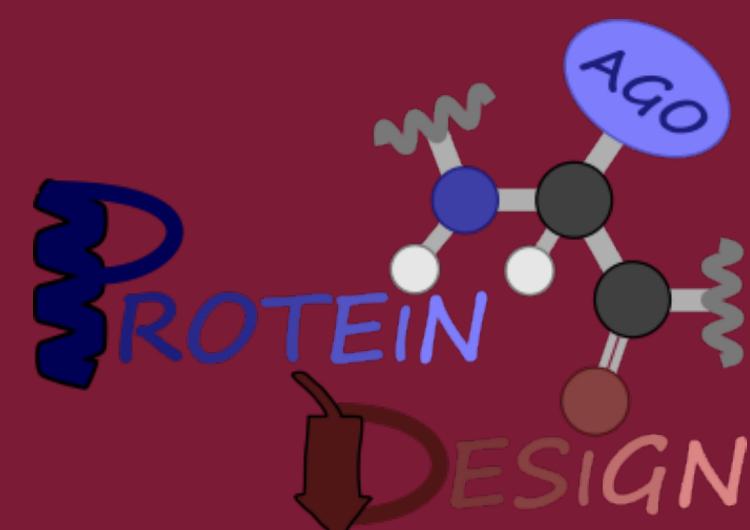
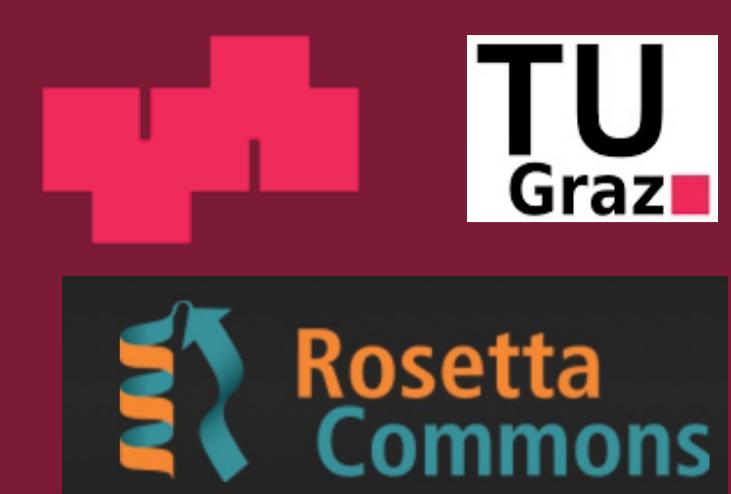


# Computational design of dimeric *de novo* heme-binding helical bundle proteins



Veronica Delsoglio\*, Gustav Oberdorfer  
Graz University of Technology, Institute of Biochemistry, Austria



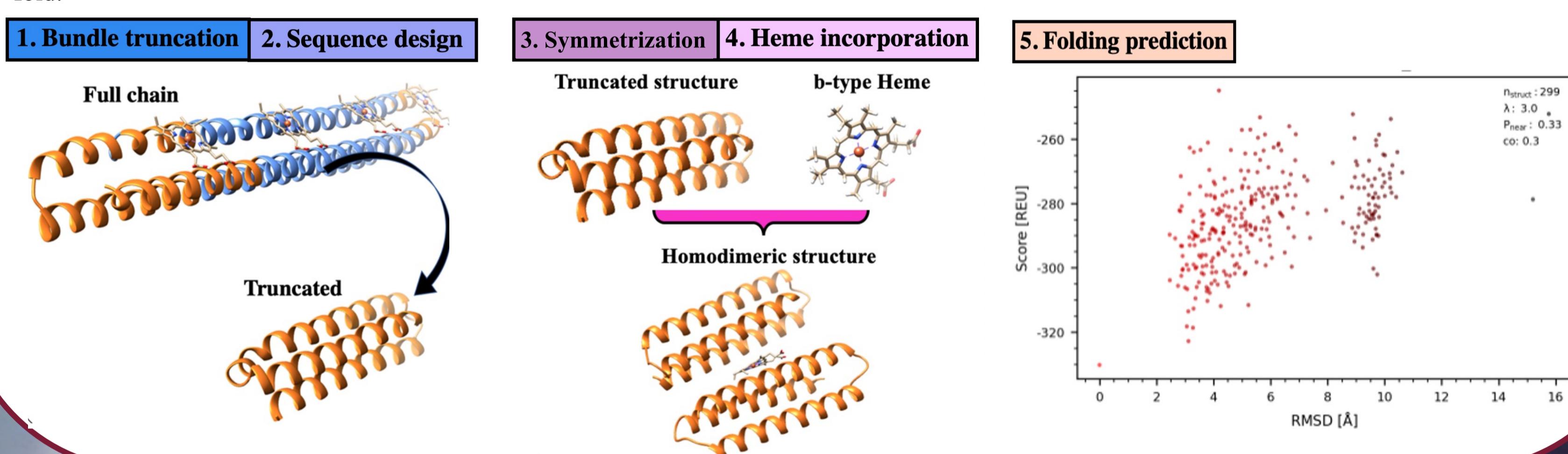
## INTRODUCTION

The comprehension of protein biochemistry, biophysics and folding is important in order to create new protein structures which could tackle important challenges in biotechnology. Indeed the amino acid sequence specifies the protein fold and the protein function. The design of proteins with altered or novel functions is challenging. In the last years, the computational methodology has advanced considerably, reaching atomic-level accuracy. With *de novo* protein design it is now possible to simulate protein sequences and predict which ones exhibit specific properties and desired functions.<sup>1</sup> Here I show my efforts towards the design and the production of dimeric helix bundles, which bind b-type hemes via a bis-histidine coordination. My work is based on two different homodimeric  $\alpha$ -helices bundles – 4\_3 and 4\_4. These structures have different amino acid sequences, but are both constituted by two antiparallel untwisted three-helix bundles of 27-residue helices and a 18-residue repeating unit (5 full circles, or ‘layers’, until they reach identity)<sup>2</sup>. Moreover, they exhibit two histidine residues which are placed to coordinate the b-type heme cofactor, so the aim is to obtain artificial proteins with a symmetric binding pocket (BP) interacting with an asymmetric ligand, a moiety often observed in nature, but never really investigated in artificial systems. The 4\_3 and 4\_4 proteins are obtained after modeling and optimization processes, using Rosetta software, interfaced to cycles of protein production and biochemical characterisation.

## RESULTS

### Modelling with Rosetta software and ab initio folding prediction

The process of protein design and optimization is simplified starting from truncated alpha helical bundle structures (1). Thousands of redesigned protein sequences are generated using PackRotamers mover. Comparing the sequences in terms of residual Rosetta Energy Units (REU) per residue, allows the best residue at each sequence position to be selected (2). Using SetupForSymmetry mover the monomeric bundle is homodimerized in a C2 fashion (3). The incorporation of the b-type heme molecule (4) is conducted using Match application, in order to find the position where the binding site can be introduced, and then EnzRepackMinimize and AddOrRemoveMatchCsts movers to design, repack and minimize the protein – ligand interface. The ab initio folding prediction (5) is conducted with RT fold.



### Protein production and biochemical characterisation

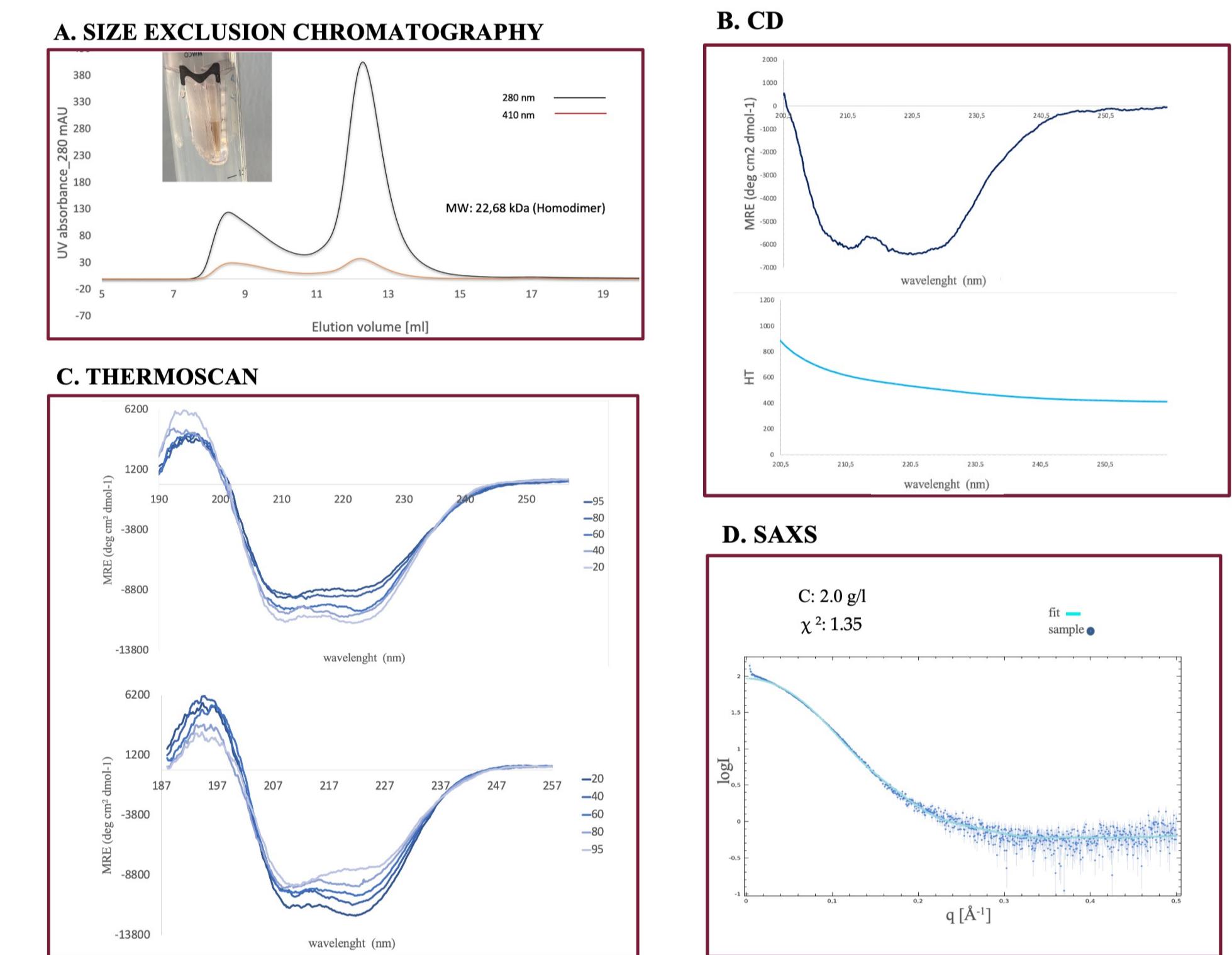
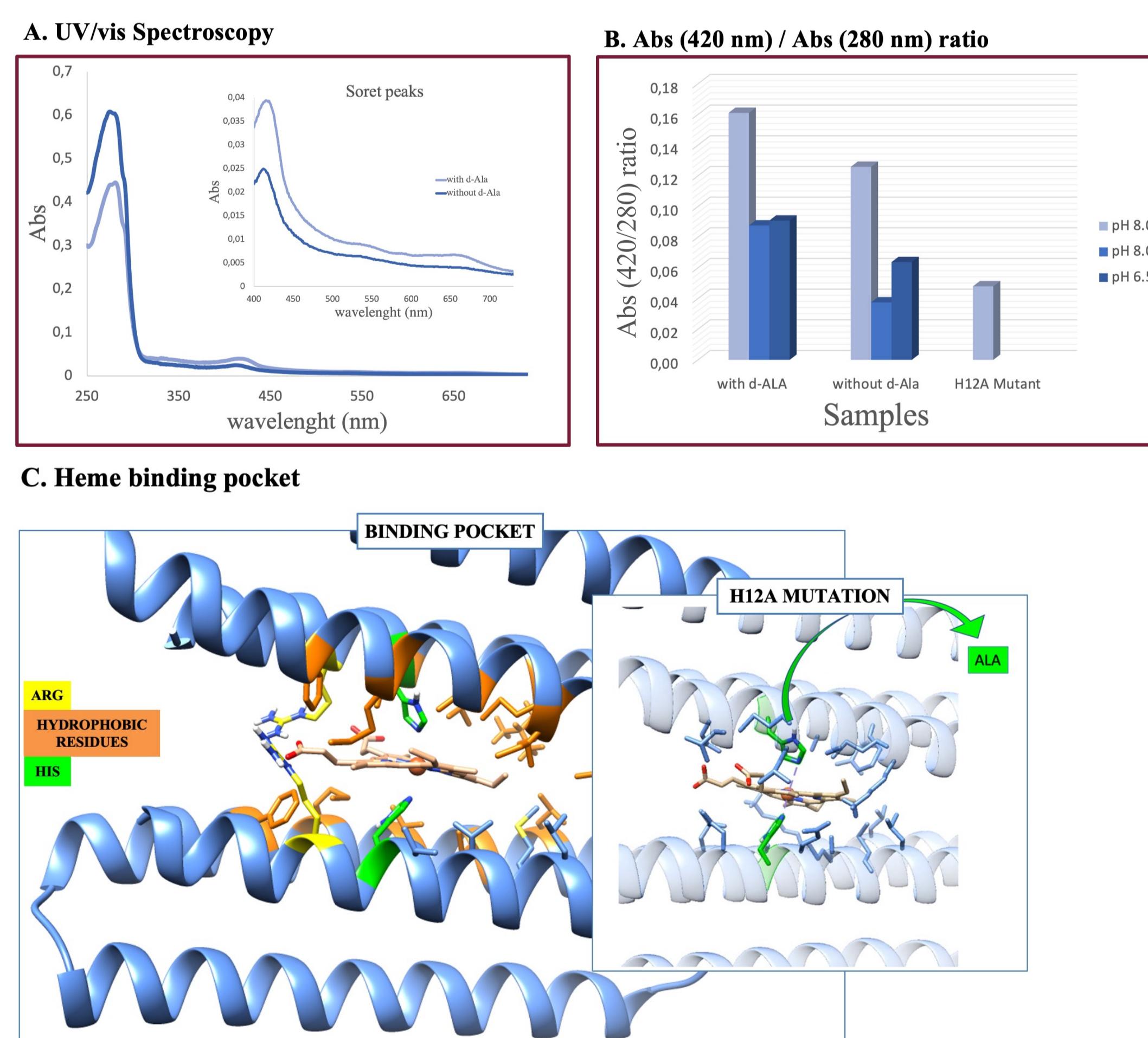


Figure 1. Structural Characterisation: homodimeric shape. (A) Size Exclusion Chromatography (SEC) profile on FPLC. The eluted peak corresponds to the homodimer which binds the b-type heme. (B-C) CD spectra and Protein Thermostability analysis. The CD spectra (B) shows two negative peaks (208 nm and 222 nm) which are typical of  $\alpha$ -helices secondary structures. The protein thermostability (C) is also tested, demonstrating the protein resistance until 95 °C. (D) Small X-Rays Scattering (SAXS) analysis. The data analysis performed using PRIMUS software shared a good fit ( $\chi^2$ ) between the experimental data and the theoretical one.

### Bis-histidine heme coordination

Figure 2. Structural Characterisation: b-type heme binding. (A) UV-vis Spectroscopy analysis. The profile shares two main peaks, one at 280 nm (protein) and one at 420 nm (Soret peak - caused by the presence of the binding heme molecule). The dark-blue profile is referred to the protein produced adding  $\delta$ -Aminolevulinic acid ( $\delta$ -ALA) – the b-type heme precursor -, the light-blue one to the same protein produced without  $\delta$ -ALA.

(B) Ratios of absorbance peaks at 420 nm and 280 nm. The histogram represents the absorbance ratio of different samples of the same proteins purified at different pH values. The pH could influence the protonation state of the histidine residues, causing shifts of Soret peak. In the histogram, the first sample starting from the right is the protein produced adding  $\delta$ -ALA, the second sample is the protein produced without  $\delta$ -ALA, the third sample is the mutated form protein, produced with  $\delta$ -ALA, where the bis-coordinating histidines are mutated in two alanines. (C) Overview of how the optimized binding pocket for b-type heme binding looks like. In green are represented the coordinating histidine residues, in orange all the hydrophobic residues of the pocket, favouring an hydrophobic environment for the heme.



## CONCLUSIONS:

- After several cycles of protein modeling - optimization and protein production - biochemical characterisation, two homodimeric proteins able to bind the b-type heme are found (4\_3 and 4\_4).
- The CD experiments confirm the designs exhibit  $\alpha$ -helical secondary structures. They remain folded also at very high temperatures. The designed models of the proteins seem also to fit well the small x-rays diffraction pattern, meaning the proteins are folded in solution.
- The Soret and protein peak ratios of  $\delta$ -ALA, not  $\delta$ -ALA and H12A mutant samples are compared, confirming the histidine residues act as the main binding residues of the heme inside the BP. However, also the residues of the BP seem to play a role, in favouring the heme binding.

## REFERENCES

1 P.-S. Huang, G. Oberdorfer, C. Xu, X. Y. Pei, B. L. Nannenga, J. M. Rogers, F. DiMaio, T. Gonen, B. Luisi, D. Baker, High thermodynamic stability of parametrically designed helical bundles, *Science* 346, 481-484 (2014).

2 <https://www.rosettacommons.org>

\* Contact: veronica.delsoglio@tugraz.at