Computational and Structural Biotechnology Journal 27 (2025) 5078–5086

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research Article

An integrative structural biology approach to identify the binding mode of a nanobody towards the pea ascorbate peroxidase

Claudia D'Ercole [a,1], Marco Orlando [b,1], Kristina Eleršič Filipič [c], Ario de Marco [a,2,*]

[a] Lab of Environmental and Life Sciences, University of Nova Gorica, Vipavska cesta 13, Nova Gorica 5000, Slovenia
[b] Department of Biotechnology and Biosciences, University of Milano Bicocca, Piazza della Scienza 2, Milano 20126, Italy
[c] National Institute of Chemistry, Hajdrihova 19, Ljubljana SI-1000, Slovenia

ARTICLE INFO

Keywords:
Nanobodies
Ascorbate peroxidase
Hydrogen-deuterium exchange coupled to mass spectrometry
Integrative structural biology
Protein cross-linking

ABSTRACT

The optimization of diagnostic devices such as biosensors often requires understanding the molecular details of the interaction between capture and target biomolecules. This can be experimentally obtained by cryo-electron microscopy, the preferred method for the analysis of large protein complexes, while NMR and x-ray crystallography are effective for determining the structure of complexes formed by relatively small molecules. Nevertheless, all these approaches are demanding in terms of time and resources and, therefore, we explored the possibility to reduce the experimental load by compensating with *in silico* modelling. Here we demonstrate that an accurate prediction of the binding mode between a nanobody and its target pea ascorbate peroxidase, an oxidative stress biomarker in plants, can be obtained by combining cross-linking mass spectrometry, hydrogen-deuterium exchange coupled to mass spectrometry and *in silico* modelling. Such model allowed to precisely design negative mutants that confirmed its accuracy. In conclusion, this study shows that an unconstrained prediction based on deep learning models is still not sufficiently reliable for new targets and difficult-to-model biomolecule classes such as nanobodies, while an experimental-guided approach can provide valuable structural information for lead optimization campaigns of such reagents.

1. Introduction

The biochemistry of forest plants has been historically neglected, probably because of the technical issues related to the study of these organisms that are not compatible with lab conditions due to their dimensions and the length of their physiological cycles. Nevertheless, in an age characterized by rapid climate changes, the study of forest trees should attract major attention, given their role as CO_2 sink and storage pool [1]. Specifically, it would be necessary to better understand how external stress factors can affect the tree metabolic equilibrium and whether enzymes known for their detoxifying activity might be useful to monitor, and possibly mitigate, altered physiological conditions due to environmental stress. Ascorbate peroxidase enzymes are a class of candidates that can be widely suited as biomarkers for such monitoring studies, considering their crucial role as scavengers of active forms of oxygen [2]. In the case of forest trees, analyses of physiological fluctuations are hindered by the lack of reliable reagents such as antibodies suitable for setting the diagnostic assays necessary to quantify the protein biomarkers. Conventional antibody discovery is time-consuming and expensive, while the isolation of antigen-specific recombinant antibodies by panning libraries has become a widespread and effective modality to recover immunoreagents applied in several research fields [3]. Among antibody types, there has been an increasing interest in nanobodies because of the biotechnological advantages they can offer in applications as different as super-resolution microscopy, assisted crystallography, cryo-EM, *in vivo* imaging, therapy or affinity chromatography [4–9] and because of the simplicity by which they can be engineered into reagents tailored for their final applications [10].

In a recent report [11], we showed a straightforward biopanning protocol for the isolation of two nanobodies (BA5 and BG12) specific for *Populus nigra* (poplar) ascorbate peroxidase (pAPX), previously shown to be a biomarker of physiological stress conditions [12]. Anyway, nanobodies obtained by panning benefit from improvement of their biophysical features, such as affinity or stability [13]. Conventional

---

[*] Corresponding author.
E-mail address: ario.demarco@ung.si (A. Marco).
[1] These authors equally contributed
[2] ORCID 0000–0001-7729–819X

https://doi.org/10.1016/j.csbj.2025.11.024
Received 8 September 2025; Received in revised form 10 November 2025; Accepted 12 November 2025

optimization approaches are based on random mutagenesis followed by further rounds of panning performed under more stringent conditions [14]. Recently, rational design based on *in silico* predictive methods and artificial intelligence (AI) systems for *de novo* redesign is taking over [15,16] but the success rate is highly dependent on the accuracy of the structure prediction of the nanobody in complex with its target [17]. Whereas X-ray crystallography is still the gold standard for determining the molecular structures of proteins, now reliable information can be obtained extremely more rapidly by leveraging AI methods trained on published structural data [18,19]. Nevertheless, in public repositories there are just few structurally characterized complexes between nanobodies and corresponding antigens [20], a condition which does not allow AI to sample the highly variable conformations of complementarity determining regions (CDRs) that usually compose most of the nanobody paratopes. The consequence is that the training possibility for nanobody-antigen complexes has been limited and still impacts negatively the accuracy of AI-based prediction methods. Therefore, alternative methods must be used to gain the necessary structural information. Integrative structural biology approaches have proven to be successful in this task, such as the combination of nuclear magnetic resonance (NMR) and cross-linking mass spectrometry (XL-MS) to build a computational model of the biparatopic anti-HER2 nanobody [21]. In the most recent years, hydrogen-deuterium exchange coupled with mass spectrometry (HDX-MS) has been successful exploited as well to map epitopes and paratopes [22–24]. This method was often paired to other techniques, such as XL-MS differential photoreactive surface modification and molecular docking [25,26]. This work reports on an integrative experimental and computational structural biology approach designed to obtain a reliable model of the interaction between the previously screened nanobodies and pAPX, with the aim to collect information instrumental to foster the development of a second generation of anti-pAPX nanobodies.

2. Materials and methods

2.1. Unconstrained pAPX-nanobody complex modelling

The amino acid sequence of pAPX from *Pisum sativum* was obtained from NCBI GeneBank (accession number: AAA33645), while those of nanobodies BA5 and BG12 were retrieved from D’Ercole at al. [11]. The sequences employed in this study are reported in the Figure S1. 3D structures of each nanobody interacting with the APX dimer were modelled by Colabfold 1.5.5 (https://github.com/sokrypton/ColabFold, [27]), using different AlphaFold2-Multimer v.2.2 (AFM) models (v1, v2, v3, [28]). For each attempt, made with a different model version, five runs with different random seeds were performed; for each seed, five different combinations of *–max-seq* and *–max-extra-seq* multiple sequence alignment subsampling options (32:64, 64:128, 128:256, 256:512, 512:1024, 1024:2048, 2048:4096) were used, as described in da Silva et al. [29] and *–num-recycles* was set to 20 *–recyle-early-stop-tolerance* to 0.5. The PDB ID: 1APX was used as template for pAPX, while no template was used for the nanobodies. All the generated models were re-ranked by HADDOCK physics-based score using HADDOCK3 suite (2024.12.0b7 version [30]) and their consistency was evaluated by applying the default FCC clustering strategy (6 Å root mean square deviation, RMSD, cut-off) to the nanobodies, after superposing the complexes along the backbone atoms of pAPX monomer interacting with the nanobody.

2.2. Protein production

The synthetic gene corresponding to pAPX was produced by Twist Bioscience (San Francisco, CA – USA) after sequence optimization for expression in *E. coli.* The resulting DNA was cloned in a modified pET-14b vector in frame with AviTag. The construct was expressed in *E. coli* BL21(DE3) cells after the addition of 0.5 mM isopropyl-β-D-thiogalactoside (IPTG). Cells were grown for 3 h, at 37 °C, pelleted (30 min at 4.500xg) and lysed in lysis buffer (Tris-HCl 50 mM, pH 8.0, NaCl 500 mM, $MgCl_2$ 5 mM) by alternating three cycles of freezing / thawing. Lysates were then sonicated on ice (6x, amplitude 80 %, 1 min pulse, 1 min off), incubated for 30 min at RT with DNase I (33 U/ml) and lysozyme (100 µg/ml) and finally centrifuged at 4 °C (30 min at 13,000xg). Proteins were purified by IMAC from the resulting supernatant using the fast liquid chromatography ÄKTA pure™ system (Cytiva - Marlborough, MA, USA). Samples were injected at a flow rate of 1 ml/min into a Talon Hi-Trap column (Cytiva - Marlborough, MA, USA), previously equilibrated with washing buffer (WB, Tris-HCl 50 mM, pH 8.0, NaCl 500 mM, imidazole 15 mM). After column washing in the same buffer, the bound proteins were eluted by increasing the imidazole concentration in the elution buffer (EB, Tris-HCl 50 mM, pH 8.0, NaCl 500 mM, imidazole 500 mM). The pAPX-AviTag construct was desalted in PBS, pH 7.4, by means of a Hi-Trap Desalting column (Cytiva - Marlborough, MA, USA) and evaluated by gel filtration (Superdex™ 75 Increase 5/150 GL or Superdex™ 200 10/300 GL, Cytiva), UV scan and SDS-PAGE as reported previously. Protein aliquots were stored at - 80 °C.

Unique nanobody clones were cloned into a modified pET-14b vector between *Nco*I and *Not*I restriction sites to obtain fusion proteins with AviTag. Primers ordered from Kemomed d.o.o. (1231 Ljubljana- Črnuče) were employed for the amplification of the insert and the vector using repliQa HiFi ToughMix® (Quantabio). The Inserts were amplified with the primers: F1.Nb (5’-tttaactttaagaaggagatataccatggctgatgtgcagctgc-3’) and R1.Nb (5’- agatgtcgttcagacctgcggccgc*tcc*gccgctggagacggtgac-3’). Vector amplification was performed using F2.pET (5’-ggagcggccgcaggtctg-3’) and R2.pET (5’-cccatggtatatctccttcttaaagttaaaca-3’). The PCR products were separated by agarose gel electrophoresis (1 % w/v agarose) and cleaned using a peqGOLD Gel Extraction Kit (PEQLAB Biotechnologie). An insert to vector ratio of 3:1 was used and mixed 1:1 with Gene Art™ Gibson Assembly MasterMix (Invitrogen™). After 20 min incubation at 50 °C, 60 µl of competent DH5α *E. coli* cells were transformed and the resulting amplified plasmids were isolated using peqGOLD Plasmid Miniprep Kit II (PEQLAB Biotechnologie) and their sequence was confirmed by Sanger sequencing (Azenta Life Sciences). Nanobody production was evaluated with a small-scale test after expression in *E. coli* BL21-DiscoTune cells induced by means of rhamnose 50 µM and IPTG 0.2 mM, at 20 °C, overnight. Large-scale nanobody purification was performed as described above. Nanobody mutants designed after *in silico* modelling were synthetized by Twist Bioscience (San Francisco, CA – USA) and then cloned, purified and evaluated for their quality using the same procedures described above for wild type BA5. Protein concentration was inferred by quantifying the sample absorbance at 280 nm and dividing the resulting value for the specific absorbance coefficient.

2.3. Surface plasmon resonance

Surface plasmon resonance (SPR) analyses were performed using a Biocore X100 instrument, pAPX as the ligand and both wild type and BA5 mutants as analytes. pAPX was immobilized on a CM5 chip by amine coupling, reaching a Rmax of 5480 RU. PBS, pH 7.4 was used as the running buffer and all measurements were performed at 25 °C. The analytes were injected at concentrations ranging from 250 to 8000 nM, using a flow rate of 30 µl/min, association time of 45 s and dissociation time of 135 s. The chip was regenerated by extensive washing with 1.5 M $MgCl_2$ dissolved in water. The results represent measurements of three replicates; wild type BA5wt was measured a fourth time at the end of experiments as a control of the system stability and chip regeneration efficiency. The results corresponded to the difference between the signals in Fc2 (analyte channel) and Fc1 (reference channel). The dissociation constant (K_D) values were determined using steady-state analysis of the SPR binding sensorgrams instead of kinetic fitting of the association and dissociation phases. Measurements were performed at equilibrium or near-equilibrium points within a defined 5 s analysis window where the signal was stable. Data analysis were performed by means of Biacore T200 Evaluation Software 3.2.1 and Origin 8, using the option for steady-state fitting, with appropriate baseline corrections.

2.4. *Hydrogen-deuterium exchange coupled with mass spectrometry (HDX-MS)*

To observe HDX kinetics of proteins by mass spectrometry, 85 μM of pAPX-AviTag alone or complexed 1:1 with either Nb-BA5-AviTag or Nb-BG12-AviTag were diluted 4.7 times into deuterated buffer (D_2O PBS, pD 7.4) to initiate isotopic exchange in the accessible regions of the proteins. The reaction was conducted at 2 °C and different reaction times were evaluated (2 s, 10 s, 90 s, 20 min, 2 h). Time points at 2 and 90 s were triplicated (with the exception of the 90-second time point for APX-BA5 condition in APX HDX-MS experiment, which was duplicated). The reactions were quenched by adding an acidic solution (glycine 0.54 M, urea 3.2 M, thiourea 1.1 M, TCEP 214 mM, pH 2.3) in a ratio of 14:1. Samples were injected on a co-immobilized Pepsin / Nepenthesin-2 protease column (SecurityGuard™ ULTRA Cartridge UHPLC Fully Porous Polar C18, 2.1 mm ID, Phenomenex, Torrance, CA). The generated peptides were trapped and desalted for 3 min at a flow rate of 0.2 ml/min using an isocratic pump delivering 0.4 % aqueous solution of formic acid. Then peptides were separated on an equilibrated (solvent A: 2 % acetonitrile / 0.4 % formic acid in water) C18 reversed-phase column (Luna® Omega 1.6 μm Polar C18 100 Å 100 × 1.0 mm) with a 10–45 % linear gradient of solvent B (95 % acetonitrile / 0.4 % formic acid) delivered by the 1290 Infinity II LC system pumping at 50 μl/min. The protease, trap column and analytical column were kept in a box at 0 °C. Peptides were detected with TimsToF Pro mass spectrometer (Bruker Daltonics). The peptide mapping based on tandem mass spectrometry was performed for non-deuterated samples using the same conditions as mentioned above. Measured data were exported by DataAnalysis v. 5.3 software (Brunker Daltonics) and isotopic distribution was analysed using DeutEx software [31]. Measured deuterium incorporation was corrected using fully deuterated controls prepared as described previously [32]. Differences were considered significant when their values for each condition differed by more than two times the average standard deviation. Moreover, statistical analysis of uncertainty and difference significance testing with a coverage factor $k = 4$ were performed based on recommendations for replicated time points [33]. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [34] partner repository with the dataset identifier PXD070410.

2.5. *Chemical cross-linking mass spectrometry (XL-MS)*

For chemical cross-linking, 20 μM of pAPX-AviTag in PBS buffer, pH 7.4, was complexed 1:1 with either Nb-BA5-AviTag or Nb-BG12-AviTag and cross-linked using 100 molar excess of either CID-cleavable reagents DSPU (Urea cross-linker – C3-arm, CF Plus Chemicals) or DSBU (Urea cross-linker – C4-arm, CF Plus Chemicals). Both cross-linkers were freshly prepared before the reaction by dissolution in anhydrous dimethylsulfoxide. The cross-linking reaction proceeded for 30 min, at 20 °C. Cross-linked samples were reduced for 30 min, at 60 °C using 10 mM dithiothreitol. Next, iodoacetamide was added to the concentration of 20 mM and incubated for 30 min, at 20 °C. Samples were then diluted 2 times by adding ammonium bicarbonate 50 mM, pH 8, and the protein digestion proceeded overnight, at 37 °C, after the addition of trypsin gold (Promega) in a 1:20 (*w/w*) ratio. Tryptic peptides were loaded onto Evotip Pure (Evosep), desalted off-line, and tips were applied to the Evosep One (Evosep) autosampler. The eluted peptides were loaded onto the reverse phase analytical column (PepSep C18 15 cm×150 μm, 1.5 μm) (Bruker Daltonics) directly interfaced with the captive spray ionization source of TimsTOF SCP mass spectrometer (Bruker Daltonics) operated in positive data–dependent mode. Acquired LC-MS/MS data were exported in DataAnalysis v. 5.3 (Bruker Daltonics) and cross-linked peptides were searched by Merox 2.0.1.4 software [35]. MeroX software was used with the following settings: precursor precision (MS1) 5.0 ppm, fragment ion precision (MS2) 10.0 ppm, cleavage after lysine and arginine, number of miscleavages 3, static modifications: carbamidomethylation of cysteine, variable modifications: oxidation of methionine, cross-linker specificity sites: lysine, serine, threonine, tyrosine, N-terminus, score cut-off: 50, FDR cut-off 1 %. The raw mass spectrometry data were deposited under ProteomeXchange Consortium via the PRIDE [34] partner repository with the dataset identifier PXD070481.

Table 1
pAPX and nanobody amino acid regions sensitive to deuterium exchange, as identified by HDX-MS.

TARGET	amino acid residues	Structural element
APX	37–77	pAPX active site
	249–275	AviTag, HisTag
Nb-BA5	20–37	CDR1
	95–105	CDR3
	127–144	AviTag, HisTag
Nb-BG12	22–34	CDR1
	100–112	CDR3
	123–151	AviTag, His,Tag

2.6. *Integrative pAPX-BA5 modeling*

pAPX dimer and BA5 monomer structures were predicted similarly as reported above but deepfold_v1 model [36] was used instead of AFM v3 in the case of BA5. The side-chain torsion angles of the best ranked model (plDDT score) were refined using DiffPack [37], followed by an energy minimization and a short NVT classical molecular dynamics (MD) simulation (8000 steps) in TIP3P water molecules (10 Å padding) with the addition of charge neutralizing ions (either Na^+ or Cl^-). This step was performed using OpenMM 7.7.0 under amber ff14SB force field, with a time step of 2.0 fs and at 30 °C. The refined models were used for experimental-guided docking exploiting the web-server implementation of HADDOCK 2.4 [38]. Formatted PDB files were uploaded to the HADDOCK server using the expert access level. Distances between Cα residues were enforced as distance restraints between 10 Å (minimum) and 30 Å (maximum) [39] to incorporate the information on intermolecular cross-links collected using DSPU and DSBU crosslinkers (Table 1, Table S1). The pAPX residues predicted by HDX-MS, those from the unconstrained pAPX-BA5 complex model and those predicted by C-port [40] were set as “active”. The same was done for the three BA5 CDRs. According to HADDOCK definition, “active” residues are directly involved in the interaction, while “passive” residues are in direct contact to the “active” residues and are automatically assigned by the system. For each docked complex, 10,000 rigid-body docking structures were generated (it0), followed by a second iteration of semi-flexible docking (it1), with only the interface being fully flexible. Finally, the best 400 docked models in terms of HADDOCK score were refined in explicit solvent (itw) according to the default protocol and clustered (FCC method) at 6 Å RMSD cut-off to generate the final rank of clusters. The centroid of the best cluster (most negative average HADDOCK score) was used for further refinement with the clustENMD method [41] of the ProDy package [42]. Three replicates were run. The model was prepared by adding hydrogen atoms at pH 7 with pdbfixer (https://github.com/openmm/pdbfixer [43]). The sampling was performed applying normal modes of an anisotropic network model followed by clustering and MD refinement. Five generations were performed, using the first three global modes and generating 50 conformers with an average C_α RMSD distance of 1 Å from the parent conformer of the previous generation. At each generation, the maximum number of sampled clusters were increased by 20, from 20 (starting

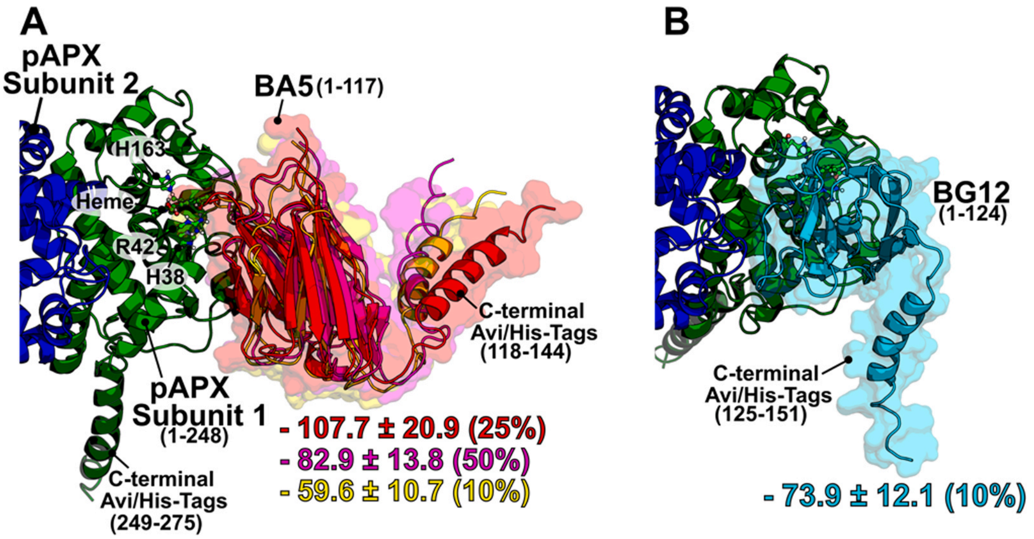


Fig. 1. Unsupervised prediction of pAPX in complex with BA5 and BG12. Cartoons reporting the visualization models for pAPX-BA5 (A) and pAPX-BG12 (B). The active site of pAPX is represented by ball-and-sticks with oxygen atoms in red, nitrogen in blue and polar hydrogens in white. Different regions are highlighted. The centroid of each cluster was visualized after superposing the complexes, using pAPX dimer as a reference, but only if there were at least two structures forming a cluster. The reported HADDOCK scores (more negative, higher affinity) are the average ± SD of all the cluster members. The figure was prepared with open-source PyMol 3.0 (https://github.com/schrodinger/pymol-open-source).

conformer) to 120 (5th generation). A short, classic MD simulation in NVT conditions (30°C) was performed after energy minimization of each sampled conformer, solvated in tip3p water molecules (10 Å padding) and neutralized by adding either Na^+ or Cl^- ions. OpenMM 7.7.0 [44] was used as an MD-simulation engine and the system was parameterized applying the amber ff14SB force field [45]. A time step of 2.0 fs was used; the number of MD steps were increased by 1000 over generations, starting from 3000 (starting conformer) to 8000 (5th generation). The sampled conformations were summarized by cluster analysis by means of the *gromos* method of Gromacs 2019.6 package [46], using backbone atom coordinates of BA5 and a RMSD threshold of 4 Å. The non-covalent interactions at tetrameric interfaces were predicted by ProLIF (https://github.com/chemosim-lab/ProLIF [47]), using standard ProLIF definitions of non-covalent interactions at 6 Å distance threshold. Only residues involved in interactions in at least 50 % of the samples were considered hotspot residues after binding energy prediction (see next paragraph).

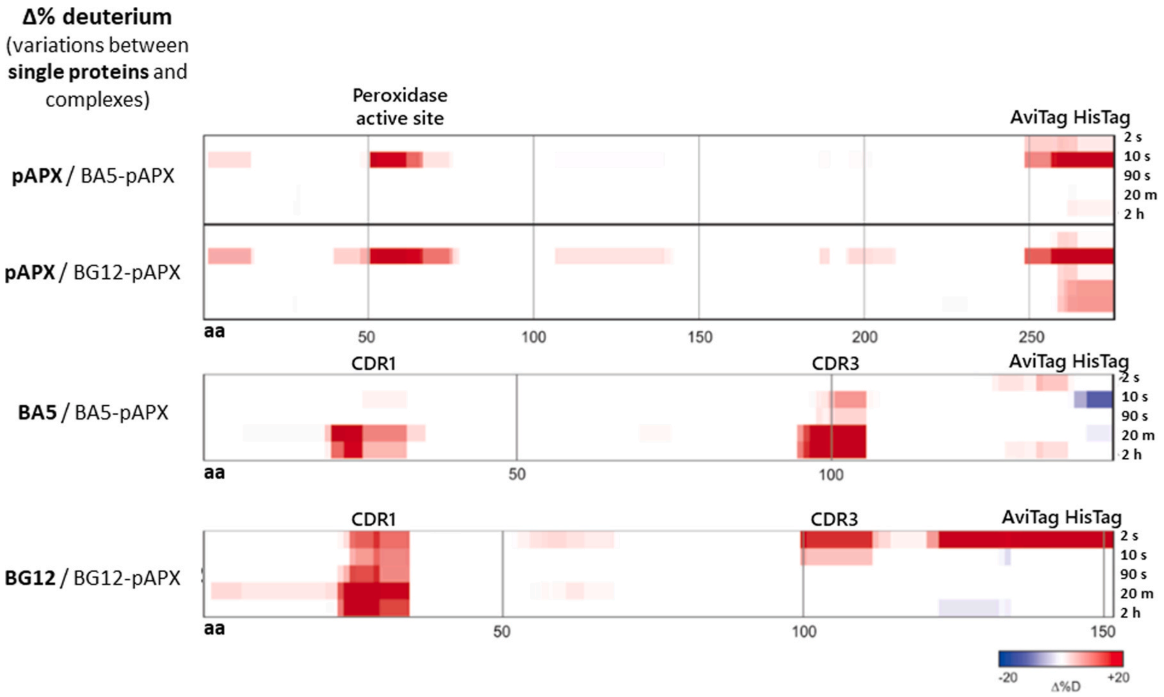


Fig. 2. HDX differential heat-map of deuterium exchange percentage of pAPX and nanobodies anti pAPX. Deuterium exchange was compared between free pAPX and pAPX complexed with anti-pAPX nanobodies. The levels of deuterium incorporation were assessed at 5 time points (2 s, 10 s, 1 min 30 s, 20 min, 2 h) in both antigen and nanobodies. The red colour represents the regions where the deuterium uptake of complexed protein is enhanced compared to free protein, whereas the blue colour represents the opposite trend.

2.7. In-silico prediction of BA5 hotspot interacting residues

The binding free energy (ΔG_{bind}) contribution of BA5 residues that participate the most to the interaction with pAPX model was calculated using the MM-GBSA method [48] of the AmberTools21 package [49], employing the per-residue free-energy decomposition protocol [50], as done in [13,21]. The single-trajectory approach was used on clustENMD conformers of the pAPX-BA5 complex, after tleap parametrization under the ff14SB force field [51]; the trajectories corresponding to the free proteins were extracted from that of the complexes. Two different generalized Born implicit-solvation models were used: i) GB^{OBC}, *GB= 5* option, setting *PBradii mbondi2;* ii) a modified GB^{HCT}, *GB= 8* option, setting *PBradii mbondi3* [52]. The accumulated mean value and standard deviation of ΔG_{bind} were obtained by averaging over clustENMD replicates and above reported solvation models. Hot-spot interacting residues were defined as those with an estimated average energy contribution of ≤ -2.0 kcal/mol.

3. Results and discussion

The interaction between anti-pAPX nanobodies BA5 and BG12 and pAPX has been proved and thoroughly described [11], but the structural characteristics of the complexes remained elusive. Therefore, we initially tried to compute those complexes in an unsupervised way, by leveraging a customized AlphaFold 2-based approach, using a more aggressive sampling and re-scoring by the HADDOCK physics-based scoring function. Results showed a different level of consistency between the nanobodies in terms of the prediction accuracy; indeed, despite in no case the iPTM score (useful to evaluate the confidence of the protein-protein interface predicted by AF2) was > 0.7, for BA5 ≈ 85 % of the predictions, grouped in three slightly different clusters, converged to a region proximal to the substrate opening (Fig. 1A), while for BG12 just 10 %, belonging to the same cluster were found (Fig. 1B). Moreover, the HADDOCK score of the best cluster was better for BA5, in agreement with the EC50 values previously estimated by ELISA assay [11]. Despite the approach provided initial indications, it was not capable to confidently dictate if the centroid of the best cluster was accurate enough and the interactions implied by the slightly different clusters were not conserved, hampering the reliable selection of mutants suitable for validation tests. Consequently, we designed an experimental approach to identify the epitope/paratope surfaces by applying two complementary techniques focusing on BA5, for which both previous experimental results and present predictions seemed to indicate being the best binder and, hence, the most appropriate candidate for a lead optimization campaign.

First, we performed HDX-MS experiments using the pAPX-Nb complexes detecting the H/D exchange profiles at five reaction times (2 s, 10 s, 90 s, 20 min, 2 h). Detailed experimental information is reported in Figs. S3-S8. We started with the identification of the pAPX sequence regions modified by the interaction with BA5 and BG12, this last nanobody used as an internal control. The difference of deuterium uptake existing between pAPX alone and in complex with each of the two selected nanobodies was quantified. In the HDX heat map (Fig. 2), the most evident differences were observed within the pAPX regions 38–77 and 249–275 when the reaction time was 10 sec (Fig. 2). The first region partially overlaps the peroxidase active site (residues 33–44, Fig. 2). The region 249 – 275 corresponds to the protein C-terminus and, specifically, to the tag (AviTag and HisTag) sequence that is present on the nanobodies as well. Its structure forms an alpha helix linked to a short loop that tends to dimerize [53,54].

Next, we exploited the same approach to map the nanobody paratopes. In this case, nanobody deuteration was measured when the binders were alone or in complex with pAPX. In the case of BA5, strong differences were identified in the residues between 20 and 37 and between 95 and 105, and they increased during the reaction (Fig. 2). These two regions roughly correspond to the hypervariable complementary-determining regions CDR1 and CDR3, respectively. Additionally, a weak signal was detected intermittently at the C-terminus of the construct (residues 127–144), corresponding to the tags. The HDX heat map (Fig. 2) showed that the weak signal detectable in correspondence of CDR3 at 10 s increased progressively in intensity at longer reaction times. The strong signal corresponding to CDR1 became evident only at the longest incubation time (Fig. 2). Our interpretation of the data has been that both CDR1 and CDR3 are involved in the interaction between BA5 and pAPX. Even though the modification of the CDR3 is the first visible event, it is difficult to conclude that it corresponds to the initial contact region, because both the affinity and the absolute exchange rate of the peptide can affect the HDX-MS data output. It is therefore possible only to speculate that a first binding event (CDR3-dependent) would have then induced a structural rearrangement favouring the interaction of the CDR1 domain with pAPX.

The preliminary prediction results indicated a different binding modality for BG12 and this was confirmed by the significant differences of deuterium uptake, identified in the sequences corresponding to the residues 22–34, 100–112 and 123–151 (Fig. 2). The first two sequences partially overlap to the CDR1 and CDR3 loops, while the third region corresponds to the tags (Fig. 2). The nanobody binding dynamic showed that the CDR3 is involved in the very early antigen-antibody recognition, but the signal progressively decreased. In contrast, the CDR1 exhibited high binding at any time during the reaction, but the trend was to a progressive reinforcement (Fig. 2).

Altogether, such analyses enabled to map the regions of pAPX, BA5 and BG12 in which the most significant differences in deuterium incorporation were evident (Table 1). The results suggested that for both complexes, pAPX / BA5 and pAPX / BG12, the nanobody hypervariable regions CDR1 and CDR3 played the major role in the interaction between pAPX and the corresponding nanobodies. CDR1 and CDR3 of single-domain antibodies are highly flexible and can undergo significant conformational changes upon antigen binding [55]. This dynamic structural rearrangement allows the optimization of the contacts with the antigen and stabilize such interaction. This first data set was informative to experimentally confirm that the BA5 binding site on pAPX is likely close to the opening region between the bulk and the active site, but we also measured signals likely due to flexible regions. This confirmed the data obtained from the unsupervised modelling approach but did not allow the identification of a preferential cluster to use for validation. Therefore, we looked for orthogonal experimental data to recover by means of XL-MS and necessary to guide towards a better refined computational model.

XL-MS was applied using two alternative MS-cleavable urea-based cross linkers [56], DSPU and DSBU, which differ because of the spacer length connecting two reactive residues (9.8 and 12.5 Å, respectively). Both linkers react with primary amines and should theoretically connect lysine residues, although some reactions with hydroxyl groups are possible. The binding promoted by such cross-linkers will involve residues belonging to nearby protein portions. Labile bonds of the CID-MS/MS (collision-induced dissociation in tandem mass spectrometry) cleavable cross-linkers yielded a characteristic fragmentation pattern in MS^2 spectra, allowing highly accurate and reliable identification of the corresponding peptides in which they are incorporated. After having covalently linked pAPX with BA5, the downstream XL-MS analysis identified several interprotein contacts, three of them labelled with both cross-linking agents (Table S1). Specifically, K31 of pAPX was crosslinked to S31 and K66 of BA5, while the latter was also crosslinked to K122 of pAPX. Interestingly, these residues cover in part the region of the pAPX active site (33 – 44), and are in agreement with the HDX-MS results reported above (Table 1).

HDX-MS and XL-MS are mostly used independently to evaluate biological structures. HDX-MS is primarily applied to collect information about protein dynamics, while XL-MS is used to define constraints at protein-protein binding interfaces [57,58]. It must be underlined that XL-MS experimental technical requirements, such as the availability of

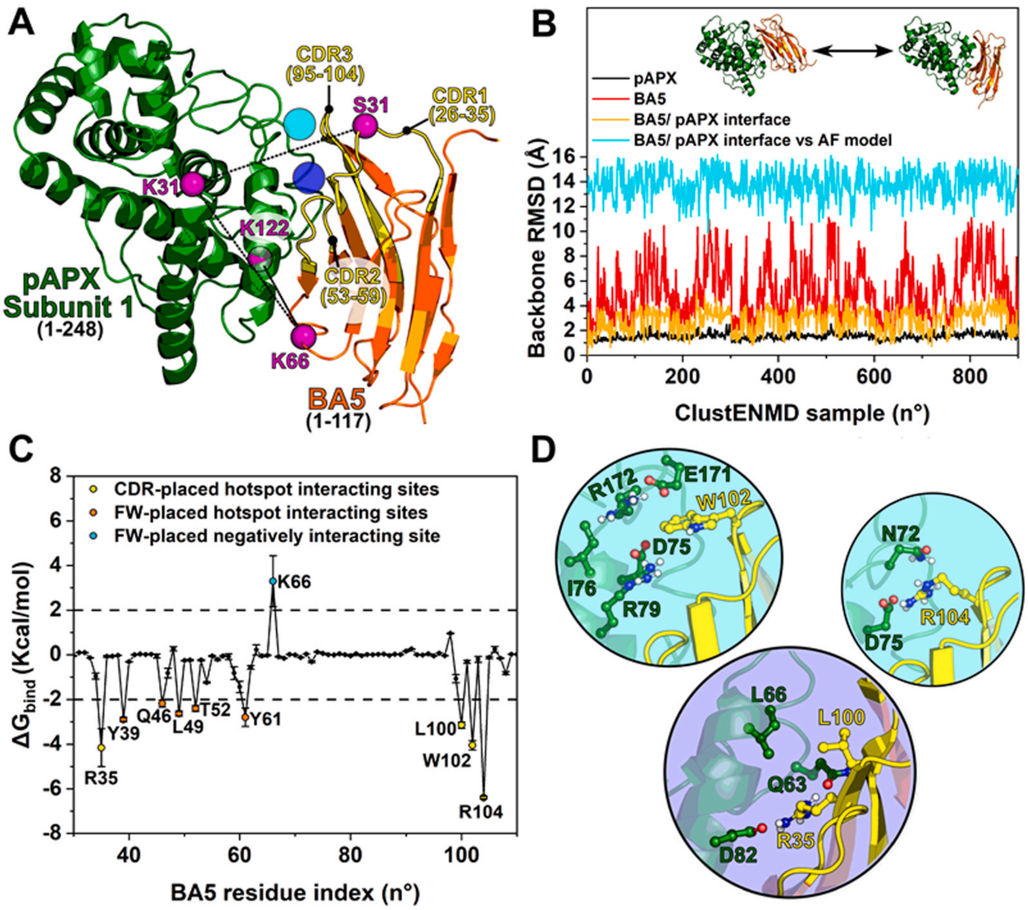


Fig. 3. Integrative modelling and dynamics sampling of pAPX-BA5 complex. (A) Cartoon reproducing the centroid of the best HADDOCK cluster of pAPX (single subunit)-BA5 complex, obtained by leveraging experimental constraints and used for ProDy sampling. The three XLs used as constraints are highlighted as magenta spheres connected by dashed lines. The cyan and blue circles represent the interfacial region involving CDRs. (B) Backbone atoms RMSD statistics along the ProDy sampled conformers. "pAPX" and "BA5" RMSDs are calculated by fitting the sampled frames along the pAPX protein of the starting HADDOCK centroid; "BA5/ pAPX interface" refers to the RMSD calculated after fitting the sampled frames along the interfacial regions and taking the starting HADDOCK centroid of the best cluster as the reference. "BA5/ pAPX interface vs AF model" differs since the reference structure is the centroid of the best HADDOCK cluster from unsupervised modelling (Fig. 1A). (C) Average per-residue ΔG of binding (ΔG_{bind}) from the per-residue decomposition of MM-GBSA binding energies, calculated over all the ProDy sampled frames and two different solvation models. Dashed lines indicate the threshold used for defining hotspot residues involved in the interaction. (D) Zoom on the main pAPX-BA5 interactions (cyan and blue circles reported in panel A) involving the hotspot residues located on BA5 CDRs. Residues are represented in ball-and-sticks with oxygen atoms in red, nitrogen in blue and polar hydrogens in white.

cross-linkable reactive amino acids and digestion sites and the impossibility to inform about the exact distance between cross-linked residues [59], differ from those specific to HDX-MS that cannot distinguish between residues involved in the direct binding and those undergoing conformational or allosteric effects [60]. XL-MS identifies spatially proximate residues that are cross-linked, but that can be minimally involved in the actual interactions because longer the peptides generated by proteolysis, lower the detail, whereas HDX-MS can help in determining protein orientation and conformation, providing complementary structural interpretation. However, the combination of the two techniques compensates for each method limitations because, when applied together, they can provide complementary information. Nevertheless, in literature there are only few studies where HDX-MS and XL-MS are combined and integrated in a single approach overcoming their respective limitations. This is the case of the characterization of the binding interface between interleukin 7 and its receptor [25,61].

In our case, the combination of the two sets of experimental data provided the most suitable information to guide effective docking, with the centroid of the best cluster having a HADDOCK-score of −227.6 ± 21 and binding at the margin of the active site opening lid loop (Fig. 3A). The model respected all the constrains even after the dynamic refinement (Fig. S2). The docking attempts performed with single experimental sets resulted in worse HADDOCK-scores (-165 ± 10.2 and −61.9 ± 3.9) and a binding mode at > 6 Å backbone RMSD (Fig. S9). Specifically, the exclusive use of XL-MS data led to a model in which BA5 bound to the other side of the active site opening lid loop, as this minimizes XL distances, while HDX-MS-only data resulted in a model not compatible with the XL-MS data restraints. The model obtained using both constraints was therefore selected for further analyses because more reliable with respect to the information recovered by either unconstrained deep learning docking approach or by employing single sets of experimental data. The RMSD of interfacial backbone atoms across sampled ProDy dynamics replicates (Fig. 3B) indicated that the binding interface formed by paratope and epitope residues, identified as > 50 % of ProDy sample (Table S2) and having an estimated $\Delta G_{bind} < -2$ Kcal/mol (Fig. 3C), is 13.7 ± 1.4 Å) apart from the unsupervised modelled counterparts. This means that the experimental-guided model provides a different binding mode. Despite a general conservation of the interfacial residues over the samples (2.8 ± 0.9 Å RMSD displacement), there is a noteworthy dynamic of the BA5 backbone coordinates (5.3 ± 2.3 Å), when fitting all the sampled frames using only pAPX backbone atoms for structural alignment. Such dynamics is due to the slight oscillation of pAPX between a closed and an open state for its substrate-binding opening during ProDy sampling, while the nanobody keeps conserved

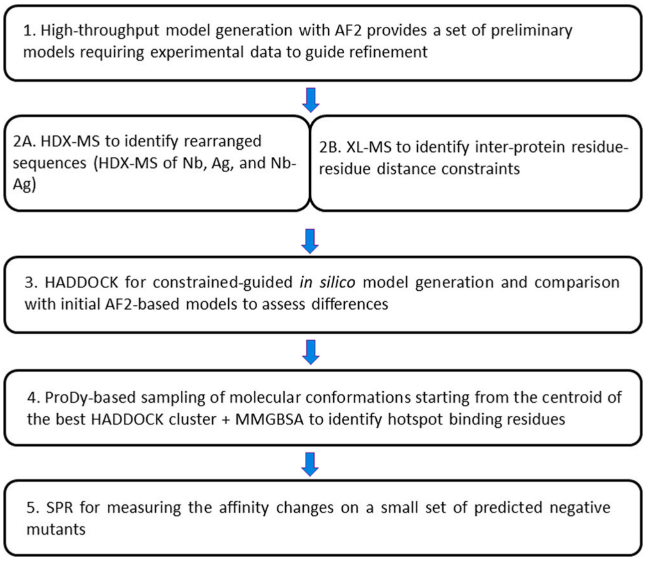
Table 2
Peroxidase activity of pAPX in presence of BA5.

[pAPX] (μM)	[BA5] (μM)	ΔAbs_{450}
5	1.25	1.432
5	2.5	1.618
5	5	1.546

pAPX activity was measured in the presence of increasing BA5 concentrations. TMB (3,3′,5,5′-tetramethylbenzidine) was used as the reaction substrate and pAPX activity was inferred by measuring the ΔAbs_{450} with respect to the control reaction performed in the absence of the enzyme.

interactions to this dynamic region (Fig. 3B). This observation suggests that BA5 binding mode might affect the dynamics of the enzyme active-site opening, and these modifications could explain the substantial variations of measured deuterium exchange corresponding to the active-site buried regions that are evidenced in the HDX-MS heatmap (Fig. 2). Therefore, we also checked if BA5 could affect the pAPX peroxidase activity at increasing concentrations but the experimental results (Table 2) suggest no effect. From the practical point of view, this indicates that enzyme quantification using BA5 as a reagent should be compatible with the contemporary measurement of pAPX activity on the same sample.

The analysis of the estimated binding energies decomposed over BA5 residues (Fig. 3C) indicated predicted interaction hotspots in both the FW and CDR regions. BA5 paratope includes four FW2 residues (Y39, Q46, L49, T52), one residue from CDR1 (R35) and three residues from CDR3 (L100, W102, R104). The residues in CDR1 and CDR3 of BA5, that according to the HDX-MS data are responsible for the binding to the antigen, interact with an epitope involving non-continuous pAPX residues (Fig. 3D). Specifically, R35 (BA5 CDR1) forms a salt bridge with pAPX D82 and a hydrogen bond with pAPX Q63; L100 (BA5 CDR3) makes stacking with pAPX Q63 and forms a partially stable hydrophobic contact with pAPX L66; W102 (BA5 CDR3) forms a hydrophobic contact with pAPX I76 and makes π-stacking with two pAPX salt bridges (E171-R172 and D75-R79); R104 (BA5 CDR3) forms a salt bridge with pAPX D75 and a hydrogen bond with pAPX N72. The validation of the model was obtained by mutating into alanine (both singularly and all three together) three of these CDR residues (R35, W102, R104), selected for being those with the lower predicted ΔG_{bind}. Nect, the evaluation was performed comparing the dissociation constants (K_D) of wild type and mutants by SPR. The wild-type protein displayed a complex binding profile, inconsistent with a simple 1:1 interaction, featuring a rapid initial association followed by a slower phase and an early rapid dissociation followed by a very slow, incomplete dissociation (Fig. 4). Such pattern suggests a multi-step binding process or higher-order complex formation (Fig. 4 and Fig. S10A) consistent with the proposed two-state binding model mechanism characterized by an initial recognition step followed by a stabilising conformational change. Affinity fitting was calculated in the low micromolar range (2.8 μM, Fig. S10B). The mutants exhibited extremely weak binding, as shown in Fig. 4, insufficient for reliable affinity measurement, but clearly indicating a significant reduction in affinity compared to wild-type. Overall, the SPR data validated *in-silico* model predictions confirming that both computationally identified sites are essential for the interaction. It must be underlined that the dissociation constant measured by SPR (2.8 μM) resulted significantly worse than the EC50 value (0.88 μM) measured previously by ELISA exploiting fluorescent nanobodies [11]. Therefore, whereas the comparison of data collected using the same methodological setting provides reliable information, such result confirms that the absolute binding values are highly dependent on the chosen technology


1. High-throughput model generation with AF2 provides a set of preliminary models requiring experimental data to guide refinement

2A. HDX-MS to identify rearranged sequences (HDX-MS of Nb, Ag, and Nb-Ag)

2B. XL-MS to identify inter-protein residue-residue distance constraints

3. HADDOCK for constrained-guided *in silico* model generation and comparison with initial AF2-based models to assess differences

4. ProDy-based sampling of molecular conformations starting from the centroid of the best HADDOCK cluster + MMGBSA to identify hotspot binding residues

5. SPR for measuring the affinity changes on a small set of predicted negative mutants

Fig. 5. Schematic representation of the integrative structural biology workflow adopted in this work. The figure summarizes the steps of the modelling process, indicating the adopted technique and the corresponding information gain.

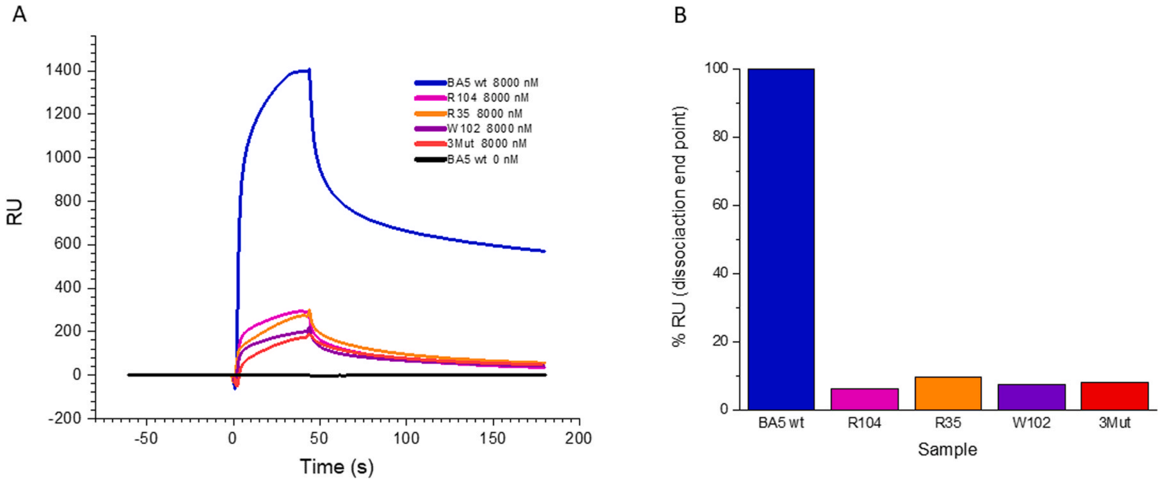


Fig. 4. Evaluation of the binding characteristics of wild type and mutants of BA5 by SPR (A)Sensorgrams of BA5 wild type (wt) and mutants R104, R35, W102, and 3Mut are shown. The sensograms correspond to the signal difference between flow cell 2 and flow cell 1 (2−1). This subtraction corrects for background and non-specific signals, isolating specific binding events. All measurements were performed at an analyte concentration of 8000 nM. (B)Corresponding RU differences (dissociation at the end point) in percentage.

[62].

4. Conclusion

An initial unconstrained computational model of the complex between pAPX and two nanobodies indicated that BA5 had superior binding potential than BG12. By combining HDX-MS and XL-MS experimental data, it was possible to refine the pAPX-BA5 complex prediction with a guided-docking approach that was validated with negative mutants (Fig. 5): this indicates the importance of integrative modelling, as both the unconstrained prediction approach and the use of either only HDX-MS or only XL-MS data sets would have provided different and less accurate models. These results represent an excellent example of a method suitable to decipher protein complex structural information that is alternative to conventional approaches such as X-ray crystallography or NMR analyses, but it can also integrate insufficient data sets obtained by conventional structural methods [21,25]. At the same time, this is also an example of how deep learning methods for structure prediction may be not sufficiently reliable when applied to class of biomolecules, like nanobodies, that are hypervariable in interacting regions, highly diverse in CDR3 length and composition and poorly represented in terms of available 3D structures in complex with diverse targets [63,64]. Finally, the defined model will contribute to nanobody engineering aimed at obtaining variants characterized by improved biophysical properties that will improve the accuracy of plant APX diagnostic devices. The only real limit we found in such approach is that it requires a mix of expertise and technologies that might be difficult to have in the same place. Nevertheless, there are several support programmes facilitating the access to specialized facilities, as the ones we successfully used.

CRediT authorship contribution statement

Ario de Marco: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Funding acquisition, Conceptualization. **Kristina Eleršič Filipič:** Writing – review & editing, Investigation, Formal analysis. **D'Ercole Claudia:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Marco Orlando:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration

The Authors declare that the proposed work has not been published, that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

Conflict of interests

The authors declare no conflict of interests

Acknowledgments

The authors wish to thank the Structural Mass Spectrometry Core Facility of CIISB, Instruct-CZ Centre, supported by MEYS CR (LM2023042) and European Regional Development Fund-Project „Innovation of Czech Infrastructure for Integrative Structural Biology" (No. CZ.02.01.01/00/23_015/0008175). This research was supported by the grants P4–0107 provided by the Javne agencije za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2025.11.024.

References

[1] Pan Y, Birdsey RA, Phillips OL, Houghton RA, Fang J, Kauppi PE, et al. The enduring world forest carbon sink. Nature 2024;631:563–9. https://doi.org/10.1038/s41586-024-07602-x.

[2] Yoshimura K, Ishikawa T. Physiological function and regulation of ascorbate peroxidase isoforms. J Exp Bot 2024;75:2700–15. https://doi.org/10.1093/jxb/erae061.

[3] de Marco A. Methodologies for the isolation of alternative binders with improved clinical potentiality over conventional antibodies. Crit Rev Biotechnol 2013;33:40–8. https://doi.org/10.3109/07388551.2012.665353.

[4] Pardon E, Laeremans T, Triest S, Rasmussen SG, Wohlkönig A, Ruf A, et al. A general protocol for the generation of Nanobodies for structural biology. Nat Protoc 2014;9:674–93. https://doi.org/10.1038/nprot.2014.039.

[5] Traenkle B, Rothbauer U. Under the microscope: single-domain antibodies for live-cell imaging and super-resolution microscopy. Front Immunol 2017;24(8):1030. https://doi.org/10.3389/fimmu.2017.01030.

[6] Lecocq Q, De Vlaeminck Y, Hanssens H, D`Huyvetter M, Raes G, Goyvaerts C, et al. Theranostics in immuno-oncology using nanobody derivatives. Theranostics 2019;9:7772–91. https://doi.org/10.7150/thno.34941.

[7] Jovčevska I, Muyldermans S. The therapeutic potential of nanobodies. BioDrugs 2020;34:11–26. https://doi.org/10.1007/s40259-019-00392-z.

[8] Wu X, Rapoport TA. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). Proc Natl Acad Sci USA 2021;118:e2115001118. https://doi.org/10.1073/pnas.2115001118.

[9] Neumair J, D'Ercole C, De March M, Elsner M, Seidel M, de Marco A. Macroporous epoxy-based monoliths functionalized with Anti-CD63 nanobodies for effective isolation of extracellular vesicles in urine. Int J Mol Sci 2023;24:6131. https://doi.org/10.3390/ijms24076131.

[10] de Marco A. Recombinant expression of nanobodies and nanobody-derived immunoreagents. Protein Expr Purif 2020;172:105645. https://doi.org/10.1016/j.pep.2020.105645.

[11] D'Ercole C, Svigelj R, Mrak T, de Marco A. Early stress detection in forest trees using a nanobody-functionalized electrochemical biosensor for ascorbate peroxidase. Plant Stress 2025;16:100844. https://doi.org/10.1016/j.stress.2025.100844.

[12] Anjum NA, Sharma P, Gill SS, Hasanuzzaman M, Khan EA, Kachhap K, et al. Catalase and ascorbate peroxidase-representative H2O2-detoxifying heme enzymes in plants. Environ Sci Pollut Res Int 2016;23(19):19002–29. https://doi.org/10.1007/s11356-016-7309-6.

[13] Orlando M, Fortuna S, Oloketuyi S, Bajc G, Goldenzweig A, de Marco A. CDR1 composition can affect nanobody recombinant expression yields. Biomolecules 2021;11:1362. https://doi.org/10.3390/biom11091362.

[14] Levisson M, Spruijt RB, Winkel IN, Kengen SW, van der Oost J. Phage display of engineered binding proteins. Methods Mol Biol 2014;1129:211–29. https://doi.org/10.1007/978-1-62703-977-2_19.

[15] Peleg Y, Vincentelli R, Collins BM, Chen KE, Livingstone EK, Weeratunga S, et al. Community-wide experimental evaluation of the PROSS stability-design method. J Mol Biol 2021;433:166964. https://doi.org/10.1016/j.jmb.2021.166964.

[16] Chu AE, Lu T, Huang PS. Sparks of function by de novo protein design. Nat Biotechnol 2024;42:203–15. https://doi.org/10.1038/s41587-024-02133-2.

[17] El Salamouni NS, Cater JH, Spenkelink LM, Yu H. Nanobody engineering: computational modelling and design for biomedical and therapeutic applications. FEBS Open Bio 2025;15:236–53. https://doi.org/10.1002/2211-5463.13850.

[18] Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w.

[19] Unsal S, Holland B, Sardag I, Timucin E. Assessing the performance of AF2 and AF3-implementations on antibody-antigen complexes. bioRxiv 2025;07(25):666870. https://doi.org/10.1101/2025.07.25.666870.

[20] Gordon GL, Greenshields-Watson A, Agarwal P, Wong A, Boyles F, Hummer A, et al. PLAbDab-nano: a database of camelid and shark nanobodies from patents and literature. Nucleic Acids Res 2025;53:D535–42. https://doi.org/10.1093/nar/gkae881.

[21] Ubbiali D, Orlando M, Kovačič M, Iacobucci C, Semrau MS, Bajc G, et al. An anti-HER2 nanobody binds to its antigen HER2 via two independent paratopes. Int J Biol Macromol 2021;182:502–11. https://doi.org/10.1016/j.ijbiomac.2021.04.032.

[22] Zhang Q, Yang J, Bautista J, Badithe A, Olson W, Liu Y. Epitope mapping by HDX-MS elucidates the surface coverage of antigens associated with high blocking efficiency of antibodies to birch pollen allergen. Anal Chem 2018;90:11315–23. https://doi.org/10.1021/acs.analchem.8b01864.

[23] Ständer S, R. Grauslund L, Scarselli M, Norais N, Rand K. Epitope mapping of polyclonal antibodies by hydrogen-deuterium exchange mass spectrometry (HDX-MS). Anal Chem 2021;93:11669–78. https://doi.org/10.1021/acs.analchem.1c00696.

[24] Tran MH, Schoeder CT, Schey KL, Meiler J. Computational structure prediction for antibody-antigen complexes from hydrogen-deuterium exchange mass

spectrometry: challenges and outlook. Front Immunol 2022;13:859964. https://doi.org/10.3389/fimmu.2022.859964.

[25] Zhang MM, Huang RY, Beno BR, Deyanova EG, Li J, Chen G, Gross ML. Epitope and paratope mapping of PD-1/Nivolumab by mass spectrometry-based hydrogen-deuterium exchange, cross-linking, and molecular docking. Anal Chem 2020;92:9086–94. https://doi.org/10.1021/acs.analchem.0c01291.

[26] Petrotchenko EV, Nascimento EM, Witt JM, Borchers CH. Determination of protein monoclonal-antibody epitopes by a combination of structural proteomics methods. J Proteome Res 2023;22:3096–102. https://doi.org/10.1021/acs.jproteome.3c00159.

[27] Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods 2022;19:679–82. https://doi.org/10.1038/s41592-022-01488-1.

[28] Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. bioRxiv 2021;10(04):463034. https://doi.org/10.1101/2021.10.04.463034.

[29] Monteiro da Silva G, Cui JY, Dalgarno DC, Lisi GP, Rubenstein BM. High-throughput prediction of protein conformational distributions with subsampled AlphaFold2. Nat Commun 2024;15:2464. https://doi.org/10.1038/s41467-024-46715-9.

[30] Giulini M, Reys V, Teixeira JMC, Jiménez-García B, V Honorato R, Kravchenko A, et al. HADDOCK3: a modular and versatile platform for integrative modeling of biomolecular complexes. J Chem Inf Model 2025;65:7315–24. https://doi.org/10.1021/acs.jcim.5c00969.

[31] Trcka F, Durech M, Vankova P, et al. Human stress-inducible Hsp70 has a high propensity to form ATP-dependent antiparallel dimers that are differentially regulated by cochaperone binding. Mol Cell Proteom 2019;18:320–37. https://doi.org/10.1074/mcp.RA118.001044.

[32] Vankova P, Salido E, Timson DJ, Man P, Pey AL. A dynamic core in human NQO1 controls the functional and stability effects of ligand binding and their communication across the enzyme dimer. Biomolecules 2019;9:728. https://doi.org/10.3390/biom9110728.

[33] Weis DD. Recommendations for the propagation of uncertainty in hydrogen exchange-mass spectrometric measurements. J Am Soc Mass Spectrom 2021;32:1610–7. https://doi.org/10.1021/jasms.0c00475.

[34] Perez-Riverol Y, Bai J, Bandla C, et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res 2022;50:D543–52. https://doi.org/10.1093/nar/gkab1038.

[35] Götze M, Pettelkau J, Fritzsche R, Ihling CH, Schäfer M, Sinz A. Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. J Am Soc Mass Spectrom 2015;26:83–97. https://doi.org/10.1007/s13361-014-1001-1.

[36] Lee JW, Won JH, Jeon S, Choo Y, Yeon Y, Oh JS. DeepFold: enhancing protein structure prediction through optimized loss functions, improved template features, and re-optimized energy function. Bioinformatics 2023;39:btad762. https://doi.org/10.1093/bioinformatics/btad762.

[37] Zhang Y., Zhang Z., Zhong B., Misra S., Tang J. DiffPack: A Torsional Diffusion Model for Autoregressive Protein Side-Chain Packing. 2024 arXiv:2306.01794v2 [q-bio.QM]. https://doi.org/10.48550/arXiv.2306.01794.

[38] Honorato RV, Trellet ME, Jiménez-García B, Schaarschmidt JJ, Giulini M, Reys V, et al. The HADDOCK2.4 web server for integrative modeling of biomolecular complexes. Nat Protoc 2024;19:3219–41. https://doi.org/10.1038/s41596-024-01011-0.

[39] Piersimoni L, Sinz A. Cross-linking/mass spectrometry at the crossroads. Anal Bioanal Chem 2020;412:5981–7. https://doi.org/10.1007/s00216-020-02700-x.

[40] de Vries SJ, Bonvin AM. CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. PLoS One 2011;6:e17695. https://doi.org/10.1371/journal.pone.0017695.

[41] Kaynak BT, Zhang S, Bahar I, Doruker P. ClustENMD: efficient sampling of biomolecular conformational space at atomic resolution. Bioinformatics 2021;37:3956–8. https://doi.org/10.1093/bioinformatics/btab496.

[42] Bakan A, Meireles LM, Bahar I. ProDy: protein dynamics inferred from theory and experiments. Bioinformatics 2011;27:1575–7. https://doi.org/10.1093/bioinformatics/btr168.

[43] 〈https://github.com/openmm/pdbfixer〉.

[44] Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, et al. OpenMM 7: rapid development of high performance algorithms for molecular dynamics. PLoS Comput Biol 2017;13:e1005659. https://doi.org/10.1371/journal.pcbi.1005659.

[45] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 2015;11:3696–713. https://doi.org/10.1021/acs.jctc.5b00255.

[46] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1-2:19–25. https://doi.org/10.1016/j.softx.2015.06.001.

[47] Bouysset C, Fiorucci S. ProLIF: a library to encode molecular interactions as fingerprints. J Chemin 2021;13:72. https://doi.org/10.1186/s13321-021-00548-6.

[48] Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 2015;10:449–61. https://doi.org/10.1517/17460441.2015.1032936.

[49] Case DA, Aktulga HM, Belfon K, Cerutti DS, Cisneros GA, Cruzeiro VWD, et al. AmberTools. J Chem Inf Model 2023;63:6183–91. https://doi.org/10.1021/acs.jcim.3c01153.

[50] Miller 3rd BR, McGee Jr TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA.py: an efficient program for end-state free energy calculations. J Chem Theory Comput 2012;8:3314–21. https://doi.org/10.1021/ct300418h.

[51] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 2015;11:3696–713. https://doi.org/10.1021/acs.jctc.5b00255.

[52] Tsui V, Case DA. Theory and applications of the generalized Born solvation model in macromolecular simulations. Biopolymers 2000;56:275–91. https://doi.org/10.1002/1097-0282(2000)56:4<275::AID-BIP10024>3.0.CO;2-E.

[53] Veggiani G, Giabbai B, Semrau MS, Medagli B, Riccio V, Bajc G, et al. Comparative analysis of fusion tags used to functionalize recombinant antibodies. Protein Expr Purif 2020;166:105505. https://doi.org/10.1016/j.pep.2019.105505.

[54] Sánchez-García L, Voltà-Durán E, Parladé E, Mazzega E, Sánchez-Chardi A, Serna N, et al. Self-assembled nanobodies as selectively targeted, nanostructured, and multivalent materials. ACS Appl Mater Interfaces 2021;13:29406–15. https://doi.org/10.1021/acsami.1c08092.

[55] Stanfield RL, Dooley H, Verdino P, Flajnik MF, Wilson IA. Maturation of shark single-domain (IgNAR) antibodies: evidence for induced-fit binding. J Mol Biol 2007;367:358–72. https://doi.org/10.1016/j.jmb.2006.12.045.

[56] Iacobucci C, Götze M, Ihling CH, et al. A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein-protein interactions. Nat Protoc 2018;13:2864–89. https://doi.org/10.1038/s41596-018-0068-8.

[57] Zheng J, Corzo C, Chang MR, Shang J, Lam VQ, Brust R, et al. Chemical crosslinking mass spectrometry reveals the conformational landscape of the activation helix of PPARγ; a model for ligand-dependent antagonism. Structure 2018;26:1431–1439.e6. https://doi.org/10.1016/j.str.2018.07.007.

[58] Mysling S, Kristensen KK, Larsson M, Beigneux AP, Gårdsvoll H, Fong LG, et al. The acidic domain of the endothelial membrane protein GPIHBP1 stabilizes lipoprotein lipase activity by preventing unfolding of its catalytic domain. Elife 2016;5:e12095. https://doi.org/10.7554/eLife.12095.

[59] Wei H, Mo J, Tao L, Russell RJ, Tymiak AA, Chen G, et al. Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications. Drug Discov Today 2014;19:95–102.

[60] Piersimoni L, Kastritis PL, Arlt C, Sinz A. Cross-linking mass spectrometry for investigating protein conformations and protein–protein interactions—a method for all seasons. Chem Rev 2022;122:7500–31.

[61] Zhang MM, Beno BR, Huang RY, Adhikari J, Deyanova EG, Li J, et al. An integrated approach for determining a protein–protein binding interface in solution and an evaluation of hydrogen–deuterium exchange kinetics for adjudicating candidate docking models. Anal Chem 2019;91:15709–17.

[62] Rich RL, Papalia GA, Flynn PJ, et al. A global benchmark study using affinity-based biosensors. Anal Biochem 2009;386:194–216.

[63] Valdés-Tresanco MS, Valdés-Tresanco ME, Jiménez-Gutiérrez DE, Moreno E. Structural modeling of nanobodies: a benchmark of state-of-the-art artificial intelligence programs. Molecules 2023;28:3991.

[64] Zhu H, Ding Y. Nanobodies: from discovery to AI-driven design. Biology 2025;14:547.