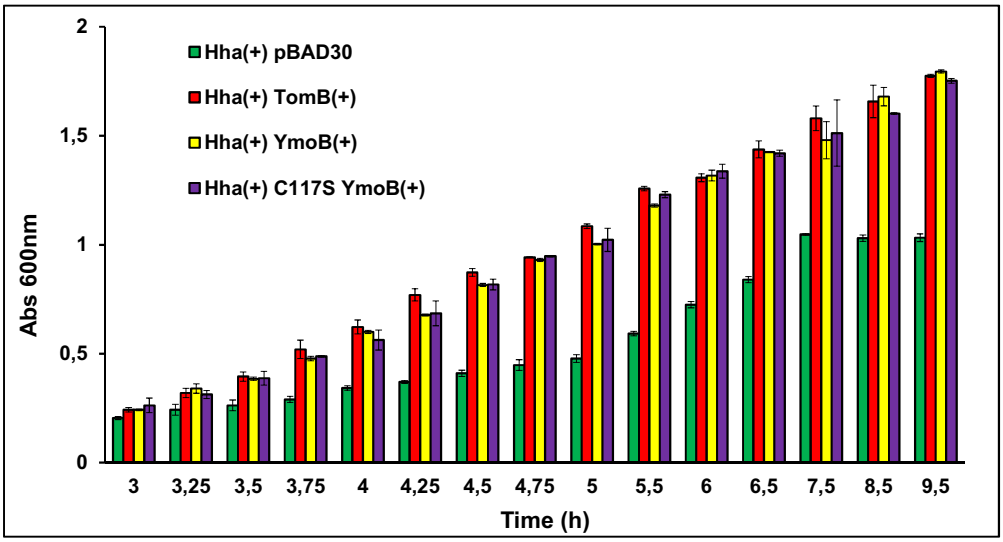


Fig 52: Corbes de creixement de *Ecoli* K12MG1655 Δhha. Les barres calorejades corresponen a les proteïnes que han estat sobre expressades; verd clar per a Hha, vermell per a la co-expressió de Hha i TomB, groc per a la co-expressió de Hha i TomB i lila per a la co-expressió de Hha i [C117S] YmoB. Els experiments es van fer per duplicat, i es mostra la desviació estàndard.

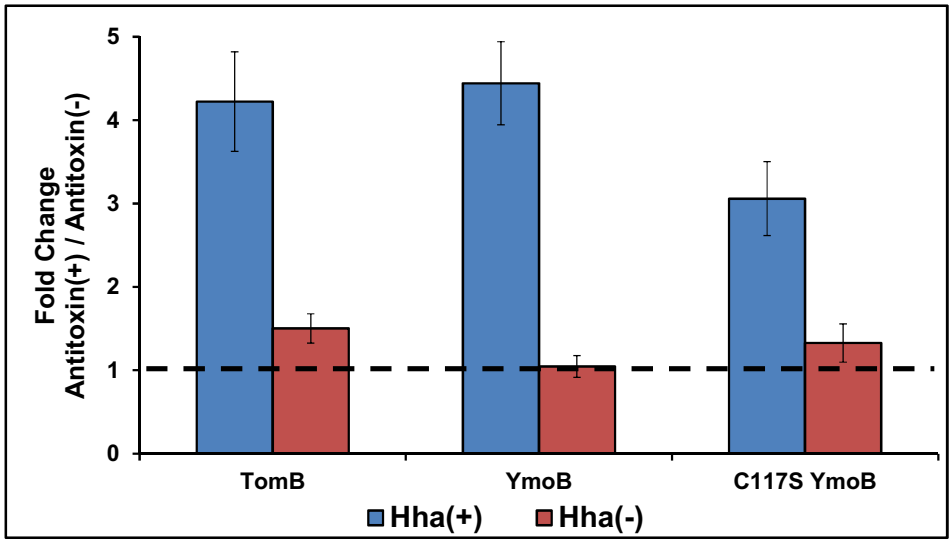


Fig 53: Efecte de les antitoxines TomB/YmoB/ [C117S] YmoB en biofilms. Mostrem el canvi d'ordre entre la formació de biofilm en presència i absència d'antitoxina. El canvi d'ordre es representa en blau quan Hha està sent sobre expressada i en vermell quan no. L'experiment es va realitzar per duplicat, i es mostra la desviació estàndard.

V.4. Determinació estructural de la construcció [C117S] YmoB mitjançant RMN

Per a la determinació de l'estructura tridimensional de [C117S] YmoB mitjançant RMN hem seguit la següent estratègia; adquisició dels experiments de RMN, assignació dels senyals de RMN als àtoms de la proteïna, obtenció de les restriccions estructurals i càlcul i refinament de l'estructura tridimensional[126, 127].

Els experiments de RMN necessaris per a l'assignació de la cadena peptídica els hem realitzat en modalitat autoservei al LRB (Laboratori de RMN de Barcelona) del CCIT (Centres Científics I Tecnològics), que pertanyen al ICTS (Institución Científico Tecnológica Singular). En canvi, degut als equilibris d'oligomerització anomenats anteriorment, els experiments per a l'assignació de les cadenes laterals i per a la obtenció de les restriccions estructurals els hem realitzat utilitzant la tecnologia Interlived-NUS-MDD, en el Swedish NMR Centre de la Universitat de Göteborg, Suècia.

Per a la obtenció de les restriccions estructurals i el càlcul de la estructura hem utilitzat el software UNIO10'[128] i per al seu posterior refinament hem utilitzat CNS (Crystallography & NMR System[129]). Hem comprovat la qualitat dels models estructurals obtinguts utilitzant PROCHECK[130].

Hem obtingut una estructura d'alta resolució, on els 20 models de menor energia presenten una mesura en els parells de RMSD[x] per sota dels 1.5Å per als àtoms pesats dels residus que conformen les hèlix α. [C117S] forma un feix de quatre hèlix (Fig 31) i la seva

[x] RMSD; Root Mean Square Deviation

superfície electrostàtica és principalment negativa, tot i que presenta dues àrees positives perfectament definides (Fig 32).

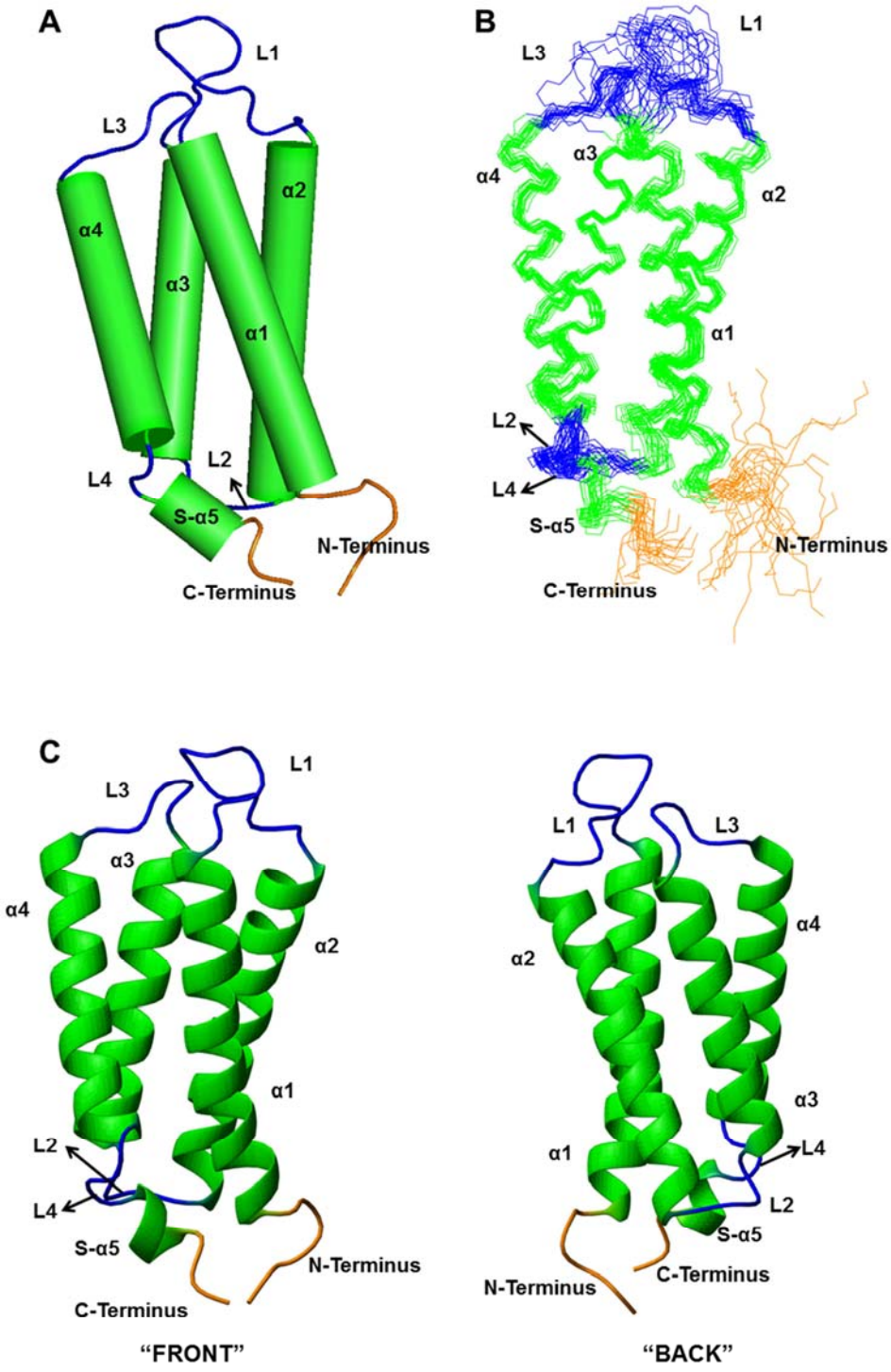


Fig 54: diferents vistes del mutant [C117S] YmoB. Les hèlix α es mostren en verd, els loops en blau i els extrems N- i C- terminals en taronja. **A:** Representació de les hèlix α en forma de cilindres per al model de més baixa energia. **B:** Superposició dels àtoms pesats localitzats a les hèlix α per als 20 models de més baixa energia. **C:** Representació de l'estructura secundària per al model de més baixa energia.

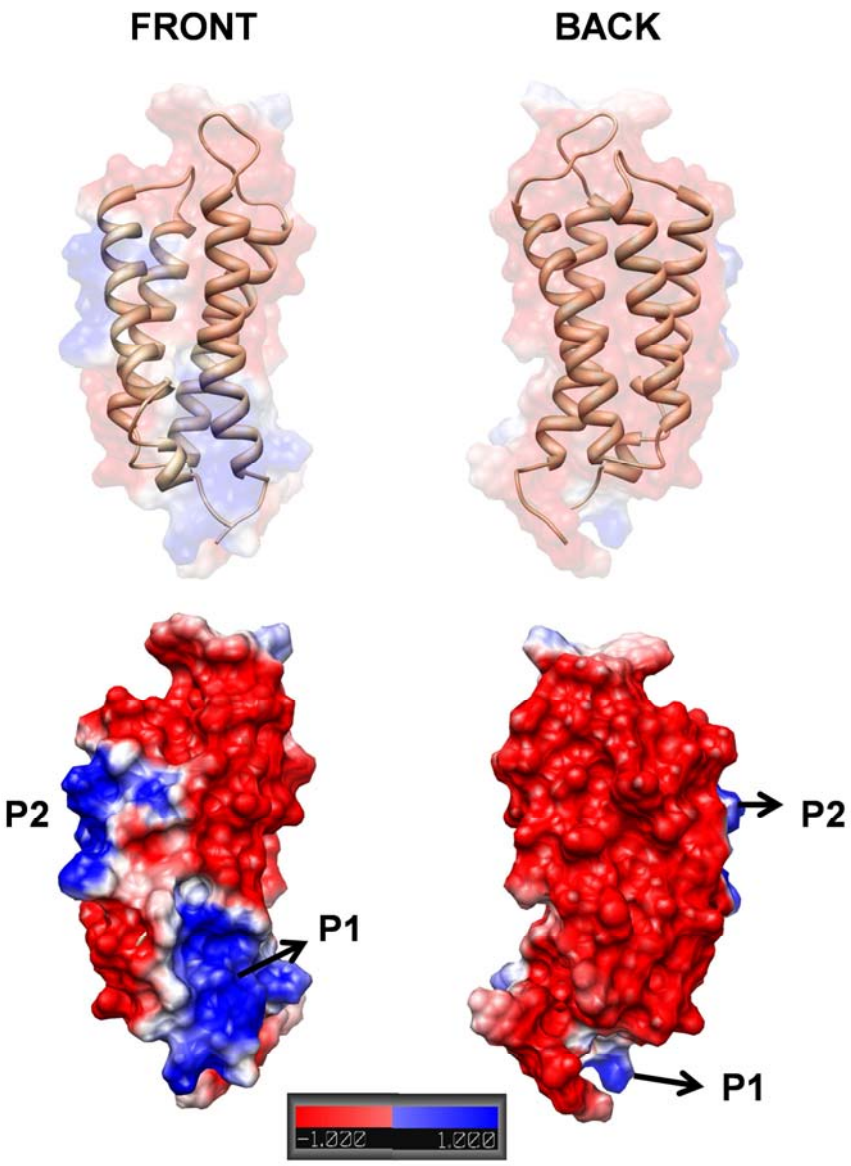


Fig 55: Potencial electrostàtic a la superfície del mutant [C117S] YmoB per al seu model de més baixa energia. Vistes frontal I posterior. En vermell es representa el potencial negatiu mentre que en blau el positiu.

V.5. Oxidació i oligomerització de YmoB: hipòtesi sobre la funció de YmoB a partir de l'estudi de la seva estructura 3D.

Durant el procés d'obtenció d'una mostra estable per a la resolució de l'estructura de YmoB hem observat que tant YmoB com el seu homòleg en *E. coli* , TomB, eren proteïnes làbils. El mutant de YmoB que conté tan sols una cisteïna a la posició 18 ha presentat la major estabilitat, tot i que han estat necessaris protocols d'adquisició d'experiments de RMN no convencionals per a poder determinar la seva estructura tridimensional.

La cisteïna 18 restant i el seu entorn estan totalment conservats entre YmoB i TomB, i aquest residu es troba totalment protegit en el cor hidrofòbic de [C117S] YmoB (Fig 56).

Hem hipotetitzat que aquesta cisteïna tan conservada i protegida deuria ser rellevant i hem estudiat en major detall tant la seva reactivitat com els processos d'oligomerització de YmoB.

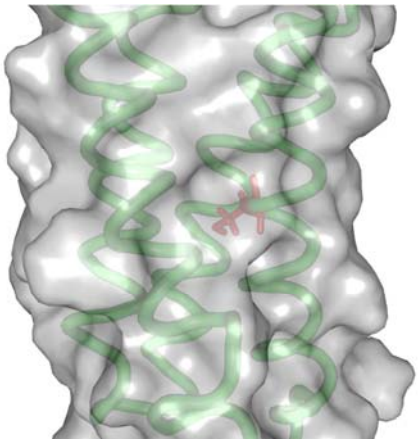


Fig 56: Expansió de l'estructura 3D del mutant [C117S] YmoB. La cisteïna 18, marcada en vermell, es troba completament orientada cap al centre del feix de 4 hèlix I totalment protegida per l'estructura terciària de la proteïna.

Per fer-ho hem utilitzat cromatografia d'exclusió molecular, espectrometria de masses, digestió amb endoproteinasa GluC seguida de nanoHPLC-ESI-MS/MS ("nano high performance liquid chromatography" acoplada a espectrometria de masses amb ionització de la mostra mitjançant nanoelectroesprai) i experiments de RMN ^{1}H-^{15}N-HSQC de mostres concentrades i diluïdes a temps 0 i 10 dies de les proteïnes YmoB i [C117S] YmoB. Hem identificat una reactivitat molt especial en C18. Aquest residu és capaç d'oxidar-se a àcid sulfènic inclús en presència de reductors fots com el TCEP. A més a més, es capaç de formar ponts disulfur amb cisteïnes d'altres molècules de YmoB, tant amb C18 com amb C117, i en el cas de la unió amb C117, també hem identificat la formació de tiosulfinatos, és a dir, ponts disulfur oxidats (C18[O]C117).

Així doncs, hem identificat entitats derivades de la proteïna YmoB que podrien actuar com a antitoxines de les toxines de la família Hha/YmoA:

a) YmoB en la seva forma monomèrica reduïda

   Per comprovar aquesta hipòtesi hem realitzat una sèrie d'experiments exploratoris de RMN per veure si YmoB interacciona amb YmoA o, com suggereix Garcia-Contreras i col·laboradors[123], amb ADN. Fins al moment no s'han obtingut resultats positius.

b) Dímers (o) formes oligomèriques de YmoB

   La forma oxidada de YmoB podria existir com a monòmer oxidat o com a dímer o formes oligomèriques majors. L'estructura del mutant [C117S] YmoB mostra

la presencia de dues àrees carregades positivament. Un model aproximat generat per l'aproximació de dos monòmers de [C117S] YmoB que permeti la minimització de les distàncies entre les C18 i les C117 mostra com les àrees positives de les dues molècules es situen en una línia contínua, suggerint la possible interacció amb molècules carregades negativament, com ara el ADN o el ARN (Fig 49).

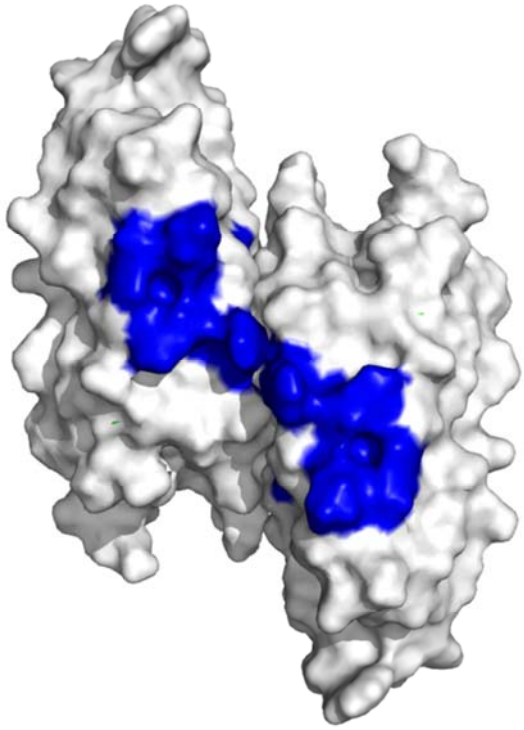


Fig 57: Model aproximat de dos hipotètics dímers de YmoB amb la distancia entre les cisteïnes C18-C117 minimitzada. S'ha representat la superfície de cada monòmer, i s'ha representat en blau la superfície corresponent als aminoàcids R99, K102, R106 i R109, responsables d'una de les àrees positives de YmoB.

c) YmoB amb la cisteïna 18 oxidada en forma d'àcid sulfènic

El fet de que la C18 s'oxidi amb tanta facilitat, inclús tan sols per l'activitat de l'oxigen molecular ens fa pensar en la seva possible activitat com a sensor d'oxigen. Dins del biofilm aquesta és una tasca important, ja que els nivells d'oxigenació dels bacteris són diferents depenent si es troben a la superfície o a l'interior del biofilm. La activitat com a antitoxina de la forma WT i del mutant [C117S] de YmoB és equivalent en bacteris *E. coli* en estat planctònic, és a dir, en situació de bona oxigenació. En canvi, en biofilms, on la falta d'oxigen pot donar-se, el mutant [C117S] perd eficiència com a antitoxina de Hha, respecte a

la forma WT. Aquesta discussió es troba en la mateixa línia de la hipòtesi de que la forma oxidada de YmoB és la antitoxina de Hha.

Seran necessaris nous experiments per determinar quina és l'espècie de YmoB rellevant que actua com a antitoxina de Hha, així com per a detallar exactament el mecanisme d'acció de les proteïnes TomB/YmoB com a antitoxines. També caldrà definir el rol que juga la C18 i la seva capacitat d'oxidació a àcid sulfènic i/o tiosulfinat.

V.6. Resum dels resultats obtinguts

En resum, els principals resultats obtinguts en aquesta tesi doctoral són:

1. Obtenció d'una mostra òptima per a la realització dels experiments de RMN; mutant [C117S] YmoB a 1.4mM, 20mM de [Na_2HPO_4 + NaH_2PO_4], 150mM NaCl, 0.2mM EDTA, 1mM TCEP, 0.01% NaN_3, pH=7.00.
2. Demostració de la funció com a Toxina-Antitoxina del parell Hha-TomB, sent Hha la toxina i TomB l'antitoxina, en cèl·lules *E. Coli* planctòniques i en biofilms. Complementació en *E. Coli* de la funció com a antitoxina de TomB per part de YmoB i de [C117S] YmoB en cèl·lules planctòniques i en biofilms.
3. Resolució de l'estructura tridimensional de la construcció [C117S] YmoB. La proteïna forma un feix de quatre hèlix i la seva superfície electrostàtica es majoritàriament negativa, mostrant dues àrees ben definides amb càrrega positiva.
4. Identificació de tres possibles espècies de YmoB com a forma activa d'antitoxina; YmoB en la seva forma monomèrica reduïda, dímers (o formes oligomèriques majors) de YmoB y YmoB amb la cisteïna 18 oxidada en forma d'àcid sulfènic.

V.7. Bibliografía

[117] T. K. Wood (2009). Insights on Escherichia coli biofilm formation and inhibition from whole-transcriptome profiling. *Environmental Microbiology* ***11****(1), 1-15.*

[118] Ghafourian, S., Raftari, M., Sadeghifard, N., & Sekawi, Z. (2013). Toxin-antitoxin Systems: Classification, Biological Function and Application in Biotechnology. *Current issues in molecular biology*, *16*(1), 9–14.

[119] C. Beloin , A. Roux , J.-M. Ghigo (2008). *Escherichia coli* Biofilms. *Current Topics in Microbiology and Immunology* ***322****, 249–289.*

[120] Yee, A., Gutmanas, A., & Arrowsmith, C. H. (2006). Solution NMR in structural genomics. *Current opinion in structural biology*, *16*(5), 611–7.

[121] Yee, A., Chang, X., Pineda-Lucena, A., Wu, B., Semesi, A., Le, B., Ramelot, T., et al. (2002). An NMR approach to structural proteomics. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(4), 1825–30.

[122] S. G. Hyberts, H. Arthanari, and G. Wagner (2012). Applications of non-uniform sampling and processing. *Topics in Current Chemistry* **316**, 125–148.

[123] García-Contreras, R., Zhang, X.-S., Kim, Y., & Wood, T. K. (2008). Protein translation and cell death: the role of rare tRNAs in biofilm formation and in activating dormant phage killer genes. *PloS one*, *3*(6), e2394.

[124] Wang, X., & Wood, T. K. (2011). Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Applied and environmental microbiology*, *77*(16), 5577–83.

[125] A. Heydorn, A. T. Nielsen, M. Hentzer, C. Sternberg, M. Givskov, B. K. Ersbøll and S. Molin (2000). Quantification of biofilm structures by the novel computer program COMSTAT. *Microbiology*, **146**, 2395–2407.

[126] P. Güntert (1998). Structure calculation of biological macromolècules from NMR data. *Quarterly Reviews of Biophysics **31**(2), 145-237.*

[127] P. Günter (2009). Automated structure determnation from NMR spectra. *Eur Biophys **38**, 129-143.*

[128] T. Herrmann (2002). UNIO'10. Automated NMR Data Analysis for Protein Structure Determination and More. http://www.unio-nmr.eu

[129] A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannuh, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren (1998). *Crystallography & NMR System*: A New Software Suite for Macromolecular Structure Determination**.** *Acta Crystallographica*, **D54**, 905-921.

[130] Laskowski R A, Rullmannn J A, MacArthur M W, Kaptein R, Thornton J M (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*, 8, 477-486.